

TR-87-006

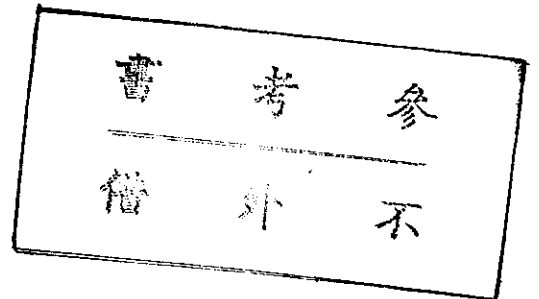
人 像 識 別

中央研究院資訊科學研究所

黃俊雄 · 陳漢明 · 曾棟炫、鍾馬龍

李東翰 · 莊一仁 · 譚志宏 · 王秋鳳

民國七十六年一月



中研院資訊所圖書室



3 0330 03 000066 0



0066

摘 要

人像識別在心理學系等實務機構皆有大量的研究，但用電腦來做此方面的研究在國內這是第一次。原因是人臉可偽裝，可變形，臉部可變黑、發白、流汗等以致要叫笨電腦來識別人臉是做不到滿意的可靠程度。本文是把人臉照像輸入電腦後，嘗試把人像上的一些特徵取出來。這些特徵是眉毛、眼睛、鼻子、嘴巴及臉之輪廓。再利用這些特徵的分類（比如眉毛可分成八字形、倒八字形、平形三類；眼睛可分成細眼、中眼、大眼等）來將人像分類。本文提供充分詳細的理論及方法。實驗結果並不理想，原因是不同人臉的影像數據變化太不規則。比如有些人頭髮與眉毛連在一起，因此在影像內要自動把眉毛切出來是不可能的。又有時嘴巴與臉的灰度值差不了多少，加上相機的雜訊大而使嘴巴切不出來。我們試圖用 HOUGH TRANSFORM 來找出臉輪廓也沒成功。成功的是眼睛及鼻子。如果相機改良好點，成功率可提高很多。

一 前言：

人像識別的研究在心理系等實務機構都有大量的研究。但用電腦來做此方面的研究是非常的困難，這牽涉到人工智慧最深奧的部份。在國內本研究是第一次嘗試此方面的試驗。起初我們是不抱很大希望能成功識別不同的人，原因是人臉可偽裝、變形，臉部可變黑、發白、流汗等以致要叫笨電腦來做識別是幾乎做不到的。因此本研究與其說是要發展識別系統倒不如說是開發一些新技術新原理。

利用電腦來做人面識別比較可行的方法是側面照像找出側面輪廓（見 Harmon [1 - 2] ），這種情形鼻子、嘴巴等之側面輪廓最為重要，可惜這地方常會變動（比如人工美容）。另外 Harmon 並沒做影像處理把輪廓找出來而是利用畫家把輪廓畫出來再輸入電腦內。探討人類心理上識別人面的各種因素曾由許多學者研究過，重要的參考資料列在 [3 - 6] 。本研究試圖從人像上取出一些特徵如眉毛、眼睛、鼻子等來加以分類（即做識別工作），這些特徵的各種類型主要是參考 Goldstein and Harmon 之文章 [4] 上之詳細分類。經過一年多的實驗照了三十張人像，發現要利用電腦自動抽取這些特徵並不是常會成功的。分析下來可歸成下面幾個原因：(1) 影像感測器雜訊太大。(2) 燈光不穩。(3) 不同的人臉變化太大不易控制在一定情況下作業，比如頭髮

會與眉毛混在一起，如此眉毛就不易找出來。(4)有些特徵根本不能用影像處理技術來取出來，比如嘴唇的顏色淺一點而與皮膚一致則就難以取出，又臉的輪廓毫無希望可取出來。

三 人像特徵：

由 Goldstein and Harmon [4] 之文章裡十一種主要的識別人臉的特徵被用來做測驗如圖 1 所示。這裏尚缺乏人臉形狀的特徵，這項遺漏太嚴重，因為衆所週知我們一看別人人臉第一眼就看出他是瘦了或胖了，而此印象是比其他特徵還深刻。人臉形狀可分成方形、蛋形、三角形、圓形。這一特徵是列爲人像識別的重要特徵之一。

在上面所提十一種特徵中我們只取五種特徵來做，這五種是：

- (1) 頭髮：分成黑、白、密、疏四類；
- (2) 眼睛：分成大、中、細三類；
- (3) 眉毛：分成濃、細、倒八字形、八字形、平形五類；
- (4) 鼻子：分成寬、窄、中、長、短五類；
- (5) 嘴巴：分成大、中、小、細唇、厚唇五類。

另外加入：(6) 臉之輪廓：方形、蛋形、三角形、圓形四類。

還有由人像面上亮度的變化可計算臉頰凸的情形及凹的地方。以上所訂的尺度皆是以眼珠的寬度爲標準。照進

WEAR	1	2	3	4	5	Col.
a. Face Shape	Square	Round	Oval	Long		20
HAIR	1	2	3	4	5	21
a. Coverage	Full	-	Receding	-	Bald	
b. Length	1	2	3	4	5	22
	Short	-	Average	-	Long	
c. Texture	1	2	3	4	5	23
	Straight	-	Wavy	-	Kinky	
d. Part	1	2	3	4		24
	Left	Middle	Right	None		
e. Shade	1	2	3	4	5	25
	Dark	Medium	Light	Gray	White	
EYEBROWS	1	2	3	4	5	26
a. Weight	Thin	-	Medium	-	Bushy	
b. Separation	1	2	3			27
	Separated	-	Meeting			
EYES	1	2	3	4	5	28
a. Opening	Slit	-	Medium	-	Wide	
b. Spacing	1	2	3	4	5	29
	Close	-	Medium	-	Wide	
c. Shade	1	2	3	4	5	30
	Light	-	Medium	-	Dark	
NOSE	1	2	3	4	5	31
a. Length	Short	-	Medium	-	Long	
b. Nostrils	1	2	3	4	5	32
	Narrow	-	Medium	-	Flared	
c. Tip	1	2	3	4	5	33
	Upward	-	Vertical	-	Downward	
d. Profile	1	2	3	4	5	34
	Concave	-	Straight	-	Hooked	
e. Projection	1	2	3	4	5	35
	Small	-	Medium	-	Large	
6. MOUTH						
a. Lip Thickness						
UPPER	1	2	3	4	5	36
	Thin	-	Medium	-	Thick	
LOWER	1	2	3	4	5	37
	Thin	-	Medium	-	Thick	
b. Lip Overlap	1	2	3			38
	Upper	Neither	Lower			
c. Width	1	2	3	4	5	39
	Small	-	Medium	-	Large	
d. Nose-to-Mouth Distance	1	2	3	4	5	40
	Short	-	Medium	-	Long	
e. Mouth-to-Chin Distance	1	2	3	4	5	41
	Short	-	Medium	-	Long	
7. CHIN						
a. Front	1	2	3	4	5	42
	Narrow	-	Medium	-	Broad	
b. Profile	1	2	3	4	5	43
	Receding	-	Straight	-	Butting	
8. EARS						
a. Length	1	2	3	4	5	44
	Short	-	Medium	-	Long	
b. Protrusion	1	2	3	4	5	45
	Flat	-	Medium	-	Sticking Out	
c. Lobes	1	2	3	4	5	46
	Attached	-	Medium	-	Not Attached	
9. CHEEKS	1	2	3	4	5	47
	Sunken	-	Average	-	Full	
10. FOREHEAD	1	2	3	4	5	48
	Receding	-	Vertical	-	Bulging	
11. FACIAL LINES	1	2	3	4	5	49-53
	Forehead	Between Eyes	Crow's Feet	Nose-Mouth	None	
	(more than one may be selected)					

Fig. 1. Typical face-feature questionnaire as filled out by a juror. (These 34 features were subsequently reduced to 22 for experimental use.)

圖 1 : Goldstein and Harmon 之人臉特徵表。

來的人像有大有小，以其眼珠寬為準來判斷其他大小。比如兩眼之距離為一判別人像之好特徵，其為寬或窄可以眼珠為準。另外臉頰的寬窄亦可以此為準來做判斷。

三人像特徵截取 (Segmentation) :

這是本文最難做到的地方，今予以一一討論。為了使像片數字平穩，照像時連續照十六張而平均起來是一比較可靠方法。

(1) 頭髮的截取：頭髮的深淺直接影響到截取的好壞。通常深黑色的頭髮可直接用統計的方法自動截取，因為其他部份都淺很多。

(2) 眼睛的截取：先把一標準人像的眼珠切下來當做樣板 (template) 然後以此樣板在要識別的人像中間去尋找眼珠，尋找方法是用樣板配對 (template matching) 即把樣板套到影像上一一比對，看那一部位最相似此樣板。相似準則是用相關係數：

$$\rho(\underline{X}, \underline{Y}) = \frac{(\underline{X} - \bar{X}) \cdot (\underline{Y} - \bar{Y})}{\|\underline{X} - \bar{X}\| \|\underline{Y} - \bar{Y}\|}$$

此處 \underline{Y} 是樣板的亮度值，排成行向量， \bar{Y} 為其平均值。 \underline{X} 為樣板對應之輸入子影像的亮度值，排成行向量， \bar{X} 為其平均值。 $\|\cdot\|$ 表示長度。

因為照像時人端正坐著，眼睛大約放在像片中間，由於打光關係眼珠有反光亮點，以此為樣板來做樣板配對可

達到很高的精確度。兩眼通常在水平位置，因此一旦找到右眼即容易找到左眼。眼珠的位置一旦確定好，我們可在以其為中心劃一長方形，大小由實驗決定。在此長方形內做 Otsu 的 local thresholding。其目的是把眼珠及眼睛切出來。由實驗結果知道眼睛上方常可完整切出來而下方就很難了，這是因為眼睛上方有睫毛比較黑，而眼睛下方常佈滿血絲不夠白而致切不出來。

(3) 鼻子的截取：鼻子在影像中比較顯著的部位而可以有有效的定位者為兩個鼻孔。兩個鼻孔在影像中呈現黑洞而附近並沒有類似的黑板塊，故使用一小塊黑樣板而在眼睛下方的四方形內做樣板配對可很快找出鼻孔。黑樣板的設計是以通常標準人像之鼻孔大小為準而其灰度值則是以本人眼珠的灰度值為主。主要原因是光在不同時間內因電壓燈絲等不穩定因素而致強弱不定因而導致所照的不同人像其灰度值亦有深淺變化。

鼻孔位置一切出來即可進一步設法切出其輪廓。方法是把右鼻孔之右邊畫出一適當大小的方格，在方格內做 Otsu 之 local thresholding 同樣地對左鼻孔如是做。由 thresholding 之結果可找出大概之輪廓。這方法有時成效很差。因此另一考慮之方法也可試著做，這是用邊際偵

側法 (edge detection) 來找鼻子的輪廓，但此法必須加上 Hough Transform 來找拋物線，如此輪廓才找得到。通常此法會受到雜訊干擾而有一些誤差，但不會太大。

有時照的像鼻孔並不顯現出來，這時就把左右鼻帽切下來做樣板。以此樣板做樣板配對亦有很大的可靠性。

(4) 嘴巴的截取：先把嘴巴的東西兩嘴角各切下一小方塊，方塊內含較黑的三角形嘴角，其中間線段較黑。以此兩方塊做樣板在輸入的影像中之鼻子下面做樣板配對則可很容易的找到嘴巴之兩角，進而用 thresholding 之方法來把嘴巴切出來。

(5) 臉輪廓的截取：臉輪廓的截取幾乎沒希望。因為臉上或多或少會有紋路，同時輪廓又與衣服領子、鬍子夾雜在一起，太亂了。我們本想用 edge detection 配上 Hough Transform，結果失敗了。用土法的 Backtracking 更不可行。我們知道人們辨認人臉輪廓是一重要特徵。而此特徵和 3 D 性質有關，因此如果要有效地取出輪廓最好是設立機械立體視覺系統。

四 人像辨認（分類）：

把人像的特徵截取後，我們就可做分類了，今說明如下：

- (1) 頭髮分類成黑、白兩類。黑的度數以眼珠的度數為主（去掉眼珠內的亮點，大約 30 左右），白的度數以背景度數為主（我們照像時背景概以白色，約 200 左右）
- (2) 眉毛分類成濃、細、倒八字（\ / ），八字（八），水平（— —）五種。

把 X - Y 座標中心移到眉毛部份的中心點。今眉毛的骨架滿足方程式 $y = a + bx + cx^2 + \epsilon$ ， ϵ 為隨機誤差。令眉毛的座標為 (x_i, y_i) ， $i = 1, 2, \dots, n$ ， n 為總點數。用矩陣表示法來估計 a ， b ， c 及 ϵ 的變異數。

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

簡寫之為：

$$\underline{Y} = X \underline{t} + \underline{\epsilon}, \quad \underline{t} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

則 \underline{t} 的估計值爲

$$\underline{\hat{t}} = \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = (X'X)^{-1} X'Y$$

由此知 \hat{b} , \hat{c} 表示眉毛的彎曲度及斜度。令誤差 ϵ 具有常態分配 $N(0, \sigma^2)$, σ 未知。則 σ^2 之估計值爲

$$\hat{\sigma}^2 = SSE / (n-3),$$

此處 $SSE = (Y - X\hat{t})' (Y - X\hat{t})$

首先檢定眉毛是否水平：即檢定假說

$$H_0 : b = c = 0$$

$$H_a : \text{not } H_0$$

在 $H_0 : y = a + \epsilon$ 時則 a 之估計值爲 $\hat{a} = \bar{Y} = \sum_{i=1}^n y_i / n$

$$\text{令 } SSE_2 = \sum_{i=1}^n (y_i - \bar{Y})^2 \text{ 則 } F = \frac{(SSE_2 - SSE) / 2}{SSE / (n-3)} \text{ 具有 } F_{n-3}^2$$

分配，所以檢定眉毛不水平 (reject H_0) if

$$F > F_{n-3}^2(\alpha),$$

α 爲第一類誤差的機率， $F_{n-3}^2(\alpha)$ 可由表查之。如果不 reject H_0 當 $\alpha = .10$ 時，我們可概括說眉毛是水平的。

由 $\hat{\sigma}^2$ 之大小我們可判定眉毛的濃、細。而由 $\hat{b} > 0$ 或 $\hat{b} < 0$ 而判知眉毛是倒八字或八字形。其次檢定兩組

人像之眉毛是否相同（同一類）。

假定第一組眉毛有 n_1 點，而滿足方程式：

$$y = a_1 + b_1 x + c_1 x^2 + \varepsilon_1, \quad \varepsilon_1 \sim N(0, \sigma_1^2)$$

第二組眉毛有 n_2 點，滿足方程式：

$$y = a_2 + b_2 x + c_2 x^2 + \varepsilon_2, \quad \varepsilon_2 \sim N(0, \sigma_2^2)$$

我們要檢定 $H_0: b_1 = b_2, c_1 = c_2$ 即兩組眉毛一樣，

$$V.S. H_a: \text{not } H_0$$

由前面計算方法可算出

$$\hat{t}_1 = \begin{pmatrix} \hat{b}_1 \\ \hat{c}_1 \end{pmatrix}, \quad \hat{t}_2 = \begin{pmatrix} \hat{b}_2 \\ \hat{c}_2 \end{pmatrix} \quad \text{及} \quad \hat{\sigma}_1^2, \quad \hat{\sigma}_2^2,$$

$$\text{而 } \text{Var}(\hat{t}) = \sigma^2 (X'X)^{-1} \quad \text{令 } W \text{ 爲 } (X'X)^{-1}$$

內對應 b, c 之子矩陣（即 $(X'X)^{-1}$ 右下方 2×2 子矩陣），

$$\text{如此得 } \text{Var} \begin{pmatrix} \hat{b}_1 \\ \hat{c}_1 \end{pmatrix} = \sigma_1^2 W_1 \quad \text{及} \quad \text{Var} \begin{pmatrix} \hat{b}_2 \\ \hat{c}_2 \end{pmatrix} = \sigma_2^2 W_2$$

則檢定假說之統計量爲

$$T_{n_1, n_2} = (\hat{t}_1 - \hat{t}_2)' (\sigma_1^2 W_1 + \sigma_2^2 W_2)^{-1} (\hat{t}_1 - \hat{t}_2)$$

reject H_0 if $T_{n_1, n_2} > \chi_{n_1 + n_2 - 6}^2(\alpha)$ 由表查，

$$\alpha = .10$$

(3) 臉面的凸凹的檢定 (ridge or valley) :

我們首先取出要檢定的部位 : 25x25 subimage 再對每一點座標 (r , c) 有亮度 f (r , c) 用下列方程式來漸近 f (r , c) :

$$f(r, c) = k_1 + k_2 r + k_3 c + k_4 r^2 + k_5 r c + k_6 c^2 + k_7 r^3 + k_8 r^2 c + k_9 r c^2 + k_{10} c^3 + \epsilon$$

此處 k_1, \dots, k_{10} 可用前面估計眉毛的方程式係數之法則估計出來 (參見 [7]) 。進一步計算最傾斜的方向 :

$$\alpha = \pm \frac{1}{2} \tan^{-1} k_5 / (k_6 - k_4)$$

$$\text{令 } r = \rho \sin \alpha, c = \rho \cos \alpha$$

$$\text{則 } f_\alpha(\rho) = A\rho^3 + B\rho^2 + C\rho + k_1 + \epsilon$$

$$\text{where } A = k_7 \sin^3 \alpha + k_8 \sin^2 \alpha \cos \alpha + k_9 \sin \alpha \cos^2 \alpha + k_{10} \cos^3 \alpha$$

$$B = k_4 \sin^2 \alpha + k_5 \sin \alpha \cos^2 \alpha + k_6 \cos^2 \alpha$$

$$C = k_2 \sin \alpha + k_3 \cos \alpha$$

$$\text{由 } f'_\alpha(\rho) = 0 \text{ 可得 } \rho = \frac{-B \pm \sqrt{B^2 - 3AC}}{3A}, \text{ 叫 } \rho_s \text{ 及 } \rho_L$$

如果 $B > 0$ 即 $f''_\alpha(\rho_s) < 0$ 判定為凸 : ridge

$B < 0$ 即 $f''_\alpha(\rho_s) > 0$ 判定為凹 : valley

$$\text{我們可計算深度 } \text{depth} = |f_\alpha(\rho_s) - f_\alpha(\rho_L)|$$

- (4) 眼睛的分類：眼睛分成大、中、細三類，這些可由眼珠的切割來判斷。眼珠切下來如果是幾近全圓，眼睛一定是大。如果切下來很細長則眼睛一定是細。其他情形為中類。
- (5) 兩眼間距離之分類：以眼珠寬為準，測量兩眼間的距離。由此距離分類成寬、窄兩類。
- (6) 鼻子的分類：分成寬、窄、中、長、短五類，同樣地以眼珠的直徑為準來測量鼻子的寬度，並測量鼻孔至兩眼珠的距離。
- (7) 嘴巴的分類：分成大、中、小、細唇、厚唇五類。先測量兩嘴角間的距離，再找出中間黑線，再找出上下嘴唇的厚度。找厚度比較困難，時常不可行。
- (8) 臉輪廓之分類：分成方形、蛋形、三角形、圓形四類。先做 edge detection 再用 Hough Transform —— 測試。Hough Transform 有圓、橢圓、拋物線、方形四種。那一種測試結果 cell count 最多即分類成那一種，很可惜試驗好幾次只有一次拋物線成功。

五 實驗結果及討論：

人像辨認是很重要的一項研究。可惜這個已牽涉到 Scene Analysis 在目前的科技層面上還沒有能力突破。我們做了一些實驗，首先人像照了進來存在 VAX 730 內

而後做人像特徵截取。圖 2 是 Young 人像，而圖 3 是其特徵截取後的情形。其兩眼距離為 136，鼻寬為 47，鼻長為 86，而嘴寬為 103。圖 4 是另一人像及其 edge detection 情形。我們設法用 Hough Transform 把臉輪廓連起來，但成功性太低原因是其他雜訊線條太多。圖 5 是人臉凸凹的檢定及其深度，鼻子最凸，在 (252, 210) 處，其凸度是 216306.8。

本文並沒有一一做識別人像實驗，主要是計算時間太長。本文強調在理論架構。現今因 COHU 相機品質不良，是以實驗做起來很吃力。等到 CID 相機一到，人像品質會更好，到時做各式實驗會更容易。人像辨認是很重要，Harmon 把臉的側面用筆劃出臉廓再輸入電腦內做識別工作，這方法是值得試因為資料簡單可及時作業。我們想免去人工筆劃完全自動辨認，這可用背後打光的技巧 (Backlighting) 值得下一回試試看。

Template Matching +
Segmentation of Image

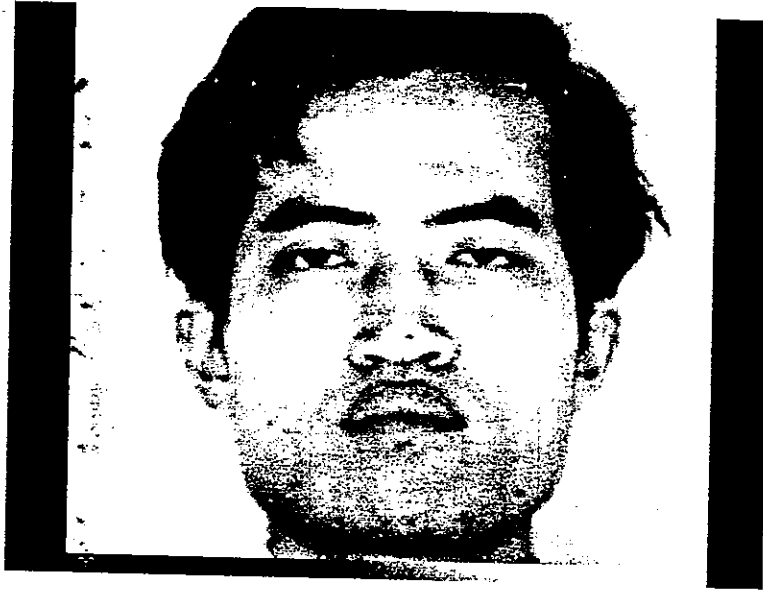


圖 2 : 楊氏人像 (Young) .

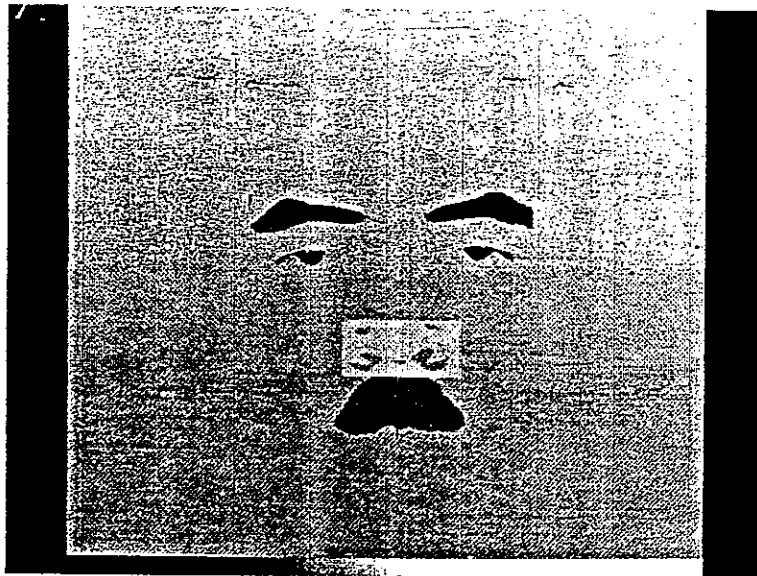


圖 3 : 圖 2 人像截取後之情形。

Edge Detection +
Hough Transform

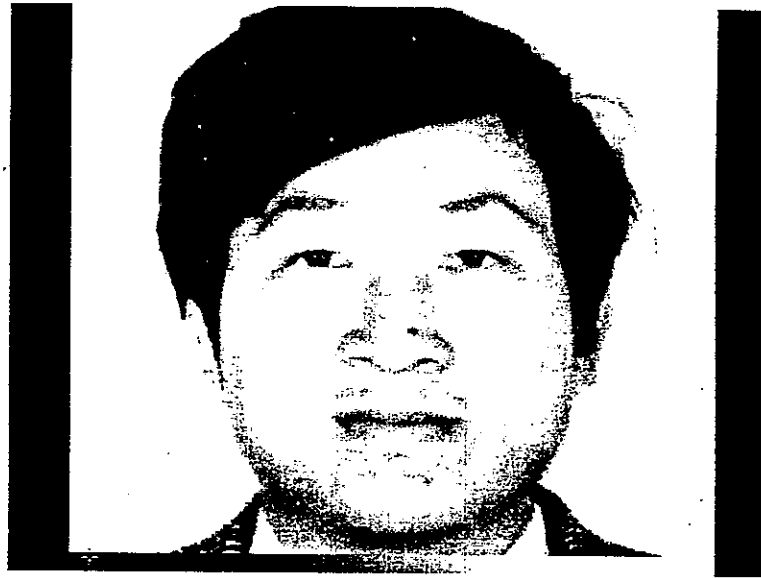


圖 4 : 人像及其下半部 edge detection .

Ridge and Valley

```
THE SURFACE IS ---->VALLEY (THIN)
THE DEPTH IS ----> 10.13490
DO YOU WANT TO CUT ANY OTHER WINDOW (Y/N) ?
25,235
INPUT POSITION OF WINDOW'S CENTER --> IX , IY
25,235
END      626
THE SURFACE IS ---->VALLEY (THIN)
THE DEPTH IS ----> 6.725494
DO YOU WANT TO CUT ANY OTHER WINDOW (Y/N) ?

INPUT POSITION OF WINDOW'S CENTER --> IX , IY
230,240
END      626
THE SURFACE IS ---->PLAIN
DO YOU WANT TO CUT ANY OTHER WINDOW (Y/N) ?

INPUT POSITION OF WINDOW'S CENTER --> IX , IY
237,245
END      626
THE SURFACE IS ---->VALLEY (THIN)
THE DEPTH IS ----> 500.0391
DO YOU WANT TO CUT ANY OTHER WINDOW (Y/N) ?
```

```
THE DEPTH IS ----> 36.32896
DO YOU WANT TO CUT ANY OTHER WINDOW (Y/N) ?

INPUT POSITION OF WINDOW'S CENTER --> IX , IY
252,210
END      626
THE SURFACE IS ---->RIDGE (FAT)
THE DEPTH IS ----> 216306.8
DO YOU WANT TO CUT ANY OTHER WINDOW (Y/N) ?

INPUT POSITION OF WINDOW'S CENTER --> IX , IY
150,120
END      626
THE SURFACE IS ---->VALLEY (THIN)
THE DEPTH IS ----> 5.648865
DO YOU WANT TO CUT ANY OTHER WINDOW (Y/N) ?

INPUT POSITION OF WINDOW'S CENTER --> IX , IY
155,120
END      626
THE SURFACE IS ---->VALLEY (THIN)
THE DEPTH IS ----> 5.267822
DO YOU WANT TO CUT ANY OTHER WINDOW (Y/N) ?
```

圖 5 : 人臉凸凹的檢定及其深度，鼻子最凸在 (252, 210)

處，凸度是 216306.8 .

參 考 資 料

1. L. D. Harmon, Automatic recognition of human face profiles, *Computer Graphics and Image Processing*, 6, 1977, 135-156.
2. L. D. Harmon, M. K. Khan, Richard Lasch and P.F. Ramig, machine identification of human faces, *Pattern Recognition*, Vol. 13, 1981, 97-110.
3. Nakaji Honda and Futoshi Sugimoto, Multivariate data representation and analysis by face pattern using facial expression characteristics, *Pattern Recognition*, Vol. 19, 1986, 85-94.
4. A. J. Goldstein and L. D. Harmon, Identification of human faces, *Proceedings of IEEE*, 59, 1971, 748-760.
5. A. J. Goldstein, L. D. Harmon and A. B. Lesk, Man-machine interaction in human-face identification, *The Bell System Technical Journal*, Vol. 51, 1972, 399-427.
6. R. J. Baron, Mechanism of human facial recognition, *Int. J. Man-machine Studies*, 1981, 137-178.
7. R. M. Haralick, Ridges and valleys on digital images, *Computer Vision, Graphics and Image Processing*, 22, 1983, 28-38.