

结构光图象的智能细化与补断算法 —在人头像三维表面重建中的应用

胡汉平 吴晓刚 李德华 彭嘉雄

(华中理工大学, 武汉 430074)

摘要 简要地介绍了三维彩色扫描系统的工作原理;对人的面部特征及其反射特性做了详细地分析;提出了一个细化与补断的智能算法,该算法具有智能性、容错性、扩展性和综合性;文末给出了实验结果及误差分析。

关键词 三维彩色扫描系统,细化,面部特征,人工智能

1 引言

获取被测物体的立体信息是计算机视觉的主要任务之一,随着计算机技术的迅猛发展,工业、军事工程、医学、影视业和广告业等领域已对计算机视觉

提出了一个越来越迫切的要求,这就是,如何快速获取物体的三维彩色模型。

我们研制的三维彩色扫描系统正是针对这一要求而提出的。它采用结构光测距的技术原理,提供了一个具有足够精度和逼真视觉效果的实用扫描装置¹⁾其中处理软件的工作过程如图1。

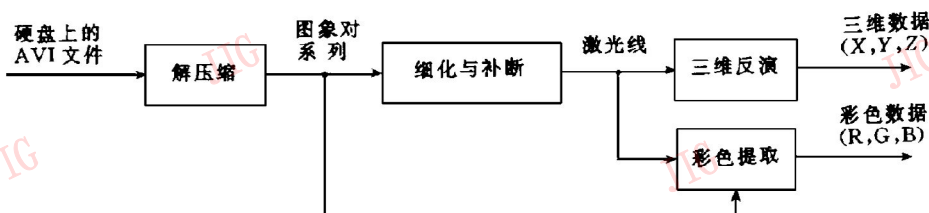


图1 处理软件总体框图

Fig. 1 Processing software

由于实际的线激光源有一定宽度(本系统采用0.4mm),被扫描物体表面存在扩散反射,以及CCD摄像机与动态图象采集卡存在量化误差及电噪声,从而导致所采图象中激光反射的并不是一条线,而是一个光带。为得到理想激光线,必须对光带进行细化。

由于本系统是以人头部为主要扫描对象的,而

人头部表面较复杂,从而导致激光反射带存在很复杂的断点情况,为了获得其三维信息,必须对光带进行补断。

2 面部特征分析

2.1 面部特征^[2]

面部识别中最常用的数字特征是积分投影:

$$(1) \begin{cases} V(x) = \sum_{y=y_1}^{y_2} I(x, y) & \text{垂直投影} \\ H(y) = \sum_{x=x_1}^{x_2} I(x, y) & \text{水平投影} \end{cases}$$

$[x_1, x_2] \times [y_1, y_2]$ 为探测窗

其次是:

(2) $I/\langle I \rangle$: 密度归一化, 对适当邻域平均灰度的比值

(3) $D(I) = |\partial_x I| + |\partial_y I|$

(4) $DD(I) = \partial_{xx} I + \partial_{yy} I$

通过识别率比较, 可以发现利用 $D(I)$ 的识别率最高。

2.2 面部对激光的反射情况

物体表面对线激光源的反射情况由三个因素决定: 表面粗糙度、表面颜色和表面水平曲率(垂直于线激光方向上的宏观平均曲率)。面部各特征区的这三个因素值均有所不同, 下面依次讨论(以下分析与第 3 节中知识库的每条规则都是密切相关的)。

2.2.1 普通区

除特征区域(以下几点将要讨论的)以外的区域都属于普通区。普通区的特点是: 粗糙度较小, 颜色浅, 水平曲率小且分布都很均匀。因此, 激光反射光带较窄, 变化缓慢, 无断点现象, 容易细化。区号 $Area=1$ 。

2.2.2 眉区

眉区定义为生长眉毛的区域, 它的特点是: 粗糙度较大, 颜色较深, 水平曲率小, 且分布都较均匀。反射光带细窄而略散乱, 细窄是由于颜色较深, 散乱是由于眉毛的散乱反射(激光射到结构细微如毛发等的物体上而产生的反射, 实际上反射的是一些散乱的细点光源)。

眉区与发区的不同在于, 眉毛稀疏, 激光反射大部分是由眉毛下的皮肤产生的, 而头发浓密, 激光反射大部分由发丝产生。因此反演眉区三维信息前还需做一次校正。区号 $Area=2$ 。

2.2.3 眼区

上眼睑与下眼睑所围区域称为眼区, 其特点为: 粗糙度、颜色和水平曲率分布都极不均匀。激光射到眼睑上时, 由于梯度陡变而产生断点, 但是该断点很窄, 并且恰好反映了眼廓信息, 因此眼睑断点并不成为问题, 反而成了可以利用的信息。激光射到眼白上反射强烈, 而射到瞳仁与瞳孔上则无反射, 这样形成

了小区域的断点, 断点恰好反映了瞳仁轮廓, 但由于断点区域并不很窄, 因此需要补断。眼睑断点可能与瞳孔断点重叠。区号 $Area=3$ 。

2.2.4 鼻区

鼻区与普通区唯一的区别在于鼻区水平曲率大且分布不均匀, 因而激光反射光带存在一大段断点, 需要补断。区号 $Area=4$ 。

2.2.5 唇区

唇区与普通区仅有的区别是, 唇区颜色深, 因而激光反射光带较窄。光带在唇廓上形成转折点, 即该转折点反映了唇廓这一重要的面部信息。区号 $Area=5$ 。

2.2.6 下颌区

下颌区与普通区在三点因素值上无区别, 但由于下颌特有的凸凹结构, 造成光带出现大段断点, 也需要补断。区号 $Area=6$ 。

2.2.7 耳区

耳区结构很复杂, 水平曲率大, 且不均匀, 断点出现并不是一大段的, 而是一小段接着一小段的。区号 $Area=7$ 。

2.2.8 发区

发区粗糙度大, 颜色因人而异, 水平曲率大且不均匀。激光反射呈散乱反射。处理发区的确很困难, 但由于头发特征要求不高, 只要抓住轮廓及大体纹理即可, 因此处理发区的算法又需要从新的角度入手, 这将在下一节中提及。区号 $Area=8$ 。

3 基于面部特征的细化与补断

3.1 总体算法

在细化与补断的总体算法中, 我们的基本思想就是“分区处理”。第一步, 必须判断当前处理的是哪一个区, 即反射区分类, 在此我们嵌入了人工智能的技术; 第二步, 根据判断结果调用相应的子算法(针对当前反射区)进行处理; 最后, 整理数据输出, 以供下一级模块使用。总体算法框图见图 2。

3.1.1 激光反射光带的提取(预处理)

预处理的目的在于提取激光反射光带, 抑制背景噪声, 消除自然光镜面反射造成的高亮度区, 以保证下一级的分类器正确分类。

在此, 我们采用同一位置上的两幅图象相减的方法来抑制噪声, 一幅图象照射激光, 另一幅图象不照射激光。

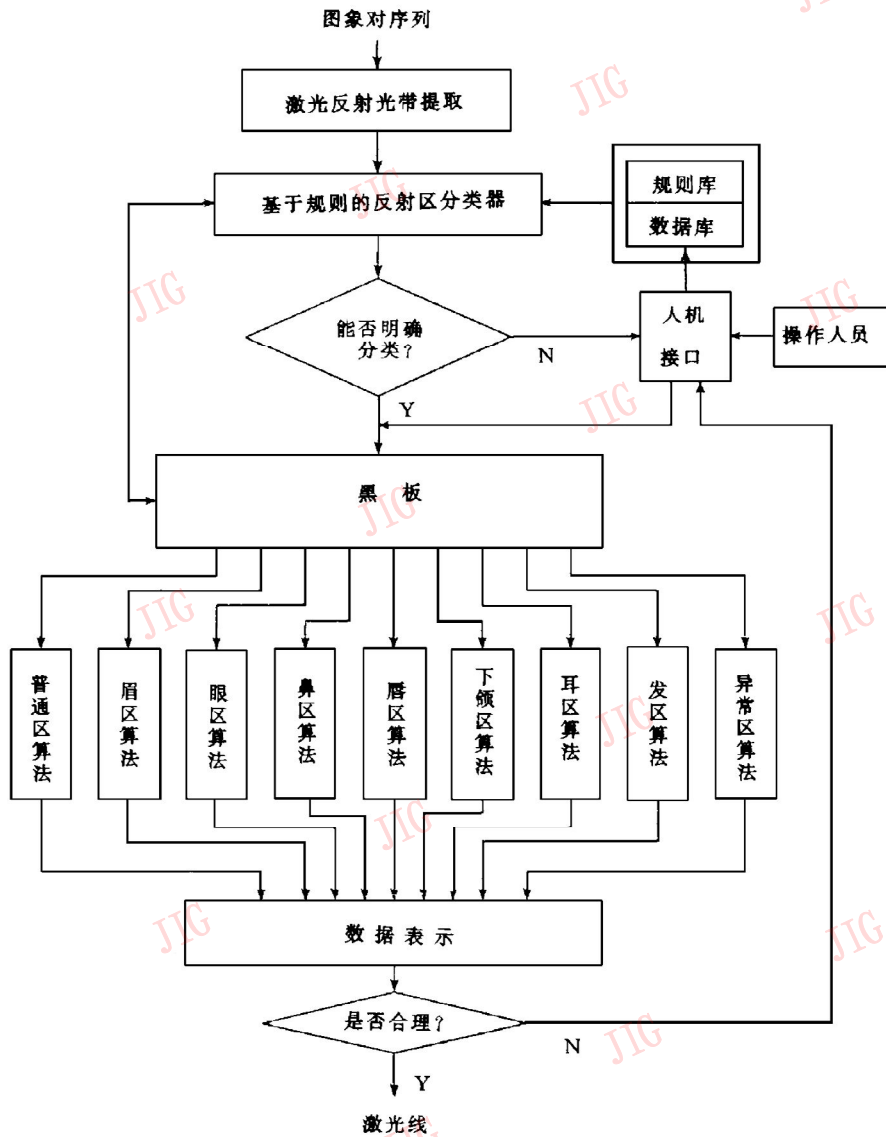


图 2 细化与补断算法框图

Fig. 2 Thinning and breakpoint joining algorithm

3.1.2 反射区分类

反射区分类单元处理的对象是预处理中提取的光带,处理的结果是给每一条光带标上分区界及分区号,处理的过程实质上是一个浮动窗(M×N)。在分类技术上我们采用了人工智能的方法,即基于规则的分类技术。

(1) 知识表示

以下规则是由第 2 节对面部反射激光的讨论及完备性分析确定的。

- ① rule_no1: IF $breakpoint() = False$,
 $HOR - Curvity() < HC_L$, $roughness() < r_L$,
 $I > I_L \wedge I < I_H$ THEN $Area(y) = 1$;
- ② rule_no2: IF $breakpoint() = False$,

- $HOR - Curvity() < HC_L$,
 $roughness() > r_L \wedge roughness() < r_H$ THEN $Area(y) = 2$;
- ③ rule_no3: IF $breakpoint() = True$, $roughness() < r_L$,
 $Area(y + \Delta y) = 2$,
 $I < I_L \vee I > I_H$
 THEN $Area(y) = 3$;
- ④ rule_no4: IF $breakpoint() = True$, $roughness() < r_L$,
 $Area(y + \Delta y) = 2$,
 $I > I_L \wedge I < I_H$
 THEN $Area(y) = 4$;
- ⑤ rule_no5: IF $breakpoint() = False$, $HOR - Curvity() < HC_L$, $roughness() < r_L$,
 $I < I_L$ THEN $Area(y) = 5$;
- ⑥ rule_no6: IF $breakpoint() = True$, $roughness() < r_L$,
 $Area(y + \Delta y) = 4$, THEN $Area(y) = 5$;
- ⑦ rule_no7: IF $breakpoint() = True$,
 $roughness() < r_L$, $Area(y + \Delta y) = 5$,
 THEN $Area(y) = 6$;
- ⑧ rule_no8: IF $breakpoint() = True$, $roughness() < r_L$,
 $Area(y + \Delta y) = 8$,
 THEN $Area(y) = 7$;
- ⑨ rule_no9:

IF $breakpoint() = True$, $roughness() > r_L$
 THEN $Area(y) = 8$

$breakpoint()$: 探测 $M \times N$ 窗口内有无断点,有则输出 True,无则输出 False;

$$roughness = \frac{1}{n} \sum_{y=y_0-\frac{M}{2}}^{y_0+\frac{M}{2}} \sum_{x=x_0-\frac{N}{2}}^{x_0+\frac{N}{2}} |x - x_0| \cdot I_D(x, y);$$

$I_0(x, y)$: 原图象; $I_L(x, y)$: 激光图象;
 $I_D(x, y)$: 差图象 $I_D = I_L - I_0$; (x_0, y_0) : 窗口中心;
 n : $I_D(x, y) \neq 0$ 的点的个数;

$$I = \frac{1}{M \cdot N} \sum_{y=y_0-\frac{M}{2}}^{y_0+\frac{M}{2}} \sum_{x=x_0-\frac{N}{2}}^{x_0+\frac{N}{2}} I_0(x, y);$$

$$HOR - Curvity = K_1 C(1) + K_2 C(2) + K_3 C(3);$$

K_i : 权系数, $C(i) = \frac{1}{2}(2x_k - x_{k-i} - x_{k+i})$,

x_k : 当前第 k 帧光带最亮值的 x 坐标;若该处为断点,则继续搜索下一帧,直至最近邻的无断点的一帧。

$Area(y + \Delta y)$: 探测当前位置的上一个特征区(即 $Area() \neq 1$ or 0)的区号;

规则以字符串数组的形式存贮在规则库文件 * .rul 中。所有参数($HC_L, r_L, r_H, I_L, I_H, M, N, K_1, \dots$)以数组[变量名,值]的形式存在数据库文件 * .dat 中,这样便于推理机查询。

(2) 推理机制

我们采用从上至下的深度优先搜索策略,推理算法框图见图 3。

图 3 中判决器“ $(Area(y) \neq Area(y - y_d))$?”实际上就是图 2 中的判决器“能否明确分类?”。其中 RN 为规则个数,在此 $RN = 9$ 。

3.1.3 子算法调用

分类完成后,若分区号发生变化,则把分区界和分区号存贮在黑板(就是一个堆栈)中,黑板存贮完一条光带的分区结果后进行一次自检,不合理则弹入人机接口。自检是通过当前分区号序列与脸部特

征区编码的比较来实现的,脸部特征区编码规则如下:

- (1) 脸部纵向各特征区区号顺序排列;
- (2) 一条纵向只有一种编码方式;
- (3) 非同一条纵向可以有同一种编码方式;

例如:“8121314151”代表“发区—普通区—眉区—普通区—眼区—普通区—鼻区—普通区—唇区—普通区—下颌区”;“871”代表“发区—耳区—普通区”。

自检通过就根据分区号来调用相应的子算法模块进行细化与补断。

完成一条光带的细化与补断后以数组 $LaserLine [i, n, 1 \text{ or } 0]$ (表示第 i 条激光线的第 n 个点的 X, Y 坐标值)的形式把激光线存在栈中。栈长度一定(我们采用 4 条激光线的容量,主要为了计算 $HOR-Curvity()$),出栈的激光线数组存在硬盘的数据文件上。

然后利用“对称性匹配程度”来检测激光线是否合理,若不合理,则返回人机接口,并修改相应规则的前件数据(如 r_L, r_H 和 HC_L 等)。

3.1.4 算法特点

总体算法具有以下三个特点:

(1) 智能性:专家系统知识库的引入使用该算法具有逻辑判断能力和一定的学习能力;

(2) 容错性:人机接口和异常区处理使算法即使在知识库不健全、推理失误的情况下也能照常运行;

(3) 扩展性:本系统虽是针对人头部提出的,但只要更换知识库、加载相应的子算法模块,即可在不需更改总算法的情况下针对其它物体进行细化和补断。

(4) 综合性:该算法是对实现符号主义与联接主义相结合的一次较为成功的尝试。

3.2 子算法

3.2.1 普通区算法

由于皮肤反射激光的过程难以精确建模,因此我们采用神经网络来细化。在此我们选用具有 BP 学习能力的多层前馈网,因为它具有结构简单,工作状态稳定的优点,而对于我们的问题(简单,要求可靠,学习时间要求不高)是非常合适的。

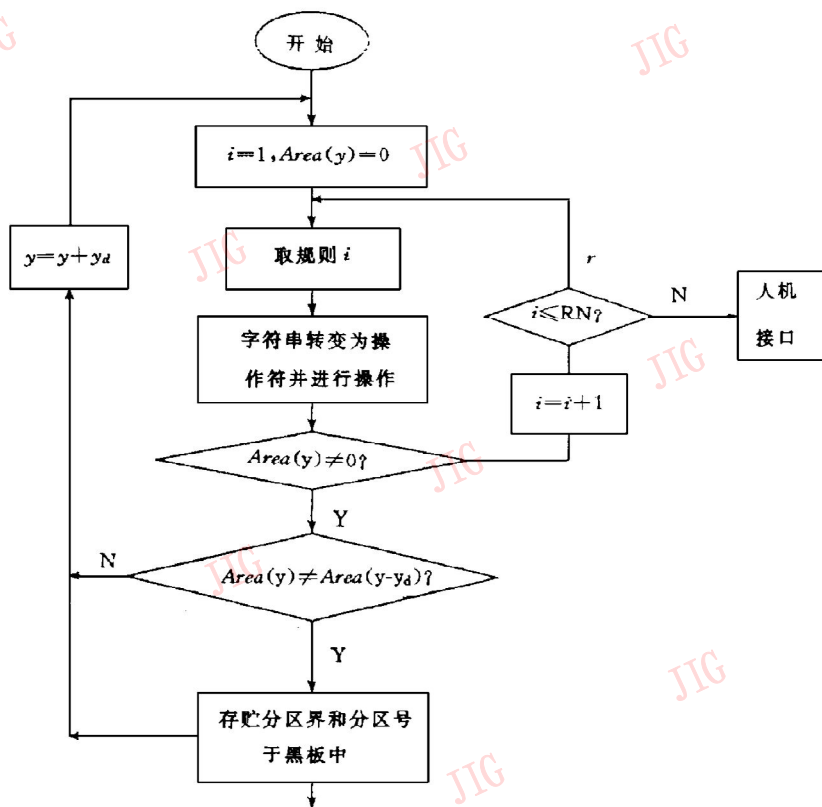


图 3 推理算法框图

Fig. 3 Reasoning algorithm

每个特征区(除 8、0 区,即发区、异常区外)的细化与补断算法实际上都是一个已经(针对各个特征区)训练好的子神经网络。各子网结构参数的选取及训练样本对的获取请参阅文献[3]。

发区算法采用现成的发形 CAD 软件来实现。

异常区算法是当操作人员在人机操作中选择了异常区($Area=0$)后才启动的。它实际上是个手动细化和补断的工作平台。它的引入给总体算法增加了很大程度的灵活性。

4 实验结果

4.1 实验结果

下面(图 4、图 5 和图 6)是 3 帧具有代表性的光带及其细化和补断的结果,参数设置如下:

参数	M	N	HC_L	r_L	r_H	L_L	L_H	K_1	K_2	K_3	γ_d
值	5	5	0.7	100	200	130	260	0.65	0.30	0.15	1

为便于观看,我们对光带图象进行了求反变换。

4.2 误差分析

在我们的数学模型中(参见文献[4]),有:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} L_1 & L_2 & L_3 & L_4 \\ L_5 & L_6 & L_7 & L_8 \\ L_9 & L_{10} & L_{11} & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

其中, (X, Y, Z) : 实物坐标; (u, v) : 成像平面坐标; $L_i (i=1, 2, \dots, 11)$: 定标系数(与光学系统结构有关)。

由于本系统中我们只关注在线激光源光平面上激光反射线所成的像,所以在此进行一次转换。设光平面过 Z 轴,与 X 轴成 θ 角,且光平面坐标 (x, y) 与实物坐标有下列关系:

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & 0 \\ \sin\theta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

所以

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} L_1 \cos\theta + L_2 \sin\theta & L_3 & L_4 \\ L_5 \cos\theta + L_6 \sin\theta & L_7 & L_8 \\ L_9 \cos\theta + L_{10} \sin\theta & L_{11} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} L_1' & L_2' & L_3' \\ L_4' & L_5' & L_6' \\ L_7' & L_8' & L_9' \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

按此数学模型进行三维反演,理论误差 $\leq 0.001\text{mm}$;但由于系统机械部分的误差,实际的平均误差 $\leq 0.3\text{mm}$ 。

值得强调的是,本系统的主要扫描对象是人的头像而不是工件,这就使扫描系统的目的不仅仅在于精度了,而在于恢复人头像的三维模型的逼真视觉效果了。因此,在此我们不是太在于追求精度,而是把重点放在了捕捉特征上。

参考文献

- 1 Sergent J. Structural processing of faces. In handbook of Research on Face Processing (A. W. Toung and H. D. Ellis, Eds.). Amsterdam: North-Holland, 1989.
- 2 Roberto Brunelli, Tomaso Poggio. Face Recognition: Features Versus Templates. IEEE Trans. on PAMI, 1993-10, 15(10).

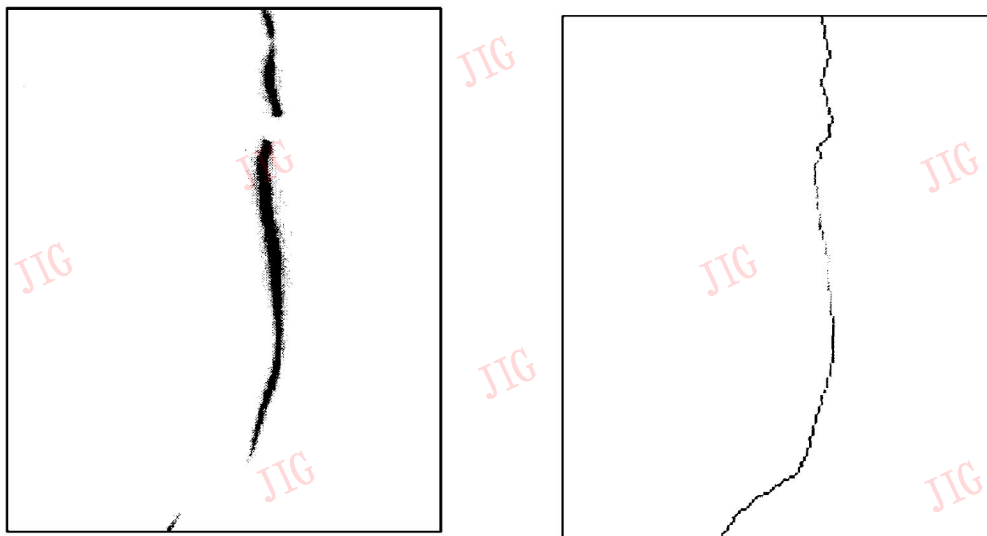


图 4 眉区和眼区

Fig. 4 Eyebrow Zone and Eye Zone

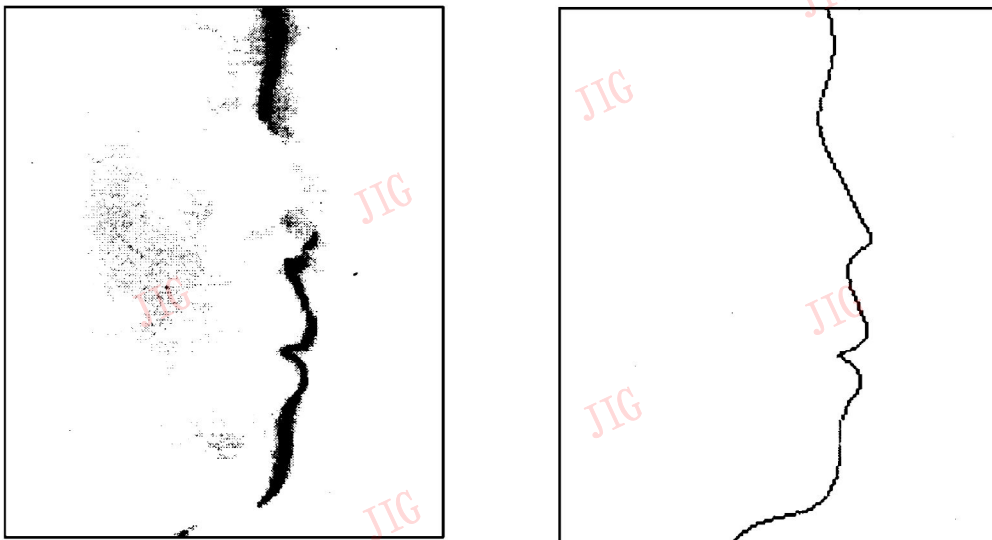


图 5 鼻区、唇区和下颌区

Fig. 5 Nose Zone, Lip Zone and Chin Zone

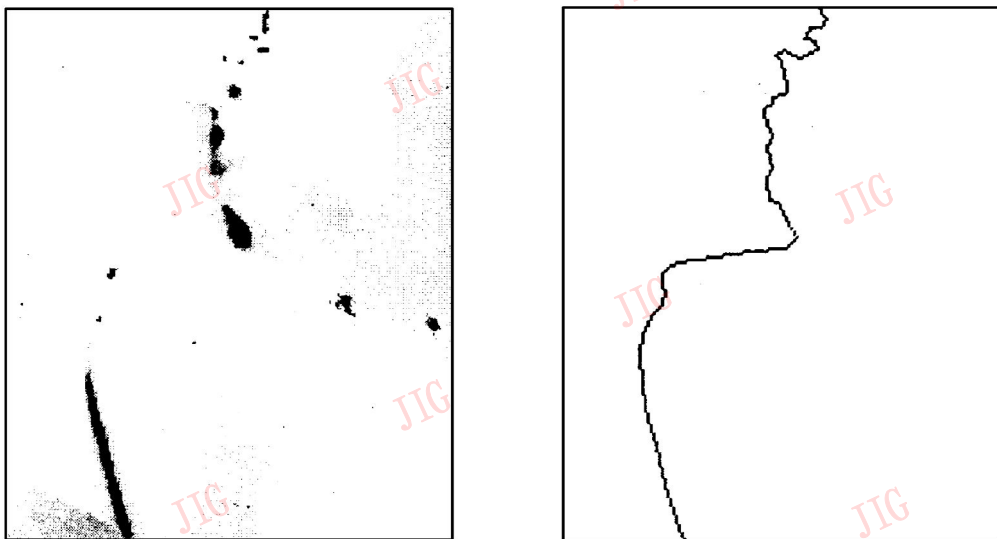


图 6 耳区

Fig. 6 Ear Zone



胡汉平,武汉市华中理工大学图象识别与智能控制研究所、博士,现主要从事人工智能、计算智能、图象采集、处理和识别等领域研究、在国内外发表论文 10 余篇

(下转 573 页)

- 3 谢俊. 神经网络在图象处理中的应用:[学位论文].
- 4 范洪. 3-D 计算机视觉系统的高精度定标和三维运动信息的准确获取:[博士学位]. 西安:北方交通大学,1993.
- 5 傅京孙,蔡自光. 人工智能及其应用.
- 6 Joseph W. GOODMAN 《STATISTICAL OPTICS》.
- 7 汤顺清. 色度学. 北京:理工大学出版社.
- 8 程景云,倪亦余. 人机界面设计与开发工具.
- 9 Deutsh E S Thinning algorithm on rectangular, hexagonal and triangular arrays. Commun. ACM 15, 1972: 827.
- 10 Koenderink J J. The structure of image. Biolog. Cybern, 1984,50:363~370.
- 11 Levine M D. A Knowledge-based computer vision. CVS, 1978.
- 12 Minsky M. A framework for representing knowledge. In The Psychology of Computer Vision (P. H. Winston, Ed.). New York: Mc Graw-Hill, 1975.
- 13 Baron R J . Mechanisms of human facial recognition. Int. J. Man Machine Studies, 1981,15:137~178.