

# 基于自适应定向正交投影分解的图象分割方法

傅 弘 阎鸿森

(西安交通大学电信学院信息与通信工程系, 西安 710049)

齐 春

(西安交通大学电信学院图象处理与识别研究所, 西安 710049)

**摘要** 将目标和背景分别对应到灰度直方图中的两个高斯分布是进行图象分割的一种常用方法,但复杂图象的直方图往往是多峰的,为了更好地拟合这种复杂图象直方图的多峰特性,提出了一种基于自适应定向正交投影分解的图象分割方法.该方法首先将这种复杂图象的直方图看作是多个高斯分布的叠加,并可通过应用自适应定向正交投影分解法来快速准确地确定每个高斯分布的权值、均值和方差,进而计算出各相邻高斯分布之间的最优阈值,以用于图象分割.在此基础上,又提出了阈值分离度的概念,并将其作为选取最终阈值的指标.应用实例结果表明,该方法能够快速有效地实现复杂图象的多阈值分割.

**关键词** 计算机图象处理(520·6040) 图象分割 自适应定向正交投影分解法 高斯分布 时频分析

**中图分类号**: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2003)03-0286-06

## Approach of Image Segmentation Based on Adaptive Oriented Orthogonal Projective Decomposition

FU Hong, YAN Hong-sen

(Dept. of Information & Communication, Xi'an Jiaotong University, Xi'an 710049)

QI Chun

(Institute of Image Processing and Recognition, Xi'an Jiaotong University, Xi'an 710049)

**Abstract** Looking upon the gray level histogram as a mixture of two Gaussian density functions is a conventional model in the image segmentation, unfortunately the histogram of the complex image often appears a multi-peak feature. In order to get a more accuracy approximation of this kind of histogram, this paper generalizes this model by considering the histogram a mixture of several Gaussian density functions, and employs a new algorithm of Adaptive Oriented Orthogonal Projective Decomposition to handle the mathematical problems involved in this process. In this proposed method, the key parameters of each Gaussian function can be calculated efficiently, which adequately leads to the determination of the optimal thresholds between different neighboring Gaussian functions. A new parameter called the Dividual Ratio of Threshold is introduced and used as the reference for the selection of the final thresholds. Experimental results show that this method can be effectively applied for the multi-threshold segmentation of complex images.

**Keywords** Image segmentation, Adaptive oriented orthogonal projective decomposition, Gaussian density function, Joint time frequency analysis

## 0 引 言

图象分割是计算机视觉的基础,已受到科研人

员的广泛重视,而阈值分割法作为一种常见的并行区域技术,由于是直接利用图象的灰度特征,因此计算方便简明,实用性强,在图象分割领域占有相当重要的地位,但阈值分割的关键和难点在于如何取得

基金项目:陕西省自然科学基金资助项目(2000X07)

收稿日期:2002-04-27; 改回日期:2002-10-08

合适的阈值. 现有的阈值选择方法大体上分为如下 3 类:一是将图象的灰度直方图看作是 高斯分布的选择法<sup>[2]</sup>;二是基于一些准则的选择法,如类间方差法<sup>[3]</sup>、熵法<sup>[4]</sup>等;三是通过直接分析图象直方图的峰值和谷底来确定阈值<sup>[5]</sup>的方法.

虽然将目标和背景分别对应到灰度直方图中的两个高斯分布(即对直方图进行高斯分解)是图象分割领域中一种基本方法,但对于具有复杂目标或背景的图象,它们的直方图通常是多峰的,这种分解方法未免显得有些粗糙.为了能对这种复杂图象进行顺利分割,基于第 1 类阈值选择法,将自适应定向正交投影分解法<sup>[6,7]</sup>的基本原理应用到图象分割领域,即将具有复杂目标或背景图象的直方图看作是多个高斯分布的叠加,通过对其分解来较好地拟合直方图的多峰特性,以便得到更为准确的分割效果.

设具有复杂目标或背景的图象的灰度直方图为  $p(z)$ ,其中  $0 \leq z \leq L-1$ ,  $L$  为图象的灰度等级.现可以把  $p(z)$ 看作是如下一组高斯分布的叠加

$$p(z) = \sum_{i=1}^N A_i g_i(z) + r(z) \quad (1)$$

其中,  $g_i(z) = (\pi\sigma_i^2)^{-0.25} e^{-\frac{(z-\mu_i)^2}{2\sigma_i^2}}$  是均值为  $\mu_i$ 、方差为  $\sigma_i^2$  的能量归一化高斯信号,  $A_i$  是它的权值,  $r(z)$  为分解后的残量,  $N$  为分量的个数.

高斯分解的目的就是要确定每个高斯分布的特征参数,即  $\mu_i, \sigma_i^2$  和  $A_i$  (代表高斯分布的序号),对于直方图  $p(z)$ ,利用式(1)求解,需要求解的参数一共是  $3N$  个.

## 1 基本原理

### 1.1 自适应定向正交投影的高斯分解法

自适应定向正交投影分解法<sup>[6]</sup>是一种新的参数化时频分析手段,由于它是利用最大投影的原理来循环搜索与原信号最匹配的基函数集,从而能以最简洁的方式重构信号.目前该方法已经成功地应用于语音信号分析<sup>[8]</sup>、引信目标特征提取<sup>[9]</sup>等领域.

利用自适应定向正交投影分解原理来对直方图进行高斯分解的原理如下:设基函数集  $G = \{g_i\}_{i=1,2,\dots,\infty}$ , 计算  $p(z)$  在其上的最大投影  $\max \langle p, g_i \rangle$  ( $\langle p, g_i \rangle$  表示  $p$  与  $g_i$  做内积运算,为表达简洁起见,以下用  $p$  代表  $p(z)$ ), 可得剩余部分的信号为

$$r = p - \max \langle p, g_i \rangle g_i \quad (2)$$

显然,  $r$  与  $g_i$  正交. 将剩余部分再在基函数集中投影,找出投影最大的基函数,依此类推,有

$$\begin{aligned} p &= \langle p, g_1 \rangle g_1 + r_1 \\ &= \langle p, g_1 \rangle g_1 + \langle r_1, g_2 \rangle g_2 + r_2 \\ &= \langle p, g_1 \rangle g_1 + \langle r_1, g_2 \rangle g_2 + \dots + \\ &\quad \langle r_{i-1}, g_{i+1} \rangle g_{i+1} + \dots + \langle r_{N-1}, g_N \rangle g_N + r_N \end{aligned} \quad (3)$$

其中

$$\begin{aligned} r_i &= r_{i-1} - \langle r_{i-1}, g_i \rangle g_i \\ \langle r_i, g_{i+1} \rangle &= \max \langle r_i, g_l \rangle_{l=1,2,\dots,\infty} \end{aligned}$$

分解结束的条件由预先设定的  $N$  决定. 自适应定向正交投影分解法中,使用的基函数为一经过伸缩、时移、频移以及线性调频的标准高斯函数

$$g_l(z) = (\pi\sigma_l^2)^{-0.25} e^{-\frac{(z-\mu_l)^2}{2\sigma_l^2} + j\omega_l(z-\mu_l) + j\eta_l(z-\mu_l)^2} \quad (4)$$

式中,  $l=1, 2, \dots, \infty$ .

一般来说,由于图象的灰度直方图以低频成分为主,其高频成分通常表现为干扰,因此,本文结合应用对象的具体特点,去掉了频率中心  $\omega_l$  和调频因子  $\eta_l$  两个参数,这相当于加了一个低通滤波器,这样式(4)可以进一步写为

$$g_l(z) = (\pi\sigma_l^2)^{-0.25} e^{-\frac{(z-\mu_l)^2}{2\sigma_l^2}}, l=1, 2, \dots, \infty \quad (5)$$

### 1.2 算法的快速实现

由于投影算法的每一次搜索都要在参数空间  $(\mu, \sigma)$  中找到最大的投影值,尽管参数空间只有 2 维,但是如果把每个值都计算一遍,那么计算量也是相当大的.例如  $\mu$  为 0 到 255 的整数,  $\sigma$  的范围是  $(1, 100)$ , 取值间隔为 0.1, 那么每次搜索就要作 256 000 次内积运算,在 Matlab<sup>TM</sup> 环境下,大约要花费 2.5min 的时间.由于这样的运算速度远远不能满足图象分割在线处理的要求,因此,有必要研究算法的快速实现方法.

设目标函数

$$h(z) = A_0 (\pi\sigma_0^2)^{-0.25} e^{-\frac{(z-\mu_0)^2}{2\sigma_0^2}} \quad (6)$$

其中,  $A_0$  为权值,且  $A_0, \sigma_0, \mu_0 > 0$ .

问题是如何从基函数集中寻找一个最佳基函数

$$\begin{aligned} g(z) &= (\pi\sigma^2)^{-0.25} e^{-\frac{(z-\mu)^2}{2\sigma^2}}, \text{使得函数} \\ H(\sigma, \mu) &= \langle h(z), g(z) \rangle = \left| \int h(z)g(z)dz \right| \\ &= A_0 \frac{\sqrt{2\sigma_0\sigma}}{\sqrt{\sigma_0^2 + \sigma^2}} e^{-\frac{(\mu-\mu_0)^2}{2(\sigma_0^2 + \sigma^2)}} \end{aligned} \quad (7)$$

达到最大值. 为方便起见,先考虑自变量  $z$  是连续的情况,实际应用时通过离散化,把积分变为求和即可.

由式(7)可知,当  $\sigma = \sigma_0, \mu = \mu_0$ ,即基函数与目标函数相同时,  $H(\sigma, \mu)$  取得最大值,那么如何找到这两个参数呢? 如前所述,由于无法在  $(\mu, \sigma)$  空间中作遍历式的二维搜索去寻找最大值,因此只能用两个一维搜索来逼近这个最佳基函数,即分别搜索  $\mu$  和  $\sigma$ . 下面对式(7)加以分析.

(1) 尽管不知道  $A_0, \sigma_0, \mu_0$  这 3 个参数,但是可以通过改变  $\mu$  和  $\sigma$  来获得整个式子的值,即用不同的基函数和直方图作内积.

(2) 令  $\sigma$  为任意正数,如  $\sigma = \sigma_1$ ,则式(7)只是  $\mu$  的函数

$$H(\mu) = K_1 e^{-K_2(\mu-\mu_0)^2} \tag{8}$$

式中,  $K_1, K_2$  为未知正数.

若式(8)达到最大值,则应有  $\mu = \mu_0$ .

(3) 当  $\mu = \mu_0$  时,式(7)变为

$$H(\sigma) = A_0 \frac{\sqrt{2\sigma_0\sigma}}{\sqrt{\sigma_0^2 + \sigma^2}} \tag{9}$$

分别令  $\sigma = \sigma_1, \sigma_2 > 0$ , 则有

$$\begin{cases} A_0 \frac{\sqrt{2\sigma_0\sigma_1}}{\sqrt{\sigma_0^2 + \sigma_1^2}} = M_1 \\ A_0 \frac{\sqrt{2\sigma_0\sigma_2}}{\sqrt{\sigma_0^2 + \sigma_2^2}} = M_2 \end{cases} \tag{10}$$

式中,  $M_1, M_2$  为内积结果.

解这个方程组,可得

$$\begin{cases} A_0 = M_1 \frac{\sqrt{\sigma_0^2 + \sigma_1^2}}{\sqrt{2\sigma_0\sigma_1}} \\ \sigma_0 = \sqrt{\frac{\sigma_1\sigma_2(b\sigma_1 - \sigma_2)}{\sigma_1 - b\sigma_2}} \end{cases} \tag{11}$$

其中,  $b = \left(\frac{M_1}{M_2}\right)^2$ .

由以上分析可得出如下搜索策略:

第 1 步,将基函数  $g(z) |_{\sigma=\sigma_1} = (\pi\sigma_1^2)^{-0.25} e^{-\frac{(z-\mu_0)^2}{2\sigma_1^2}}$  与直方图作内积,找到与内积最大值对应的  $\mu$ , 即  $\mu_0$ . 这个内积最大值就是  $M_1$ .

第 2 步,用  $\sigma = \sigma_2, \mu = \mu_0$  的基函数,得到  $M_2$ ,再根据式(11)计算出  $A_0$  和  $\sigma_0$ .

第 3 步,把最佳基函数从直方图中去除,对剩余

部分

$$r(z) = p(z) - A_0(\pi\sigma_0^2)^{-0.25} e^{-\frac{(z-\mu_0)^2}{2\sigma_0^2}}$$

重复前两个步骤搜索,直到满足结束条件为止.

用这种方法寻找最佳基函数,能够有效地降低运算量. 例如  $\mu$  仍取 0 到 255 的整数,  $\sigma$  的范围还是 (1, 100), 取值间隔为 0.1, 那么在每一步搜索中仅需作 257 次内积运算,比遍历式搜索的计算量少了近 1000 倍. 另外,在实际计算时,在第 1 步搜索中,可用一个  $\sigma = \sigma_1, \mu = \mu_1$  的基函数  $g(z) |_{\substack{\sigma=\sigma_1 \\ \mu=\mu_1}} = (\pi\sigma_1^2)^{-0.25} e^{-\frac{(z-\mu_1)^2}{2\sigma_1^2}}$  与直方图做相关函数来找到与相关函数最大值对应的坐标  $z_{max}$ , 则  $\mu_0 = L + \mu_1 - z_{max}$ . 这样做与前面所述方法的结果一致,但是却可以省去产生基函数所用时间,也就提高了运算速度.

### 1.3 阈值分离度

直方图的高斯分解完成以后,下一步的问题就是如何确定阈值. 如式(1)所述,直方图被分为  $N$  个具有不同幅度、均值和方差的高斯分布

$$A_1g_1(\mu_1, \sigma_1), A_2g_2(\mu_2, \sigma_2), \dots, A_Ng_N(\mu_N, \sigma_N)$$

其中一个高斯峰对应一个灰度目标. 在进行高斯分布划分时,首先将这些高斯分布按照均值  $\mu_i$  由小到大排序,然后求取相邻两个分布  $A_i g_i$  和  $A_{i+1} g_{i+1}$  之间的最优阈值  $t_i$  (以两个目标的错分概率最小为准则进行计算,具体方法见参考文献[1]),共  $N-1$  个. 由于如果直接用  $t_i$  来进行分割,则很有可能会因两个目标的交叠部分过大,而造成没有必要的分割,因此,需要在其中挑选合适的数值作为最终的阈值,而  $t_i$  可以称为“准阈值”. 但如何挑选合适的阈值,可通过定义一个“阈值分离度”参数来选择;对于“准阈值”  $t_i$ , 它的阈值分离度  $d_i$  为

$$d_i = \frac{\mu_{i+1} - \mu_i}{\sigma_{i+1} + \sigma_i} \tag{12}$$

上式的分子是两个高斯分布的均值间距,均值间距越大,表示它们的灰度间隔越大;分母是两个高斯分布标准差之和,这个值越大,表示它们可能重叠的部分越多. 由此可见,  $d_i$  越大,就表明这两个目标在灰度上的重叠部分越小,可分离性越大.

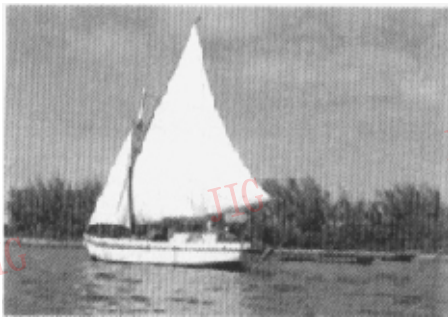
这样,在获得“准阈值”  $t_i$  后,再求出它们对应的阈值分离度  $d_i$ , 然后即可按照实际需要确定最终阈值  $T_i$ . 如果要进行单阈值分割,就选择阈值分离度最大的那个准阈值;如果要进行多阈值分割( $m$  个阈

值,  $m \leq N - 1$ ), 则可以按照阈值分离度从大到小的顺序选取前  $m$  个准阈值.

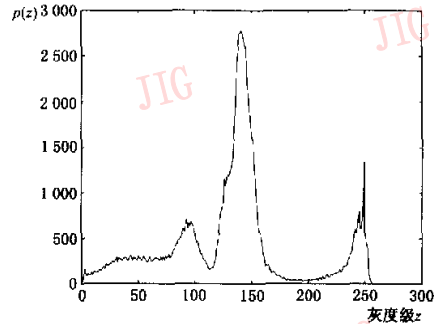
### 2 实例与讨论

图 1(a) 是一幅具有复杂背景的自然图象, 图 1(b) 是它的灰度直方图. 预先设  $N = 6, \sigma_1 = 5, \sigma_2 = 6$ , 利用本文方法可将原直方图分解为 6 个高斯函数(如图 2 所示), 其特征参数列于表 1. 如果将这 6 个高斯函数相加, 然后与原直方图对照, 则可见其叠加后的曲线不但较好地反映了原直方图的基本走

向, 而且消除了细微毛刺的影响. 表 2 所列为此 6 个高斯函数之间的准阈值及其分离度, 由表 2 可见, 高斯函数 4、5 分量之间的最优阈值 187 的分离度远高于其他准阈值的分离度, 反映到直方图上, 就是高斯函数 5、6 分量与其他高斯分量相距甚远, 而高斯函数 5、6 分量正好对应于图象中的目标“帆船”. 由此可见, 用分离度最大的阈值 187, 可以将目标与背景分离, 图 3(a) 为单阈值分割的结果. 若再增加一个阈值 115, 则可对背景进行进一步划分, 如图 4(a) 所示. 在 Matlab™ 环境下, 当  $N = 6$  时, 用该方法选取阈值仅需 0.44s.

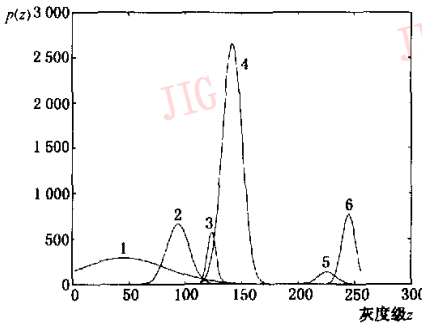


(a) 原图

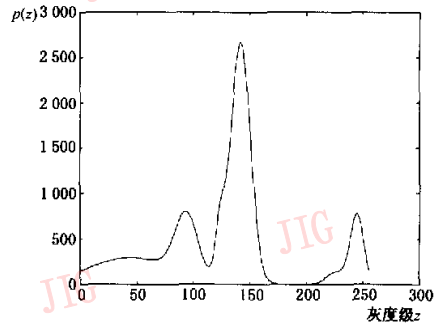


(b) 原直方图

图 1 待分割图象及其灰度直方图



(a) 分解后所得高斯函数各分量



(b) 直方图的重构

图 2 直方图的高斯分解与重构

表 1 高斯分解的特征参数

分量序号 $i$	权值 $A_i (\times 10^4)$	均值 $\mu_i$	标准差 $\sigma_i$
1	0.223 5	45	34.865 5
2	0.283 2	94	10.560 9
3	0.164 5	124	3.564 0
4	1.045 2	142	8.420 3
5	0.048 4	226	7.884 9
6	0.248 8	245	6.142 9

表 2 准阈值及其分离度

序号 $i$	准阈值	阈值分离度
1	187	5.151 7
2	115	2.123 9
3	127	1.502 0
4	232	1.354 5
5	74	1.078 7

为了验证本文方法的效果,用 Otsu 法与本文方法进行了对比测试,图 3~图 5 是分别用本文方

法和 Otsu 法<sup>[3]</sup>(穷举法)进行单阈值、双阈值和三阈值分割的结果,各次分割的计算参数列于表 3。

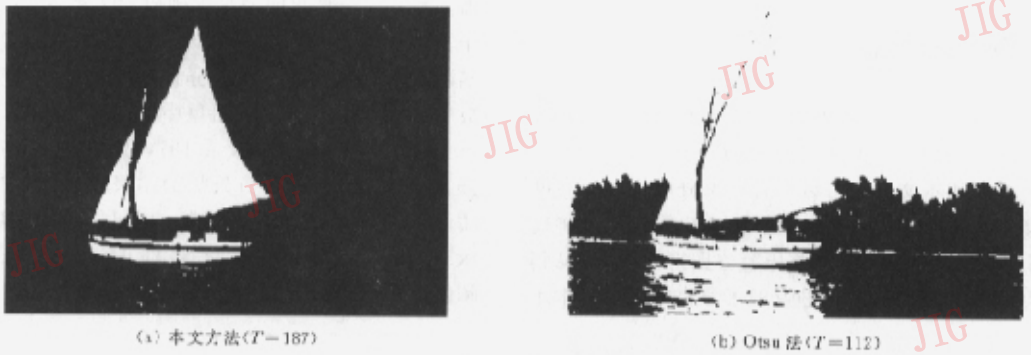


图 3 单阈值分割结果

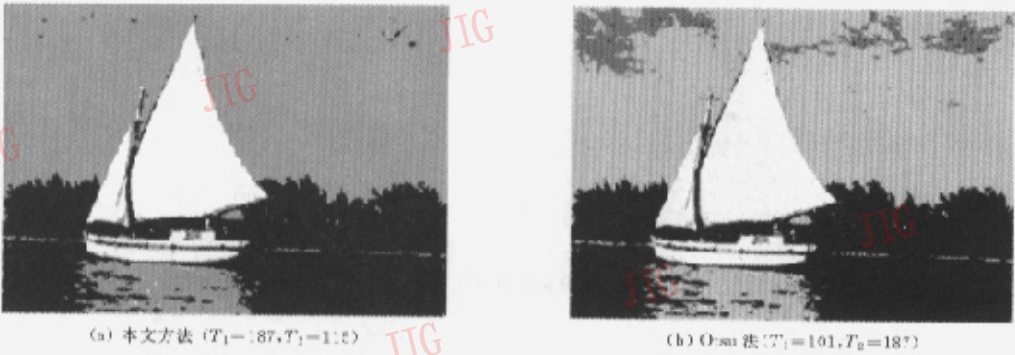


图 4 双阈值分割结果

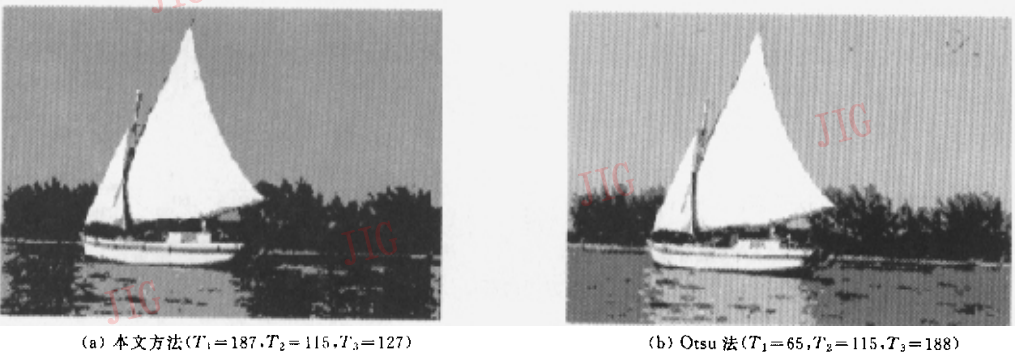


图 5 三阈值分割结果

表 3 本文方法和 Otsu 法分割参数比较

分割类型	本文方法		Otsu 法	
	阈值 $T$	计算时间(s)	阈值 $T$	计算时间(s)
单阈值	187	0.44	112	0.11
双阈值	187, 115	0.44	101, 187	18.90
三阈值	187, 115, 127	0.44	65, 115, 188	2 211.60

从分割效果上看,在单阈值的情况下,本文方法只用一个阈值 187 就成功地把目标“帆船”与背景分离开来,而 Otsu 法则把“天空”、“帆船”和“倒影”混在一起;在双阈值情况下,两种方法都获得阈值 187,均成功地分离了目标“帆船”,而且都把背景划分为“天空+倒影”和“植物+水面”两类,虽然两种

方法都出现类之间的混淆,但是本文方法的混淆程度要轻一些;在三阈值情况下,两种方法具有两个相近的阈值(115 和 187),不同的是,本文方法把“天空”做了进一步的划分,而 Otsu 法则区分了“植物”和“水面”。从计算时间上看:在单阈值情况下,Otsu 法比本文方法要快一些,但是随着阈值数目的增多,Otsu 法所需时间急剧增加,尤其是到三阈值时,运算速度几乎难以接受;而本文方法由于一次计算就给出了全部准阈值,因运算与阈值个数无关,故运算时间一直保持在 0.44s。

由于本文方法是以灰度直方图为依据进行分割的,因此适用于能够用直方图分割的图象。一般来说,这些图象应有以下两个特点:(1)图象中的目标与直方图中的高斯函数具有较好的对应性,即不同目标反映到直方图上是不同的高斯函数;(2)直方图中各高斯函数之间重叠部分较少。

参 考 文 献

- 1 章毓晋. 图象处理和分析[M]. 北京:清华大学出版社,1999:197~198.
- 2 Kittler J, Hlingworth J. Minimum error thresholding [J]. Pattern Recognition, 1989,19(1):41~47.
- 3 Ostu N. A threshold selection method from gray-level histogram [J]. IEEE Trans. on Systems, Man and Cybernetics, 1978, SMC-9 (1), 62~66.
- 4 Pun T. Entropic thresholding, a new approach[J]. Computer Graphics Image Processing, 1981,16(3):210~239.
- 5 Sezan M I. Peak detection algorithm and its application to histogram-based image data reduction [J]. Computer Vision Graphics Image Processing, 1990, 49(1):36~51.

- 6 殷勤业,倪志芳,钱世得等. 自适应旋转投影分解法[J]. 电子学报, 1997, 25(4):52~58.
- 7 殷勤业,倪志芳. 自适应旋转投影分解法的快速实现[J]. 通信学报, 1998, 19(1):22~28.
- 8 倪志芳,殷勤业. 用自适应旋转投影分解法进行语音信号的分析[J]. 信号处理, 1997,13(3):54~60.
- 9 张春. 时频分析在引信目标特征提取与 ISAR 成像中的应用[D]. 西安:西安交通大学信息工程研究所, 1997.

傅 弘 1977 年生,硕士研究生. 主要研究方向为图象处理、模式识别及非平稳信号处理技术。



阎鸿森 1945 年生,西安交通大学电子信息与工程学院副院长,教授. 主要研究方向为图象信号处理、移动通信、宽带通信网络。



齐 春 1955 年生,博士,西安交通大学图象处理与识别研究所副教授,1982 年获西安电子科技大学学士学位,1989 年获国防科技大学硕士学位. 主要研究方向为图象检测与跟踪、模式识别、高阶累积量信号处理等. 发表论文 10 余篇。

