

# 基于区域骨架信息的目标识别方法

赵 锋 赵荣椿

(西北工业大学计算机系, 西安 710072)

**摘 要** 提出了一种基于棋盘距离的骨架定义方法, 此骨架结构更接近于区域的实际形状, 骨架的计算更为简单。传统的利用骨架信息识别目标的方法仅使用骨架的结构信息, 而本文提出了骨架结构信息和骨架深度信息相结合的目标识别方法, 在骨架的基础上提出了数种形态特征, 以这些特征作为神经网络的输入进行目标识别得到较好的识别结果。

**关键词** 骨架, 棋盘距离, 骨架深度

## 1 新的骨架定义

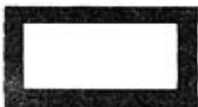
在传统的骨架定义<sup>[1]</sup>中距离  $D(\cdot)$  的定义为欧式距离:

$$D[(i, j), (u, v)] = \sqrt{(i-u)^2 + (j-v)^2} \quad (1)$$

本文的骨架定义采用的距离定义  $D(\cdot)$  为棋盘距离:

$$D[(i, j), (u, v)] = \min(|i-u|, |j-v|) \quad (2)$$

新的骨架定义更接近与区域的实际形状, 例如对于矩形区域及飞机形状区域生成的传统骨架和本文定义的骨架分别如下:



原始区域图象

传统骨架

本文骨架

图 1

显然新的骨架定义较传统骨架定义更接近与区域的实际形状。

新的骨架生成过程也较为简单, 图象邻域定义为 8 连通邻域, 先在区域图象中确定区域图形外的

边界点(这些点是区域外的点,但和区域边界相邻),将这些点的八连通的邻点均变为非区域内的点,即将区域边界向内收缩一个棋盘距离单位;然后再寻找非区域图形的边界点,将这些边界点的邻域点再变为非区域内的点,直至区域内的点均为单厚度。

该骨架较传统骨架的缺点是,不能完全地不失真地恢复原始图象,这是由于距离定义的问题造成的。

定义实际骨架图象  $S$ :

$S(i, j)$  为骨架图象  $S$  的像素点,若  $S(i, j) > 0$  则表示  $(i, j)$  为骨架上的点,  $S(i, j)$  表示该点的骨架深度,即为该点距最近的区域边界的棋盘距离;若  $S(i, j) = 0$  则表示  $(i, j)$  点不为区域骨架上的点。

定义归一化骨架图象  $S_1$ :

在骨架图象的基础上进行归一化处理,可以得到归一化的骨架图象  $S_1$ :

$$S_1(i, j) = S(i, j) / S_{\max} \quad (3)$$

其中  $S_{\max}$  为骨架图象  $S$  中的最大深度值,即:

$$S_{\max} = \max[S(i, j)], (i, j) \in S \quad (4)$$

定义二值化骨架图象  $S_2$ :

对骨架图象  $S$  进行二值化处理,可以得到二值化的骨架图象  $S_2$ ,二值化过程如下:

$$S_2(i, j) = 1, \quad \text{if } S(i, j) > 0 \quad (5)$$

$$S_2(i, j) = 0, \quad \text{if } S(i, j) = 0 \quad (6)$$

骨架  $S_2$  仅仅反映了骨架的结构信息,不包括骨架的深度信息。

## 2 骨架特征

### 2.1 骨架的平均深度

对于骨架图象  $S$  的平均深度值  $A$ , 定义为:

$$A_s = \frac{\sum \sum S(i, j)}{A_k} \quad (7)$$

其中  $A_k$  为骨架图象中骨架深度不为零的像素点数。对于非归一化的骨架则要对  $A_s$  进行归一化:

$$A_s = A_s / A_{\max} \quad (8)$$

其中  $A_{\max}$  为骨架图象中最大的骨架深度。

### 2.2 骨架深度的均方差

对于骨架  $S$  的深度均方差值  $B_s$  为:

$$B_s = \frac{\sum \sum |S(i, j) - A_s|^2}{A_k} \quad (9)$$

其中  $A_k$  为骨架图象中骨架深度不为零的像素

点数。对于非归一化的骨架则要对  $B_s$  进行归一化:

$$B_s = B_s / A_{\max}^2 \quad (10)$$

其中  $A_{\max}$  为骨架  $S$  中最大的骨架深度值。

### 2.3 骨架深度的分布状况

将骨架深度划分为  $N$  个区间, 区间分布为:

$$(0, 1/N], (1/N, 2/N], \dots, (1-2/N, 1-1/N], (1-1/N, 1] \quad (11)$$

第  $i$  个区间内的像素点分布概率  $P_i$  为:

$$P_i = A_i / A_k \quad (12)$$

其中  $A_i$  为深度值在第  $i$  个区间内的像素点数量,  $A_k$  为骨架图象中骨架深度不为零的像素点数量。由这一系列的分布概率特征可以从某方面描述骨架, 本文中选取  $N$  为 4。

### 2.4 一种新的骨架特征的定义

$$M_A = \frac{\sum S(i, j)^m S_1(i, j)}{\sum S(i, j)^m S_2(i, j)} \quad (13)$$

其中  $S$  为实际骨架图象,  $S_1$  为归一化骨架图象,  $S_2$  为二值化骨架图象, 此特征值也为归一化特征量。此特征一定程度地反映了骨架的深度值分布信息, 其分子可以表示骨架深度的  $m$  次方的统计分布, 其分母的作用是进行归一化处理,  $M_A$  的取值范围就在  $[0, 1]$  之间, 其中  $m$  的取值可以取 2, 3, 4, 若太大则此特征的分辨率不高。

仅利用骨架的深度分布信息, 还不能全面地描述目标区域, 因此结合区域的结构信息, 可全面地描述目标特征。

### 2.5 矩特征

设一图象的灰度分布为  $f(x, y)$ , 则其二维  $(p+q)$  阶矩定义为:

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad (14)$$

中心矩定义为:

$$u_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q \times f(x, y) dx dy \quad (15)$$

$$\text{其中 } \bar{x} = m_{10} / m_{00}, \bar{y} = m_{01} / m_{00} \quad (16)$$

对于数字图象  $m_{pq}, u_{pq}$ , 的计算为:

$$m_{pq} = \sum_{i=0}^M \sum_{j=0}^N i^p j^q f(i, j) \quad (17)$$

$$u_{pq} = \sum_{i=0}^M \sum_{j=0}^N (i - \bar{i})^p (j - \bar{j})^q f(i, j) \quad (18)$$

显然中心矩具有平移不变性

HU 利用代数不变量证明了绝对不变量具有旋转不变性,比例不变性以及平行投影的仿射不变性。以下比例为不变量:

$$\eta_{pq} = \frac{u_{pq}}{u_{00}^2}, \gamma = (p+q)/2 + 1 \quad (19)$$

以下为绝对不变量:

$$\begin{aligned} M_1 &= \eta_{20} + \eta_{02} \\ M_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ M_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ M_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ M_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[3(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ M_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ M_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (20)$$

对图象区域计算区域的  $M_1, M_2, M_3, M_4$  4 个不变矩特征,作为目标描述的特征。

### 2.6 最佳拟和椭圆

对于骨架  $S$  来说,其最佳拟和椭圆的长轴  $A$ 、短

轴  $B$  以及椭圆长轴方向  $\theta$  为:

$$A = \left( \frac{u_{20} + u_{02} + [(u_{20} - u_{02})^2 + 4u_{11}^2]^{1/2}}{u_{00}/2} \right)^{1/2} \quad (21)$$

$$B = \left( \frac{u_{20} + u_{02} - [(u_{20} - u_{02})^2 + 4u_{11}^2]^{1/2}}{u_{00}/2} \right)^{1/2} \quad (22)$$

$$\theta = \frac{1}{2} \tan^{-1} \left( \frac{2u_{11}}{u_{20} - u_{02}} \right) \quad (23)$$

计算椭圆的长轴短轴比  $C$  以及目标区域面积和椭圆面积的比  $D$ :

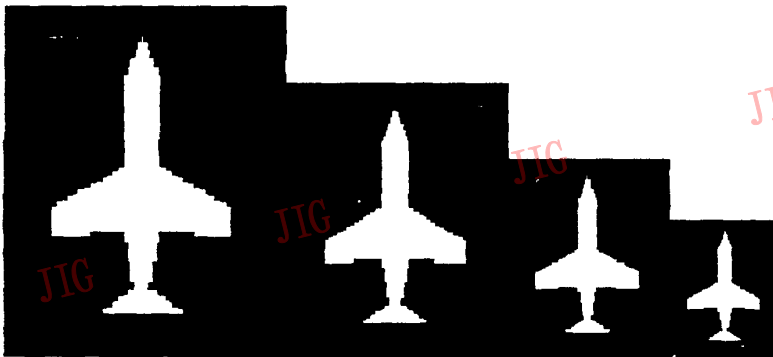
$$C = A/B \quad (24)$$

$$D = SS_1/SS_2 \quad (25)$$

其中  $SS_1$  为目标区域的面积,  $SS_2$  为目标区域最佳拟和椭圆的面积。

### 2.7 各个特征的不变性

$A_i, B_i, P_i (i=1, \dots, N), M_A$  等特征的计算均依赖于骨架的深度值分布信息,并都进行了归一化,显然这些特征具有平移、旋转不变性是确定无疑的,这里我们实验证明它们的比例不变性,对下面一系列比例图象分别计算以上特征值,可以发现这些特征的比例不变性。



(a) (b) (c) (d)

图 2

图 2	$A_s$	$B_s$	$P_1$	$P_2$	$P_3$	$P_4$	$M_A (M=3)$
(a)	0.405 7	0.049	0.276 8	0.452 4	0.172 6	0.098 2	0.536 5
(b)	0.416 6	0.053	0.281 3	0.462 2	0.161 1	0.093 7	0.543 1
(c)	0.390 8	0.047	0.262 7	0.444 4	0.179 9	0.100 3	0.5501
(d)	0.420 1	0.046	0.290 1	0.467 9	0.161 6	0.095 4	0.547 8

对目标区域的不变矩  $M_1, \dots, M_7$  的不变性是早已证明的。

显然区域的最佳拟和椭圆的长短轴长度比  $C$  和目标区域与最佳拟和椭圆的面积比  $D$  也同样具有旋转、平移、比例不变性。因此用这些特征组成的特征向量具有旋转、平移、比例不变性,基于此特征向量的目标识别是可行的。

本文将  $A_i, B_i, P_i (i=1, \dots, N, N=4), M_A, M_1, M_2, M_3, M_4, C, D$  一系列特征的部分组合起来生成目标识别的特征向量,以此特征向量作为分类系统的输入。

### 3 神经网络识别

本文采用 B-P 网进行目标的识别分类,所识别

的目标主要是航片中军事目标,主要包括:飞机、军舰、坦克、油库等,此 B-P 网的输入是前面提到的区域描述特征向量,实例中采用  $A_i, B_i, P_i (i=1, \dots, N, N=4), C$  组成特征向量,B-P 网采用三层结构,输入层神经元个数由输入特征向量的维数决定,为 7 个,中间层由 12 个神经元组成,输出层的神经元的个数由识别目标的分类数决定,为 4 个,输入层神经元的特征函数为直通函数即输入和输出相等且输入只有一个,中间层神经元的特征函数为 S 型特征函数,输出层神经元的特征函数为线性特征函数,神经网络的输出只有一个位为 1,其余为零,为 1 的位表示目标的类别。经训练后 B-P 网的权值  $W_1, W_2$  及阈值  $B_1, B_2$  如下:

$$W_1 = \begin{bmatrix} -0.4967 & -0.5259 & 0.3579 & 0.7342 & -0.4038 & -0.1266 & -0.5534 \\ -0.7882 & -0.1205 & 0.6247 & 0.4482 & 0.1288 & -0.1291 & 0.8377 \\ 0.2945 & -0.1204 & -0.3246 & 0.2931 & -0.0121 & 0.6863 & 0.6421 \\ 0.1961 & 0.2324 & -0.4447 & -0.3578 & 1.3067 & -0.1498 & -0.2358 \\ -0.1504 & 0.1177 & 0.3378 & 0.1841 & 0.4515 & -0.0946 & -0.8750 \\ -0.0775 & 0.5651 & -0.1785 & 0.5680 & -0.4460 & -0.1342 & 0.3428 \\ 0.2509 & 0.4326 & 0.0336 & -0.4677 & 0.8197 & -0.0670 & 0.2862 \\ -0.7517 & 0.0369 & 0.3911 & -0.0659 & -0.0704 & 0.5823 & 0.1011 \\ -0.4162 & -0.4956 & 0.4604 & 0.1208 & -0.1043 & 0.0979 & 0.3231 \\ -0.0285 & 0.4282 & -0.3094 & 0.0337 & -0.6545 & -0.2188 & 0.4821 \\ 0.2881 & -0.1051 & -0.3222 & -0.3671 & 0.4119 & 0.4242 & 0.3310 \end{bmatrix}$$

$$B_1 = [0.7961 \quad -0.4584 \quad -0.3821 \quad -0.2006 \quad 0.0680 \quad 0.1591 \quad 0.7036 \quad -0.4800 \quad 0.6186 \quad -0.6012 \quad -0.2043 \quad 0.0479]$$

$W_2 = \text{Columns 1 through 7}$

$$\begin{bmatrix} 0.7547 & 0.3331 & 0.1987 & 0.2016 & -0.5969 & -0.1149 & 0.3095 \\ -0.2584 & 0.4306 & 0.0414 & 0.1802 & 1.1382 & -0.3230 & 0.0622 \\ -0.2143 & -0.8322 & 0.2090 & 0.3225 & -0.3173 & -0.1428 & -0.1631 \\ -0.5344 & -0.0386 & 0.0414 & -0.2223 & -0.4016 & -0.5938 & -0.2187 \end{bmatrix}$$

$\text{Columns 8 through 12}$

$$\begin{bmatrix} -0.5448 & 0.2594 & -0.2561 & 0.3920 & -0.4424 \\ 0.6210 & 0.0420 & -0.2337 & -0.3963 & -0.1240 \\ 0.0777 & -0.2861 & 0.0813 & -0.3134 & 0.0270 \\ -0.2117 & 0.0478 & 0.0441 & 0.5707 & -0.1034 \end{bmatrix}$$

$$B_2 = [0.0258 \quad -0.2670 \quad 0.6072 \quad -0.2452]$$

应用此神经网络进行小范围的目标识别,识别效果较好,识别的正确率为 100%。作者将此方法和

传统的基于区域不变矩特征的目标识别方法进行了比较,传统的不变矩识别方法中,特征的计算较为

复杂,特征的归一化较为困难,特征的分辨率较低,同样识别本文的 4 类目标,识别的结果较本文的方法差,如对某飞机图象作为输入时网络的输出分别为:(未进行最大,值选择前)

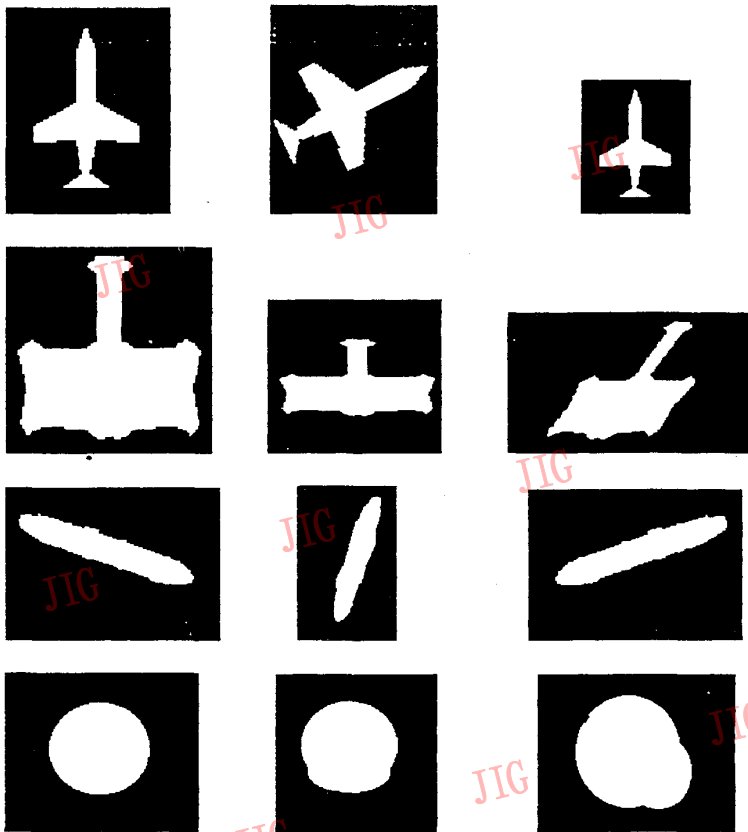
本文方法 [0.967 8,0.012 9,0.023 6,0.034 7]

传统方法 [0.6134,0.3283,0.4359,0.3871]

显然本文的方法识别能力较强,并且采用本文方法神经网络的权值调整过程较传统不变矩方法时

间缩短一半左右。不过仅采用本文提出的第 1 至第 4 种特征进行目标识别对骨架深度分布相近,但骨架形状相异的目标不具有识别能力,因此在此类目标识别过程中应适当加入传统不变矩特征进行识别(本文的识别实例未使用矩特征)。

#### 4 训练样本



#### 5 结论

本文在棋盘距离的基础上定义骨架,此骨架结构更接近于区域的实际形状,骨架的计算更为简单。将骨架结构信息和骨架深度信息相结合,提出了一组进行目标识别的特征,以此特征进行组合生成特征向量作为目标识别的依据可以得到较好的识别结果。实验发现仅依据骨架的深度分布信息提取的特征在很多情况下也可以得到正确的分类结果。

#### 参考文献

- 1 赵荣椿等. 数字图象处理导论. 西安:西北工业大学出版社.
- 2 米航宇等. 机器视觉中矩特征的快速抽取方法. 模式识别与人工智能,1989,2(4).
- 3 金琪,戴汝为. 矩 不变矩 角度矩及傅里叶描绘子. 模式识别与人工智能,1993,6(3).
- 4 Cho-Huak The and Roland T. Chin . On digital approximation of moment invariants. Computer Vision, Graphics, and Image Processing 33, 318~326 (1986).
- 5 Jeffrey Wood. Invariant pattern recognition: a review. Pattern Recognition,1996,29(1):1~17.



**赵 锋** 博士生,1992年毕业于北京理工大学电子工程系,获学士学位,1995年毕业于西安电子科技大学微波电信工程系,获硕士学位,1998年毕业于西北工业大学计算机系,获博士学位。研究兴趣:图象处理、模式识别、数据库系统、网络系统集成。

**赵荣春** 教授,博士生导师,1960年毕业于中国人民解放军军事工程学院航空系,1982.2—1990.8曾到英国 Sarrey 大学进修访问。从事信号、图象处理、模式识别和计算机视觉等方面研究。

## A New Target Recognition Method Based on The New Feature of Region Skeleton

Zhao Fei, Zhao Rongcun

(Dept. of Computer Science & Engi Northwestern Polytechnical University, Xi'an 710072)

**Abstract** In the paper, A new skeleton definition of region based on chessboard distance is submitted, which is more close to the true figure of region than traditional skeleton, the algorithm is brief. Instead of only using the structure information of skeleton in traditional recognition method based on the skeleton information, The depth information and the structure information of skeleton are combined in the algorithm. A series of new feature to recognize the target is submitted in this paper also. The feasibility of the algorithm is proved by the experiment results.

**Keywords** Skeleton, Chessboard distance, The depth of skeleton

## 艾尔莎康柏再度携手

德国著名图形卡厂商艾尔莎(ELSA)日前宣布,康柏电脑公司将与其再度携手,扩展获奖的 ELSA GLoria Synergy AGP 图形加速卡在康柏新型专业工作站 AP500 中的应用。这是继康柏在其 AP400 和 AP200 专业工作站中集成艾尔莎图形解决方案之后的又一次合作。

GLoria Synergy 是一种高性能的 2D 和 3D 图形加速卡,在 Windows NT3.51 和 4.0 环境下支持多显示器。它与康柏专业工作站的结合将为计算机辅助设计(CAD)、数字内容创建(DCC)、电子设计自动化(EDA)、几何图形信息系统以及金融市场带来无限生机。ELSA GLoria Synergy 前不久荣获了 CADALYST 杂志评出的五星奖,被誉为“2D 图形卡的最新成就”。

除了众所周知的硬件性能外,艾尔莎在软件方面也颇具优势。它为康柏专业工作站产品在 AutoCAD、3D Studio MAX/VIZ、Adobe Photoshop、MicroStation、OpenGL 和金融等方面的应用提供了高性能和高效率。艾尔莎的 OpenGL 技术支持抗锯齿功能,CDRS-03 性能测试值为 32.3,而 Dynamics Picture 202 为 14.94;ATI Rage Pro Turbo 仅为 14.1,相差一倍以上。而且,ELSA POWERdraft 是全球唯一的支持 AutoCAD R13 和 R14 硬件加速的驱动软件;ELSA MAXtreme 针对 3D Studio MAX R2、3D StudioVIZ R2 可以为用户提高 3D 图形性能;ELSA Desktop 更是可为用户需要的金融、CAD 和图形处理应用提供灵活机动的多显示器的实用程序。(汪虹)