

一种航空照片中的小目标识别方法

王秋让* 赵荣椿** 郑南宁*

(* 西安交通大学人工智能与机器人研究所, 西安 710049, ** 西北工业大学计算机科学与工程系, 西安 710072)

摘要 提出了一种能够从具有复杂背景的航空照片中识别出小目标的模式识别方法。此方法先对图象进行二值化, 其中二值化所需的门限值是通过将图象的灰度直方图进行形态滤波并结合小波变换而得到的, 然后从二值图象中提取出各非零区域, 接下来对原始图象中与这些区域位置对应的灰度区域进行第二次门限化, 最后对各小二值图象块进行平移、尺度和旋转不变性变换, 求出每个变换图象块的 n 阶几何矩做为其不变性特征输入 BP 网完成各目标的识别。

关键词 形态滤波, 小波变换, 自动门限化, 几何矩

1 引言

要从复杂的环境中目标识别出来, 最大的困难不是来自识别本身, 而是来自预处理, 即如何完整、准确地将每个独立的、面积足够大的模式从复杂的背景中分割出来。目前从分割出来的模式集合中识别出目标的技术相对来说比较成熟。

一个基于门限化的自动模式识别系统必须能够自动地进行门限选择, 而一般的门限自动选择方法都假定背景和前景各占一个灰度级区间, 而实际上图象的灰度级可能集中在多个不同的区间上, 目标可能占据其中的一个或多个区间, 因此, 将遇到所谓多门限化图象分割的问题。这里, 我们提出了一种新的多门限选择方法, 这种方法先对灰度直方图进行形态滤波, 再进行小波分解从而找出直方图的每个极小值点作为各区间的分割点。

二值化必然造成图象信息的丢失, 为了将二值化的这一影响减少到最低程度, 我们采取的策略是: 先对整幅原始图象进行门限化, 称之为粗门限化, 其目的在于分割出与周围灰度明显不同的图象区域, 然后对相应位置上的原始灰度图象区域分别进行门限化, 以充分利用图象的局部信息, 我们称这次门限化为细门限化, 其目的在于识别。这一做法实际上将

门限化操作大大地向后推延了。

一个模式识别方法想要能投入实际应用, 就应该具有平移、尺度和旋转不变性。本文提出的方法先对细门限化的结果图象块依次进行平移、尺度和旋转不变性变换, 再求变换图象的 $1 \sim n$ 阶几何矩, 以此作为 BP 网的输入完成各模式的识别。用这种方法来识别复杂环境中的模式, 其鲁棒性较不变矩特征法强, 速度较高阶神经网络法快。

2 灰度图象的数学形态学

数学形态学最初是针对二值图象的处理而提出的, 现在它已被成功地应用于灰度图象的处理。

数学形态学是进行图象处理的一种方法, 其基础是集合理论, 它用集合论的观点来看待形态。二值图象的数学形态学把形态看成是二元组的集合, 每一个二元组代表一个非零象素点的坐标, 一幅图象就是非零象素点的一个集合。而灰度图象的数学形态学则把形态看成是三元组的集合, 每个三元组代表一个象素点的坐标和这一象素点的灰度值。每个形态操作都是一个二元运算, 两个操作数分别称为原始数据和结构元素, 结构元素相当于卷积运算或傅立叶变换的核。数学形态学有 4 种基本运算: 扩

张、侵蚀、开启和闭合。

设一图象的灰度函数为 $f(x, y)$, 结构元素为 $k(x, y)$, 则 4 种基本操作定义如下:

扩张: $f \oplus k(x, y) = \sup[f(x-i, y-j) + k(i, j)]$,

其中 $i, j \in k, x-i, y-j \in f$;

侵蚀: $f * k(x, y) = \inf[f(x+i, y+j) - k(i, j)]$,

其中 $i, j \in k, x+i, y+j \in f$;

开启: $f \cdot k(f * k) \oplus k$;

闭合: $f \cdot k(f \oplus k) * k$;

对灰度图象的形态操作来说, 常用的灰度结构元素有球体、圆锥体、圆柱体和抛物曲面体等。图 1 为一个用球形结构元素对一灰度图象进行形态操作的示意图, 为了方便表示起见, 用定义在 X 坐标上的一维函数来代表定义在 $X-Y$ 平面上的二维灰度图象 $f(x, y)$, 灰度曲面用一维函数 $z = 8\sin(\frac{\pi x}{12}) +$

$10\cos(\frac{\pi x}{25}) + 4\sin(\frac{\pi x}{25}) + 7\cos(\frac{\pi x}{50}) + 10\sin(\frac{\pi x}{100})$ 来表示, 球形结构元素的半径为 10, 实线表示原函数, 虚线表示形态操作的结果。由图可知, 扩张操作的结果为圆球在图象的灰度曲面上滚动时圆心的轨迹, 而侵蚀操作的结果为圆球紧贴灰度面的下表面滚动时圆心的轨迹; 闭合运算的结果为圆球在图象的灰度曲面上滚动时各个时刻球面上最低点的集合, 其作用在于补谷, 而开启运算的结果为圆球紧贴灰度曲面的下表面滚动时各个时刻球面上最高点的集合, 其作用在于削峰。显然, 开启和闭合的这种作用可以用来对一曲线或曲面进行平滑, 实际上, 形态滤波就是对信号进行开启—闭合运算以滤掉高频分量。形态滤波也被称为滚球滤波。

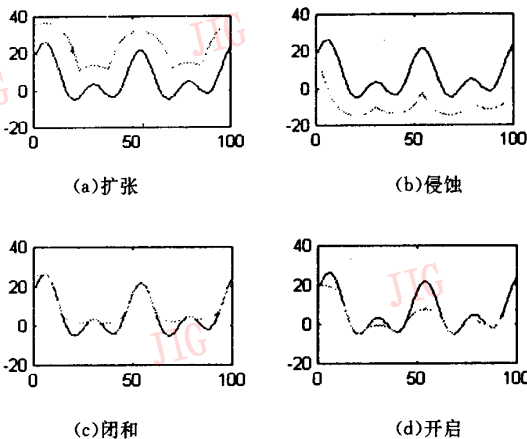


图 1 灰度图象的形态学操作示意图

二值图象可以认为是一种特殊的灰度图象, 因此为灰度图象定义的形态学 4 种基本操作仍然适用于二值图象, 只不过 $f(x, y)$ 和 $k(x, y)$ 的取值变为 0 或 1 而已。

对二值图象的操作, 用区域形状的观点比用灰度值的观点更容易理解。用区域的观点来看形态学操作, 扩张与侵蚀分别引起区域的膨胀与收缩; 开启运算起的是分离作用, 比结构元素小的孤立部分都将被过滤掉, 而其情节保持不变; 闭合运算起着连通补缺的作用, 其主要情节保持不变。

3 小波变换结合数学形态学的门限化方法

迄今为止, 人们已提出了许多门限选择方法, 大多数方法都需要事先确定门限化的级数, 即需要人的监督, 有的方法尽管不需要事先确定门限化级数, 它们也往往要么得到伪门限或要么丢掉真门限, 二者之间很难协调。因为我们的门限化分两个阶段进行, 在粗门限化阶段, 对门限化精度要求不高, 在这种情况下, 我们提出了一个小波变换结合数学形态学的粗门限化方法, 它能快速地、无监督地按图象灰度级的分布找到多个门限值, 把象素相对集中的各灰度区间分开。

此方法的思想是找到图象灰度直方图包络线的所有极小值点, 它们的横坐标(即其对应的灰度值)便是各区间的端点。具体来说, 就是先对灰度直方图进行闭合运算得到其包络线, 再通过开启—闭和运算对此包络线进行平滑, 注意, 此处闭合和开启两运算的次序不可颠倒。尽管对其进行了形态滤波, 此包络线并不光滑, 所以再对其进行小波分解, 其低频分量便是光滑的了。设包络线低频分量的极小值为 s , 小波分解的阶数为 n , 则门限值 $T = s \times 2^n$, 门限化的精度为 2^n 个灰度级。在粗门限化阶段对精度要求并不高。

如果省掉形态滤波这一步, 只要对灰度直方图进行小波分解的阶数足够高, 所得的低频分量也会是光滑的, 但这样做会造成门限值定位不准, 例如若对灰度直方图的包络线进行 5 阶小波分解之后, 其低频分量才变得光滑, 则门限选择的精度为 32, 即门限值只能是 32 的倍数, 这一精度在大多数情况下是不可接受的。

4 图象分割

此处的图象分割指的是将图象中与周围灰度明显不同的物体提取出来。为了及早地将不可能的模式排除掉,这一阶段我们将不把与目标大小相差太远的模式抽取出来,例如,假设我们的目标是照片中的飞机和轮船,那么图象中具有明显不同灰度的区域如果像素数少于某一值(如 20),那么我们就不予理睬,实际上,就算其是一目标模式,也很难将其识别出来。

最简单的分割方法是对灰度图象直接门限化,将各灰度区间内的图象区域提取出来。可是如果图象中存在较目标小得多的区域,则直接门限化也会把它们抽取出来,这是我们所不希望的。

设第三节的门限自动选择方法将原始图象的灰度区间分割成 n 个子区间,则我们的方法先对原始灰度图象进行 n 次二值化得到 n 幅二值图象:第一次二值化将灰度值位于第一个灰度子区间的像素二值化为 1,其余像素二值化为零,得到第一幅二值图象;…;第 i 次二值化将灰度值位于第 i 个灰度子区间的像素二值化为 1,其余像素二值化为零,得到第 i 幅二值图象,…;我们称这些二值图象为原始二值图,然后分别对每一幅二值化图象进行开启操作得到一幅开启图象,开启操作的结构元素要选得足够大,足以去掉所有较目标模式尺寸小得多的细碎区域,但至少能使较大的区域留下一个像素点,结构元素如果选得太大,则可能将目标也滤掉,如果选得太小,则会保留很多的噪声区域。结构元素是根据图象中目标的最小尺寸事先选定的。用开启图象定位每个大区域在原始二值图中的位置,开启图象中的每个区域只需给出其中任意一点的坐标位置。在原始二值图中,以这些位置上的点为初始点进行区域扩张,便得到一幅除去了细碎区域的“干净”图象。将这幅图象中的各区域用最小的矩形框起来,得到各矩形的位置和尺寸,最后用这些矩形将原始灰度图象中相应的区域截下来供以后识别之用。

5 小灰度图象块的门限自动选择

前一步得到的矩形图象块所包含的内容比较单一,像素主要分布在 2 个灰度区间上,在这种情况下,传统的门限自动选择方法能够找到合适的门限值。在这里我们采用的是 P. Sahoo 等人提出的

Renyi 熵门限自动选择方法^[1]。

在二值化过程中会遇到这样一个问题:由于原始灰度图象中噪声的干扰,如果直接用每一点本身的灰度值与门限值比较来决定二值化为 0 还是 1,则二值化图象中也会包含许多小麻点。我们的办法是用每个像素某一邻域的灰度平均值与门限值做比较来决定二值化为 0 还是 1,这一措施很好地抑制了噪声的干扰问题。

6 平移、尺度和旋转不变性目标识别

迄今为止,人们已经提出了许多具有平移、尺度和旋转不变性的模式识别方法,最常用的是高阶网法和不变矩特征法。

高阶网模式识别方法必须解决权值组合爆炸的问题。在现有的解决组合爆炸问题的方法中,万里青等人^[6]用 Log-Spiral 变换将尺度和旋转转化为平移的方法较为简单实用。我们的系统没有采用高阶网的方法,因为图象块的尺寸不确定,这使得网络输入层结点数必须选得足够大以满足最大尺寸的图象块,因此会造成效率下降。

不变矩特征法都先要将图象映射到一个单位区域上,然后再求各阶不变矩。如几何不变矩法要将图象映射到 $\{(x, y) | -1 \leq x \leq 1, -1 \leq y \leq 1\}$ 上,而 Zernike 矩法的单位区域为 $\{(x, y) | x^2 + y^2 \leq 1\}$ 。这个映射操作对噪声极为敏感,因为映射时需要找出距离重心最远的点将其映射到单位区域的边界上,如果最远的点为一噪声点,则之后得出的不变矩将会有很大的误差。

本文的方法是对每个图象块依次进行平移、尺度和旋转不变性变换,求出变换图象块的若干阶几何矩作为 BP 网的输入完成模式的识别。

一幅大小为 $M \times N$ 的二维图象 $f(x, y)$ 的 $(p+q)$ 阶几何矩定义为:

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y); \text{ 其中 } p, q = 0, 1, 2, \dots, n$$

阶几何矩共有 $n+1$ 个,分别为 $m_{0n}, m_{1(n-1)}, m_{2(n-2)}, \dots, m_{n0}$ 。

$f(x, y)$ 的重心 (\bar{x}, \bar{y}) 定义如下:

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

则平移变换定义为:

$$T(f(x, y)) = f(x - \bar{x}, y - \bar{y})$$

尺度变换 S 的操作对象为平移之后的图象,其

定义为:

$S(T(f(x,y)))=f(\frac{x-\bar{x}}{\alpha}, \frac{y-\bar{y}}{\alpha})$, 其中 α 为尺度因子。

尺度变换的目的是使变换后图象的目标像素数为一定数 β , 目标像素指的就是灰度值为 1 的像素。 α 与 β 的关系为 $\alpha=(\beta/m_{00})^{\frac{1}{2}}$ 。[3] 我们选择 $\beta=1000$ 。

旋转变换利用了这样一个事实: 给定一矢量集, 其协方差矩阵的最大特征值所对应的特征向量指向方差最大的方向。设模式的最大方差方向与 X 轴的夹角为 θ , 则旋转变换只需将坐标系逆时针旋转 θ 或 $\theta+\pi$ 就可使模式的最大方差方向与 X 轴重合。因此旋转变换定义为:

$$R(S(T(f(x,y))))=f(\cos\varphi\frac{x-\bar{x}}{\alpha}-\sin\varphi\frac{y-\bar{y}}{\alpha}, \sin\varphi\frac{x-\bar{x}}{\alpha}+\cos\varphi\frac{y-\bar{y}}{\alpha})$$

其中, $\varphi=\theta$ 或 $\theta+\pi$, 即每个图象块经过旋转变换后产生两个变换图象块, 二者方向相反。

因为 θ 的推导过程较为冗长, 此处不再列出, 详细推导请参见[3]。

经过不变性变换之后所求得的几何矩具有平移、尺度和旋转变换不变性。将每幅变换图象的 $0\sim n$ 阶共 $\frac{(n+1)(n+2)}{2}$ 个几何矩作为 BP 网络的输入可以快速准确地识别出目标。

7 实验结果

我们以图 2 所示的一幅航空照片为例来说明本方法的有效性。

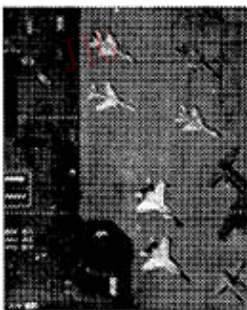


图 2 机场的航空图象

在门限选择时, 我们选取的小波基为双正交小波, 其滤波器 h 及 \tilde{h} 的系数如表 1 所示。

表 1 滤波器 h 及 \tilde{h} 的系数(分别为 9 拍和 7 拍)

n	0	± 1	± 2	± 3	± 4
$h(n)/\sqrt{2}$	0.848 5	0.353 6	-0.070 7	0	0
$\tilde{h}(n)/\sqrt{2}$	0.858 6	0.368 7	-0.075 8	-0.015 2	—

图 3 为对图 2 进行小波变换结合形态滤波的门限选择结果。图 3(a)为其灰度直方图, 灰度区间为 $[0, 255]$; 图 3(b)为以半径为 6 的球状结构元素对 3-1 进行形态滤波的结果; 图 3(c)为对图 3(b)进行二阶小波分解后的低频分量。从图 3(c)中找出的极小值点的横坐标为 14, 故门限值为 $14 \times 4 = 56$, 这个结果是相当准确的。

图 3(a)所示的灰度直方图具有特殊性, 表现为象素在 $[130, 255]$ 这一很长的灰度区间上呈相对均匀的分布, 这很可能是由于图象中某类模式占据了此灰度区间引起得, 因此门限选择方法也应能提取出这一区间。一种方法是求出灰度直方图低频分量各点 $f(i)$ 的梯度 $\nabla f(i), i=0, 1, \dots, 256/2^n-1$, 梯度的过零点对应着图 3(c)中的极值点, 左正右负为极大值点, 左负右正为极小值点。在求各点梯度时, 一旦 $|f(i)-f(i-1)| < \alpha$, 则置 $\nabla f(i)=0$, 其中 α 为一常数, 则使梯度左负右正的点及使梯度保持为零的区间端点乘以 2^n 便为所求的灰度门限值。图 3(d)为图 3(c)所示的灰度直方图低频分量的梯度曲线, 其中仅在横坐标 14 处梯度由负变正, 梯度在 $[33, 63]$ 上保持为零, 故门限值为 56、132 和 252, 其中 252 距最大灰度 255 极近, 有理由认为其不是一真正的门限值。

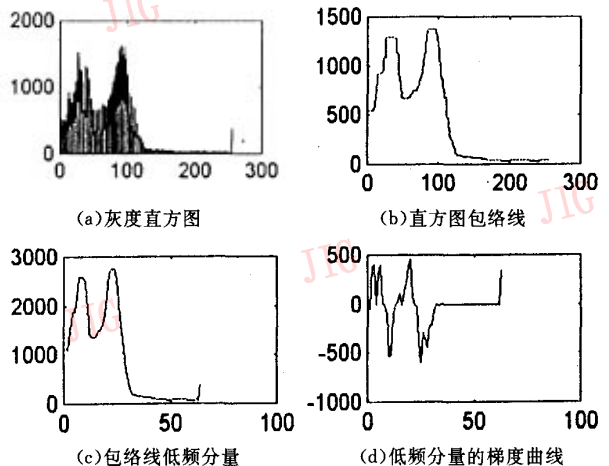


图 3 小波变换结合形态滤波的门限选择

接下来我们以 56, 132 为门限得到 3 副二值图象, 分别对其进行分割以确定目标的位置及其在原

始图象中的位置。图 4 为依据灰度区间 $[132, 255]$ 所对应的二值图象,从图 2 中分割出的矩形图象块。限于篇幅,此处略去依据另外 2 幅二值图象分割出的图象块。

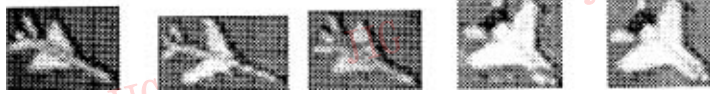


图 4 从图 2 中分割出的矩形图象块



图 5 小图象块的二值化结果

最后用 BP 网对变换后的图象块进行识别。我们选择 BP 网的输入层结点数为 45(分别对应 0~8 阶几何矩),隐层结点数为 20,输出层结点数为 5(分别对应一种型号的飞机),以图 6 中各图象块为样本对神

图 5 为图 4 中各灰度图象块的二值化结果。

图 6 给出了图 4 中各图象经过不变性变换之后的结果(每个图象块在图 5 只给出了其中一个变换结果,另一个因为仅仅与之方向相反,故略)。

经网络进行训练。若直接对图 5 中的图象块进行识别,正确率为 100%。因为缺少合适的照片,所以训练网络的样本很少,也未做非样本空间中目标的识别率实验。



图 6 图 5 各图象块不变性变换结果

8 结 论

本文提出了一种全新的目标识别的体系结构,这种体系结构特别适于复杂背景中的目标识别,它利用形态滤波和小波变换进行全局门限的自动选择,利用门限化后的二值图象对原始图象分割,对分割出的图象块进行第二次门限化,对所得的二值图象进行平移、尺度和旋转不变性变换,将变换后二值图象的几何矩作为图象的特征输入 BP 网最终完成了对目标的识别。本文的方法非常强调预处理阶段的工作,因为我们认为较好的预处理结果对识别结果是至关重要的。

参 考 文 献

- 1 Sahoo P, Wuukins C, Yenger J. Threshold Selection Using Renyi's Entropy. Pattern Recognition, 1997, 30(1).
- 2 万里青,赵荣椿,孙隆和. 不变性目标识别方法研究信号处理, 1996, 12(2).

- 3 Cem Yuceer, Kemal Oflazer. A Rotation, Scaling, and Translation Invariant Pattern Classification System. Pattern Recognition, 26, 687~710.
- 4 Jackway Paul T, Mohamed Deriche. Scale-Space Properties of the Multiscale Morphological Dilation-Erosion. IEEE Trans, 1993, PAMI-18: 38~51.
- 5 Alireza Knotanzad, Lu Jiin-Her. Classification of Invariant Image Representations Using a Neural Network. IEEE Trans, 1990, ASSP-38: 1028~1038.
- 6 Perantonis Stavros J, Lisboa Paulo J G. Translation, Rotation, and Scale Invariant Pattern Recognition by High-Order Neural Networks and Moment Classifiers. IEEE Trans. Neural Network, 1992, 3: 241~251.



王秋让 1993年毕业于哈尔滨科技大学,1998年于西北工业大学计算机系获博士学位。现于西安交通大学人工智能与机器人研究所做博士后研究。主要从事计算机视觉、图象处理和模式识别方面的研究工作。



赵荣椿 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。

郑南宁 教授, 博士生导师, 1985年获日本庆应大学博士学位, 1996年被评为中国青年科学家, 曾发表论文多篇。现主要从事计算机视觉和模式识别、神经网络、智能系统并行处理的研究。

Recognition of Small Targets in Aerial Photos by two-time Thresholding

Wang Qiurang*, Zhao Rongchun**, Zheng Nanning*

(* Inst. of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049)

(** Dept. of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072)

Abstract A system which can recognize compact targets from complex environments independently of objects' position, size and orientation is proposed. The system first selects threshold(s) automatically for the original image to get several binary images. In the simplest case, one threshold and, therefore, two binary images are got, one of which is got by setting 1 the pixels whose gray value is larger than the threshold, setting others pixel 0, and the other is completely opposite to this one. The binary images are used to determine the patterns' position. The system then uses the position data to extract rectangle areas from the original gray image. For every extracted area, a threshold is selected to get two binary images which are respectively transformed independently of position, size and orientation. Finally the geometrical moments are calculated for the transformed binary image to input to a BP as the image's features to recognize it.

Keywords Morphology filtering, Wavelet transform, Automatic thresholding, Geometrical moments

VTEL 公司首家推出全系列产品支持 H. 323 标志网络化数字视讯时代的到来

全球网络化数字视频通讯系统的领导者—VTEL 美国视讯公司宣布它已成功地开发出支持在 Internet 协议 (IP)、电话或其它混合网络上进行视讯传输的软件。这项技术使 VTEL 从桌面到大型会议室的所有视讯产品都能同时支持 H. 323 和 H. 320 (ITU 国际电信联盟协议), 同时也可以运行于 ATM 和帧中继网, 使客户能够在各种网络上安装视讯系统。目前, VTEL 是世界上唯一拥有这一技术的公司。这项革命性的突破将于 10 月底正式启用。