

基于 RBF 神经网络的复杂背景下的运动目标检测^{*}

赖作镁^{1,2} 王敬儒¹ 张启衡¹

(中国科学院光电技术研究所国家 863 计划光束控制重点实验室 成都 610209)¹

(中国科学院研究生院 北京 100039)²

摘要 由于运动摄像机的存在使得复杂背景下的运动目标检测问题更加复杂,根据场景中目标与背景具有不同的运动、任意场景可以分成不同的运动区域这一基本事实,提出一种新的基于 RBF 神经网络的运动目标检测算法。运动补偿后求参考帧与补偿后的当前帧之间的光流,联合当前像素坐标及其灰度值得到五维特征向量作为 RBF 网络的输入,RBF 网络学习算法通过最小化由 Bayesian 理论和能量最小化理论导出的损失函数实现。学习矢量量化方法修正网络的中心,收敛后网络的输出就是运动目标区域。试验结果证明了算法的有效性。

关键词 目标检测,光流,径向基函数网络,学习矢量量化,运动补偿

Moving Object Detection under Complex Background Based on Radial Basis Function Neural Network

LAI Zuo-Mei^{1,2} WANG Jing-Ru¹ ZHANG Qi-Heng¹

(Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209)¹

(Graduate School of the Chinese Academy of Sciences, Beijing 100039)²

Abstract It is well known that moving object detection under complex background becomes more difficult because of the existence of moving cameras. According to the fact that background and objects have different motion, the moving scene can be decomposed into different regions with respect to their motion information by means of an radial basis function(RBF) network learning scheme. After robust global background motion compensation, five-dimensional (5-D) feature vectors include pixel intensities, current pixel coordinates and pixel dense optical flow field extracted from image sequences are treated as the inputs of the RBF network. The learning algorithm for the network minimizes a cost function derived from the Bayesian estimation theory. Learning vector quantization method modifies the center of network, and each output unit of the network is associated to a moving object. Experimental results indicate the algorithm's validity after many sequences are tested.

Keywords Object detection, Optical flow, Radial basis function network, Learning vector quantization, Motion compensation

1 引言

复杂背景下的运动目标检测有许多相似的应用,比如视频监控,视频对象分割等。由于运动摄像机的存在使得复杂背景下的运动目标分割问题变得更加困难,复杂背景一般包括山川、河流、树木等,背景本身也有可能运动。这时,鲁棒的全局背景运动补偿技术通常是目标检测的第一步。

很多文献对运动目标检测问题进行了研究,运动分割^[1]方法是比较常用和研究得较多的方法之一。然而,如果仅仅采用运动这一时域特征,很难得到令人满意的结果。而且运动分割方法有分割结果不精确的缺点,它基本能把目标定位出来,一般来说分割得到的目标不完整,目标边缘特征不明显等,所以它仍然是个急待研究的课题。运动分割中引入 Bayesian 理论^[2]的方法弥补了很多运动分割的固有缺点,也

满足了不少具体的应用需求。也有文献充分利用 RBF 网^[3,4]具有很好的函数拟合能力和模式分类能力的优点,将 RBF 网成功应用于杂波背景下弱小目标的检测。

本文提出一种新的基于 RBF 网络的复杂背景下的运动目标检测算法。为了满足更普遍的应用,当存在摄像机运动情况下,先通过鲁棒的背景运动补偿算法补偿背景的运动,得到稳定的补偿后的当前帧,然后在补偿后的当前帧与参考帧之间求光流。结合当前像素点的灰度以及坐标值得到五维特征向量作为 RBF 网络的输入,RBF 网络最小化由 Bayesian 理论和能量最小化理论导出的损失函数达到平滑光流场和光流场聚类的目的。RBF 网络的参数估计采用经典的学习矢量量化方法^[5],经过多次迭代以后,RBF 网络的输出就是运动目标的轮廓。系统的流程图如图 1 所示。

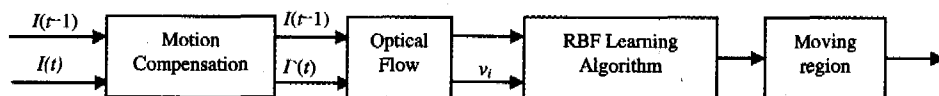


图 1 系统流程图

^{*}基金项目:国家 863 计划项目资助课题(2003AA823050)。赖作镁 博士研究生,从事图像处理目标的识别与跟踪研究;王敬儒 研究员,博士生导师,主要研究方向为光电探测、图像处理、目标识别与跟踪等;张启衡 研究员,博士生导师,主要研究领域为目标探测与信号处理。

2 特征向量的提取

首先采用 Simon Baker 和 Iain Matthews^[6]提出的运动补偿算法对背景进行补偿,这是一种基于 Levenberg-Marquardt 迭代的分辨率方法。初始运动估计采用特征点法,摄像机

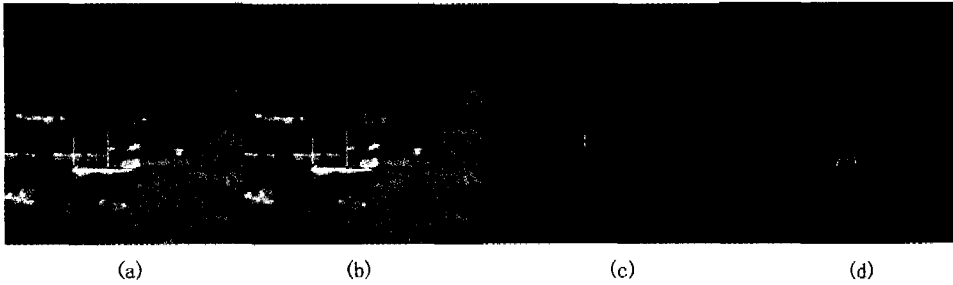


图2 运动补偿效果图

理论上,补偿后参考帧图像 I_{t-1} 与补偿后的当前帧图像 I'_t 应该相对静止。这时可以采用摄像机静止情况下的目标检测算法如帧间差分法来检测目标,由于传感器噪声和算法本身的误差等影响,很难保证所有帧间差分后都能很好地突出目标。这里必须采用标准 H-S 光流算法^[7]求补偿后残余运动场,即求补偿后的当前帧与原始帧之间的光流。H-S 光流在基本光流方程(2)上增加全局约束(1)。

$$E_s = \iint_D (u_x^2 + u_y^2) + (v_x^2 + v_y^2) dx dy \quad (1)$$

$$E_c = \iint_D (I_x u + I_y v + I_t)^2 dx dy \quad (2)$$

最小化 $E_c + \lambda E_s$ 对每个像素点,得到两个方向的运动 u_i 和 v_i ,再加上像素坐标 x_i, y_i 和像素点的灰度值 I_i 一起组合成五维特征向量,作为 RBF 网络的输入。五维特征向量可以表示成:

$$u_i = [u_i, v_i, x_i, y_i, I_i] \in \begin{cases} s_k & \text{静止区域中的特征向量} \\ m_k & \text{运动区域中的特征向量} \end{cases} \quad (3)$$

$k=1, 2, \dots, L$

假定待检测场景含有 L 个运动区域,接下来的问题就转化为特征向量的分类问题。

3 Bayesian 理论与能量最小化

假定 $(\tilde{m}_k, \tilde{s}_k)$ 为特征向量 (m_k, s_k) 的估计,根据 Bayesian 理论,运动目标的分类问题可以通过最大化以下后验概率密度函数:

$$p(\tilde{m}_j, \tilde{s}_j | I'_t, I_{t-1}) > p(\tilde{m}_k, \tilde{s}_k | I'_t, I_{t-1}) \quad j=1, 2, \dots, L, j \neq k \quad (4)$$

上式又可以写成:

$$p(\tilde{m}_j, \tilde{s}_j | I'_t, I_{t-1}) = \frac{p(I'_t | I_{t-1}, \tilde{m}_j, \tilde{s}_j) p(\tilde{m}_j, \tilde{s}_j | I_{t-1})}{p(I'_t | I_{t-1})} \\ = \frac{p(I'_t | I_{t-1}, \tilde{m}_j, \tilde{s}_j) p(\tilde{m}_j | \tilde{s}_j, I_{t-1}) p(\tilde{s}_j | I_{t-1})}{p(I'_t | I_{t-1})} \quad (5)$$

由于式(5)的分母与运动无关,可以忽略不计。第一项分子为分割的先验概率,体现了运动补偿的精度。第二项是含有运动信息的运动分割概率。第三项是含有灰度和像素位置等特征的分割概率。再根据能量最小化理论,以上概率函数可以进一步转化为以下能量函数的最小化:

$$E_j = E(I'_t | I_{t-1}, \tilde{m}_j, \tilde{s}_j) + E(\tilde{m}_j | \tilde{s}_j, I_{t-1}) + E(\tilde{s}_j | I_{t-1}) \quad (6)$$

4 RBF 神经网络学习算法

RBF(Radial Basis Function)神经网络是一种由输入层、

运动模型采用仿射变换模型。运动补偿算法的有效性表现在图 2, (1a)是参考帧图像, (1b)是补偿后的当前帧图像, (1c)和 (1d)分别是补偿前后的差分并阈值化后的结果,这里取了同一阈值 10。

隐含层、输出层组成的神经网络,它的隐含作用函数由高斯函数或者其他辐射状函数所构成,而输出层作用函数通常为简单的线性函数,其结构如图 3。

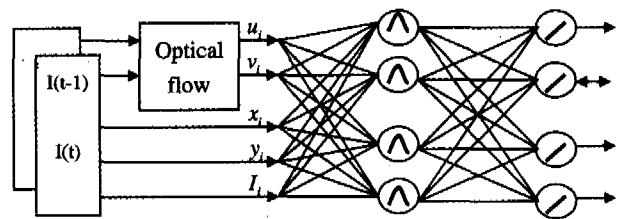


图3 用于运动分割的 RBF 网络结构示意图

输入层是非线性变换层,实现了 X 到 $\Phi_j(X)$ 的非线性映射:

$$\Phi_j(X) = \exp\left[-\frac{(u_j - X)^T (u_j - X)}{\sum_j} - H(X)\right] \quad (7)$$

其中, $X = \{x_1, x_2, \dots, x_N\}$ 为输入向量, $\Phi_j(X)$ 是第 j 层高斯内核函数对输入向量 X 的响应。 u_j 为第 j 个高斯变换单元的“中心”向量,是 RBF 学习算法中最重要的变量。 \sum_j 是第 j 个高斯变换单元的宽度。

$$y_k(X) = \sum_{j=1}^L \lambda(k, j) \Phi_j(X) \quad (8)$$

$$Y_k(X) = \frac{1}{1 + \exp(-y_k(X))} \\ = \frac{1}{1 + \exp(-\sum_{j=1}^L \lambda(k, j) \Phi_j(X))} \quad (9)$$

输出层的作用是将变换层线性加权合并。其中, L 是隐含层单元的数目, Y_k 是第 k 层网络的输出,这里选用 sigmoidal 函数把输出限制在 $[0, 1]$ 。

根据式(6)提供的能量函数更新 RBF 参数。式(6)中第一项体现了由运动信息 $(\tilde{m}_k, \tilde{s}_k)$ 得到的运动补偿的精度,通过 RBF 神经网络学习机制能进一步提高补偿精度和平滑光流场,即:

$$E(I'_t | I_{t-1}, \tilde{m}_j, \tilde{s}_j) = H(X) \quad (10)$$

对每次学习,首先采用经典的学习矢量量化方法估计隐含层的 u_j, \sum_j 两个参数,然后估计用最小均误差法(LMS法)完成加权系数 $\lambda(k, j)$ 的学习,经过多次迭代收敛后, RBF 神经网络的输出就是运动区域。学习算法按如下步骤进行:

- 1) 初始化类别数 P , 中心初值 u_j , 初值随机给定。
- 2) 对于每个输入向量 X , 采用式(11)求出与它距离最近

的一个中心。

$$\|u_j - X\|^2 = \min \|u_k - X\|^2 \quad (11)$$

3)分别按照式(12)和(13)对中心和宽度进行修正。

$$\widetilde{u}_j(t+1) = \widetilde{u}_j(t) + \frac{1}{n_j} (\widetilde{u}_j(t) - X) \quad (12)$$

$$\widetilde{\sigma}_j(t+1) = \frac{n_j - 2}{n_j - 1} \widetilde{\sigma}_j(t) + \frac{(u_j - X)^T (u_j - X)}{n_j - 1} \quad (13)$$

4)计算中心变换平方和,如果 $E_t < \epsilon$,说明 RBF 网络的中心已经变化很小,令 $t = t + 1$,转到第 2 步对下一个输入矢量进行学习,否则转到第 3 步继续修正。

$$E_t = \sum_{j=1}^k \|\widetilde{u}_j(t+1) - \widetilde{u}_j(t)\|^2 \quad (14)$$

5)每次中心修正完后,采用最小均方差法完成对权值的学习。

经过上述步骤,便能较准确地确定 RBF 网络的中心。当中心已经能够准确地反映出输入样本聚类时,则中心的修正量趋近于零,从而稳定地收敛于聚类点。

5 试验结果分析

所提算法经过许多图像序列验证,图 4 是两个典型的结果。第一个序列图像是从美国电影“断剑”中截取的片断,由

11 架飞机在一个沙滩平面上飞过,有些飞机的对比度很低,甚至很难用肉眼辨认出来。第二个图像序列是从实验室红外相机拍摄的图片,一辆卡车自西向东行驶,背景有很多树木,且传感器噪声较大,用一般的检测方法很难全自动地检测出目标。本文所提出的全自动的检测方法,对不同的场景不需要调整阈值,实现了全自动目标检测,取得很好的效果。(3a)(3e)是原始图像序列中的一帧,(3b)(3f)是补偿稳定图像后连续帧间的光流场,(3c)(3g)经过 RBF 神经网络后的光流输出。为了达到更好的视觉效果,由于光流值一般比较小,这里都把光流幅值归一化到[0,255]之间,(3d)(3h)是最终的检测结果。

图中的白色边缘为补偿后包裹图像的结果。在第二个图像序列中由于帧间平动较大,包裹表现得非常明显。对比使用 RBF 神经网络前后的光流输出,基于 RBF 神经网络的光流分割结果提高了信噪比和平滑了光流场,更加有利于检测运动目标,当然这种方法比一般的邻域聚类方法也有明显的优势。从图中可以看出,它还能检测出肉眼很难辨识出的灰度值很低的目标。由于是基于运动的检测方法,飞机的机翼也能很好地检测出来。

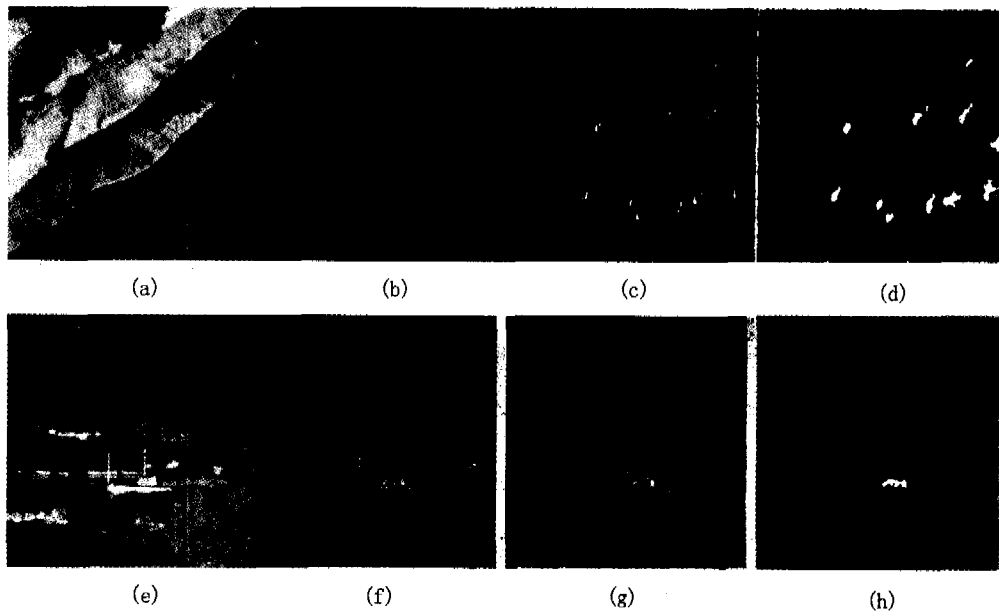


图 4 仿真结果

结论 本文提出一种新的基于 RBF 神经网络的运动目标检测算法,在不要求非常精确分割目标的领域,能很有效地检测运动目标,定位目标的大致轮廓和位置。本文虽然引入了 Bayesian 理论,但是运动分割的许多固有缺点还没有很好地解决,这也是未来研究的方向。

参考文献

- 1 Wang J Y A, Adelson E H. Representing moving images with layers. *IEEE Trans on Image Processing*, 1994, 3(5):625~638
- 2 Konrad J, Dubois E. Bayesian estimation of motion vector fields. *IEEE Trans on PAMI-14*
- 3 Leung H, Dubash N, Xie N. Detection of Small Objects in Clut-

- ter Using a GA2RBF Neural Network. *IEEE Trans on Aerospace and Electronic Systems*, 2002, 38(1)
- 4 Bors A G, Gabbouj M. Minimal topology for a radial basis function neural network for pattern classification. *Digital Signal Processing: a Review Journal*, 4: 173~188
- 5 Xu L, Krzyzak A. Rival Penalized Competitive Learning for Clustering Analysis, RBF net, and Curve Detection. *IEEE Trans. on Neural Networks*, 1993
- 6 Baker S, Matthews I. Lucas-Kanade 20 Years On: A Unifying Framework. *International Journal of Computer Vision*, 2001
- 7 Horn B K P, Schunck B G. Determining Optical Flow. *Artificial Intelligence*, 1981, 17:185~203