

doi: 10.3788/gzxb20124108.0914

基于脉冲耦合神经网络融合的压缩域运动目标分割方法

王慧斌, 沈俊雷, 王鑫, 张丽丽

(河海大学 计算机与信息学院, 南京 210098)

摘 要:针对 H.264 压缩域内运动目标分割算法所存在的弱自适应性和抗噪能力差等问题, 本文提出了一种基于脉冲耦合神经网络的压缩域运动目标分割方法. 该方法采用时空域矢量均值滤波对运动矢量进行预处理, 减少运动目标丢失率, 并设计了前后向矢量累积方法, 增强运动矢量的可靠性. 基于脉冲耦合神经网络设计的融合模型可以将累积后的运动矢量和宏块模式进行融合处理, 增强分割算法的抗噪能力, 保证加快分割速度的同时兼顾运动区域的分割准确度. 另外, 采用最小交叉熵作为点火终止判断条件, 实现了最佳分割模板的自适应获取. 仿真实验表明, 本文算法在自适应性和抗噪能力方面均有较好表现, 可以准确分割出监控视频中的运动目标.

关键词:运动目标分割; 脉冲耦合神经网络融合模型; H.264 压缩域; 视频监控

中图分类号: TP391.9

文献标识码: A

文章编号: 1004-4213(2012)08-0914-8

0 引言

运动目标分割是实现图像分析理解的基础和前提. 目前, 运动目标分割方法主要是基于像素域进行处理, 如采用空间分割法^[1-2]和时间分割法^[3-4]等. 然而在高实时性要求的视频监控应用中, 采用像素域方法对每一帧图像进行的目标分割处理, 必须是对视频完全解码之后, 不可避免地带来时滞问题. 如果在压缩域中直接提取视频对象或进行相关处理操作, 可免除相对耗时的解码操作. 因此, 基于压缩域内的运动目标分割研究引起了国内外学者的关注, 并成为重要研究方向之一.

采用压缩域内运动目标分割方法的思路是通过将部分解码获取的运动矢量、离散余弦变换 (Discrete Cosine Transform, DCT) 系数等作为初始信息, 利用 DCT 系数或者运动矢量的相似性提取运动区域, 从而实现运动目标的分割. 近年来, H.264 编码标准在较高的压缩比下仍能提供连续、流畅的画面, 大大缓解了视频监控系统对数据存储和网络带宽的压力, 已渐成为视频监控领域的主流. 目前, 围绕 H.264 标准的压缩域内运动目标分割研究也取得了部分成果. 文献[5-6]提出直接从稀疏运动矢量场中分割运动对象的方法; 文献[7]提出了一种融合 K 均值(K-means)和期望最大化 (Expectation

Maximization, EM) 聚类算法区分出运动宏块的运动目标分割方法; 文献[8]提出一种基于马尔可夫随机场的最大后验概率模型, 通过求解模型标记运动区域; 文献[9]提出一种利用改进的蚁群聚类算法将具有运动相似性的运动矢量聚类到若干个区域, 分割出运动目标和背景区域的方法. 然而, 这些研究成果仍然存在以下若干问题值得关注: 1) 仅采用运动矢量进行分割导致算法抗噪能力差、分割准确度低; 2) 分割算法在不同的场景中自适应性弱, 而且对分割结果没有一个评价指标, 容易导致“过分割”和“欠分割”; 3) 分割过程中引入的经典算法模型复杂, 计算量大.

另外, 现有压缩域目标分割算法在兼顾实时性和准确性时难度大. 使用 K-means 聚类算法分割运动矢量, 分割速度快, 但是准确度不高; 而采用马尔可夫随机场 (Markov Random Field, MRF) 模型^[8]、蚁群^[9]等经典算法, 模型又过于复杂, 虽然分割准确度得到了提高, 但牺牲了分割速度.

针对目前 H.264 压缩域内运动目标分割算法存在的自适应性弱、抗噪能力差以及模型过于复杂的问题, 本文围绕视频监控应用需求, 提出了一种基于脉冲耦合神经网络 (Pulse Coupled Neural Network, PCNN) 融合模型的压缩域运动目标分割方法. 基于 PCNN 的双通道调制以及脉冲耦合特

基金项目: 江苏省高校科研成果产业化推进项目 (No. JHB2011-6) 资助

第一作者: 王慧斌 (1967-), 男, 教授, 博士, 主要研究方向为信息融合和图像处理. Email: hbwang@hhu.edu.cn

收稿日期: 2012-01-09; 修回日期: 2012-04-12

性,通过融合运动矢量和宏块模式信息,并设计评价指标判断寻找最优阈值,最终实现最佳分割。

1 H. 264 压缩域 PCNN 融合方法

1.1 压缩域问题描述

实际上,H. 264 压缩域内的问题可以归纳为两种类型的不确定性行为:

1) 仅仅采用了运动矢量作为分割初始信息,当编码器本身的缺陷、码流传输过程中出现噪音或者部分背景无规则抖动,极易产生运动矢量噪音很难被处理掉,最终导致分割准确度降低。如图 1(a)~(c),目标区域的阈值分割过程中始终不能滤除场景下面的噪音区域。

2) 运动矢量的强度具有不确定性,而通常采用的阈值分割算法是在不同的场景下需要人为设定参量,并且缺少对分割结果的评价指标,导致分割准确度不高,造成“过分割”与“欠分割”现象。如图 1(a)~(c)和(d)~(f),不同目标((a)~(c)中的汽车,(d)~(f)中的行人)其阈值的大小不一样,而且非最佳阈值下过分割或欠分割现象明显。

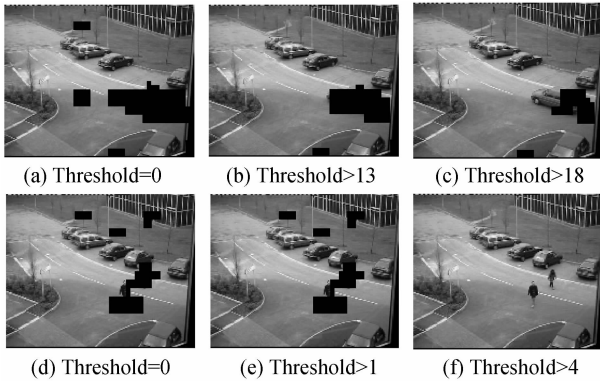


图 1 运动矢量的阈值不确定性

Fig. 1 Uncertainty of the motion vector threshold

考虑到 H. 264 压缩域内的宏块编码模式在一定程度上能够大体区分背景和运动区域,因此,如果分割算法采用宏块编码模式信息和运动矢量的融合,将会一定程度上提高分割算法的抗噪能力。而在运动矢量不确定的情况下,可以考虑采用自适应分割算法,即根据不同场景下的运动矢量联合宏块编码模式信息自适应寻找最优阈值。因此,兼顾上述情况的关键是设计合理的融合处理模型。

Eckhorn 在研究猫的大脑视觉皮层实验中观察到与特征有关的神经元同步行为现象,继而提出了 PCNN 模型理论^[10]。该模型利用神经元的线性相加、非线性相乘调制耦合以及相似神经元脉冲传播特性,弥补了输入数据的空间不连贯和幅度上的微小变化,能够较好地解决像素域图像分割问题以及分割问题中的“空洞”现象^[11]。

压缩域内的运动矢量场是不连续的,存在过多的“空洞”。采用 PCNN 模型能在分割过程中,利用相似神经元脉冲传播特性,将存在不连续运动矢量的宏块同时“点火”,保证分割后区域的连续性,因此,在解决“空洞”问题时具有很好的优势。再者,由于 PCNN 模型是一种无需预先训练的神经网络分割算法,这使其计算复杂度降低,有利于提高目标分割速度。另外,利用 PCNN 模型所具有的双通道调制特性的无耦合、有耦合两种链接机制,可对压缩域内运动矢量和宏块模式信息的进行融合控制,有效抑制噪音区域的运动矢量,提升分割算法的分割准确度。因此,基于 PCNN 可以设计出一种融合模型,该模型能够融合压缩域中运动矢量和宏块编码模式两种信息,能稳定快速的进行分割的同时有很强的抗噪性能。同时,该融合模型可以通过多次迭代获取最优阈值,且会根据评价指标获取当前帧最佳分割模板,因而可以实现对分割准确度的控制。

1.2 基本 PCNN 模型

PCNN 是由若干个 PCNN 的神经元互连所构成的反馈型网络,其每一神经元由三个部分组成:接收域,调制部分和脉冲产生部分,如图 2。

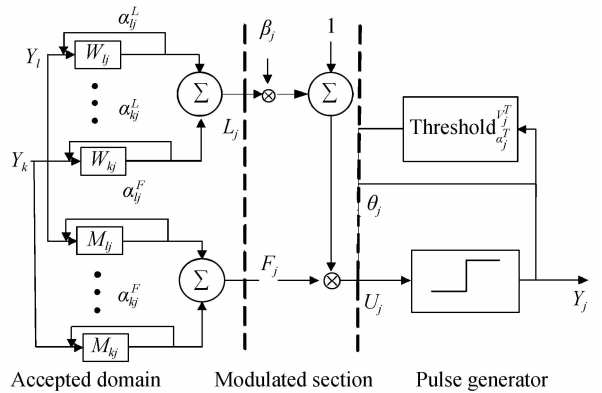


图 2 单个 PCNN 神经元模型

Fig. 2 Single-PCNN neuron model

接收域是接收神经元的外部激励,形成馈送输入 F_{ij} 和链接输入 L_{ij} 两部分。其中, F_j 和 L_j 的计算公式为

$$L_j = \sum_k L_{kj} = \sum_k [W_{kj} e^{-\alpha_{kj}^L t}] \otimes Y_k(t) + J_j \quad (1)$$

$$F_j = \sum_k F_{kj} = \sum_k [M_{kj} e^{-\alpha_{kj}^F t}] \otimes Y_k(t) + I_j \quad (2)$$

式中, W_{kj} 与 M_{kj} 表示突触联接权; α_{kj}^L 与 α_{kj}^F 为时间常量; I_j 与 J_j 表示输入常量。

调制部分是将来自 L 通道的信号 L_j 加上一个正的偏移量,然后与来自 F 通道的信号 F_j 进行相乘调制,计算出内部行为 U_j 。模型中偏移量归整为 1, β_j 为联结强度,其计算公式为

$$U_j = F_j (1 + \beta_j L_j) \quad (3)$$

脉冲产生部分是由阈值可变比较器与脉冲产生

器组成。

阈值可变比较器:当神经元输出一个脉冲,神经元的阈值就通过反馈迅速得到提高.其计算公式为

$$\frac{d\theta}{dt} = -\alpha_j^T \theta_j + V_j^T Y_j(t) \quad (4)$$

式中, V_j^T 与 α_j^T 分别表示阈值的幅度系数与时间常量, θ_j 表示阈值, $Y_j(t)$ 表示 t 时刻的脉冲。

脉冲产生器:当神经元的阈值 θ_j 超过 U_j 时,脉冲产生器关掉,停止发放脉冲;当阈值低于 U_j 时,脉冲产生器打开,神经元被点火,输出一个脉冲或脉冲序列.其数学公式为

$$Y_j = \text{Step}(U_j - \theta_j) \quad (5)$$

但是, Eckhorn 提出的神经网络模型在图像处理应用上存在一定的局限性,为此, G. Kuntimad 简化了 PCNN 模型^[12], 保留了原始 PCNN 模型的脉冲传播和脉冲耦合等重要特性, 解决了参量选取复杂等难题. 简化的二维 PCNN 神经网络如图 3.

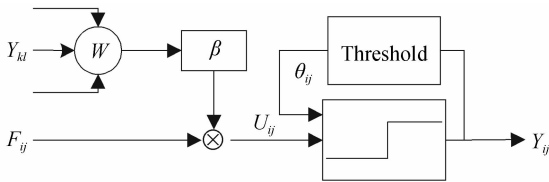


图3 简化的单层二维 PCNN 神经网络

Fig. 3 Simplified two-dimensional single-layer PCNN

1.3 压缩域内 PCNN 融合模型

1.3.1 基于 PCNN 融合模型设计

在 L 通道引入宏块编码模式信息, 利用双通道调制特性, 让其作为 F 通道的外部激励, 使宏块编码模式权值大的区域(运动区域)优先产生脉冲, 而权值小的区域(背景区域)得到抑制. 其中, 本文对链接权系数 W 进行修改, 引入宏块模式权值, 从而实现背景和运动区域的控制。

对链接强度系数 β 进行修改, 引入宏块编码模式信息, 在权值大的地方采用有耦合的方式, 在权值小的地方采用无耦合的方式. 这样也能在脉冲产生的过程中对背景区域进行抑制。

同时, 为了使融合模型在分割过程中兼顾实时性和准确性, 在解决如何减少迭代次数和评价指标的选取问题时采用如下设计: PCNN 模型中阈值比较器和整个分割过程中的脉冲产生次数相关, 阈值衰减的快, 整个点火次数就少, 但分割准确度降低; 阈值衰减过慢, 分割准确度提高, 点火次数就增加. 因此, 设计采用联合宏块模式信息动态调整阈值比较器的衰减幅度, 达到在运动区域减缓衰减幅度, 实现精分割; 在背景区域增大衰减幅度, 实现粗分割。

最小交叉熵用于度量 2 个概率分布之间信息量的差异, 通过搜索获取使分割前后图像信息量差异

最小的阈值. 而压缩域内运动矢量分布的特点就是存在运动矢量的地方往往就是运动区域, 所以使分割后的图像与分割前的图像运动矢量分布的特点差异最小, 往往就是最优的分割模板, 因此, 设计采用最小交叉熵准则作为分割算法的评价指标。

1.3.2 融合模型结构

H. 264 压缩域的 PCNN 融合模型采用二维单层网络, 如图 4.

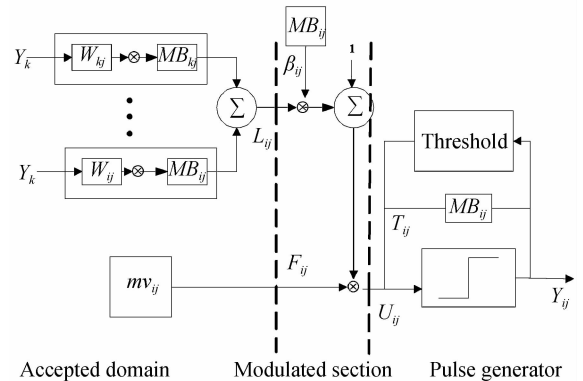


图4 H. 264 压缩域的 PCNN 融合模型

Fig. 4 Fusion-PCNN model based on H. 264 compressed domain

接收域部分: 接收的外部激励为上一次点火输出的脉冲 $Y_{pj}(n-1)$ 和累积后的 x 方向和 y 方向的运动矢量绝对值和 mv_{ij} . 在链接权系数 $W_{i,j;p,q}$ 设为

$$W_{i,j;p,q} = \frac{1}{(i-p)^2 + (j-q)^2} \times MB_{i,j;p,q} \quad (6)$$

式中, (p, q) 为 (i, j) 8 邻域内的任一宏块坐标, $MB_{i,j;p,q}$ 为宏块大小模式的权值信息, 其中权值定义公式为

$$MB = \begin{cases} 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8 = 1 \\ 4 \times 8, 8 \times 4, 4 \times 4 = 2 \\ \text{Intra} 4 \times 4, \text{Intra} 16 \times 16, \text{Skip} = 2 \end{cases} \quad (7)$$

式中, Intra $n \times n$ 表示各种大小的帧内预测宏块模式, Skip 表示 Skip 模式, $n \times n$ 表示各种大小的帧间预测模块模式. 其中, Intra 4×4 、Intra 16×16 、Skip 这几种模式本身不存在运动矢量, 大多数情况下都是采用邻域插值的方法插值获取其运动矢量; Skip 模式一般情况下运动矢量为 0, 也有可能是大目标区域中的不移动的区域, 插值后会存在运动矢量; 而 Intra $n \times n$ 几种模式是存在运动矢量的, 往往代表新出现的运动区域, 故本文综合考虑上述情况, 把这几类归为一类, 赋予较大权值。

调制部分: 接收 F 通道和 L 通道的数据产生内部行为 U_{ij} , 其中, 由 β 链接强度控制点火快慢, 故本文对 β 做如下设计

$$\beta_{ij} = \begin{cases} 0.3 \times MB_{ij} & (MB_{ij} = 2) \\ 0 & (MB_{ij} = 1) \end{cases} \quad (8)$$

式中, MB_{ij} 为宏块大小模式的权值, 0.3 为比例系数. 当权值为 1 时, β 处于无耦合状态, 点火比较慢; 当权值为 2 时, β 处于强耦合状态, 点火比较快.

脉冲产生部分: 采用获取的全局最大的 U_{ij} 为阈值为初始阈值 T_0 . G 是衰减幅度, 在权值大的地方, 衰减幅度小, 实现精分割; 在权值小的地方, 增大衰减幅度, 实现粗分割, 故本文的衰减幅度公式设计为

$$G_{ij,n} = \begin{cases} 1 & (n=0) \\ 0.8^{(3-MB_{ij})} \times G_{ij,n-1} & (n>0) \end{cases} \quad (9)$$

式中, MB_{ij} 为宏块大小模式的权值, 0.8 为比例系数.

最后输出的 Y 表示脉冲序列, 当内部耦合变量 $U > T$ 时, 那么对相应的宏块赋予 1, 否则, 赋予 0, 最终得到一幅由脉冲序列生成的二值图.

PCNN 融合模型的分割算法的计算公式为

$$F_{ij} = mv_{ij} \quad (10)$$

$$L_{ij}(n) = \sum W_{i,j;p,q} \times Y_{pq}(n-1) \quad (11)$$

$$U_{ij}(n) = F_{ij}(n)(1 + \beta_j L_{ij}(n)) \quad (12)$$

$$T_{ij}(n) = \begin{cases} G_{ij,n} T_0 & Y_{ij}(n) = 1 \\ T_0 & Y_{ij}(n) = 0 \end{cases} \quad (13)$$

$$Y_{ij}(n) = \begin{cases} 1 & U_{ij}(n) > T_{ij}(n) \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

1.3.3 迭代终止评价准则

本文采用最小交叉熵准则. 首先, 利用 PCNN 模型公式计算出第一次点火后的二值模板. 对二值模板计算最小交叉熵, 其计算公式为

$$D(p_1 : p_0) = p_1 \times \ln \frac{p_1}{p_0} + p_0 \times \ln \frac{p_0}{p_1} \quad (15)$$

式中, p_1 为模板中点火的运动矢量的概率, p_0 为模板中未点火的运动矢量的概率.

然后, 持续迭代点火获取分割模板, 并计算每次得到的 $D(p_1 : p_0)$, 直到 $D(p_1 : p_0) = D_{\min}$ 时, 停止迭代. 此时获取的二值模板为最优分割模板.

2 基于 PCNN 融合的压缩域分割算法

本文提出的 H.264 压缩域运动目标分割方法, 首先将从 H.264 编码端提取到的运动矢量 mv 进行归一化, 初始去噪, 滤除掉明显的噪音区域; 然后采用时空域矢量均值滤波方法, 在滤波的同时利用时空域信息尽可能保证弱小运动目标信息不被滤除, 滤波后再次去除滤波过程中引入的噪音; 再对滤波后的运动矢量进行前后向迭代投影矢量累积; 最后将稠密后的运动矢量和宏块模式信息作为初始信息, 利用 PCNN 融合模型进行分割, 并采用最小交叉熵判断点火终止条件, 直至获取最佳分割模板. 基于 PCNN 融合模型的 H.264 压缩域运动分割算法

如图 5.

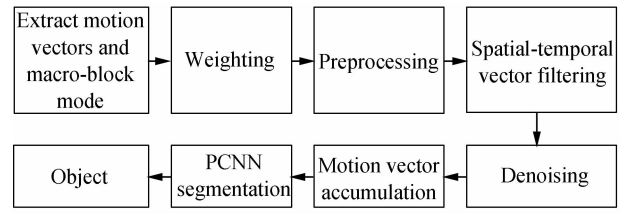


图 5 基于 PCNN 融合模型的 H.264 压缩域分割算法框图
Fig. 5 Segmentation algorithm based Fusion-PCNN model in H.264 compressed domain

2.1 运动矢量预处理

在 H.264 压缩域中运动目标分割最重要的信息是运动矢量, 且存在着噪音, 必须采用预处理技术使其真实可靠, 才能保证后续分割算法的准确性. 本文采用了时空域均值滤波和前后向迭代投影矢量累积法两种比较有效的手段进行运动矢量预处理.

2.1.1 时空域矢量均值滤波

通过判断邻域的非零运动矢量的数目, 简单区别噪音区域与运动区域, 只对噪音区域进行空域滤波, 并在空域滤波基础上, 利用时域滤波补偿被滤除掉的弱小运动目标区域的运动矢量.

噪音区域的空域滤波公式为

$$\begin{cases} mv_{x,y}^i = mv_{x,y}^j & (j \in n, N \geq 7) \\ mv_{x,y}^i = \text{med}\{mv_{x,y}^j\} & (j \in n, N < 7) \end{cases} \quad (16)$$

式中, $mv_{x,y}^i$ 表示当前宏块的运动矢量 x 或者 y 方向, n 表示邻域大小, N 为非零运动矢量的数目, $\text{med}\{mv_{x,y}^j\}$ 为 n 邻域内 4×4 大小的宏块区域的 x 或者 y 方向的运动矢量均值滤波.

时域滤波公式为

$$mv_t = (mv_{t-1} + mv_t + mv_{t+1}) / 3 \quad (17)$$

式中, mv_{t-1} 、 mv_t 、 mv_{t+1} 分别表示 $t-1$ 、 t 、 $t+1$ 时刻的运动矢量.

2.1.2 前后向迭代投影矢量累积

前后向迭代投影矢量累积的方法是: 利用 t 时刻的运动场 MF_t 和 $t+1$ 时刻的运动场来重建预测运动场 $PMF_{t+1,t}$, 利用 t 时刻的运动场 MF_t 和 $t-1$ 时刻的运动场来重建预测运动场 $PMF_{t-1,t}$, 然后将 t 时刻的运动场和 $t \pm 1$ 时刻的预测运动场、 $t \pm 2$ 时刻的预测运动场, 以及 $t \pm n$ 时刻预测运动场累积起来获得累积运动场. 若 $n=1$, 本文的累积公式为

$$AMF_t = MF_t + PMF_{t+1,t} + PMF_{t-1,t} \quad (18)$$

式中, AMF_t 为 t 时刻的累积运动场, MF_t 为 t 时刻的运动场, $PMF_{t+1,t}$ 为由 MF_{t+1} 和 MF_t 重建的预测运动场, $PMF_{t-1,t}$ 为由 MF_{t-1} 和 MF_t 重建的预测运动场.

$PMF_{t \pm 1,t}$ 每个块的运动矢量可由式(19)求出.

$$V_i^p = \sum_{j=1, s_j \in B_i}^n V_j \times \frac{S_j}{\sum_j S_j} \quad (19)$$

式中, V_i^p 为 $PMF_{t \pm 1}$ 中任意块 B_i 的运动矢量, 共有 n 个块覆盖在块 B_i 上面, 每个块的运动矢量为其在 MF_t 中对应块的运动矢量 V_j , S_j 为这些块覆盖在块 B_i 上的面积.

2.2 压缩域 PCNN 分割

H. 264 压缩域 PCNN 分割算法主要步骤为:

- 1) 定义初始变量. β 公式中比例系数 0.3; 衰减系数 G_0 为 1, 衰减比例系数为 0.8;
- 2) 根据 PCNN 融合模型式(11)和(12), 计算出 L 通道、U 通道的值;
- 3) 根据阈值判断条件, 当 U 大于等于阈值 T, 则该对应的宏块区域就点火; 否则, 不点火;
- 4) 对点火后的模板 B 按照式(15)计算最小交叉熵 $D(p_1 : p_0)$, 和上一帧分割模板的最小交叉熵进行比较. 不满足终止条件, $D(p_1 : p_0) = D_{min}$, 即不是最小交叉熵值, 则对阈值按照衰减式(9)和(13)进行衰减, 获取一个新的阈值 T' , 循环计算步骤 2)、3)、4); 直到满足终止条件, 否则, 直接跳出循环体;
- 5) 当前模板为最优模板 B_{opt} , 分割算法结束.

H. 264 压缩域 PCNN 融合模型算法的流程图如图 6.

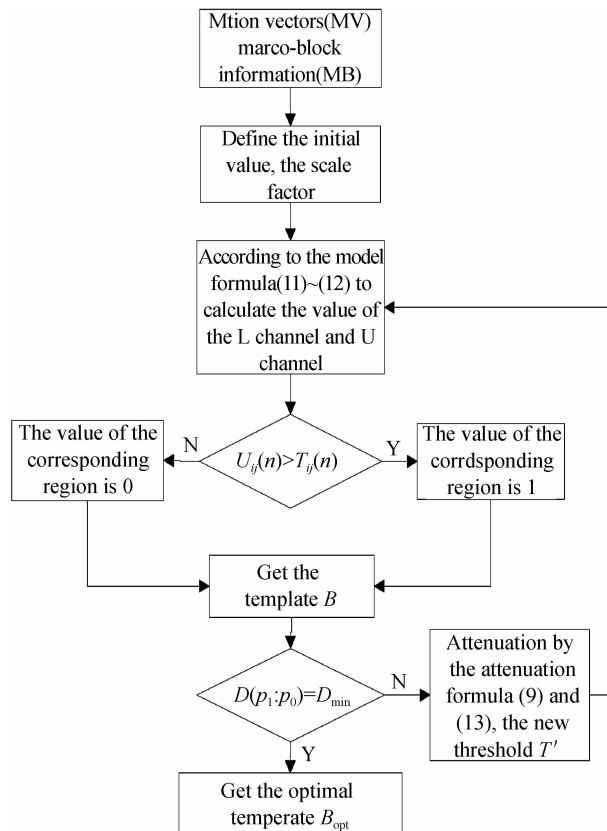


图 6 H. 264 压缩域 PCNN 融合模型算法流程图
Fig. 6 The diagram based Fusion-PCNN model in H. 264 compressed domain

3 实验分析

本文所选用的运动矢量场均从 JM8.6 版本的 H. 264 编码软件压缩后的码流中提取. H. 264 编码配置如下: Baseline Profile, IPPP..., 只有第一帧为 I 帧, 其余都是 P 帧, 1 个参考帧, 运动估计搜索范围为 $[-16, 16]$, 量化参量为 28, 帧大小为 176×144 .

3.1 室内视频监控序列的实验分析

为了验证本文算法的有效性, 采用国际标准测试序列主观验证本文算法. 图 7 为 Hall 标准测试序列采用上述 PCNN 分割算法分割后的结果图. 图 7(d) 为第 18 帧的 PCNN 分割结果, 图中较强的运动矢量区域已经分割出来, 并根据该图的最小交叉熵自适应分割出的最强区域, 基本能标示出该人的运动区域, 而且能解决目标遮挡的问题; 图 7(e) 为第 43 帧的 PCNN 分割结果, 在无遮挡的情况能准确的标示出该运动人物的运动区域, 能达到比较满意的分割结果; 图 7(f) 为第 108 帧的 PCNN 分割结果, 当目标局部存在运动时, 只能分割出该运动的区域, 而对非运动的腿部或者运动矢量小的区域会造成分割的目标残缺.

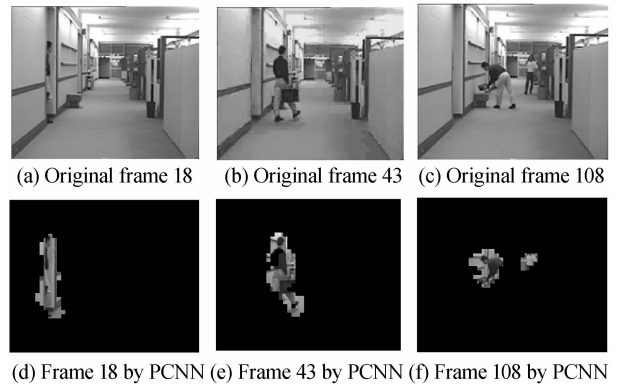


图 7 Hall 标准测试序列分割结果
Fig. 7 Hall test sequence segmentation results

3.2 室外视频监控序列的实验分析

本文选取了监控视频序列 PETS-2000 和 PETS-2001 进行测试. 图 8(a) 为原始帧序列中第 531 帧一个行人和一辆车运动的画面; 图 8(b) 为第 830 帧多行人和一辆车的画面, 右边 3 个行人紧靠在一起并行行走; 图 8(c) 第 2 558 帧两辆车和一个行人的画面.

图 8(d) 为第 531 帧中行人目标由于过于微小, 运动区域仅几个宏块, 而且相对于车子的运动矢量很少, 从图 8(d)~(f) 可以看出, 本文方法由于采用了时空域滤波技术和前后向矢量累积的手段, 能很好地保留这一运动区域, 并且 PCNN 融合模型能很好地抑制其周围的噪音, 所以分割结果比较理想. 图 8(g)~(i) 是文献[6]方法的分割结果, 行人被当成

噪音,造成了漏检.图 8(j)~(l)是采用文献[9]方法的目标分割结果,其分割算法对背景的抑制不如本文算法,造成过分割.对于图 8 中第 830 帧,采用本文的方法能很好地分割出运动目标区域,而对右边 3 个行人紧靠在一起的运动目标区域不能有效地分割.从文献[6]和文献[9]的算法中也可以看出都不能有效地区分出运动目标,主要原因是压缩域内分

割的基本单位是宏块,只能获取到宏块级的分割准确度,如 4×4 、 8×8 等.图 8 第 2 558 帧对左边刚进入画面的行人分割过程中,可以看出本文的算法对其周围的背景区域能很好的抑制,文献[6]和文献[9]都不能抑制其噪音,对车辆的运动区域分割,文献[6]分割准确度低,文献[9]存在过分割,而本文的算法分割效果相对于文献[6]和文献[9]更好.

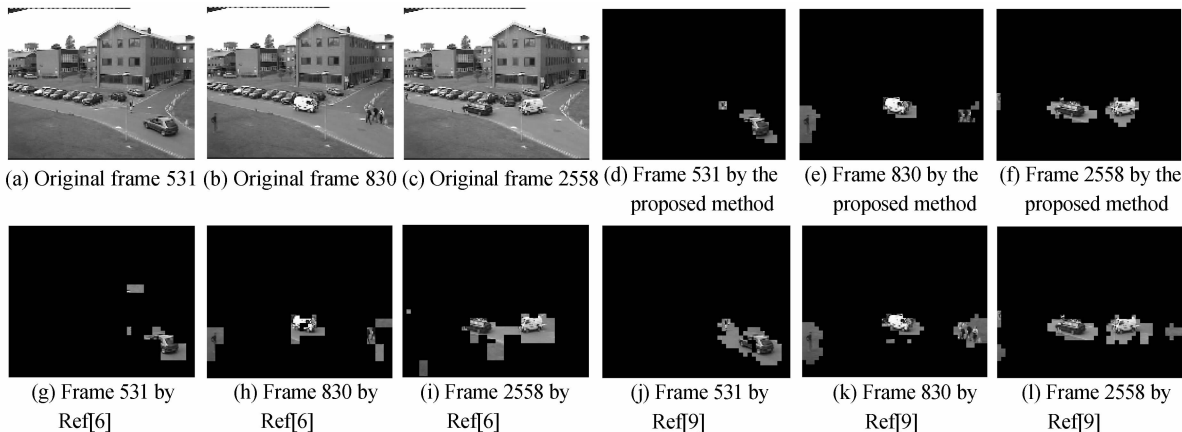


图 8 PETS-2001 视频序列实验结果及比较
Fig. 8 PETS-2001 sequence Experimental results and comparison

图 9(a)为原始序列中的第 425 帧,一辆车进入画面,图 9(b)为第 644 帧,一个行人和一辆即将停下的车,图 9(c)为第 1 287 帧两个行人在画面中运动.

对于图 9 中的第 425 帧,本文算法能很好地分割出运动目标,而文献[6]和文献[9]存在过分割现象,对背景中的噪音未能很好地抑制.图 9 第 644 帧中,当行人的速度和即将停下的车辆速度接近时,本文的算法能很好地分割出行人目标,但不可避免地

引入由风吹等因素引起的画面左下角草坪的运动,被误以为是运动目标,原因就是当时人、车的运动矢量和风动引起的运动矢量接近,而导致不能有效地分割,文献[6]和文献[9]也不能有效地抑制其噪音,原因在于初始的运动矢量场已经引入了噪音.图 9 第 1 287 帧中人的运动矢量大于环境引起的局部运动矢量时,本文的算法能很好地抑制其噪音,而文献[6]却不能.

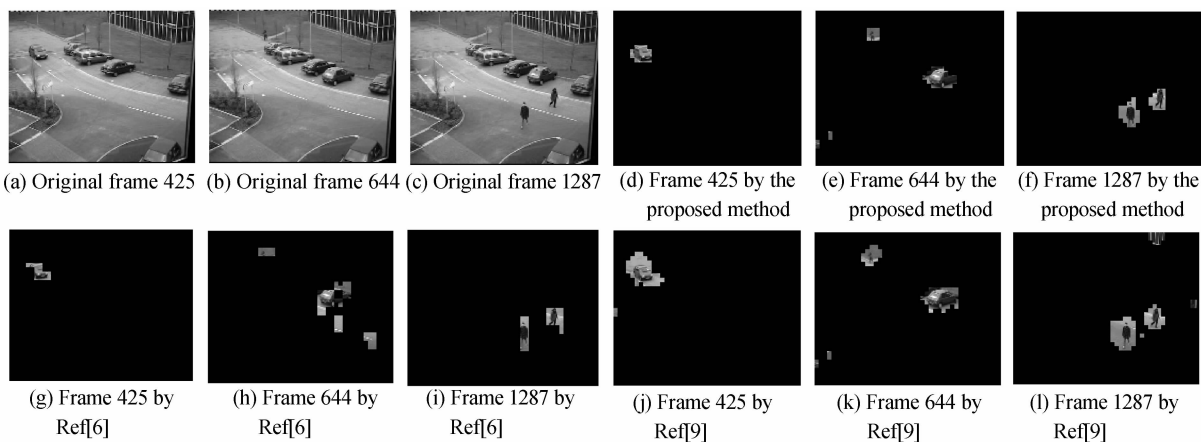


图 9 PETS-2000 视频序列实验结果及比较
Fig. 9 PETS-2000 sequence Experimental results and comparison

从图 8 和图 9 中可以看出,本文的算法能较好地分割出运动目标区域,对背景区域的噪音得到了抑制,能减少很多过分割现象.文献[6]存在欠分割现象,而且整体分割准确度不高.而文献[9]存在很多过分割现象,对噪音的抑制也不如本文算法.

对图 9 中的 PETS-2000 监控视频流的本文算法和文献[6]、文献[9]算法结果用 Precision 和 Recall 2 个指标进行统计的结果.其中,Precision 和 Recall 的在压缩域内的计算公式

$$\text{Precision} = \frac{\text{正确分割宏块}}{\text{正确分割宏块} + \text{虚警宏块}} \quad (20)$$

$$\text{Recall} = \frac{\text{正确分割宏块}}{\text{正确分割宏块} + \text{漏警宏块}} \quad (21)$$

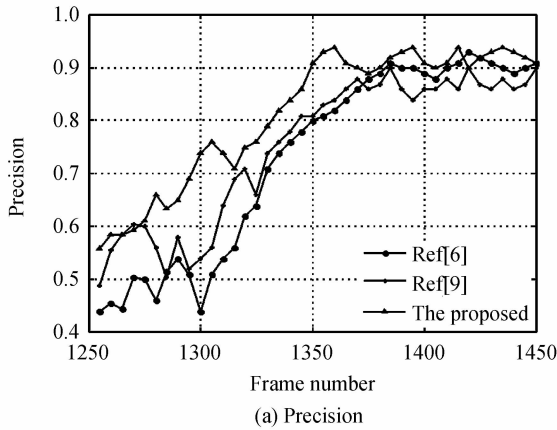


图 10(a)和(b)中横坐标表示视频的帧数,纵坐标表示 Precision 指标和 Recall 指标的百分比,其中选用的数据是 PETS-2000 第 1 250 帧至 1 450 帧,

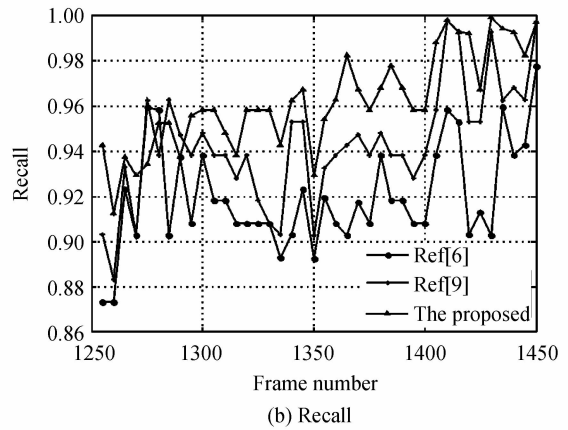


图 10 PETS-2000(1250-1450)的 Precision 和 Recall 指标
Fig. 10 Values of Precision and Recall of PETS-2000(1250-1450)

共 200 帧数据. 其中,前面 50 帧中由于目标离监控摄像机较远,运动矢量较低,分割准确度不高;当后面第 1 300 帧~第 1 400 帧中,目标渐渐靠近摄像头,运动矢量增强,分割准确度提高. 从图 10 中可以看出,本文的算法在查全率和查准率上都略优于文献[6]和文献[9].

从表 1 的数据和图 10 曲线中可以看出,本文的算法略优于文献[6]和文献[9]. 在查准率方面,比文献[6]平均提高了 20%,比文献[9]平均提高了 12%;在查全率方面,比文献[6]平均提高了 5%,比文献[9]平均提高了 3%.

表 1 本文算法与文献[6]、文献[9]的 Precision 与 Recall 指标

Table 1 Values of Precision and Recall in the proposed algorithm and algorithms in Refs. [6] and [9]

Frame	Precision			Recall			
	Proposed	Ref. [6]	Ref. [9]	Proposed	Ref. [6]	Ref. [9]	
PETS-2000	140~260	89.61%	70.74%	75.94%	98.14%	93.62	95.13%
	1250~1450	81.38%	72.88%	74.93%	96.25%	91.98%	94.33%
PETS-2001	445~645	83.41%	63.34%	70.95%	95.29%	90.75%	92.84%
	730~830	85.24%	65.61%	73.5%	97.31%	93.08%	96.89%

另外,为了考察算法的运算效率,本文通过 PETS-2000 视频序列中第 1 250 帧至 1 450 帧中 200 帧视频数据的分割时间进行了仿真实验与统计,结果见图 11. 从图 11 中可以看出,本文算法分割一帧图像平均耗时为 0.2 s,文献[9]算法的平均耗时为 1.8 s,文献[6]算法的平均耗时为 0.1 s,本

文算法比文献[9]算法耗时减少了 89%,与文献[6]算法相差不多.

4 结论

本文提出了一种基于 PCNN 融合模型的压缩域运动目标分割方法. 在运动矢量的预处理过程中采用时空域的权值均值滤波与前后向矢量累积技术,在滤除噪音的同时保留弱小运动目标区域,结合 H. 264 压缩域内的特点对 PCNN 模型进行设计,提升了分割算法的抗噪能力,并通过判断最小信息交叉熵判断点火终止条件获取最优分割模板. 这种方法有别于传统的阈值、聚类分割方法,在保证实时性的同时仍能取得较为满意的分割结果. 实验表明,在监控视频中得到比较满意的分割结果.

参考文献

[1] DUARTE M A, ALVARENGA A V, AZEVEDO C M, et al. Automatic micro calcifications segmentation procedure based on otsu's method and morphological filters[C]. Rio de

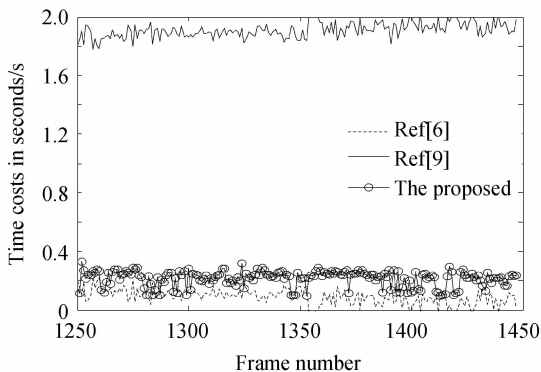


图 11 PETS-2000(1 250~1 450)耗时曲线
Fig. 11 PETS-2000(1 250~1 450)time costs

- Janeiro: Health Care Exchanges (PAHCE), 2011 Pan American, 2011: 102-106.
- [2] WANG D, KWOK N M, JIA X, *et al.* A cellular automata approach for superpixel segmentation [C]. Shanghai: International Congress on Image and Signal Processing, 2011: 1122-1126.
- [3] EVANGELIO R H, SENST T, SIKORA T. Detection of static objects for the task of video surveillance [C]. Kona: IEEE Applications of Computer Vision (WACV), 2011: 534-540.
- [4] YU Hai-bin, HE Zhi-wei, LIU Yuan-yuan, *et al.* A dynamic texture-based method for multidirectional motions segmentation of crowd [C]. Shanghai: International Congress on Image and Signal Processing, 2011: 1132-1136.
- [5] ZENG W, DU J, GAO W, *et al.* Robust moving object segmentation on H. 264/AVC compressed video using the block based MRF model [J]. *Real Time Imaging*, 2005, **11** (4): 290-299.
- [6] LIU Zhi, LU Yu, ZHANG Zhao-yang. Real time spatiotemporal segmentation of video objects in the H. 264 compressed domain [J]. *Visual Communication and Image Representation*, 2007, **18**(3): 275-290.
- [7] WANG Pei, WANG Jun. Block characteristic based moving object segmentation in the H. 264 compressed domain [C]. Shanghai: IEEE International Conference on Signal Processing Systems (ICSPS), 2010: 643-647.
- [8] CHEN Ming-sheng, LIANG Guang-ming. Abstraction moving objects from H. 264 bit stream with MAP-MRF [C]. Chengdu: IEEE Wireless Communications Networking and Mobile Computing, 2010: 1-4.
- [9] WU Zhi-xia, WANG Pei. Moving object segmentation in H. 264/AVC compressed domain using ant colony algorithm [C]. Dalian: IEEE International Conference on Signal Processing Systems (ICSPS), 2010: 716-719.
- [10] ECKHORN R, REITBOECK H J, ARNDT M. Feature linking via synchronization among distributed assemblies: Simulation of results from cat cortex [J]. *Neural Computer*, 1990, **2**(3): 293-307.
- [11] GU Xiao-dong, YU Dao-heng. PCNN's principles and applications [J]. *Journal of Circuits and Systems*, 2001, **6** (3): 45-50.
- [12] POPPE C, BRUYNE S D, PARIDAENS T. Moving object detection in the H. 264/AVC compressed domain for video surveillance applications [J]. *Journal of Visual Communication and Image Representation*, 2009, **20** (6): 428-437.

Moving Object Segmentation Based on Fusion-PCNN in Compressed Domain

WANG Hui-bin, SHEN Jun-lei, WANG Xin, ZHANG Li-li

(College of Computer and Information Engineering, Hohai University, Nanjing 211100, China)

Abstract: Aiming to solve problems of the weak ability on adaptation and noise resistance in object segmentation, a novel PCNN based moving object segmentation method is presented in H. 264/AVC compressed domain. First, a spatial-temporal vector filtering is used as the preprocessor to reduce the target loss rate. Then, a forward-backward vector cumulative method is proposed to enhance the reliability of motion vectors. Finally, a Fusion-PCNN model is designed to fuse the cumulative motion field and the macro-block coding mode, which enhances the ability of noise resistance in object segmentation and limits the complexity. Moreover, the minimum cross-entropy is used to determine the firing conditions for an optimal self-adaptive segmentation template. Experimental results show that the proposed algorithm is outperformance and has the ability of self-adaptation and noise resistance in object segmentation. More accurate results are presented by the surveillance video.

Key words: Moving object segmentation; Fusion-PCNN model; H. 264 compressed domain; Video surveillance