

压缩感知目标跟踪

郭岩松, 杨爱萍, 侯正信, 何宇清

GUO Yansong, YANG Aiping, HOU Zhengxin, HE Yuqing

天津大学 电子信息工程学院, 天津 300072

School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China

GUO Yansong, YANG Aiping, HOU Zhengxin, et al. Object tracking using compressive sensing. *Computer Engineering and Applications*, 2011, 47(32): 4-6.

Abstract: In video analysis, it usually needs to decode and reconstruct the video sequence before any higher level processing such as classification or detection. However, sometimes a video analysis needs to be proceeded without revealing sensitive information, e.g. the identity of people. This paper proposes a new encoding scheme which enables object-tracking without reconstructing the video sequence. According to compressive sensing theory, encoding a video sequence into a few pseudo-random projections of each frame is reasonable due to the sparsity in the interested information. Then the decoder will exploit the sparsity of background-subtracted image to recover the foreground target. Taking a prior knowledge as target position which is estimated by a particle filter, it is possible to improve the reconstruction of foreground target. As expected, this encoding scheme has capability of privacy-protection as well as security.

Key words: compressive sensing; object tracking; particle filtering; reweighted ℓ_1 minimization

摘要: 视频分析通常在分类或检测等高级任务之前解码并重构视频序列。但是, 有时希望只进行视频分析而不暴露敏感信息, 例如人员身份。提出了一个能够跟踪目标而不需要重构视频序列的编码方案。根据压缩感知理论, 用每帧的少量伪随机投影编码一个视频序列。解码器利用背景消除图像的稀疏性重构前景目标。以粒子滤波器估计的目标位置作为先验知识, 可以改进前景目标位置的重构。该编码方案同时具有隐私保护和加密功能。

关键词: 压缩感知; 目标跟踪; 粒子滤波; 重新加权 ℓ_1 最小化

DOI: 10.3778/j.issn.1002-8331.2011.32.002 文章编号: 1002-8331(2011)32-0004-03 文献标识码: A 中图分类号: TP391.4

随着视频获取和存储的电子器件的成本持续下降, 视频监控已广泛应用于机场、银行、交通以及私人住宅的场景监控。这些系统采集的细节信息甚至超过人类的感知能力。

通常对监控视频进行视频分析需要在像素域解码和重构视频序列。然而在某些应用场合, 希望在完成这些分析任务时不暴露敏感数据。例如当系统的目标是计数人群中人数时, 进行这项任务就不必捕捉个人身份^[1]。这时需要做的工作是分割人群, 找到平均人体尺寸, 并以适当的方式跟踪人的运动。Dufaux 等人^[2]基于 Motion JPEG 2000 视频提出了一个有效的隐私保护技术, 其中包含获取视频序列中感兴趣区域的变换域系数, 采用机器学习机制来识别要保护的区域。这种视频分析是在伪随机基上进行的, 一旦解码器得到密钥, 就能够重构完整场景。

本文提出一个适合隐私保护的视频监控应用的编码方案, 能够跟踪视频目标而不需要重构视频序列。该方案利用每帧有限数量的伪随机投影编码一个视频序列, 解码器利用背景消除图像所具有的稀疏性并重构前景目标的位置。该方案采用粒子滤波器预测目标位置, 再用估计的位置作为先验

知识改进前景目标位置的重构。隐私保护体现在只采用编码的随机投影不可能重构原始视频内容, 安全性体现在如果没有用来产生随机投影的种子, 就无法解码。

1 压缩感知基本理论

压缩感知理论^[3-5]认为只要信号能够用一个在某个基展开的少量非零系数表示, 就能够通过有限数量的不相干非自适应线性测量完全重构。令 $\theta \in \mathbb{R}^N$ 为一个 K 稀疏向量, 即 θ 的 N 个元素中只有 K 个非零元素。信号 $x \in \mathbb{R}^N$ 可表示为 $x = \Psi\theta$, Ψ 是信号的稀疏基。令 $y \in \mathbb{R}^n$, $n < N$, 是 x 通过线性随机投影(测量)得到的, 即 $y = \Phi x$ 。如果测量矩阵 Φ 满足受限等度量性质(RIP)^[4-5], 则求解下列优化问题:

$$\min \|\theta\|_1 \text{ 满足 } y = \Phi\Psi\theta \quad (1)$$

相当于寻找 $y = \Phi\Psi\theta$ 的最稀疏解, 只要测量数量满足 $n \geq CK \log(N/K)$ 。实际上, 只要矩阵 Φ 的列与信号的稀疏基 Ψ 不相干, 就满足 RIP 性质。当采样矩阵 Φ 的元素服从零均值、方差为 $1/n$ 的高斯分布时, 测量基以极大概率与任何指定

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.61002027)。

作者简介: 郭岩松(1976—), 男, 博士生, 主要研究领域为数字图像与视频处理、压缩感知等; 杨爱萍(1977—), 女, 博士, 副教授; 侯正信(1945—), 男, 教授, 博士生导师; 何宇清(1973—), 男, 博士, 讲师。E-mail: guoys2005@126.com

收稿日期: 2011-06-08; **修回日期:** 2011-07-25

的基不相干。

在大多数实际应用中, 测量中含有噪声(例如量化噪声)。考虑含有噪声的测量 $y = \Phi x + z$, 其中 z 是有界范数噪声, 即 $\|z\|_2 \leq \sigma$ 。原始信号 x 的近似可以通过求解如下问题而得到:

$$\min \|\theta\|_1 \text{ 满足 } \|y - \Phi \Psi \theta\|_2 \leq \sigma \quad (2)$$

式(2)是一个二阶锥规划问题(SOCP)^[6], 可以在 $O(n^3)$ 复杂度下求解。已经提出一些快速算法寻找式(2)的解。本文拟采用 SPGL1 算法^[7]进行求解。

Candès 等人^[8]最近的工作表明, 通过对式(2)中的目标函数进行适当的加权, 可以在加强重构的同时减少所需测量数。该方法的原理是通过使用一些关于支撑集和稀疏信号值的先验知识, 朝着实际非零值来引导重构过程。此时需要求解如下问题:

$$\theta = \arg \min \|\mathbf{W} \theta\|_1 \text{ 满足 } \|y - \Phi \Psi \theta\|_2 \leq \sigma \quad (3a)$$

令 $z = \mathbf{W} \theta$, 则 $\theta = \mathbf{W}^{-1} z$, 上式可以写成:

$$z = \arg \min \|z\|_1 \text{ 满足 } \|y - \Phi \Psi \mathbf{W}^{-1} z\|_2 \leq \sigma \quad (3b)$$

其中 \mathbf{W} 是一个对角线矩阵, 对角线上的权值 $w = [w_1, w_2, \dots, w_N]$ 反比于期望的信号幅度。这些加权值由粒子滤波器估计的边界框进行设置。

2 系统结构

本章提出的视频跟踪方案的框图如图1所示。

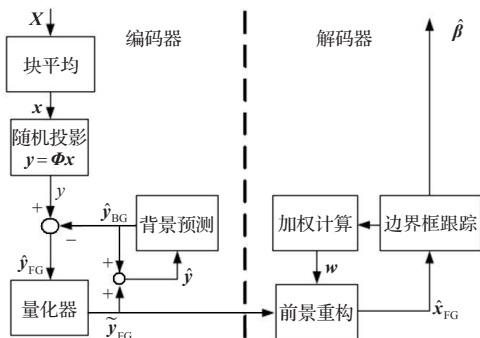


图1 视频跟踪方案框图

编码器产生量化的前景投影 \hat{y}_{FG} , 其中各方框的功能如下。

(1) 分块平均: 原始帧 X 被分割成 N 个大小为 $B \times B$ 的块。计算每块的亮度成分的平均并存储为向量 $x \in \mathbb{R}^N$ 。

(2) 随机投影: 对 x 进行线性随机投影 $y = \Phi x$, $y \in \mathbb{R}^n$, $n < N$ 。矩阵 $\Phi \in \mathbb{R}^{n \times N}$, 其元素抽样于高斯分布 $\mathcal{N}(0, 1/n)$, 在解码器端是已知的。

(3) DPCM 编码: 随机投影 y 用背景预测器 \hat{y}_{BG} 进行预测并进行 DPCM 编码, 得到量化的前景投影 \hat{y}_{FG} , 并将其发送到解码器。量化器是步长为 Δ 的最佳均匀标量量化器, 量化范围自适应地依赖于输入信号的方差, 使得失真率是固定的, 而速率是可变的。

解码器各方框的功能如下。

(1) 前景重构: 求解下列优化问题得到重构的前景 \hat{x}_{FG} :

$$\hat{\theta} = \min \|\mathbf{W} \theta\|_1 \text{ 满足 } \|\hat{y}_{FG} - \Phi \Psi \theta\|_2 \leq \sigma \quad (4)$$

其中 \mathbf{W} 是在对角线上权值为 $w = [w_1, w_2, \dots, w_N]$ 的对角线矩阵, 参数 σ 依赖于量化步长 Δ 和投影数量 n ^[9], Ψ^T 是 2D 小波变换矩阵, $\theta = \Psi^T x_{FG}$ 是 x_{FG} 的小波系数。重构的前景 $\hat{x}_{FG} = \Psi \hat{\theta}$ 。

(2) 限位框跟踪: 用粒子滤波器跟踪背景消除目标剪影的

限位框 $\hat{\beta}$, 用前景估计 \hat{x}_{FG} 作为观测数据以计算更新步骤中的粒子似然。

(3) 加权计算: 除了提供运动目标的跟踪, 限位框 $\hat{\beta}$ 的粒子滤波预测还作为从随机投影重构前景估计的先验信息。根据限位框位置提供的信息计算式(4)中的加权矩阵 \mathbf{W} 。令 w_k 为 2D 小波系数 θ 的第 k 个系数的加权因子, 加权值计算为:

$$w_k = \frac{1}{r_k + \varepsilon} \quad (5)$$

其中, 引入参数 $\varepsilon > 0$ 以提供稳定性, 在理想情况下 r_k 应正比于小波系数 θ_k 的绝对值^[9]。由于小波系数 θ_k 是未知的, 本方案根据相对于限位框位置的第 k 个小波系数的像素位置来设置 r_k 的值。如果这个像素位置在限位框之内, 就设置 $r_k = 1$; 当像素位置在限位框之外并远离限位框时, r_k 光滑地衰减至零。首先分别计算像素位置到最近的垂直和水平限位框边界的距离 d_x 和 d_y , 然后关联到这个像素的系数为 $r = e^{-(\alpha_x d_x + \alpha_y d_y)}$ 。其中 $\alpha_x > 0$ 和 $\alpha_y > 0$ 控制窗口衰减率, 这里设为 0.1, 可以保证窗口衰减既不过快也不过慢。

3 背景模型

令 $\{x(1), x(2), \dots, x(t)\}$ 是从一个固定摄像机得到的帧序列, \hat{x}_{BG} 是场景静态背景的估计。在每个时刻 t , 前景图像为 $\hat{x}_{FG}(t) = x(t) - \hat{x}_{BG}$ 。背景模型不应当是固定的, 必须适应亮度和运动变化, 因此必须连续更新。一个简单和计算高效方法是滑动平均法^[10]:

$$\hat{x}_{BG}(t+1) = \alpha x(t) + (1-\alpha) \hat{x}_{BG}(t) \quad (6)$$

其中 α 是学习速率。当摄像机直接获取随机投影而不收集整个图像时也可以采用这种方法。令 $\hat{y}_{BG}(t) = \Phi \hat{x}_{BG}(t)$ 为背景投影, 容易计算前景投影为:

$$\hat{y}_{FG}(t) = \Phi \hat{x}_{FG}(t) = \Phi(x(t) - \hat{x}_{BG}(t)) = \Phi x(t) - \Phi \hat{x}_{BG}(t) = y(t) - \hat{y}_{BG}(t) \quad (7)$$

背景投影仍可以用滑动平均法更新:

$$\hat{y}_{BG}(t+1) = \alpha y(t) + (1-\alpha) \hat{y}_{BG}(t) \quad (8)$$

4 限位框跟踪

可以将跟踪问题看作一个随机滤波问题, 滤波的目标是在给定已知观测 $\{\eta(1), \eta(2), \dots, \eta(t)\}$ 时, 估计随机变量 $\zeta(t)$ 在时刻 t 的最佳当前状态。这个问题能够用粒子滤波的方法高效地解决, 这里采用顺序重要性重新采样(SIR)粒子滤波器^[11]加以实现。设当前状态只与前一个状态有关, 当前观测只与当前状态有关, 则可以定义一个转移模型 $P(\zeta(t)|\zeta(t-1))$ ^[12]和一个似然函数 $P(\eta(t)|\zeta(t))$ 。在跟踪问题中 $\zeta(t) = [c(t), s(t), v(t)]^T$ 是表示目标状态的六维矢量, 其元素分别是限位框中心位置 $c(t) = [c_x(t), c_y(t)]^T$, 尺寸 $s(t) = [s_x(t), s_y(t)]^T$ 和速度 $v(t) = [v_x(t), v_y(t)]^T$, 并以重构的前景 $\hat{x}_{FG}(t)$ 作为观测 $\eta(t)$ 。

粒子滤波器的预测步骤采用如下转移模型 $P(\zeta(t)|\zeta(t-1))$:

$$\begin{cases} \hat{c}_i(t) = c_i(t-1) + v_i(t-1) \Delta_T + \xi_c \\ \hat{s}_i(t) = s_i(t-1) + \xi_s \\ \hat{v}_i(t) = v(t-1) + \xi_v \end{cases} \quad (9)$$

其中 $[c_i(t), s_i(t), v_i(t)]^T$ 是粒子 i 在时刻 t 的状态向量, $[\hat{c}_i(t), \hat{s}_i(t), \hat{v}_i(t)]^T$ 是预测的粒子 i 在时刻 t 的状态向量, Δ_T 是时间采样间隔, ξ_c, ξ_s, ξ_v 是为系统提供多种假设的随机项。以如下方式估

计限位框的中心位置、尺寸和速度:

$$\begin{cases} c(t) = (1 - \alpha_c)(c(t-1) + v(t-1)\Delta_T) + \alpha_c \sum \omega_i(t) \hat{c}_i(t) \\ s(t) = \sum \omega_i(t) \hat{s}_i(t) \\ v(t) = (1 - \alpha_v)v(t-1) + \alpha_v(c(t) - c(t-1)) \end{cases} \quad (10)$$

其中 $\omega_i(t)$ 是粒子 i 在时刻 t 的加权, 表示各个限位框样本在给定观测下的可能性, $\sum_i \omega_i(t) = 1$, α_c 和 α_v 是适应速率。

根据如下似然函数计算粒子加权:

$$\omega_i(t) = \frac{E_i(t)}{E(t)} (D_i(t))^2 \quad (11)$$

其中 $E_i(t)$ 是重构前景 $\hat{x}_{FG}(t)$ 在粒子 i 限位框内的能量; $E(t)$ 是重构前景 $\hat{x}_{FG}(t)$ 的能量; $D_i(t)$ 是粒子 i 限位框的密度, 表示目标特征, 定义为粒子 i 限位框内大于某个阈值的元素数量除以粒子 i 限位框内元素总数。

5 实验结果

用不同的随机投影数与原始像素数之比 $\delta = n/N$ 测试了跟踪系统。按照第3章的方法, 将图像分为不重叠的 8×8 块, 通过计算每块的平均亮度构成向量 x 。 x 包含 $N = 1584$ 个像素, 取 $\delta = 0.2$, 产生 $n = 317$ 个投影, 选择量化步长 Δ 使得平均速率 $R_{DPCM} = 3$ 比特/投影, 一幅图像总的数据量为 $nR_{DPCM} = 951$ 比特, 信噪比为:

$$SNR = 10 \lg \frac{\|\hat{y}_{FG}\|_2^2}{\|\hat{y}_{FG} - \tilde{y}_{FG}\|_2^2} \approx 40 \text{ dB} \quad (12)$$

已经验证, 对于相同的失真, 标准PCM编码将需要平均速率 $R_{PCM} = 7$ 比特/投影, 因此DPCM方法允许节省大约4比特/投影。PCM编码的对象是没有经过背景消除处理的压缩测量 y , 编码结果是 \tilde{y} , y 中所含信息比 y_{FG} 多, 所以码率更高。

为了评估跟踪性能, 对重构的前景 \hat{x}_{FG} 设置阈值, 并比较估计的前景目标位置与真实的前景目标位置。通过改变阈值 τ , 可以建立给定测量数 n 和像素速率 $r_{DPCM} = \delta R_{DPCM}$ 比特/像素时的ROC曲线。定义 TPR 为检测概率:

$$TPR = \frac{\text{有目标且检测到目标}}{\text{有目标且检测到目标} + \text{有目标但未检测到目标}}$$

FPR 为误检率:

$$FPR = \frac{\text{无目标但检测到目标}}{\text{无目标但检测到目标} + \text{无目标且未检测到目标}}$$

通过画出随 τ 变化的 TPR 和 FPR 关系图建立ROC曲线。

ROC曲线表明利用先验信息可以使跟踪性能显著提高。粒子滤波限位框先验跟踪系统的ROC曲线非常接近准确限位框先验跟踪系统的ROC曲线。这意味着在第4章描述的粒子滤波器方案提供了前景限位框的一个非常好的估计。粒子滤波器限位框先验与理想先验之间仍存在一个小的差别, 今后可以进一步研究其他更好的解决方案加以改进。

图2(a)示出了大厅监控序列在 $\delta = 0.2$ 时的ROC曲线。不同曲线对应于CS加权向量 w 的不同计算方法。

(1) 无先验: 求解式(4)的如下优化问题而不使用任何跟踪系统:

$$\hat{\theta} = \min \|\theta\|_1 \text{ 满足 } \|\tilde{y}_{FG} - \Phi \Psi \theta\|_2 \leq \sigma \quad (13)$$

(2) 粒子滤波限位框先验: 使用本方案所提出的跟踪系统。

(3) 准确限位框先验: 不用任何跟踪系统, 从准确手工限位框开始计算加权窗, 即利用准确限位框计算的加权窗重构前景, 再估计前景目标位置。

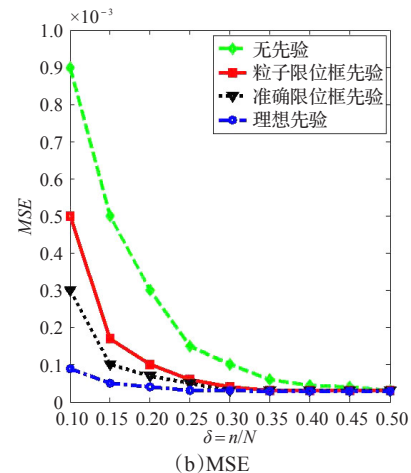
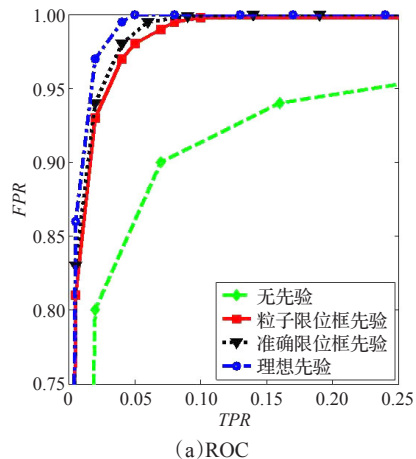


图2 采用不同先验重构时的系统性能

(4) 理想先验: 不用任何跟踪系统, 直接用实际前景的小波系数 $\theta = \Psi^T x_{FG}$ 计算加权窗。令 θ_i 为 θ 的第 i 个元素; 相应的加权 w_i 为:

$$w_i = \frac{1}{|\theta_i| + \epsilon} \quad (14)$$

图3示出了第50帧的压缩感知重构和粒子滤波跟踪结果。图3(a)是块平均的原始帧, 示出了直接从原始块平均序列估计的限位框和用重构的序列估计的限位框; 图3(b)示出了在小波域各子带中的 r_k ; 图3(c)为直接从原始块平均序列

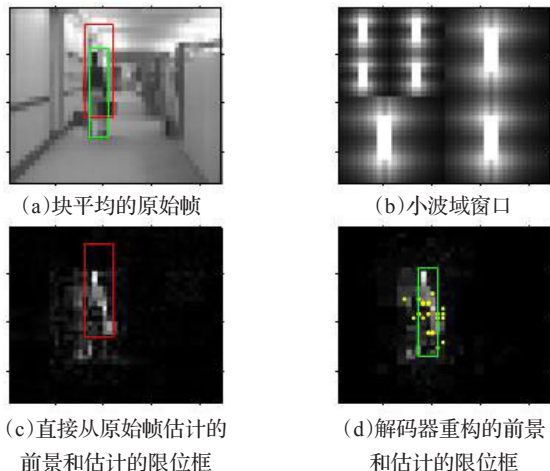


图3 第50帧重构和跟踪结果

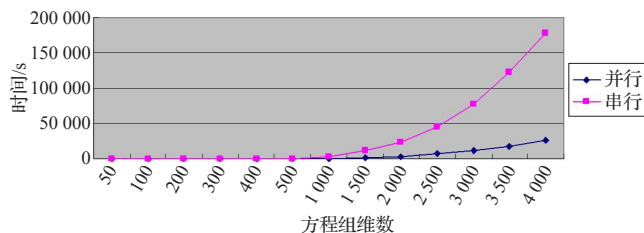


图7 时间比较图

运算时间,而在方程组维数规模超过300时,在GPU平台上的计算速度已经超过了在CPU上的计算速度。当计算维度超过1000的时候,计算加速比稳定在6.4~7.0左右。注意到,在维数为2000的时候,加速比有一个下滑,这是因为线程数量在达到一定的瓶颈之后,线程切换的时间在整体时间中的比例有一个增大的过程,但随着维数的增加,这个切换时间在整体时间消耗中的占用比例又会有所减小,表现为后面的加速比相对比较稳定。

4 结束语

在Gauss-Jordan消去法的基础上,本文提出了一种CUDA架构下的大规模稠密线性方程组的并行求解方法,并在GeForce GT240显卡上针对于不同规模的算例进行了测试,计算结果表明,该算法在线性规模小于100的时候,要远弱于测试平台的串行方法,而在计算规模超过300维的时候,其计算速度将超过串行方法,同时对于较大规模的线性方程组求解,如3000维以上,加速比比较稳定,在一定程度上摆脱了在GPU

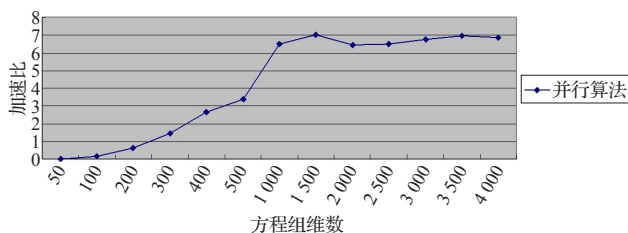


图8 加速比

上进行大规模线性方程组求解时,由于大量线程切换而带来的时间消耗在整体运行时间中的高比重问题。

参考文献:

- [1] 张舒,褚艳利,赵开勇.GPU高性能运算之CUDA[M].北京:中国水利水电出版社,2009.
- [2] Baader G M.Asynchronous iterative methods for multiprocessors[J]. J of the ACM,1978,25(2):226-244.
- [3] Quintana-Orti G, Igual F D, Quintana-Orti E S, et al.Solving dense linear systems on platforms with multiple hardware accelerators[J].ACM SIGPLAN Notices,2009,4(44):121-130.
- [4] 葛振,杨灿群,吴强,等.线性系统求解中迭代算法的GPU加速方法[J].计算机工程与科学,2009,31(A1):179-182.
- [5] 吴恩华.图形处理器用于通用计算的技术、现状及其挑战[J].软件学报,2004,15(10):1493-1540.
- [6] Kirk D B, Wen-mei W H.Programming massively parallel processors: a hands-on approach[M].San Francisco, CA, USA: Morgan Kaufmann Publishers Inc,2010.

(上接6页)

估计的前景和限位框;图3(d)为解码器估计的前景和限位框,圆点表示限位框粒子的中心位置。

6 结论

本文提出了一个具有隐私保护功能的视频跟踪系统。该系统分析固定摄像机捕捉的视频帧,并检测场景中的运动目标。提出的算法不需要重构原始视频序列就能够进行跟踪,从而实现隐私保护。利用压缩感知能够实现数据压缩,同时具有一定的安全性。

可以扩展系统功能,在多个目标和有遮挡的情况下进行跟踪。为了提高解码性能,需要研究其他重构策略,从而进一步减少所需随机投影数。例如,可以研究利用估计的限位框位置或根据目标剪影生成加权窗的不同方法,也可以修改加权方案,考虑小波系数在不同子带中的衰减。可以进一步研究如何在投影域中直接利用帧间的时间相关性。在编码级,可以研究更复杂的背景消除技术,以处理多峰背景分布,并保留在投影域中直接运算的可能性。

参考文献:

- [1] Cevher V, Sankaranarayanan A, Duarte M, et al.Compressive sensing for background subtraction[C]//Proc European Conf on Computer Vision,Marseille,France,2008.
- [2] Dufaux F,Ouaret M,Abdeljaoued Y,et al.Privacy enabling technology for video surveillance[C]//Proceedings of SPIE,2006,

6250:205-216.

- [3] 石光明,刘丹华,高大化,等.压缩感知理论及其研究进展[J].电子学报,2009,37(5):1070-1080.
- [4] Candes E.Compressive sampling[C]//International Congress of Mathematicians,Madrid,Spain,2006.
- [5] Donoho D.Compressed sensing[J].IEEE Transactions on Information Theory,2006,52(4):1289-1306.
- [6] Boyd S, Vandenberghe L.Convex optimization[M].[S.l.]: Cambridge University Press,2004.
- [7] Berg E V,Friedlander M P.In pursuit of a root,TR-2007-19[R].Department of Computer Science,University of British Columbia,2007.
- [8] Candes E,Wakin M B,Boyd S P.Enhancing sparsity by reweighted ℓ_1 minimization[J].Journal of Fourier Analysis and Applications,2008.
- [9] Candes E,Romberg J,Tao T.Stable signal recovery from incomplete and inaccurate measurements[J].Communications on Pure and Applied Mathematics,2006,59:1207-1223.
- [10] Piccardi M.Background subtraction techniques: a review[C]//IEEE International Conference on Systems, Man and Cybernetics, 2004:3099-3104.
- [11] Ristic B,Arulampalam S,Gordon N.Beyond the Kalman filter: particle filters for tracking applications[M].[S.l.]: Artech House, 2004.
- [12] Rowe D,Rius I,Gonzles J,et al.Robust particle filtering for object tracking[C]//International Conference on Image Analysis and Processing,2005,3617:1158-1165.