

一种自适应的时空视频去噪算法

凌波 叶秀清 顾伟康 杜歆
(浙江大学信息与通信工程研究所 杭州 310027)

摘要: 该文提出了一种基于运动估计的自适应时空视频去噪算法。算法直接嵌入在视频压缩编码中。空域上通过局部相似性寻找候选滤波点避免由噪声干扰形成的伪相似点,提高滤波点的有效性。时域上对各像素点进行运动预测,既增加了滤波点数量,减弱帧与帧之间的闪烁现象,又避免了因帧间运动而带来的运动模糊。另一方面,通过噪声强度和图像边缘信息自适应调节滤波窗口大小及滤波权重大小,实现了滤波强度自适应和空间细节的保护。

关键词: 自适应空时滤波; 运动预测; 噪声估计; 相似性测度; 边缘信息

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2008)10-2287-05

An Adaptive Spatio-Temporal Video Denoising Algorithm

Ling Bo Ye Xiu-qing Gu Wei-kang Du Xin

(Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027, China)

Abstract: A robust spatio-temporal video denoising algorithm is proposed based on the motion estimation, which is embedded into the video encoder. On one hand, this algorithm can be used in the space domain to eliminate the false similar pixels from neighbor pixels and improve the efficiency of pixels; it also can be used in the time domain to reduce the flicker among frames, and avoid motion blur. On the other hand, the algorithm can realize the adaptive strength of the filter and protect the special detail, when the size or the weight of the filters is adaptively adjusted by the strength of the noise and the edge information.

Key words: Adaptive spatio-temporal filter; Motion estimation; Noise estimation; Similarity measure; Edge information

1 引言

视频序列由于采集源、光线等的作用往往会受到噪声的干扰,使得画质下降。在带宽有限的网络中传输的视频需经过压缩,压缩算法通过时域上的运动搜索和补偿来去除时域上的冗余,而噪声的存在使得运动估计的准确率降低,从而使编码效率降低,码流增大,对传输造成很大影响。因此,去除视频序列中的噪声是很有必要的。

视频序列去噪的方法,基本可按空域,频域,小波域,时域等来进行划分。鉴于变换域对时间要求较高,选择在时空域上进行视频去噪。Mahmoudi 等人^[1]通过局部灰度平均,梯度等方法,在全图上搜索与当前像素相似的像素点,取得了很好的效果。但其处理时间长,且没有考虑时域信息,容易导致滤波后相邻帧之间有明显闪烁。Zlokolica 等人^[2]提出了一种新的基于运动和细节自适应的滤波方法,但其经验阈值过多,滤波窗口大小固定,对不同噪声强度适应性不好。Chan 等人^[3]提出的空时滤波算法,在时域上仅仅考虑前面一帧相同位置上的像素点,可能参与滤波的像素点比较

少,在噪声较强时,效果较差。

本文提出了一种直接嵌入编码器中的基于时空滤波的新算法。首先估计噪声强度并提取图像边缘信息,通过上述信息自适应选择滤波窗口的大小,再利用编码时得到的运动估计参数估计下一帧各像素点的运动参数,获取时域上候选滤波点的位置,利用时空相关性进行去噪滤波,从而很好地解决了边缘保持、运动模糊和滤波强度自适应等问题。

2 噪声强度估计

噪声强度的估计,是自适应滤波的一个重要环节。噪声强度 σ 估计得过高或过低,都会导致滤波后图像质量的严重下降。Amer 等人^[4]采用一种基于方向结构的算法,利用8个方向结构计算一个能表征图像块亮度均匀程度的测度,寻找一系列均匀块,通过这些均匀块,对视频的噪声强度进行估计。该方法处理速度快,对有大量纹理的图像效果也不错。

通过该方法处理高斯白噪声时,估计值与真实值差别不大。但是,在实际的视频序列中,由于很多经过了伽马矫正,在高亮区和低亮区噪声不是很明显,在这种情况下,找到的均匀块通常是在这些高亮度或者低亮度区域,而造成噪声强度估计不准确。因此对块的平均亮度要进行适当的限制。

2007-09-25 收到, 2008-04-22 改回

国家自然科学基金(60502006)和浙江省科技计划项目(2007C21007)资助课题

通过实验发现，将参与噪声估计的块的亮度平均值限制在 [25,230] 时，能很好地解决这个问题。

3 视频滤波去噪

利用视频信号在空间上的相关性和时间上的连续性，在空域和时域上寻找一些与目标像素相似的像素点来估计目标像素点的像素值。整个视频去噪算法流程如图1所示。

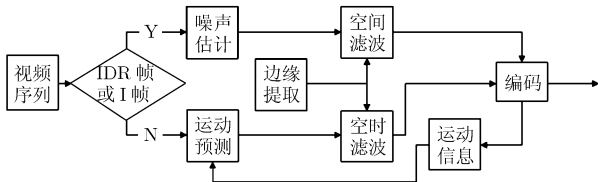


图1 视频去噪算法流程图

假设 $p(i)$ 为像素 i 的估计值， $p(j)$ 为与其相似的候选像素点，那么有

$$p(i) = \sum_{j \in R} w(i, j) \cdot p(j) \quad (1)$$

R 为当前帧 t 和前一帧 $t-1$ 内所有滤波候选点构成的集合。 w 是一个权重函数：

$$w(i, j) = 1/Z(i) \cdot \exp(-D^2(i, j)/\sigma^2) \quad (2)$$

$Z(i)$ 为归一化因子， σ 为噪声强度， $D(i, j)$ 是一个距离函数，定义为两像素邻域均值之差。设 ave 为 3×3 邻域均值函数，有

$$D(i, j) = |ave(i) - ave(j)| \quad (3)$$

下面讨论如何选择自适应窗口大小以及如何在空域和时域上选择滤波候选点。

3.1 自适应窗口大小的选择

滤波窗口大小直接反应噪声强度的大小。噪声强度大，滤波窗口大，反之，窗口小。同时，对于视频图像边缘部分，滤波窗口应该适当的小一些，可以避免滤波造成的边缘模糊。

边缘信息通过 Sobel 算子提取。为了防止噪声点的干扰，先通过一个 3×3 的低通滤波来减弱噪声的影响，然后获取亮度图上的边缘信息。如图2所示。边缘信息二值化的门限是对标准序列 Foreman, Akiyo, Carphone, News, Salesman, Suzie 等在噪声强度为 2 到 12 的范围内通过大量的实验统计获得的，它是关于噪声强度 σ 的线性函数。如式(4)所示。

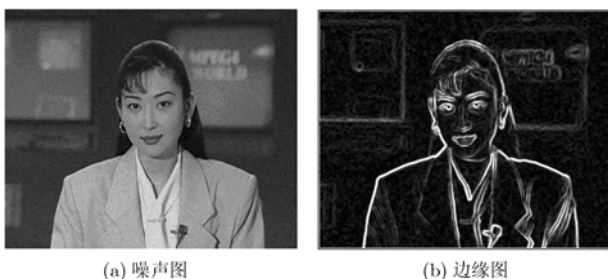


图2 $\sigma = 7$ 的 Akiyo 序列第1帧

$$th_edge = (\sigma + 1) \cdot 7 \quad (4)$$

根据边缘信息和噪声强度 σ 来确定自适应滤波窗口的大小。在非边缘点，滤波窗口的大小等于噪声强度的大小，为了计算方便，取奇数大小的滤波窗口。实验发现，当噪声强度 $\sigma \leq 5$ ，滤波窗口选择 3×3 的大小能达到最佳的滤波效果。在边缘点，在 $\sigma > 5$ 时，滤波窗口选择 5×5 ，当 $\sigma \leq 5$ 时，滤波窗口选择 3×3 。因此，自适应滤波窗口的大小是分段设计的。

$$win_size = \begin{cases} ((\sigma >> 1) << 1) + 1, & \text{非边缘点, } \sigma > 5 \\ 5, & \text{边缘点, } \sigma > 5 \\ 3, & \sigma \leq 5 \end{cases} \quad (5)$$

3.2 空间滤波点的选择

根据滤波窗口中像素点与中心像素点的相似程度来选择滤波候选像素点。一种简单有效的方法是以像素点的邻域均值作为该像素的邻域特征来衡量像素点相似程度。

假设当前处理的像素点为 $p(i)$ ，其 3×3 邻域均值为 $ave(i)$ ，滤波窗口中像素点 $p(j)$ 的邻域均值为 $ave(j)$ ，则相似性测度 $M(i, j)$ 定义为

$$M(i, j) = ave(i)/ave(j), \quad j \in \text{region} \quad (6)$$

region 是以当前像素为中心，大小为 win_size 的区域。设 α 为相似性测度因子，则有

$$(1 - \alpha) < M(i, j) < (1 + \alpha) \Rightarrow p(j) \in R \quad (7)$$

3.3 运动估计

在视频编码时，每个宏块或者块都要进行运动估计，来获取邻近帧的最佳匹配宏块或者块。由于相邻帧时间间隔很短，根据运动的连续性，运动的物体将保持前一帧的运动趋势，将前一帧的运动矢量块，按照其具有的运动矢量，投影到当前帧。当前帧中被其覆盖区域的整像素点将具有与其一样的运动矢量。如图3所示。由于在 H.264 标准中，块的最小单位为 4×4 ，因此，运动估计的最小单位也是 4×4 。考虑计算速度，运动矢量选择整像素精度。

3.4 时域滤波点的选择

运动估计后，当前帧的所有像素点分为两大类：一类为具有运动矢量的像素点，一类为没有被任何块覆盖的无运动矢量的像素点。根据这两种情况分别选择时域候选滤波点。同时考虑到运动估计的局部性以及一些运动估计快速算法的使用，运动估计搜索到的并不一定是最佳匹配，因此，需要通过一定的比较和限制来去除可能存在的影响。

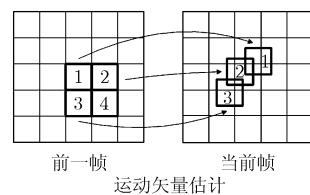


图3 基于块的运动预测

(1)像素点具有运动矢量 设 x, y, t 分别表示像素点在图像上的横坐标, 纵坐标, 以及时间轴上的时序坐标。当前帧像素点为 $p(x, y, t)$ 。 Δx 和 Δy 为其估计的运动矢量, ave 为邻域均值函数。

$$\begin{aligned} D_0 &= |\text{ave}(x, y, t) - \text{ave}(x, y, t - 1)| \\ D_1 &= |\text{ave}(x, y, t) - \text{ave}(x + \Delta x, y + \Delta y, t - 1)| \end{aligned} \quad (8)$$

D_0 为当前帧像素点 $p(x, y, t)$ 与前一帧相同位置像素点 $p(x, y, t - 1)$ 的相应邻域平均值的差值。 D_1 为当前帧像素点 $p(x, y, t)$ 与前一帧通过运动估计得到的相应像素点 $p(x + \Delta x, y + \Delta y, t - 1)$ 的相应邻域平均值的差值。如果 $D_0 \geq D_1$, 则在帧 $t - 1$ 中, 将以 $p(x + \Delta x, y + \Delta y, t - 1)$ 为中心, 大小为 win_size 的窗口作为滤波窗口, 反之, 则将以 $p(x, y, t - 1)$ 为中心, 大小为 win_size 的窗口作为滤波窗口。

(2)像素点没有运动矢量 直接将 $p(x, y, t - 1)$ 为中心, 大小为 win_size 的窗口作为滤波窗口。

确定了时域滤波窗口后, 对窗口中的每个像素, 利用式(6)和式(7)选择时域滤波候选点。

4 实验结果

实验中, 对一些加入高斯白噪声的标准序列, 以及实时采集下来的序列进行了测试。并同几个算法进行了比较。噪声强度估计中窗口为 5×5 , 窗口滑动时不重叠, α 取 0.1。

图 4 中, 将提出的算法和 SIGMA 滤波算法^[5], 自适应 KNN 算法^[2]进行了比较。图 4(a)为 CIF 大小 Foreman 序列第 15 帧的原图, 图 4(b)是高斯白噪声强度为 $\sigma = 10$ 的噪声图, 图 4(c)是本文提出的算法去噪以后的结果图, 图 4(d), 图 4(e)分别是自适应 KNN 算法, SIGMA 滤波算法的结果图, 从图中可以明显看出, 本文提出的算法, 能更好地去除噪声的影响。

图 5, 图 6 分别是 CIF 大小, $\sigma = 7$ 的 Akiyo 序列第 75 帧和 CIF 大小, $\sigma = 5$ 的 Carphone 序列第 164 帧的噪声图和去噪以后的结果图。在测试中, 当噪声强度逐渐减小到 $\sigma \leq 5$ 时, 滤波以后的图像信噪比基本没有提升, 有的甚至下降。分析其原因, 发现在弱噪声的影响下, 图像本身的信噪比还是比较高的, 比如 $\sigma = 5$ 时, 平均的信噪比在 34dB 左右, 而如果用 5×5 的滤波窗口对图像进行滤波去噪, 平滑的影响大于噪声本身对图像的影响。因此, 在弱噪声的情况下, 需要采用小的滤波窗口。所以, 当 $\sigma \leq 5$ 时, 滤波的窗口都为 3×3 。

图 7 是在实验室通过普通的摄像头采集的序列。由于受到灯光的影响, 序列中某些区域的亮度趋于饱和, 对噪声强度的估计影响很大。通过对区域亮度平均值的限制, 很好地解决了这个问题。该序列估计的噪声强度 $\sigma = 2$, 采用 3×3 的滤波窗口。

峰值信噪比是客观上评价图像序列去噪效果的重要尺度之一。分别对噪声图与原始图, 及滤波后的去噪图与原始

图求得峰值信噪比。图 8 是 Foreman, Akiyo, Carphone 去噪滤波前后峰值信噪比的对比图。可以明显的看出, 去噪后的序列峰值信噪比都有很大的提高。

表 1 为 Foreman 序列在噪声强度为 10, Akiyo 序列噪声

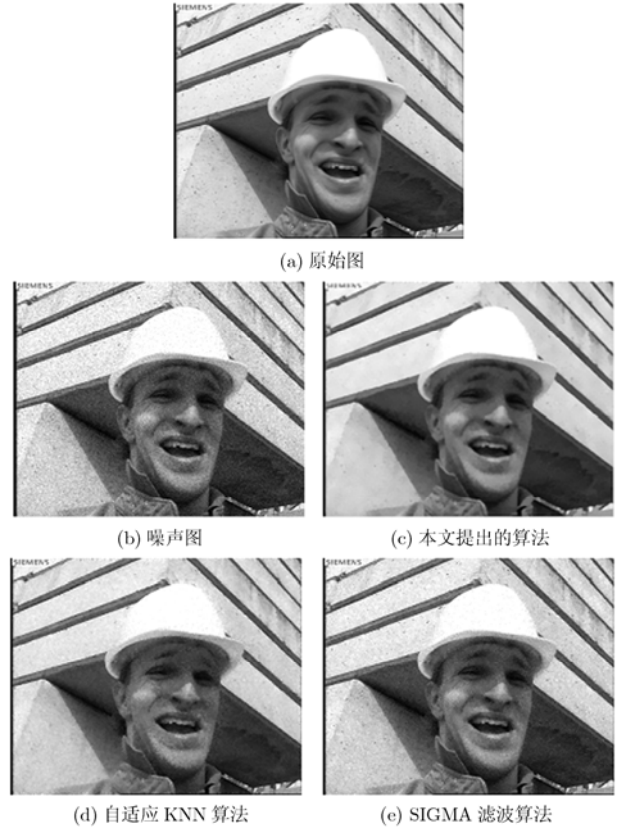


图 4 Foreman 序列第 15 帧, CIF, 噪声强度 $\sigma = 10$ 的原始图, 噪声图, 去噪结果图



图 5 Akiyo 序列第 75 帧, CIF, $\sigma = 7$

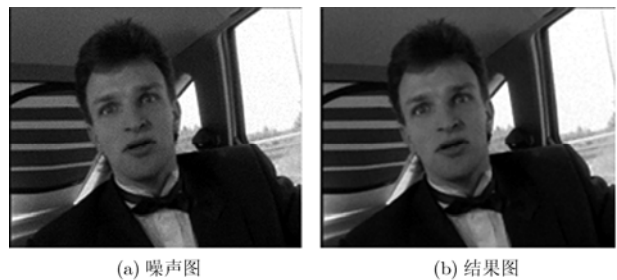


图 6 Carphone 序列第 164 帧, CIF, $\sigma = 5$



图 7 实时采集序列 LB, QVGA, $\sigma = 2$, 窗口为 3×3

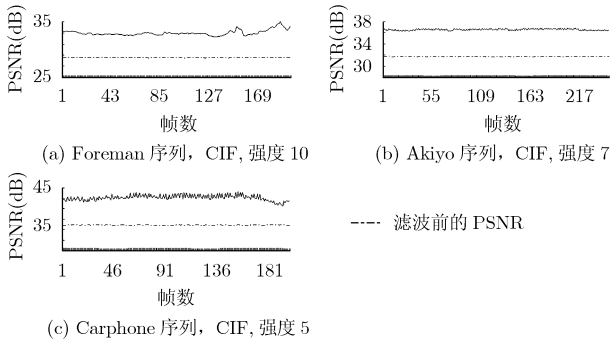


图 8 滤波前后峰值信噪比 PSNR 的对比

表 1 不同序列不同噪声强度下滤波前后的平均峰值信噪比 PSNR(dB)

序列	噪声强度	噪声图	SIGMA	KNN	本文算法
Foreman	10	28.49	29.69	29.71	33.32
Akiyo	7	31.73	34.21	34.29	36.84
Carphone	5	34.84	38.04	38.32	40.12

强度为 7, Carphone 序列噪声强度为 5 的情况下, 经过几种不同的去噪滤波算法后得到的平均峰值信噪比。从表中可以看出, 在 PSNR 客观评价上, 本文提出的滤波算法平均提高了 2dB 左右的峰值信噪比。

经过去噪滤波以后的视频序列, 不仅质量得到了提高, 同时, 编码后的码流也减小了很多。表 2 是 4 个序列去噪前后编码码率的对比。符合 H.264 的 baseline 标准。采用快速搜索算法, 量化参数为 10。从表中可以看出, 通过去噪处理后, 编码后的码流减小了很多。

5 结束语

本文提出的基于运动估计的保留细节的自适应时空联

表 2 去噪前后编码码流大小(M)比较

图像大小	噪声强度	帧数	去噪前码流	去噪后码流	去噪后/去噪前	
Foreman	CIF	10	200	11.36	5.16	45.40%
Akiyo	CIF	7	250	7.9	3.69	46.71%
Carphone	CIF	5	200	7.7	4.54	58.96%
LB	QVGA	2	391	9.45	5.53	58.52%

合视频去噪方法能够很好地去掉视频序列中的噪声干扰。该算法结合空域的相似性, 时域连续性, 实现了时空联合滤波, 提高了滤波的质量。同时, 通过噪声强度和边缘信息自适应地调节滤波窗口的大小以及滤波权重的大小, 实现了滤波强度自适应以及细节的保护。实验表明, 通过该算法去噪以后能提供很好的视频画质。在以后的工作中, 将深入进行两个方面的研究: 算法计算时间的优化; 针对目标区域进行的去噪滤波。

参考文献

- [1] Mahmoudi M and Sapiro G. Fast image and video denoising via nonlocal means of similar neighborhoods [J]. *IEEE, Signal Processing Letters*, 2005, 12(12): 839-842.
- [2] Zlokica V and Philips W. Motion and detail adaptive denoising of video [C]. *IS&T/SPIE 16th Annual Symposium on Electronic Imaging*, San Jose, USA, January, 2004: 403-412.
- [3] Chan Tai-wai, Au O C, and Chong Tak-song, *et al.* A novel content-adaptive video denoising filter [C]. *Acoustics, Speech, and Signal Processing*, Philadelphia, USA, March 18-23, 2005: 649-652.
- [4] Amer A and Dubois E. Fast and reliable structure-oriented video noise estimation [J]. *IEEE Trans. on Circuits and Systems for Video Technology*, 2005, 15(1): 113-118.
- [5] Ghazal M, Amer A, and Ghraryeb A. Homogeneity-based directional sigma filtering of video noise [C]. *IEEE International Conference on Image Processing*, Genova, Italy, Sept 11-14, 2005: 97-100.
- [6] Guo Li-wei, Au O C, and Ma Meng-yao, *et al.* An encoder-embedded video denoising filter based on the temporal LMMSE Estimator [C]. *IEEE International Conference on Multimedia and Expo*, Toronto, Canada, July, 2006: 841-844.
- [7] Lee Seong-won, Vivek M, and Jang Ji-hoon, *et al.* Noise-adaptive spatio-temporal filter for real-time noise removal in low light level images [J]. *IEEE Trans. on Consumer Electronics*, 2005, 51(2): 648-653.
- [8] Rabie T. Robust color video denoising [C]. *IEEE International Conference on Computer Systems and Applications*, Dubai/Sharjah, UAE, March 8-11, 2006: 792-798.
- [9] Xu R and Pattanaik S N. A novel Monte Carlo noise reduction operator [J]. *Computer Graphics and Applications*, 2005, 25(2): 31-35.

凌波: 男, 1980年生, 博士生, 研究方向为图像处理、视频压缩、视频信号处理。
 叶秀清: 女, 1936年生, 教授, 长期从事图像处理、视频信号处理、计算机视觉和机器人导航研究。
 顾伟康: 男, 1939年生, 教授、博士生导师, 浙江大学“通信与信息系统”国家重点学科学术带头人, 主要研究方向为模式识别、视觉信息处理、视频信息传输及智能机器人系统。
 杜歆: 男, 1975年生, 博士, 副教授, 主要研究计算机视觉、实时图像处理。