

# 背景提取与前景滤波相结合的时空联合视频降噪<sup>\*</sup>

周占民

(航空光学成像与测量技术研究所 中国科学院长春光学精密机械与物理研究所 长春 130033)

**摘要:** 针对视频图像的高斯型随机噪声,提出一种背景提取与前景滤波相结合的时空联合视频降噪算法。结合图像膨胀处理和背景差分法将视频图像分为背景和前景部分,前景部分和背景部分分别采用基于 Non-local means filter 的时空联合视频降噪算法和时域平均算法进行降噪处理,并将处理之后的前景和背景相加,得到最终的视频图像序列。最后,给出了 Non-local means filter 方法和本文降噪方法降噪效果的对比试验。实验结果表明,Non-local means filter 和本文降噪方法降噪后 2 个测试序列的 PSNR 分别为 33.0043、29.0365 和 35.8340、31.5261。这说明对于背景固定的监控类视频,该算法在降低算法复杂度、提高实时性的基础上,有效的处理和保留了视频图像的低频信息和高频细节。

**关键词:** 视频降噪;背景差分;Non-local means;时空联合

**中图分类号:** TN911.81 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Spatio-temporal combination method for video denoising with the combining of background extraction and foreground filtering

Zhou Zhanmin

(Research Department of Airborne Optical Imaging and Measurement Technology, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China)

**Abstract:** To solve the random gauss noise in the video, this paper proposed a spatio-temporal combination method with the combining of background extraction and foreground filtering. Video images in this paper are divided into two parts: The foreground and the background, using background difference method combined with image expanding. Then deal with the two parts using spatio-temporal combination method based on the Non-local means filter and direct meaning in time domain. And then the output video can be got by adding the foreground and the background which are proposed. At last, this paper give the comparison of the Non-local means filter and the method in this paper. Experimental results indicate that the PSNR of the tow test video after denoising reaches 33.0043, 29.0365 and 35.8340, 31.5261. Experiment shows that for the monitoring video with fixed background, the arithmetic in this paper effectively deal with the low-frequency information and the high-frequency details, on the basis of reducing the complexity and improving the real-time ability of the method.

**Keywords:** Video denoising; Background difference; Non-local means; Spatio-temporal combination

## 1 引言

视频图像信号在采集和传输的过程中,难免会受到污染,使视频图像产生噪声,影响视频图像视觉效果和主观质量,进而影响对图像信息的判断<sup>[1-2]</sup>。对视频图像进行降噪处理是行之有效的手段之一。视频降噪方法一般可以分为时域降噪法,空域降噪法和变换域降噪法几种<sup>[3-6]</sup>,而视频图像之所以能够进行降噪处理,是因为视频序列的图像信息在时域、空域或变换域上具有相关性,噪声在这些域中是不相

关的<sup>[3]</sup>。一般来说对于画面景物相对固定的视频图像,时域降噪的效果会比较好<sup>[7]</sup>,所以摄影师对静景拍摄时总是摆定位置连拍几张,然后做平均处理,这样就可以得到噪声相对较小的图像。空域滤波方法是针对单帧图像进行处理,比较典型的方法有高斯滤波,中值滤波以及由它们衍生出来的其他改进算法等<sup>[8-9]</sup>。典型的变换域降噪方法有傅里叶变换滤波、小波变换滤波<sup>[10-12]</sup>等,通过对信号多尺度的细化分析,小波变换方法更有利于保留图像的边缘和细节,但是由于小波

收稿日期:2015-03

<sup>\*</sup> 基金项目:吉林省科技发展计划(20126016)项目

基的选取比较困难,应用受到了一定的限制<sup>[12-13]</sup>。

Non-local means filter<sup>[14-15]</sup>是一种性能优良的空域滤波器,与传统的时域滤波相比,这种方法在性能上有了很大的提高,在其基础上也发展出了很多改进的方法<sup>[15-17]</sup>,他的权值的确定跟高斯滤波等方法不同,权值的大小与像素块相似程度相关在很大程度上降低了由于单个点存在比较严重的噪声而又以此点作为参考点带来的影响。但是单纯的 Non-local means 空域滤波方法的算法复杂度高,不能保证视频的实时性,同时,若降噪强度选用过大,还会产生图像细节模糊的现象,而且单纯的空域滤波还容易造成单帧图像效果较好,但视频连续播放时闪烁的现象。针对这些问题对 Non-local means 方法进行了改进。首先采用基于噪声估计的背景差分法将视频图像分为背景和前景部分;然后将典型的 Non-local means 空域滤波方法进行改进,扩展为时空联合滤波方法,并应用在已经分离出来的视频图像的前景上;然后将背景部分采用时域平均的方法进行降噪处理后,与前景部分相加后输出得到最终的视频图像。通过实验验证,本文算法取得了较好的效果。

## 2 视频降噪工作原理

### 2.1 算法整体框架

图1为本文设计的视频图像降噪算法流程。

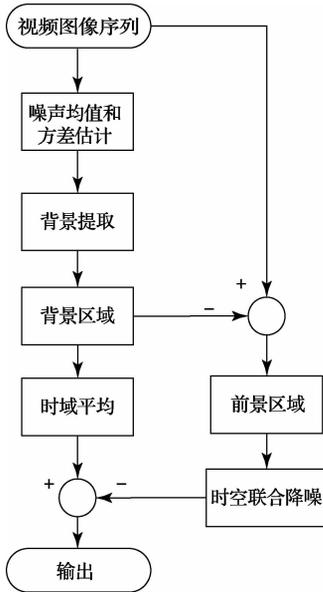


图1 算法整体框架

首先对输入视频图像序列进行噪声均值和方差的估计,然后设定阈值将视频图像分为背景和前景部分,对前景部分采用基于 Non-local means filter 的时空联合降噪方法进行降噪处理,背景部分采用时域平均的方法进行降噪处理。最后将降噪处理之后的背景和前景相加后得到最终的视频图像序列。

### 2.2 基于噪声估计的背景差分法

光流法、帧间差分法、背景差分法是当前比较常用的3种背景提取算法<sup>[18]</sup>,它们也经常被用来进行运动目标的检测。但是光流法实时性差、稳健性低,帧间差分法较难检测出完整的前景目标。选用被普遍采用的背景差分法,这种方法简单有效,实时性好。

设第  $k$  帧视频图像像素点  $x$  处的观测值为:

$$f(x, k) = p(x, k) + n(x, k) \quad (1)$$

式中:  $p(x, k)$  为第  $k$  帧图像中  $x$  处的值,  $n(x, k)$  表示第  $k$  帧图像中  $x$  处的噪声值。

针对背景固定的监控视频,背景部分短时间内可以近似认为:

$$p(x, 0) = p(x, 1) = \dots = p(x, k) \quad (2)$$

则,可以得到:

$$Y(x, k) = f(x, k) - f(x, k-1) = p(x, k) - p(x, k-1) + n(x, k) - n(x, k-1) = n(x, k) - n(x, k-1) \quad (3)$$

视频图像中高斯噪声是较为常见的一种噪声,根据高斯噪声的特点,假设待处理视频图像的噪声为具有独立同分布特性的加性噪声。则有:

$$E(Y) = 0 \quad (4)$$

$$\sigma_Y^2 = 2\sigma_n^2 \quad (5)$$

式中:  $E(Y)$  和  $\sigma_Y^2$  分别为  $Y$  的均值和方差,  $\sigma_n^2$  为  $n$  的方差。那么  $Y(x, k)$  的均值和方差就可以用下面两个公式来估计。

$$\hat{u} = \frac{1}{M \times N} \sum_{x \in B_{M \times N}} Y(x, k) \quad (6)$$

$$\hat{\sigma}_Y^2 = \frac{1}{M \times N} \sum_{x \in B_{M \times N}} [Y(x, k) - \hat{u}]^2 \quad (7)$$

具体计算时,在图像的背景区域取一个  $M \times N$  像素的区域块  $B$ ,然后在当前像素块内将当前帧及其前一帧视频图像做差得到  $Y(x, k)$ ,  $x \in B$ ,然后再按照式(6)和(7)进行视频图像噪声和方差的估计。

根据分析可知,当视频图像中景物未发生剧烈的运动时,差值  $Y(x, k)$  表示的主要为视频图像的前、后两帧视频图像之间噪声的差。当视频图像中景物发生剧烈的运动时,差值  $Y(x, k)$  表示的除了视频图像的前后两帧视频图像之间噪声的差以外,主要为视频图像中景物运动产生的差值。

根据理论分析,当视频图像中景物未发生剧烈的运动时,  $Y(x, k)$  表示的主要为视频图像的前后两帧视频图像之间噪声的差。当视频图像中景物发生剧烈的运动时,  $Y(x, k)$  表示的除了视频图像的前后两帧视频图像之间噪声的差以外,主要为视频图像中景物运动产生的差值。根据数理统计原理,背景  $|Y(x, k)|$  的值会以很高的概率小于前景  $|Y(x, k)|$  的值。这样根据正态分布的性质,可以设定阈值  $T = 3\hat{\sigma}_Y$ ,那么  $Y(x, k)$  将以 99.74% 的概率落在  $[-3\hat{\sigma}_Y, +3\hat{\sigma}_Y]$  内。

具体进行背景提取时,首先设定提取背景所需要的视频图像的帧数  $N$ ,并且对应每个像素点初始化一个变量  $num(x) = 0$ ;设定阈值  $T = 3\hat{\sigma}_Y$ ,比较  $|Y(x, k)|$  与阈值  $T$  的大

小,如果  $|Y(x,k)| < T$ ,则将第  $k$  帧图像  $f(x,k)$  对应像素记为背景区域,对应的将  $num(x)$  的值加 1,而将  $|Y(x,k)| \geq T$  对应的像素区域记为前景区域;最后将记录的变量  $num$  的值等于  $N-1$  的所有像素点,按式(4)~(8)进行处理;

$$f(x) = (f(x,k) + f(x,k-1) + \dots + f(x,k-N+1))/num(x) \quad (8)$$

最后将所有满足背景条件的像素点输出,就得到了当前的背景。

在背景提取时,因为背景复杂或阈值设定不合适就会导致背景提取不完整,视频图像不均匀的现象,在实际提取时,对提取出的背景进行了适当的膨胀处理。背景提取的效果如图 2 所示。



图 2 mother-daughter 序列背景提取效果

由实验效果可以看出,采用背景差分的方法确实可以有效地将整个视频画面的前景部分(如图 2 中的白色部分)和背景部分(如图 2 中的黑色部分)有效的分离开来,而且相对比较完整,这样就为接下来的视频降噪奠定了基础。

### 2.3 基于 Non-local means filter 的时空联合视频降噪算法

假定视频图像的待处理区域为:

$$v = \{v(x) | x \in B\} \quad (9)$$

式中: $x$  为  $B$  内的像素点,则经过 Non-local means filter 降噪之后的图像为:

$$NL[v](x) = \sum_{y \in B} w(x,y)v(y) \quad (10)$$

式中: $NL[v](x)$  是降噪之后的图像表达式, $w(x,y)$  是表征像素点  $x$  和像素点  $y$  之间的相似程度的权值,满足  $0 \leq w(x,y) \leq 1$  且  $\sum_y w(x,y) = 1$ 。

Non-local means filter 在很大程度上解决了由于个别点存在比较严重噪声的情况下,又以此作为参考点对降噪效果带来的影响。

但是即使每一帧视频图像的处理效果都比较好,连续播放时仍会出现闪烁的现象,因此在 Non-local means filter 基础上做了时域的改进,如图 3 所示,改进的视频降噪算法

不仅考虑当前视频帧,而且还考虑了当前帧的前一帧和最后一帧。具体的计算公式如下。

$$g(x) = \sum_{i=-1}^1 \sum_{y \in B_{i+1}} w_{i+1}(x,y)v_{i+1}(y) \quad (11)$$

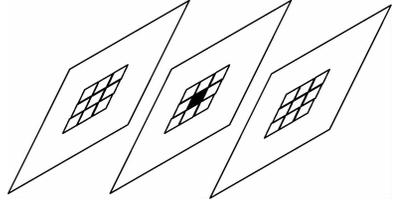


图 3 时空联合视频降噪算法示意

式中: $i=0$  表示当前帧, $i=-1$  前一帧, $i=1$  表示后一帧。式中的  $w(x,y)$  按下式计算。

$$w(x,y) = \frac{1}{Z(x)} e^{-\frac{\|v(N_x) - v(N_y)\|_{2,a}^2}{h^2}} \quad (12)$$

式中: $\|v(N_x) - v(N_y)\|_{2,a}$  是加权欧式距离计算式, $a > 0$ ,是一个高斯核的标准差。 $h$  为参数,可以通过改变其大小来控制滤波强度。 $Z(x)$  为归一化函数,其表达式为:

$$Z(x) = \sum_y e^{-\frac{\|v(N_x) - v(N_y)\|_{2,a}^2}{h^2}} \quad (13)$$

可以看出当搜索范围内 2 个点的相似程度高,则权值  $w(x,y)$  就高,相反权值就低。

### 3 实验结果

为了验证该文设计算法的效果,进行了实验验证。实验算法在自行设计的某监控系统上实现,硬件 CPU 采用飞思卡尔公司 i.MX535 系列芯片,嵌入 Linux 系统,实现了 4 通道视频采集与处理<sup>[19]</sup>。实验硬件条件如图 4 所示。

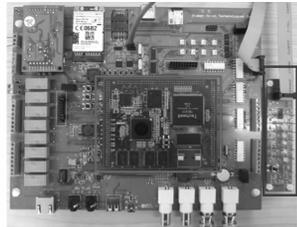


图 4 实验硬件条件

实验系统视频采集和降噪处理过程如图 5 所示,首先 4 通道模拟视频信号通过硬件滤波后进入 TW2835 芯片,通过 TW2835 芯片的 A/D 转换等一系列操作后将数字信号送入 i.MX53 的视频图像处理单元,在 i.MX53 中进行视频图像信号的降噪处理操作,最后通过 Linux 系统 V4L2

接口 API 传送给上层采集与显示应用程序。

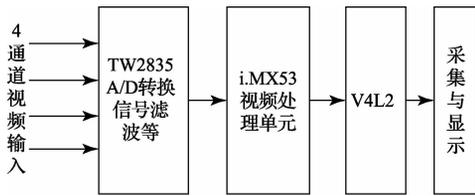


图5 实验系统视频处理过程

本文采用标准视频序列和实验视频序列分别进行降噪处理的效果对比如图6和图7所示。



(a)噪声图像



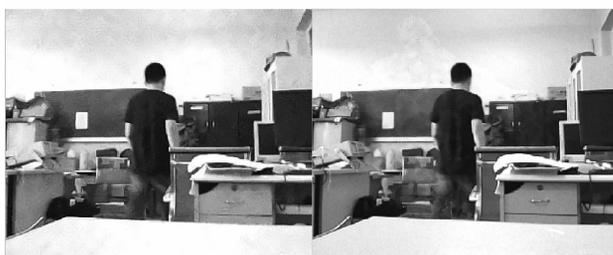
(b)NL降噪

(c)本文算法

图6 标准序列降噪对比



(a)噪声图像



(b)NL降噪

(c)本文算法

图7 实验序列降噪对比

实验分为2组,一组实验采用标准视频序列 mother-daughter;另一组实验为实验系统采集得到的视频序列 test。2组实验的结果分别为:图6(a)~图6(c)和图7(a)~图7(c)。其中图6(a)和图7(a)为噪声图像,图6(b)和图7(b)为采用NL方法降噪后的图像,图6(c)和图7(c)位采用本文方法降噪后的图像。表1为实验结果的PSNR值对比。可以看出,本文算法在视频图像降噪效果的客观评价<sup>[20]</sup>上具有较明显的优势。

表1 实验结果 PSNR 值

	高斯噪声 图像	NL 降噪	本文算法
mother- daughter	29.2452	33.0043	35.8340
test	25.6320	29.0365	31.5261

## 4 结 论

提出了一种背景提取与前景滤波相结合的视频降噪方法。对背景部分采用时域平均滤波,保留了时域平均滤波对于固定景物的良好降噪效果;对前景部分采用基于 Non-local means filter 的时空联合视频降噪,既保留了NL方法的降噪效果,同时加入了时域降噪的部分,保留了更多的图像细节的同时,又消除了单纯进行空域滤波容易造成的闪烁现象。实验结果表明,这种方法对于背景相对固定的监控类视频具有明显的优势,既能有效的处理视频图像的低频信息,又能通过调节前景滤波强度较好的保留高频部分的细节,同时对于视场内动目标比较少的视频图像,该算法还能有效的降低算法的复杂度,提高算法的实时性。

## 参考文献

- [1] LU Q B, LU Z B, TAO X Q, et al. A new Non-local video denoising scheme using low-rank representation and total variation regularization [J]. International Symposium on Circuits and Systems (ISCAS), 2014: 2724-2727.
- [2] 高辉. 数字视频监控实时降噪算法研究[D]. 广州:暨南大学, 2012: 1-5.
- [3] 周占民. 通信基站监控系统视频降噪算法研究与实现[D]. 哈尔滨:哈尔滨工业大学, 2013.
- [4] 龙红梅, 田逢春, 谭洪涛, 等. 一种基于运动估计的3D视频降噪算法[J]. 计算机与数字工程, 2009(6): 122-124.
- [5] CHEN G Y, BUI T D, KRZYSAK A, et al. Adaptive video denoising using block matching 3-D filtering[J]. IEEE Canadian Conference on Electrical & Computer Engineering (CCECE), 2012(25): 1-4.
- [6] REN J, ZHUO Y, LIU J Y, et al. Illumination-

- invariant Non-local means based video denoising[J]. IEEE International Conference on Image Processing (ICIP), 2012(19):1185-1188.
- [7] GHAZAL M, SU C, AMER A. Motion and region detection for effective recursive temporal noise reduction [C]. Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on. IEEE, 2007, 1: I-677-I-680.
- [8] LU J, ZOU Y, YE Z. Enhanced fractal-wavelet image denoising [C]. ISECS International Colloquium on Computing, Communication, Control, and Management, 2008. CCCM08. IEEE, 2008, 1: 115-119.
- [9] WALKER J S, CHEN Y J. Image denoising using tree-based wavelet subband correlations and shrinkage [J]. Optical Engineering, 2000, 39(11): 2900-2908.
- [10] DAI J, AU O C, PANG C, et al. Video denoising based on transform domain minimum mean square error[C]. 2011 18th IEEE International Conference on Image Processing (ICIP). IEEE, 2011: 2573-2576.
- [11] BALSTER E J, ZHENG Y F, EWING R L. Combined spatial and temporal domain wavelet shrinkage algorithm for video denoising[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(2): 220-230.
- [12] 杨文波, 马天玮, 刘剑. 非局部变分修复法去除高密度椒盐噪声[J]. 中国光学, 2013, 6(6): 876-884.
- [13] YIN W Y, ZHAO H W, LI G P. et al. A block based temporal spatial Non-local mean algorithm for video denoising with multiple resolution [J]. International conference on Signal Processing and communication systems(ICSPCS), 2012(6):1-4.
- [14] BUADES A, COLL B, MOREL J M. A non-local algorithm for image denoising[C]. CVPR 2005. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. IEEE, 2005(2): 60-65.
- [15] WANG H Z, CHEN L, XU S L. Improved video denoising algorithm based on spatial-temporal combination[J]. International Conference on Image and Graphics (ICIG), 2013(6):64-67.
- [16] 叶庆伟, 谢永昌, 狄红卫. 基于 Non\_local means 的时空联合视频降噪算法[J]. 微型电脑应用, 2012(5): 32-35.
- [17] 李高西, 曹军, 张福元. 基于视觉灵敏度及粗集的彩色图像滤波算法[J]. 电子测量与仪器学报, 2014, 28(2):211-217.
- [18] 文丽丽, 罗洪艳, 张少祥, 等. 数字人脑切片图像中小脑组织的连续自动分割[J]. 仪器仪表学报, 2013, 34(1):133-139.
- [19] 杨思, 蔡晓东, 李长俊. 移动可视化智能监控系统[J]. 国外电子测量技术, 2013, 32(4): 60-62.
- [20] 安雪晶, 田媛. 无参考图像质量评价方法的设计原则[J]. 中国光学, 2009, 4(2):140-144.

## 作者简介

周占民, 1989 年出生, 博士, 2013 年于哈尔滨工业大学获得硕士学位, 主要研究方向为航空光学成像技术、数字图像处理技术与嵌入式系统开发。

E-mail: 405932@qq.com