<u>EE392J Project Report</u>

Motion-Compensated Noise Reduction in B&W Motion Picture Films

ZHU Xiaoqing
SUID: 4883211
03/15/2002

==

Abstract

Blotchy noise is one of the most common and visually annoying noises noticed in digitized black-and-white motion picture films. This work investigates presents a three-stage blotchy noise reduction scheme combining efforts of temporal median filtering and motion-compensated filtering after joint motion/noise detection. Implementation of each module of the scheme is discussed and finally performance of the scheme is tested with several short sequences and results are discussed.

1. Introduction

With the growing popularity and power of digital video processing today, many of the conventional media materials, like old black-and-white (B&W) motion picture films, are being digitized for convenience of storage and manipulation. However, direct conversion of these films usually yields in very poor quality due to presence of noise and degradation of the original analogue storage material over a long period of time, therefore it is desirable to process these converted sequences afterwards and provide some visual enhancement for them.

This project looks into noise reduction of B&W motion pictures and mainly works on the removal of blotchy noise, since it is the most common and visually noticeable degradation of digitized videos. I experimented with different denoising filters and noise detection methods, as well as proposing my own scheme of blotchy noise reduction, which combines the effort of temporal median filtering of the entire sequence and motion-compensated filtering of selected areas after joint motion/noise detection. Finally a brief evaluation of the scheme and possible ways for further improvement are discussed.

2. Blotch Noise Features

It is generally noted that artifacts of digitized old B&W films include blotches (“dirty spots”), scratches (vertical thin lines), spatial instabilities (e.g. jerky displacement of adjacent frames due to misalignment of film in the projector) and intensity instabilities (illumination fluctuation and gray shadowing on a large area of the frame). In literature detection and removal of line scratches have been studied at length and removal of these lines are treated as missing data interpolation problems following an underling Autoregressive (AR) model or Bayesian framework.. On the other hand remedy for blotchy noises are less addressed explicitly as an individual problem. They are, more or less, viewed as an “easier” kind of noise that are supposed be automatically removed during some general denoising procedures for digital video.

However, a closer look at these blotches reveals that they appear as dirty patches and affect a continuously connected area of the frame, that they are of arbitrary shape and size, with varying range of intensity (not necessarily purely black or white), yet showing high contrast against the background, and instead of staying at the same place over multiple frames (as scratch lines do) blotches tend to flicker around at random positions from frame to frame. All these characteristics are very different from commonly used noise models such as the popular additive white gaussian noise or salt-and-pepper noise. Presumably then general denoising methods such as joint spatio-temporal (motion-compensated) filtering wouldn't work well. We thus need to treat them more specifically.

3. Blotch Removal Scheme

There are several challenges and constraints in trying to remove blotchy noises. First, it is noted that, while blotchy noises do not fit into any existing popular noise models for digital video, trying to find a new model for them may also be extremely difficult since they present little mathematical tractability. Therefore model-based detection methods, which proved to be successful for line scratch removals cannot be adopted here. Secondly, presence of blotchy noise together with other artifacts like intensity instability might lead to inaccurate or spurious motion-estimation results which are crucial if we want to incorporate motion-compensated filtering for noise reduction. Thirdly, since there are huge amount of video data converted from this source and since for the project algorithms are implemented within MATLAB computational complexity shouldn't be too huge. And lastly, there shouldn't be obvious artifacts or blurring introduced after blotchy noises are removed.

To overcome the above difficulties and constraints, the proposed scheme process the sequence following the three steps as shown in Figure 1: temporal median filtering, joint motion/noise detection and motion estimation/compensated filtering. For each module in the block diagram different implementation options have been experimented, their differences are compared and discussed in each subsection below.

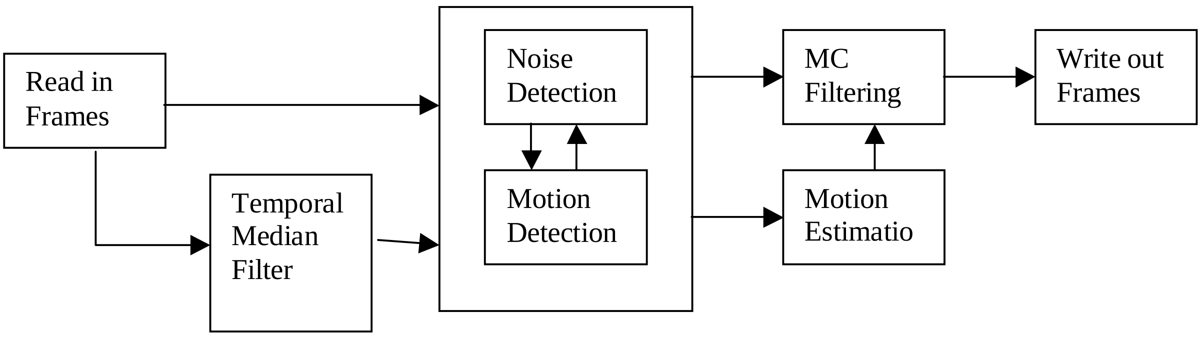

Figure 1 Block diagram

3.1 Pre-denoising filter

First step is to denoise the whole sequence with some simple filter so that we might have a "cleaner" version of the original sequence to help in the later motion estimation/compensation stage. Only median filters are considered here due to the simplicity of their implementation as well as the fact that they introduce less blurring effects than other linear filter. I tried with:

a) purely spatial filtering (3*3 cross median filter, frame by frame process)
b) purely temporal filtering (median window length=3)
c) purely temporal filtering (median window length=5)

Figure 2 Original Frame

Figure 3: Purely spatial filtering

Figure 4: Purely temporal filtering (Window length=3)

Figure 5: Purely temporal filtering (Window length=5)

Sample resultant frames from each filter are shown in figures 2-5. We can tell that all the filters provide some smoothing effect and have generally denoised the frame due to their low-pass nature but details (e.g. the single line in the sky) are preserved. In terms of blotch removal, it is obvious that purely spatial median filter is almost help less, whereas temporal median filter of length 3 has alleviated the blotches substantially, and in the result of 5-tap temporal median filter blotches are barely distinguishable from the background. This result is expected since blotch noises extend to certain areas in space but do not persist over time. As a matter of fact in the resultant sequence of 5-tap temporal median filtering, almost all blotchy noises are already removed.

This simple comparison motivates the choice of 5-tap temporal median filter over the other two. And as the temporal median filter smoothes out sharp transitions in intensity at each pixel position, it not only denoised the whole frame and removed blotches, but also helped in stabilizing the illumination fluctuations. This is beneficial both visually and for motion estimation performance in the next stage.

However, it is worth noting that when tested with another sequence containing fast moving parts, artifacts arise due to lack of motion compensation. It is further noted that artifacts occur with moving object as temporal median filtering shall "shrink" the size of that object, and for extreme cases like a moving single line it might disappear completely in the filtered sequence. Figure 6 below illustrates the effect of temporal median filtering sans motion compensation. Nevertheless, we can try to recover these artifacts by applying motion compensated filtering again on the original frame with help from this much 'cleaner' yet partially artifacted version of the sequence.

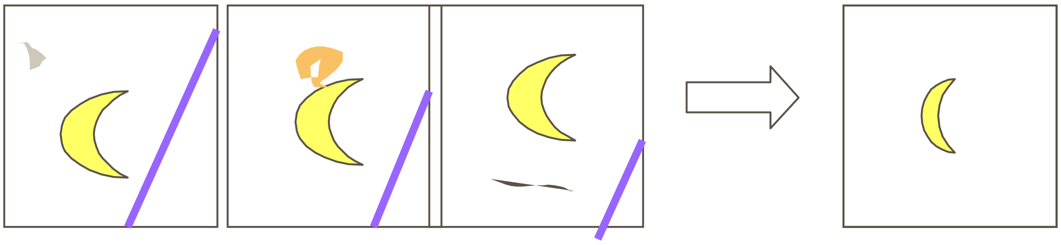

Figure 6 Effect of temporal median filtering w/o MC

3.2 Joint Motion/Noise Detection

Note that the result from temporal median filtering is already very satisfactory for regions without motion[1].In order to save computation, therefore, in the second stage we try to detect whether motion exists so as to exclude unnecessary motion estimation in the post-processing stage. It is also noted that motion may only exist in part of the frame, so the detection is carried out in a section-wise manner and are spatially adaptive to different regions of each frame.

Each frame is divided into 8*8 non-overlapping sub-sections that are scanned in turn. The detection step is carried out in a "sandwiched" and parallel manner, comparing each section with

[1] To push the statement to the extreme, I tried with a synthetic sequence consisting a repeated still image with added noise to randomly selected frames.The 5-tap median filter achieved almost perfect restoration of the original noise-free sequence.

its counterparts in the previous and next frames, as well as the three corresponding sections in the median-filtered version, and decision is made out in two steps:

A. Detect whether there is significant change between the original and filtered version.
If there is no comparatively small change other than noise, we do not need to post-process that part at all. This detection is simply checking whether the mean absolute difference (MAD) of the current section in the original frame (denoted as A) and filtered frame(B). In order to exclude the case where there is too much noise scattering around, we add one more criterion, namely the number of "significantly" differing pixels should be more than 5% of the overall section. The threshold of being "significantly different" can be adjusted to favor either sensitivity to motion or immunity to noise. In the project it is chosen to be 15 empirically.

B. Decide whether the detected change is due to motion or noise.
Criterion is that if there is blotchy noise instead of motion, the section difference for adjacent frame compared should be consistent for both the unfiltered and filtered versions. After trial with several formulas to represent this similarity, I realized that while the denoised version may have a lower average difference value than the noisy one, the ratio between section difference of adjacent frame and that of consecutive odd/even frames should be approximately the same as for the original case if there is no interruption of blotchy noise. So the following criterion is used for rejecting blotchy noise:

Abs((ErrLA/ErrLNA)/(ErrLB/ErrLNB)-1)<Threshold
& Abs((ErrNA/ErrLNA)/(ErrNB/ErrLNB)-1)<Threshold

Here ErrLA is MAD between the last section and current section for original frames, ErrNA is MAD between next and current sections, ErrLNA is MAD between last and next sections, and ErrLB, ErrNB, ErrLNB are corresponding values for the filtered version. Again the value of Threshold can be adjusted to trade off between computational complexity of motion estimation as price paid for false alarm against penalty of losing spatial texture in regions of small motions with insufficient/sloppy detection. In the project the value is chosen to be 10%.

3.3 Post-processing with MC-median filter

Once motion is detected, estimation of motion vector field is carried out for every pixel in that section. Full search block matching method is used with search range (-16, +16) and not surprisingly this is the computational bottleneck of the whole program. For each pixel its neighboring 9*9 block of pixels is assumed to have the same vector field. This implicitly takes advantage of the smoothness assumption of motion vector field and result in a smoothly changing motion vector field for the processed section.

Since we have been checking each frame in a sandwich manner and already have access to its two temporal neighbors, it is convenient to do bi-directional motion-estimation and compensation for a more robust result plus a benefit of averaging along motion-trajectory for further noise reduction. The price to pay, of course, is twice as much calculation in motion estimation.

In order to improve the accuracy and robustness of motion trajectory estimation, optimization criterion is chosen as weighted mean of absolute frame difference. The weight is reciprocal of difference between median filtered pixel and original pixel plus a constant term 1:

WeightMap =1./(ErrFrame+1)

where ErrFrame records difference between original and filtered frame and the constant term prevents the weights from blowing up to infinity where frame difference is 0. The reciprocal relationship incorporates the knowledge from the "cleaner" frame to level off influence of outliers in the summed error term.

Now we arrive at the final step of actually modifying the median-filtered sequence to remove artifacts without reintroducing blotchy noise. First bi-directional motion estimation and compensation is calculated for each section detected as having motion. Then we apply a median filtering over the three different versions of that frame: original, median-filtered, and motion-compensated. There are a number of ways to combine these candidates together and try combine the strength of all three versions and come up with a new frame that is neither noisy nor blurry.

Linear filtering is possibly a bad idea since small blotches may show up again as part of the weighted sum, and in regions of motion discrepancy of the median-filtered version and the original one might lead to fancy boundaries of moving object. And it can be seen that spatial filtering shall introduce more blur than the temporal filter, and since we now already have a denoised version of the sequence, the blurring caused by spatially filtering again is both unwanted and unnecessary. Therefore direct median filtering is adopted and this time we can apply it in a motion-compensated flavor. Note that for frames or sections in a frame where no motion compensation is involved, this step of filtering again is not necessary. And for regions with motion, we should try to maintain the original spatial texture meanwhile rejecting the potential blotches. I have tried with several strategies of selecting from the different versions and it seems that taking the median value out of the three turns out to have pretty good results.

4. Test Results

I obtained several B&W motion picture sequences from the website Internet Moving Images Archive: Movie Collection (http://www.archive.org/movies/) and tested my algorithm with parts from the sequences. In order to evaluate the performance from several aspects, I deliberated selected sequences according to their level of noisiness and how much motion they contain. We therefore arrive at four different cases in combination:

A. <u>Almost-static sequence with severe blotchy noise: (view of Golden Gate Bridge)</u>

For such kind of sequences the temporal median filter step should already produce satisfactory step and in the later motion detection phase we should seldom sink into the tedious work of motion estimation and compensation. I had a "MotionMarker" flag the code which indicates whether ME/MC needs to be carried out, and as expected, for this sequence of 70 frames, we only calculated motion vector field for 74 sections altogether. The saves the computation by a factor of 70*8*8/74=60.

B. <u>Complicated scene with lots of motion and heavy noise (scene of people walking around on Golden Gate Bridge)</u>

It can be seen that most of the serious blotches are removed after the process, and since the original video was already at low resolution, there is no noticeable degradation in motion rendering of the scene after blotch removal. In terms of computational cost, out of ?? frames altogether ?? sections are processed with motion estimation.

C. Scene containing some apparent motion and also obvious noise (closing up show of man and a flying flag)

Since the original video is of better quality and less severe noise, close examination each frame reveal some remaining distortion of the pre-processing median filter. However, when viewing the processed video, the difference is somewhat less obvious. Blotchy noises are still removed successfully. Out of 10 frames motion estimation was carried out for 57 sections. This may indicate the saving factor in general as around 10*8*8/57>= 10.

D. Scene with lots of motion but less noticeable noise (children playing with swings and running)

This is the hardest trial on the algorithm and due to the presence of fast motion; distortion of direct temporal filtering is pretty serious. As in Appendix Figure ?, the running man in the front almost disappeared into the background. After motion compensated filtering he is partially restored into the scene but still with the whole sequence one can notice the difference. Removal of blotchy noise is still achieved. Although there was on dark line which persisted for several frames in the sequences.

5. Conclusion and Further Discussion

To conclude, we still lose object details when it is moving very fast, yet for old time motion pictures the original frame is also very blurry too, so perhaps a high quality rendering for such kind of situation is hard to achieve anyway. As for near-static scenes or those with slow camera panning/zooming, which is typically applied photographing skills in the old days, the results is very satisfactory. Frequently appearing blotches that flicker around the scene almost disappear and as a by-product of temporal median-filtering the whole scene is denoised in general and instability of scene illumination, which is another annoying artifact of digitized B&W motion pictures, is also alleviated to some extent.

In terms of computational cost, the block matching full search is obviously bottleneck of the system. However our joint motion/noise detection stage helped to reduce the number of needed motion estimation by a factor of ?? for static scene and ?? in general. One further saving might be achieved by using log search methods or apply the redundancy between forward and backward motion vector fields and predict from each other, or by applying iterative OFE methods with the prediction as an initial guess.

It should also be pointed out that the block-matching method still seem to yield inaccurate motion estimation result and thus affect the recovery of distortions by temporal median filtering. One way to improve it is to try half-pixel accuracy motion vector values, which shall cost more computation or another way is to apply original frame pixels directly on positions where distortion occur. I have tried joint edge detection of the difference frame between original and filtered version as well as the difference between temporally adjacent frames and then imposing original pixel values on the median filtered ones at detected regions. Initial results have better recovery of the blurring effect but quality of the processed video rely heavily on the performance of edge detection. Further research might be conducted to find better edge detection and maybe try with a combination method again to reduce the distortion.

6. References

[1] Joyeux L., Buisson O. et al. Detection and Removal of Line Scratches in Motion Picture Films, IEEE 548-553, April 1999.

[2] Kokaram Anil C., Removal of Line Artefacts for Digital Dissemnation of Archieved Film and Video, IEEE 245-249, September 1999

[3] Chong M.N. Liu P. et al. A New Spatio-Temporal MRF Model for the Detection of Missing Data in Image Sequences, IEEE 2977-2980, 1997

[4] Bretschneider T., Miller C., Kao O., Interpolation of Scratches in Motion Picture Films

[5] Ozkan Mehmet K., Sezan M.Ibrahim, Tekalp A.Murat, Adaptive Motion-Compensated Filtering of Noisy Image Sequences, IEEE Transactions on Circuits and Systems for video Technology, Vol. 3, No. 4, 277-290, August 1993.

[6] Wang Yao, Ostermann Jorn, Zhang Ya-Qin, Video Processing and Communications, Prentice Hall, 2002.

[7] Tekalp M. Digital Video Processing, Prentice Hall, 1995

[8] Kokaram Anil C., Morris Robin D., et al. Detection of Missing Data in Image Sequences, IEEE Tansactions on Image Processing, Vol. 4, No 11, November 1995

[9] Kokaram Anil C., Godsill Simon J., MCMC for Joint Noise Reduction and Missing Data Treatment in Degraded Video, IEEE Transactions on Signal Processing, Vol. 50, No. 2, February 2002.