

Article

Using Noise Level to Detect Frame Repetition Forgery in Video Frame Rate Up-Conversion

Yanli Li *, Lala Mei, Ran Li  and Changan Wu

School of Computer and Information Technology, Xinyang Normal University, Xinyang 464000, China; mll2106@163.com (L.M.); liran@xynu.edu.cn (R.L.); wchangan@sina.com (C.W.)

* Correspondence: liyl@xynu.edu.cn

Received: 7 August 2018; Accepted: 21 August 2018; Published: 24 August 2018



Abstract: Frame repetition (FR) is a common temporal-domain tampering operator, which is often used to increase the frame rate of video sequences. Existing methods detect FR forgery by analyzing residual variation or similarity between video frames; however, these methods are easily interfered with by noise, affecting the stability of detection performance. This paper proposes a noise-level based detection method which detects the varying noise level over time to determine whether the video is forged by FR. Wavelet coefficients are first computed for each video frame, and median absolute deviation (MAD) of wavelet coefficients is used to estimate the standard deviation of Gaussian noise mixed in each video frame. Then, fast Fourier transform (FFT) is used to calculate the amplitude spectrum of the standard deviation curve of the video sequence, and to provide the peak-mean ratio (PMR) of the amplitude spectrum. Finally, according to the PMR obtained, a hard threshold decision is taken to determine whether the standard deviation bears periodicity in the temporal domain, in which way FR forgery can be automatically identified. The experimental results show that the proposed method ensures a large PMR for the forged video, and presents a better detection performance when compared with the existing detection methods.

Keywords: frame rate up-conversion; frame repetition; video forensics; noise level; periodicity detection

1. Introduction

The abundant temporal redundancy in digital video is easily used by fakers, leaving digital video open to the threat of tampering. Video frame rate up-conversion (FRUC) [1] is one of the commonest operations for tampering with digital video in the temporal domain, e.g., in order to obtain a high-definition label, some forgers often increase the bits of video by FRUC operation before uploading a video to a video-sharing website, which damages the economic interests of webmasters and audiences [2]. To guarantee the authenticity of video data, some digital forensic techniques are needed to detect the FRUC forgery.

As a popular FRUC operation, frame repetition (FR) is often used by commercial video-editing software (e.g., VideoEdit Magic [3]) because of its intuitive architecture and low complexity. To detect FR, there are currently two types of forensic techniques, i.e., residual detection and similarity detection. The basic idea of residual detection is to find those repetitive frames by observing the residual variation between adjacent frames. Existing works use different methods to measure residual variation, e.g., Bestagini et al. [4] used mean square error (MSE) and Xia et al. [5] achieved it by analyzing texture details. When a video undergoes lossy compression, the qualities of repetitive frames degrade to varying levels, which randomly changes the statistics of residuals. Therefore, lossy compression can affect the performance of residual detection. Compared with residual detection, similarity detection can reduce the performance loss from lossy compression, and its innovation lies in finding a FR forgery by measuring the structural similarity between adjacent frames. Bian et al. [6] achieved

that through the periodicity of the structural similarity (SSIM) index [7]. Yang et al. [8] found the periodicity of inter-frame similarity by computing the distance between features of adjacent frames. Moreover, some existing works also used the bad effects in the interpolated frames to trace FR forgery, e.g., Yao et al. [9] explored the edge intensity to measure evidence of tampering; Ding et al. [10] found the relationship between statistics of residual and interpolation to reveal tampering; Ding et al. [11] used the artifact-indicated map and tchebichef moments to identify the frame interpolation. The aforementioned methods avoid measuring the pixel residual, so they can reduce the bad effects of lossy compression. However, the inter-frame similarity is easily influenced by noise. Once the forged video is polluted by some noise, similarity detection suffers some performance degradation. Therefore, an FR forensic technique, which is robust to noise, is required.

Due to thermal noise in cameras and natural noise from the outside, noises always exist in video sequences. While the video is forged by FR, the statistics of noise in video will change. Therefore, by detecting the noise varying in the temporal domain, it can be determined whether the video is forged by FR. The objective of this paper is to use Gaussian distribution to model the noise in a video frame, and find the FR forgery by observing the temporal variation of noise level. The proposed method directly uses the noise statistics to detect fake traces, avoiding the performance loss from noise. The main contributions of this work can be summarized as follows:

- This paper uses Gaussian distribution to model the noise in a video frame, and find the FR forgery by observing the temporal variation of the noise level.
- The proposed method directly uses the noise statistics to detect fake traces, avoiding the performance loss from noise.

Experimental results indicate that the proposed noise-level detection method can effectively identify the up-converted video sequences by FR, and guarantee high detection accuracy.

The rest of this paper is organized as follows. Section 2 briefly reviews residual detection and similarity detection. Section 3 describes the proposed noise-level detection method in detail. Section 4 presents the experimental results. Finally, Section 5 concludes this paper.

2. Background

FR directly inserts the new frames by duplicating the adjacent frames, and it can increase the frame rate of a video, i.e., supposing the t -th frame is the to-be-inserted frame, which is given by,

$$f[t] = f[t + 1] \quad (1)$$

where $f[t]$ is the to-be-inserted frame, $f[t + 1]$ is the adjacent original frame. As shown in Figure 1, the original frame rate can be multiplied by periodically inserting interpolated frames, which means that there always exist original frames which are consistent with interpolated frames in content. Therefore, FR forgery in video will be detected only if some repeated frames are detected in a certain time domain. Existing works use two types of detection methods, i.e., residual detection and similarity detection, which are briefly introduced as follows.

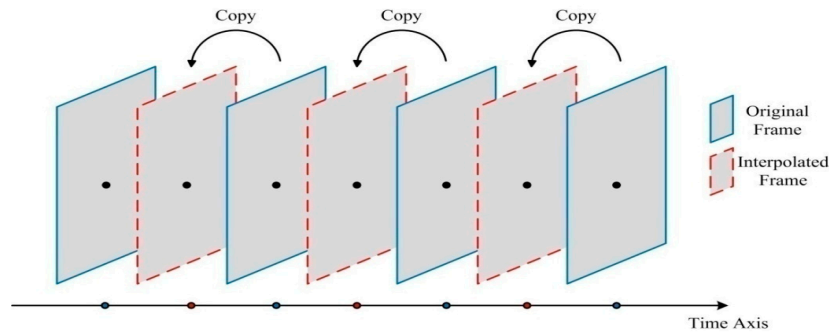


Figure 1. Illustration of frame repetition (FR) forgery.

2.1. Residual Detection

Residual detection calculates residual variation between two adjacent frames, which is given by,

$$v[t] = \|f[t] - f[t + 1]\|_2^2 \quad (2)$$

where $\|\cdot\|_2$ is the l_2 norm. If $f[t]$ is a repetitive frame, $v[t]$ is zero; otherwise $v[t]$ is not zero. Once these repetitive frames appear periodically, the residual variation will be zero periodically, so the fact that the residual variation periodically decays to zero can prove that the video is forged by FR. In general, when the original frame rate of video is increased by FR, in order to reduce the amount of data, H.264 [12] or high-efficiency video coding (HEVC) [13] is used to compress the forged video, which would cause some error between the compressed frame and the video frame as follows,

$$\hat{f}[t] = f[t] + e[t] \quad (3)$$

$$\hat{f}[t + 1] = f[t + 1] + e[t + 1] \quad (4)$$

where $\hat{f}[t]$ and $\hat{f}[t + 1]$ denote the compressed frames of the t -th frame and the $(t + 1)$ -th frame respectively, $e[t]$ and $e[t + 1]$ respectively denote the errors between $\hat{f}[t]$, $\hat{f}[t + 1]$ and their corresponding original frames. The residual variation between compressed frames can be expressed as follows,

$$\hat{v}[t] = \|\hat{f}[t] - \hat{f}[t + 1]\|_2^2 = v[t] + \|e[t] - e[t + 1]\|_2^2 \quad (5)$$

From Equation (5), it can be seen that when a forged frame suffers from lossy compression, the residual variation between compressed frames involves both the original residual variation and the error term. Because each compressed frame degrades at a different level, the error term can randomly change over time. So it can be seen that lossy compression can interfere with the periodicity of residual variation, thereby degrading detection performance.

2.2. Similarity Detection

Similarity detection calculates the SSIM value between two adjacent frames as follows,

$$s[t] = \text{SSIM}(f[t], f[t + 1]) = \frac{(2\mu_t\mu_{t+1} + c_1) + (2\lambda_{t,t+1} + c_2)}{(\mu_t^2 + \mu_{t+1}^2 + c_1)(\lambda_t^2 + \lambda_{t+1}^2 + c_2)} \quad (6)$$

where μ_t and μ_{t+1} are the means of $f[t]$ and $f[t + 1]$ respectively, λ_t and λ_{t+1} are the variances of $f[t]$ and $f[t + 1]$ respectively, $\lambda_{t,t+1}$ is the covariance between $f[t]$ and $f[t + 1]$, c_1 and c_2 are constants. The SSIM value ranges in $[0, 1]$, if it is set to be 1, showing that the two frames are the same, and the smaller the value is, the less similar the two frames are. Periodically inserting repetitive frames, the SSIM value

will be 1 periodically, so it can be proved that the video is forged by FR once the SSIM value is seen to increase to 1 periodically. When transmitting a forged video, the noise is mixed inevitably as follows,

$$\bar{f}[t] = f[t] + n[t] \quad (7)$$

$$\bar{f}[t + 1] = f[t + 1] + n[t + 1] \quad (8)$$

where $\bar{f}[t]$ and $\bar{f}[t + 1]$ denote the noisy frames of the t -th frame and the $(t + 1)$ -th frame respectively, $n[t]$ and $n[t + 1]$ denote noise terms respectively. The SSIM value between noisy frames can be calculated as follows,

$$\bar{s}[t] = \text{SSIM}(\bar{f}[t], \bar{f}[t + 1]) = s[t] + \delta(n[t], n[t + 1]) \quad (9)$$

where $\delta(n[t], n[t + 1])$ is the interference caused by $n[t]$ and $n[t + 1]$. According to Equation (9), when a forged frame is attacked by noise, the SSIM value between noisy frames would involve not only the SSIM value of the original frame but also the interference terms. So when a forged video is mixed with too much noise, the periodicity of the SSIM would be influenced, thereby degrading the detection performance.

According to the central limit theorem [14], $e[t]$ and $n[t]$ can be modeled as Gaussian noise with zero mean whether the video frame is attacked by lossy compression or external noise or not. Because there naturally exists noises in video sequences, the noises would inevitably be introduced while repeating frames. Therefore, by detecting the noise variation in the temporal domain, the fake traces can be detected. Compared with the residual detection and similarity detection, the advantage of noise detecting is that the forgery evidence directly comes from noise term, which avoids noise interference and, therefore, can improve the detection performance.

3. Proposed Noise-Level Detection

The flow chart of the noise-level detection is shown in Figure 2. The noise level of each video frame is first estimated, i.e., the standard deviation of Gaussian distribution, and generates the standard deviation sequence $\sigma[t]$. Then, the fast Fourier transform (FFT) [15] is used to calculate the amplitude spectrum of $\sigma[t]$, and the periodicity of $\sigma[t]$ is determined by detecting whether there is a large spike in the amplitude spectrum. Finally, in order to realize automatic detection, we calculate the ratio of the maximum and mean value of the amplitude spectrum, and set the threshold to determine whether the video is forged by FR. The following describes the proposed noise-level detection in detail.

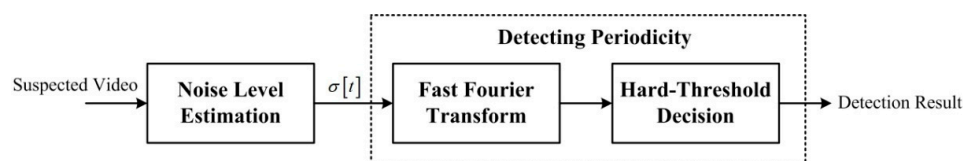


Figure 2. Flow chart of noise-level detection.

3.1. Noise-Level Estimation

The standard deviation is used to measure the level of Gaussian noise in video frame, and the existing methods can accurately estimate the standard deviation of a noisy frame, e.g., Tai et al. [16] used the Laplace operator and adaptive edge detection, and Zoran et al. [17] estimated the noise level based on the scale invariance of a natural image. The high-dimensionality of video requires that noise level estimation should be fast and effective, and the median absolute deviation (MAD) [18,19] based on wavelet transform meets that requirement. It is indicated by [19] that the standard deviation of

Gaussian noise is proportional to the MAD value of image wavelet coefficients. According to this conclusion, the wavelet coefficients of the t -th frame are first calculated as follows,

$$\mathbf{y}_t = \boldsymbol{\psi}^T \mathbf{f}_t = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_L]^T \mathbf{f}_t \quad (10)$$

where \mathbf{f}_t is the column vector of the t -th frame $f[t]$ through raster scanning, $\boldsymbol{\psi}_l$ is Daubechis-4 orthogonal wavelet basis and L is the total number of wavelet bases, $\boldsymbol{\psi}$ is the representation matrix whose l -th column is $\boldsymbol{\psi}_l$, and \mathbf{y}_t is the column vector composed of wavelet coefficients. The MAD value of wavelet coefficients can then be calculated as follows,

$$m_t = \text{median}(|\mathbf{y}_t - \text{median}(\mathbf{y}_t)|) \quad (11)$$

where $\text{median}(\cdot)$ denotes the median of input vector. Finally, the standard deviation of the video is generated as follows,

$$\sigma[t] = \begin{cases} |m_t - m_{t+1}| & |m_t - m_{t+1}| > 10^{-4} \\ 1 & |m_t - m_{t+1}| \leq 10^{-4} \end{cases} \quad (12)$$

Equation (10) can be implemented by fast algorithm [20], so the noise-level estimation has a high execution speed.

FR forgery inserts interpolated frames periodically, which makes the standard deviation of the forged video change periodically. As shown in Figure 3, we up-convert the *Foreman* video with CIF format from 15 fps to 30 fps, and use the proposed method to estimate the standard-deviation curves of the original video and its forged version. It can be seen that the original standard-deviation curve changes slowly and smoothly, while the forged one changes periodically. Therefore, the periodicity of the standard-deviation curve can be used as evidence to identify FR forgery.

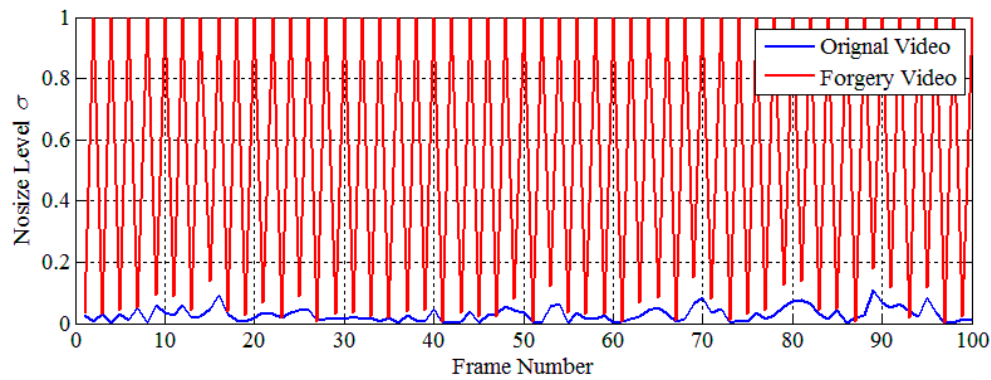


Figure 3. Illustrations of standard-deviation curve for both original and forged *Foreman* sequences.

3.2. Periodicity Detection

The FR forgery is automatically identified by using spectrum analysis to detect the periodicity of the standard-deviation curve. The discrete Fourier coefficient $\zeta[k]$ of $\sigma[t]$ is first computed as follows,

$$\zeta[k] = \sum_{t=0}^{N-1} \sigma[t] e^{-j \frac{2\pi k t}{N}} \quad 0 \leq k \leq N-1 \quad (13)$$

where N is the video length, and Equation (13) is implemented by FFT. Then, the amplitude spectrum is obtained as follows,

$$\eta[k] = |\zeta[k]| \quad (14)$$

finally, the DC components of $\eta[k]$ are removed as follows,

$$\alpha[k] = \eta[k] - \frac{1}{N} \sum_{k=0}^{N-1} \eta[k] \quad (15)$$

as shown in Figure 4, by Equations (13)–(15), we analyze the standard-deviation curves of the original and forged *Foreman* sequences in Figure 3. It can be seen that the amplitude spectrum of the original standard-deviation is small and distributed evenly while there is a large spike on the spectrum amplitude of the forged one. The spike is much higher than other coefficients, which can prove the periodicity of the standard-deviation curve, therefore determining whether the video is forged by FR according to this spike.

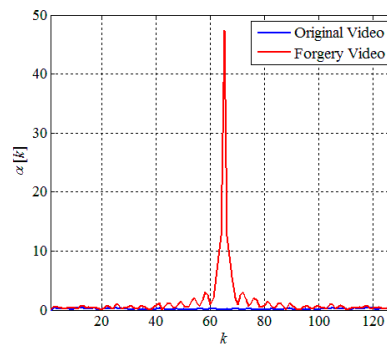


Figure 4. Illustrations of Fourier spectrum for both original and forged *Foreman* sequences.

To realize automatic detection, we need to measure the two Fourier spectrums shown in Figure 4. So the peak-mean ratio (PMR) of the amplitude spectrum is defined as follows,

$$R = \frac{\max\{\alpha[k] \mid k = 0, 1, \dots, N-1\}}{\frac{1}{N} \sum_{k=0}^{N-1} \alpha[k]} \quad (16)$$

where $\max\{\cdot\}$ denotes the maximum value of input set. The amplitude spectrum of the forged video has an isolated large spike, so the R value is large. The amplitude spectrum of the original video is small and evenly distributed, and therefore there is a small R value. PMR can be used as an index to determine whether the video is forged by FR. By setting an appropriate threshold, automatic detection can be realized by the following hard threshold decision,

$$\text{FR} = \begin{cases} \text{on} & R \geq \text{Thr} \\ \text{off} & R < \text{Thr} \end{cases} \quad (17)$$

where Thr is a preset threshold, *on* represents the existence of FR forgery, and *off* denotes no FR forgery. PMR can also be used to evaluate the performance of the detection algorithm. A large PMR value indicates that the detection algorithm provides a significant spike, which can improve the detection accuracy. Conversely, a small PMR value indicates that the detection algorithm provides a spike with small amplitude, which means a possibility of mistakes in detection. In the proposed FR forensic method, the threshold Thr is only a parameter, and it is important for the accuracy of detection. Based on the PMR distribution on the testing video database, the cross-validation proposed by [6] is used to select a proper threshold. According to the experimental results in Section 4.2, Thr is set to be 5.

4. Experimental Results and Analyses

Fifty uncompressed YUV sequences in 30 *fps* constitute the basic group of the testing video database, in which there are 10, 24, 8 and 8 video sequences with QCIF, CIF, 720P and 1080P formats, respectively. To evaluate the performance of the proposed method, the basic group is used to construct the negative set (NS) and the positive set (PS). All test videos are mixed with Gaussian noise, and the standard deviation of each video frame is evenly distributed in [0, 5]. The NS is composed of original test videos, and PS videos are obtained by down-converting video sequences in NS to 15 *fps* and then raising to 30 *fps* by FR forgery. We first evaluate the subjective performance of the proposed method by presenting the amplitude spectrum of the standard-deviation curve, and then evaluate the objective performance by comparing our method with residual detection and similarity detection. The PMR, false positive rate (*FPR*) and false negative rate (*FNR*) are used as objective criteria. *FPR* and *FNR* are defined as follows,

$$FPR = E_{NS} / C_{NS} \quad (18)$$

$$FNR = E_{PS} / C_{PS} \quad (19)$$

where E_{NS} and E_{PS} are the counts of mistakes when detecting NS and PS respectively, C_{NS} and C_{PS} are the capacities of NS and PS respectively. In addition, the detection accuracy (*DA*) is defined as follows,

$$DA = 1 - \frac{FNR + FPR}{2} \quad (20)$$

4.1. Subjective Performance Evaluation

Figure 5 shows the amplitude spectrums of standard-deviation curves for test video sequences in different formats. It can be seen that the spectrum amplitudes of the original videos are small and distributed evenly while all forged videos have significant spikes. For QCIF format, the forged video not only has a large spike, but also its spectrum amplitude changes rapidly, owing to the limited spatial resolution of QCIF format. With insufficient pixel samples, the Gaussian distribution cannot fully describe the noise statistics, which can weaken the periodicity of the standard deviation. For other formats, sufficient pixel samples guarantee a high accuracy of standard-deviation estimation, so the spectrum amplitude is uniformly distributed except the spike. It can be seen from the above results that the proposed method is suitable for high-definition video sequences. Moreover, the variations of motion and color in video sequences have little effect on the performance of the proposed method, e.g., the proposed method generates large spikes for both the *Soccer* sequence with fast motions and the *Tractor* sequence with image zooming. To sum up, the proposed method has a stable detection performance for the video sequences in different contents and formats.

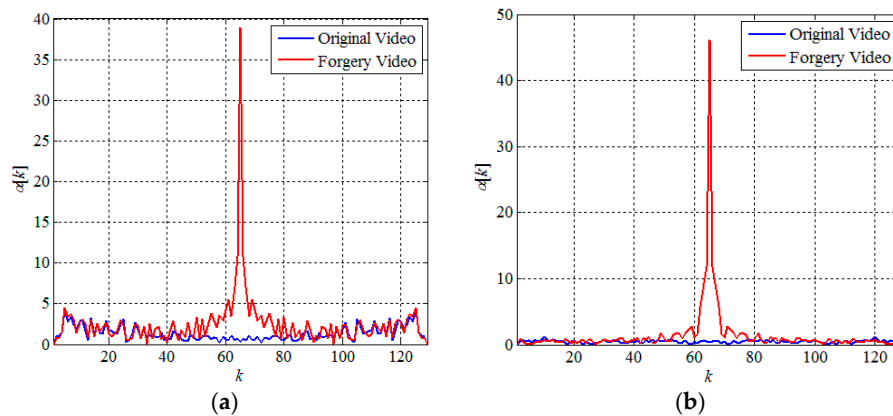


Figure 5. Cont.

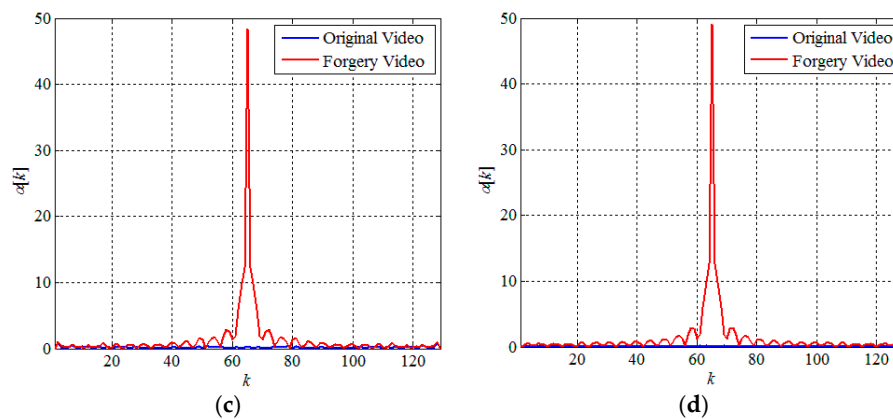


Figure 5. Illustrations of Fourier spectrums for both original and forged video sequences: (a) *Soccer* sequence in QCIF format, (b) *Mobile* sequence in CIF format, (c) *Mobcal* sequence in 720P format, (d) *Tractor* sequence in 1080P format.

4.2. Objective Performance Evaluation

Table 1 lists the average PMR of the amplitude spectrum on NS and PS for different detection methods. On videos in any format in NS, the average PMR of the noise-level detection is higher than those of residual detection and similarity detection. On videos in QCIF, CIF and 1080P format in PS, the average PMR of the noise-level detection is also the highest among all methods. Considering all the instances in NS and PS, the average PMR on NS is 8.63 and 8.79 for residual detection and similarity detection, respectively; however, the average PMR on NS is only 2.62 for noise-level detection; the average PMR on PS of noise-level detection has 6.87 and 3.62 PMR gains when compared with those of residual detection and similarity detection. Table 1 also lists the relative difference Δ between the average PMR values on NS and PS. A large Δ value means the amplitude spectrum has a significant spike. For videos in any format, the Δ value of the noise-level detection is higher than those of other detection methods. Considering all the instances in NS and PS, the Δ value of the noise-level detection is 0.91, but the Δ values of residual detection and similarity detection are 0.61 and 0.65, respectively. The results indicate that the noise-level detection obtains a large spike when compared with residual detection and similarity detection. The average PMR on NS of noise-level detection ranges in [2.42, 2.92] while the average PMR on PS of noise-level detection ranges in [27.32, 30.17]. There is obviously a big gap between the above two ranges, so the automatic detection can be realized by a hard threshold decision. According to the cross-validation proposed by [6], the threshold Thr is set to be 5, and the FNR, FPR and DA of different detection methods are calculated as shown in Table 2. For residual detection and similarity detection, some mistakes occur on NS and, especially for videos in 1080 P format, the FNR value is up to 1.00, i.e., all original videos are misjudged as forged. For the noise-level detection, the FNR values are 0.00 for videos in QCIF, CIF and 720P format, and the FNR value is only 0.25 for videos in 1080P format. These results indicate that noise-level detection can accurately identify the original video. For videos in any format in PS, FPR values of noise-level detection are 0.00, which indicates that noise-level detection can accurately identify the forged video. Considering all the video sequences, the FNR and FPR values of noise-level detection are 0.04 and 0.00, respectively, and they are much lower than those of residual detection and similarity detection. The DA of the noise-level detection is 0.98 which is, respectively, 0.39 and 0.36 higher than those of residual detection and similarity detection. To summarize the above, the proposed noise-level detection guarantees a better detection performance when compared with residual detection and similarity detection.

Table 1. Average peak-mean ratio (PMR) values of amplitude spectrums for different detection methods.

| Detection Method | QCIF | | | CIF | | | 720P | | | 1080P | | | Total | | |
|-----------------------|------|-------|----------|------|-------|----------|-------|-------|----------|-------|-------|----------|-------|-------|----------|
| | NS | PS | Δ | NS | PS | Δ | NS | PS | Δ | NS | PS | Δ | NS | PS | Δ |
| Residual detection | 6.76 | 17.24 | 0.61 | 7.67 | 19.88 | 0.61 | 10.04 | 29.58 | 0.66 | 11.73 | 23.88 | 0.51 | 8.63 | 21.98 | 0.61 |
| Similarity detection | 7.36 | 21.91 | 0.66 | 7.63 | 24.31 | 0.69 | 10.17 | 29.63 | 0.66 | 11.90 | 26.30 | 0.55 | 8.79 | 25.23 | 0.65 |
| Noise-level detection | 2.42 | 27.32 | 0.91 | 2.45 | 30.17 | 0.92 | 2.93 | 28.32 | 0.89 | 2.92 | 27.95 | 0.889 | 2.62 | 28.85 | 0.91 |

Denote R_1 by the average PMR on NS and R_2 by the average PMR on PS, and the relative difference between R_1 and R_2 is $\Delta = 1 - R_1/R_2$.

Table 2. FNR, FPR and DA of different detection methods.

| Detection Method | QCIF | | | CIF | | | 720P | | | 1080P | | | Total | | |
|-----------------------|------|------|------|------|------|------|------|------|------|-------|------|------|-------|------|------|
| | FNR | FPR | DA | FNR | FPR | DA | FNR | FPR | DA | FNR | FPR | DA | FNR | FPR | DA |
| Residual detection | 0.70 | 0.10 | 0.60 | 0.71 | 0.00 | 0.64 | 0.90 | 0.00 | 0.55 | 1.00 | 0.00 | 0.50 | 0.80 | 0.02 | 0.59 |
| Similarity detection | 0.70 | 0.00 | 0.65 | 0.67 | 0.00 | 0.67 | 0.80 | 0.00 | 0.60 | 1.00 | 0.00 | 0.50 | 0.76 | 0.00 | 0.62 |
| Noise-level detection | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.25 | 0.00 | 0.88 | 0.04 | 0.00 | 0.98 |

5. Conclusions

In this paper, a noise-level based detection method is proposed to identify FR forgery. To solve the problem of noise where the existing detection methods are inadequate, the proposed method directly uses the varying noise level in time to identify whether the video is forged by FR. The MAD based on wavelet transform is first used to estimate the standard deviation of each video frame. Next, FFT is employed to calculate the amplitude spectrum of the standard-deviation curve, and the PMR of the amplitude spectrum is calculated. Finally, according to the PMR of the suspected video, a hard threshold decision is performed to automatically identify whether the video is forged by FR. Experimental results show that the PMR of the original video is much smaller than that of the forged video, and when the threshold Thr is set to be 5, compared with residual detection and similarity detection, the proposed noise-level detection presents a better detection performance on the NS and PS composed of test videos in different formats.

In the proposed method, the threshold is set according to the experimental data, which cannot guarantee stable detection performance when changing the testing video database. Therefore, an adaptive threshold setting needs to be studied in the future.

Author Contributions: Y.L. contributed to the concept of the research, performed the data analysis and wrote the manuscript; L.M. helped perform the data analysis with constructive discussions and helped perform the experiment. R.L. conducted the experiment and collected the data. C.W. revised the manuscript, and provided the necessary support for the research.

Funding: This work was supported in part by the National Natural Science Foundation of China, under Grants nos. 61501393, and in part by Nanhu Scholars Program for Young Scholars of XYNU.

Conflicts of Interest: The authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. Tsai, T.H.; Shi, A.T.; Huang, K.T. Accurate frame rate up-conversion for advanced visual quality. *IEEE Trans. Broadcast.* **2016**, *62*, 426–435. [CrossRef]
2. Bian, S.; Luo, W.; Huang, J. Exposing fake bit rate videos and estimating original bit rates. *IEEE Trans. Circuits Syst. Video Technol.* **2014**, *24*, 2144–2154. [CrossRef]
3. Video Edit Magic. Available online: <https://video-edit-magic.en.softonic.com> (accessed on 5 December 2017).
4. Bestagini, P.; Battaglia, S.; Milani, S.; Tagliasacchi, M.; Tubaro, S. Detection of temporal interpolation in video sequences. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 3033–3037.
5. Xia, M.; Yang, G.; Li, L.; Li, R.; Sun, X. Detecting video frame rate up-conversion based on frame-level analysis of average texture variation. *Multimed. Tools Appl.* **2017**, *76*, 8399–8421. [CrossRef]
6. Bian, S.; Luo, W.; Huang, J. Detecting video frame-rate up-conversion based on periodic properties of inter-frame similarity. *Multimed. Tools Appl.* **2014**, *72*, 437–451. [CrossRef]
7. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]
8. Yang, J.; Huang, T.; Su, L. Using similarity analysis to detect frame duplication forgery in videos. *Multimed. Tools Appl.* **2016**, *75*, 1–19. [CrossRef]
9. Yao, Y.; Yang, G.; Sun, X.; Li, L. Detecting video frame-rate up-conversion based on periodic properties of edge-intensity. *J. Inf. Secur. Appl.* **2016**, *26*, 39–50. [CrossRef]
10. Ding, X.; Yang, G.; Li, R.; Zhang, L.; Li, Y.; Sun, X. Identification of motion-compensated frame rate up-conversion based on residual signal. *IEEE Trans. Circ. Syst. Video Technol.* **2018**, *28*, 1497–1512. [CrossRef]
11. Ding, X.; Zhu, N.; Li, L.; Li, Y.; Yang, G. Robust localization of interpolated frames by motion-compensated frame-interpolation based on artifact indicated map and tchebichef moments. *IEEE Trans. Circ. Syst. Video Technol.* **2018**. [CrossRef]
12. *Advanced Video Coding for Generic Audio-Visual Service*; ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC); ITU-T and ISO/IEC JTC 1; International Telecommunication Union: Geneva, Switzerland, 2003.

13. Sullivan, G.J.; Ohm, J.R.; Han, W.J.; Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Trans. Circ. Syst. Video Technol.* **2012**, *22*, 1649–1668. [[CrossRef](#)]
14. Papoulis, A.; Pillai, S.U. *Probability, Random Variables and Stochastic Processes*, 4th ed.; McGraw-Hill: New York, NY, USA, 2002.
15. Bergland, G.D. Fast Fourier transform hardware implementations-A survey. *IEEE Trans. Audio Electroacoust.* **1969**, *17*, 109–119. [[CrossRef](#)]
16. Tai, S.C.; Yang, S.M. A fast method for image noise estimation using Laplacian operator and adaptive edge detection. In Proceedings of the 2008 3rd International Symposium on Communications, Control and Signal Processing, St Julians, Malta, 12–14 March 2008; pp. 1077–1081.
17. Zoran, D.; Weiss, Y. Scale invariance and noise in natural images. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; pp. 2209–2216.
18. Leigh, A.; Wong, A.; Clausi, D.A.; Fieguth, P. Comprehensive analysis on the effects of noise estimation strategies on image noise artifact suppression performance. In Proceedings of the 2011 IEEE International Symposium on Multimedia, Dana Point, CA, USA, 5–7 December 2011; pp. 97–104.
19. Donoho, D.L. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory* **1995**, *41*, 613–627. [[CrossRef](#)]
20. Mallat, S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).