CS 6870 Term Project Assignment Report

Detection of Tampered Videos (12)

Abhijith C S (CS11B003)
Department of Computer Science and Engineering, IIT Madras

***Abstract* - The development of video editing techniques enables us to create realistic synthesized videos. Such synthesized or tampered videos may contain frames which are duplicated from the same video to hide some unfavorable object or contain objects from different sources embedded into the video. The attempt is to detect whether videos contain evidence of tampering by locating forged regions in a video using correlation of noise residue. The block-level correlation values of noise residual are extracted as a feature for classification. The distribution of correlation of temporal noise residue in a forged video has been modelled as a Gaussian mixture model (GMM).**

I. Introduction

The technology has been enhanced much in the recent years and we are able to create realistic synthesized videos using several available tools. It's not very easy to say whether a given video contains any forgery. Videos have got wide applications in today's world. Say for example, the shot videos can be used as evidences for crime investigations. In such situations originality of the video might be questioned. Advanced video editing technologies let people corrupt/tamper/modify a video in several ways. Thus it's important to detect 'synthesized' part of a video in such cases. Ways of tampering a video include resampling [1], copy and paste, slicing [2][3], and double compression [4]. Detecting the forgery in a given video is a broad area with huge scope for research.

This project aims to implement an algorithm to check whether a given video or set of frames contains any signs of tampering/forgery/synthesized-frames in it.

II. Algorithmic Description

This project aims to address passive forgery detection in a digital video based on the statistical property of noise residue. Idea is to analyze the temporal correlation of block-level noise residue to locate the tampered regions of a video. The distributions of temporal noise correlation values of video blocks in forged and normal regions are modelled using a GMM model[5][6], where the GMM model parameters are estimated using the Expectation-Maximization (EM) algorithm[7].

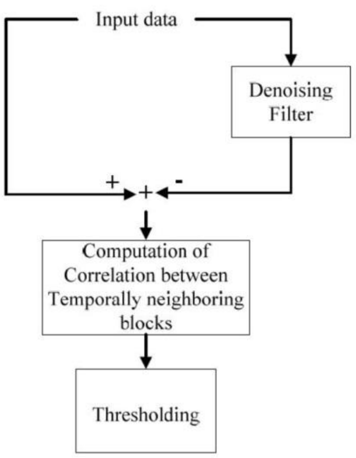

Fig. II.1 Flow of Processes

The whole process of detecting forgery [Fig. II.1] in the video include three parts namely, *Extraction of Noise Residue*, *Calculation of Block-level Noise Correlation Values* and *Forgery Detection by Statistical Analysis of Noise Residue.*

Initially, the process same as the one proposed in [2], the noise residue of each video frame is extracted by subtracting the original frame from its noise-free version. The *wavelet denoising filter* suggested in [8] is used to obtain the noise-free image at the first hand.

Later, each video frame is partitioned into non-overlapping blocks of size $N \times N$. The correlation of the noise residue between the same spatially indexed blocks of two consecutive frames is then computed.

And finally the goal is to locate tampered blocks by analyzing the statistical properties of block-level noise correlations.

A. *Extraction of Noise Residue*

In the process of extracting the noise residue, the method used in [2][8] is used. Assumption is that the high-frequency wavelet coefficients can be modeled as the sum of a stationary white Gaussian noise and a noise-free image. The wavelet decomposition is performed on a noisy image in order to obtain its wavelet coefficients. Once the decomposition is done, only the high frequency coefficients are used for later computations. Local variations of these wavelet coefficients are estimated as in [9] by defining a window size of *WxW* where $W = \{3,5,7,9\}$. Then the wiener filter with optimal algorithm as suggested in [10] with the following profile [Fig. A.1] is used for denoising process. For each wavelet coefficient, repeat previous steps until the process converged. Finally, the inverse wavelet transform is used to obtain the noise-free image.

$$c_{\text{den}}(i,j) = c(i,j)\frac{\hat{\sigma}^2(i,j)}{\hat{\sigma}^2(i,j)+\sigma_0^2}$$

Fig. A.1 Profile used for Wiener Denoising

After the noise-free image is obtained, the noise residual n(i,j) can be easily extracted by subtracting the original image from its noise-free version. The noise residue n(i,j) consists PNU and high-spatial-frequency details of image content. In [2], the PNU was extracted from noise residual n(i,j) via an averaging operation.

B. *Calculation of Block-level Noise Correlation Values* (Not Completed)

Let $n_{i,j}$ denote the noise residual at pixel coordinate (i,j). The correlation value *r* between previous frame and current frame on each block can be defined as in [Fig. B.1] and computed.

$$r = \frac{\sum_i \sum_j \left(n_{i,j}^t - \overline{n}^t\right)\left(n_{i,j}^{t-1} - \overline{n}^{t-1}\right)}{\sqrt{\sum_i \sum_j \left(n_{i,j}^t - \overline{n}^t\right)^2 \sum_i \sum_j \left(n_{i,j}^{t-1} - \overline{n}^{t-1}\right)^2}}$$

Fig. B.1 Correlation Value Calculation

The idea is that, if a region is forged, the correlation value of temporal noise residue in the region is usually changed, either increased or decreased, depending on the forgery scheme used.

C. *Forgery Detection by Statistical Analysis of Noise Residue* (Not Done)

The tampering process usually changes the temporal statistical property of sensor residue. So that we can easily distinguish from the tampered regions and the non-tampered regions by simply analyzing the statistical properties of block-level noise correlation. After the correlation values of every two temporally neighboring blocks are obtained, the parameters of the distributions of normal blocks and forged blocks are estimated respectively via the maximum-likelihood estimation. The area of forged regions is usually much smaller than that of normal region. Hence, a simple pre-classification scheme to quickly determine whether a video frame has been forged.

III. Output

The first step, the *Extraction of the Noise Residue* was carried out in order to obtain a frame-set which is denoised. Calculation of block level noise correlation values was also done.

A. *Observation* (Partial)

Denoised frame-set which are used for the calculation of block level noise correlation values was obtained successfully. Calculation of correlation values was initiated but not finished.

IV. Conclusion

It's a digital video forgery detection scheme using temporal noise correlation without the need of embedding any prior digital signature in the compressed video. This method achieves promising detection accuracy for fine-quality videos.

References

[1] A.C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of re-sampling," *IEEE Trans. Signal Process.*, vol. 53, no.2, pp. 758-767, Feb. 2005.

[2] S. Bayram, H. T. Sencar, and N. Memon, "Source camera identification based on CFA interpolation," in *Proc. IEEE Int. Conf. Image Processing*, vol.3, no., pp. III-69-72, 11-14, Sept. 2006.

[3] Y. Q. Shi, C. Chen, and W. Chen, “A natural image model approach to splicing detection,” in *Proc. ACM Multimedia Security Workshop*, pp. 51-62, Sept. 2007, Dallas, Texas.

[4] Z. Fan and R. L. de Queiroz, “Identification of bitmap compression history: JPEG detection and quantizer estimation,” *IEEE Trans. Image Process*., vol. 12, no. 2, pp. 230–235, Feb. 2003.

[5] MLA Andersen, Torben G., and Bent E. Sørensen. "GMM estimation of a stochastic volatility model: a Monte Carlo study." *Journal of Business & Economic Statistics* 14.3 (1996): 328-352

[6] Bond, Stephen, Anke Hoeffler, and Jonathan Temple. GMM estimation of empirical growth models. Vol. 3048. *London: Centre for Economic Policy Research*, 2001.

[7] Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal statistical Society* 39.1 (1977): 1-38.

[8] M. K. Mihcak, I. Kozintsev, and K. Ramchandran, “Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising,” in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 6, pp. 3253–3256, Mar. 1999, Phoenix, AZ

[9] Kivanc Mihcak, M., et al. "Low-complexity image denoising based on statistical modeling of wavelet coefficients." *Signal Processing Letters*, IEEE 6.12 (1999): 300-303.

[10] Chen, Jingdong, et al. "New insights into the noise reduction Wiener filter." *Audio, Speech, and Language Processing, IEEE Transactions* on 14.4 (2006): 1218-1234.