Fast Sensor Fingerprint Extraction Method Using Wavelets Transform

Rajveer Kaur* and Er. Sukhwinder Sharma**
BBSBEC, Fatehgarh Sahib, Punjab

Abstract- In this paper we have developed a new method to detect video source in wireless noisy scenario which is very fast compared to state-of-art algorithm. The algorithm is developed by using wavelet domain filtering of gray scaled image and a running average to compute sensor noise under the assumption that AWGN has zero mean and thus if averaged over N frames, the resultant noise will cancel (mean noise will be zero) and sensor noise has a pattern which will not cancel out remaining the sensor noise only. Another advancement to use frequency domain and matrix dot multiplication to fastly compute correlation for N sources to detect any forgery or intrusion in the wireless camera feed.

I. Introduction

Forensic knowledge Analysis (FDA) could be a branch of Digital forensics. It examines structured knowledge with relevancy incidents of monetary crime. The aim is to get and analyse patterns of dishonorable activities. Knowledge from application systems or from their underlying knowledgebases is remarked as structured data. Unstructured knowledge in distinction is taken from communication and workplace applications or from mobile devices. This knowledge has no overarching structure and analysis thence suggests that applying keywords or mapping communication patterns. Analysis of unstructured knowledge is typically remarked as pc forensics.

In arithmetic, a riffle series could be a illustration of a square-integrable (real- or complex-valued) perform by an exact orthonormal series generated by a riffle. Nowadays, riffle transformation is one among the foremost in style of the time-frequency-transformations. this text provides a proper, mathematical definition of associate orthonormal riffle and of the integral riffle remodel.

Source identification could be a major part of video forensics. once presenting a video clip as proof during a court of law, characteristic the supply (acquisition device) of the video is as vital because the video itself. as an example, if a police work camera captured the scene of a suspect’s alibi, it’s necessary to prove that the video was really recorded by the claimed camera. Otherwise it are often solid or derived from Associate in Nursing devious supply, that makes the proof invalid. within the industry, important revenue loss is caused each year by surreptitious recording in motion-picture show theaters and therefore the resulting illegal distribution. Video supply identification is utilized to trace down such piracy crimes. whereas web allows video sharing at an outsized scale, it conjointly opens doors for propagation of illegal or inappropriate materials, like the video containing kid creation or racial emotion.

On the opposite hand, wireless communication has seen an amazing growth within the recent years. Wireless cameras have additionally become more and more well-liked. within the security camera market, wireless video cameras still replace their wired counterparts attributable to the benefit of readying. In plan of action networks, wireless cameras area unit wide used as video sensors. they sometimes don’t have native storage; video is captured and wirelessly streamed to a sink. thanks to the inevitable packet loss and unpredictable transmission delay in wireless streaming, interference and blurring often seem within the received frames. For such videos, experiments show that the present supply identification
strategies suffer from important performance deterioration or perhaps fail to figure. Interference and blurring caused by the lossy wireless channel severely tamper with the sensing element fingerprint recognition.

In this paper, we have proposed a new method for video source identification which make use of lossy video also and perform wavelet decomposition to extract source fingerprint. The concept is developed on gray scale image which made the proposed system very fast. Also, since it is very fast, the system can be incorporated in real-time wireless feed detection system where the fingerprint of each camera is stored already. In case of wireless feed change, the fingerprint mismatch will occur which can be detected in real time using a proposed N-D matrix with N sources in frequency domain (fft application) dot multiplying with another N-D matrix (N sources computed fingerprint in frequency domain to generate correlation values. These values can be thresholded to detect spoofing.

II. Related Work

Chen et al in [1] also planned to include wireless channel signatures and selective frame process into supply identification, that considerably improve the identification speed. They conducted in depth real-world experiments to validate our technique. The results show that the supply identification accuracy of the planned theme for the most part outperforms the prevailing ways within the presence of video block and blurring. Moreover, our technique in an exceedingly position to spot the video supply in a near-real-time fashion, which might be wont to sight the wireless camera spoofing attack.

Liang et al in [2] proposed a video condensation technique for video forensics to condense a protracted amount of police investigation video into a brief video clip underneath keeping vital info. The projected technique is comprised of the subsequent 3 steps: image sequences transformation, motion orientating filter, and ribbon carving. First, the first video sequence is reborn into X-T and Y-T image sequences. Then, a motion orientating filter operator supported Gaussian gradient method is applied to extract moving-object maps in a very specific direction from the X-T and Y-T image sequences. Finally, they applied a ribbon carving approach to realize the aim of video condensation. Preliminary experiments on realistic video information in contestible the relevancy of the projected technique.

Zamani et al in [3] proposed associate degree sweetening analysis formula supported super-resolution. Hence, they bestowed associate degree answer that could be a single-frame answer for super-resolution. For that purpose, their projected technique incorporates distributed cryptography with Non-Negative Matrix resolution so as to boost hallucination of probes in video. distributed cryptography is used in learning a localized part-based mathematical space that synthesizes higher resolution with relation to overcomplete patch dictionaries. we have a tendency to look at their projected technique and compare with progressive strategies particularly resampling and super-resolution technique, by enhancing probes in exhibit videos. They measured the image quality victimization peak-signal-to-noise-ratio. The result shows that their projected technique outperforms state-of-the-art strategies when enhancing probes in exhibit videos.

Hsu et al. in [4] proposed associate economical cross-camera vehicle chase technique via affine invariant object matching. a lot of specifically, they developed the matter as invariant image feature matching among completely different viewpoints of cameras. To realize vehicle matching, the 1st extract invariant image feature supported ASIFT (affine and scale-invariant feature transform) for every detected vehicle in a very camera network. Then, to enhance the accuracy of ASIFT feature matching between pictures from completely different viewpoints, they planned to expeditiously match feature points supported their ascertained spatially invariant property of ASIFT, moreover because the min-hash
technique. As a result, cross-camera vehicle chase will be expeditiously and accurately achieved. Experimental results demonstrate the effectiveness of the planned algorithmic rule and also the feasibleness to video forensics applications.

Tou et al. in [5] proposed a spacial structural analysis (SSA) methodology to extract the de-blocking filter traces, by examining element intensity changes. At first, the Social Security Administration computes the intensity variations on the horizontal and vertical direction on the U channel of the compressed video. The Social Security Administration then locates all perpendicular turning points that were fashioned by a try of intersected and perpendicular edges that seem to unendingly share identical intensity distinction within the same direction. Finally, a standard spacial distance are determined supported the very best incidence of spacial distance between the horizontal and vertical lines fashioned on the perpendicular turning points detected. The space are accustomed confirm if compression has been conducted for the video, wherever the common spacial distance computed is the image of the dimensions of the macroblocks on each horizontal and vertical direction. The experimental results showed that the determined common spacial distance is eight within the videos compressed with H.264/AVC that matches the dimensions of the 8x8 macroblocks. Videos compressed while not de-blocking filter showed stronger peak within the accumulated spacial distance count compared to those with de-blocking filters applied.

Yin et al. in [6] proposed, designs, and implement a completely unique large-scale Digital Forensics Service Platform (DFSP) that may effectively observe black-market content from net videos. additional specifically, they planned a distributed design by taking advantage of Content Delivery Network (CDN) to enhance measurability, which might method huge variety of net videos in real time. They planned CDN-based Resource-Aware planning (CRAS) algorithmic rule, that schedules the tasks expeditiously within the DFSP consistent with resource parameters, like delay and computation load. They deploy the DFSP system within the net, that integrates the CDN-based distributed design and CRAS algorithmic rule with a large-scale video detection algorithmic rule, and judge the deployed system. Our analysis results demonstrate the effectiveness of the platform.

Dang et al. in [7] proposed a unique rhetorical technique to tackle the matter of characteristic pc generated (CG) from real humans in videos. It exploits the temporal data inherent of a video sequence by analyzing the spatio-temporal look of facial expressions in each CG and real humans. notwithstanding rendering facial features has reached outstanding performances, CG face look over time still presents some underlying mechanical properties that greatly dissent from the natural muscle movements of real humans. They build Associate in Nursing economical classifier on a group of options describing facial dynamics and spatio-temporal changes throughout smiling to tell apart CG from human faces. Experimental results demonstrate the effectiveness of the planned approach.

III. Proposed Work

This present work has developed a new method to compute sensor noise which uses less number of frame averaging. The method is known as running average sensor noise prediction. To achieve this, we define the following steps:

In step 1, we take the image as input and convert it to grayscale and then compute its 4th level DWT. This step saved the computation of DWT for each channel. Also, since grayscaling will not change the overall noise structure of the image, this step has still preserved the sensor noise fingerprint.

In step 2, we applied a wiener filter with varying window size $D = \{3, 5, 7, 9\}$ to each sub-band (H, V and D) of the decomposed image for each level. This operation serves a very important purpose, i.e. it performed denoising stage.
In step 3, we used these denoised sub-bands for each level to reconstruct the original image. This new image is denoised image.

In step 4, we applied the same method of frame subtraction i.e. to subtract the denoised frame from original frame. This left us with only the sensor noise + AWGN.

In final step 5, we perform k-level folding and averaging of the noise only frame. This step completely removed AWGN and the system is left with only sensor noise. The k-fold concept is made fast by using 8 connected neighbor folding.

Now we performed computation of absolute summation over row (or column) to check for data. Instead of folding the system here, we computed the fft to find out the resonant frequency and periodicity, which came out to be periodic in case of blocking frames. This step completely preserved the time.

After these steps, standard running average is computed for streaming frame which gave a robust sensor noise value after n frames n being in order of 10-20 frames i.e. a sample of about 1 second with 25 fps. is enough to compute sensor noise.

To test, the correlation of the two sensor data is performed and a threshold is used to define that the data is taken from same image. Now, if all the sensor fingerprints are stored, this correlation can be performed in single step by performing 2D correlation on N-D matrix with N being number of wireless cameras in network. One step matrix dot multiplication in fft domain will yield correlation values for all the cameras and if any correlation value is less than threshold, we can detect the intrusion immediately.

IV. Results and Discussions

The work is developed in MATLAB™ R2013B which supports many newly introduced features to handle video data and wavelet transforms. We started by implementing the core steps of loading a video data into MATLAB workspace.

After loading the video, we started a loop which looped from 1 to 20. In this loop we fetched a single frame (next frame from 1 to 20). We performed steps like image normalization (converting a uint8 image (0-255) in a double image (0-1 floating point values). This allows us to perform mathematical operations over images which generates floating point values. Then we performed steps that are dwt with 4-level decomposition, denoising using wiener denoising filter, reconstruction from denoised coefficients of wavelet, frame subtraction, frame blocking detection using FFT and finally averaging over n frames to compute final sensor fingerprints.

This fingerprint is then used to compute correlation between an unknown footage source and the source detection is successful as judged by truth value of the video source. The time taken to process 20x3 = 60 frames came out to be approximated 57 seconds which gives an average of 25 fps video streaming which is completely real time.

We took two cameras and took two footages. One extra footage from Camera-0 is taken to verify the accuracy detection of the system. Results are presented as follow:

![Figure 1: Computing the Sensor Noise for Camera-0 in this step.](image)

In figure 1, the sensor noise from camera-0 is computed using proposed method. It took approximately 19 seconds to complete the processing on 20-frames in total.

![Figure 2: Computing the Sensor Noise for Camera-1 in this step](image)
In this figure 2, Sensor noise for camera-1 is computed, this noise is computed using our proposed method and it also took approximately 19 seconds for 20 frames of data.

![Figure 3: Computing the Sensor Noise for Test Video Footage.](image)

In this figure 3, the test footage is taken and its noise is computed. This noise is then compared with the both Camera-0 noise and Camera-1 noise to check for correlation. This step also took an approximate of 19 seconds for 20 frames of data.

![Figure 4: The result of Comparison is shown and total run-time is shown.](image)

This figure 4 shows that the test footage is detected to be taken from Camera-0 which it indeed is which verifies that the accuracy of the system. Further total run-time of the system is shown which is about 57 second on average which gives us real time performance with about 22 fps.

![Figure 5: Sensor Noise for Camera-0 and Camera-1](image)

After removing normal noise, the only thing left is the sensor noise. After 20x of intensity scaling we can barely see some scratchy noise print on the image. This is the pure sensor noise left behind.

![Figure 6: Error reduction with number of frames to compute sensor noise](image)

In this figure 6, we can see that the green line (our proposed method) is more steep at around 20 frames and is the optimal value to compute sensor noise with minimum error value, the other method (cyan plot) never reaches this point even after 32 frames as it that method requires more than 100 frames to reach that value.

V. Conclusion

In this paper, we have proposed a very fast sensor fingerprint extraction method with lesser amount of data (2-3 second of video with 25 fps is enough). We have developed over the work of [1] and proposed a very fast and robust method which can compute fingerprint from as small as 20 frames of a video which is less than 1 second of video with 25 fps average frame rate.

The results show that it is very fast method and works with real-time streaming data like streaming videos from wireless camera and can be utilized to detect image forgery or source forgery in case of intrusion very efficiently.

In the future work, we can advance over this work to work with large resolution image as the time taken by high resolution image with faster fps is very large compared to current scenario.

References


