Signal Processing: Image Communication 57 (2017) 1–7

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

Dealing with video source identification in social networks

Irene Amerini [a,*], Roberto Caldelli [b], Andrea Del Mastio [a], Andrea Di Fuccia [c], Cristiano Molinari [c], Anna Paola Rizzo [c]

[a] Media Integration and Communication Center (MICC), Università degli Studi di Firenze, Viale Morgagni 65, 50134 Firenze, Italy
[b] National Interuniversity Consortium for Telecommunications - CNIT, Parma, Italy
[c] Polo Tecnologico, Presidenza del Consiglio dei Ministri, Rome, Italy

ARTICLE INFO

Keywords:
Video source identification
Social networks
Fingerprint
PRNU

ABSTRACT

Certainly detecting the source of a digital video it is a crucial task to be tackled by the image forensic scientific community; in fact, knowing the brand and model of the device used for the video acquisition could be very useful to focus investigations in a specific direction. Nowadays, videos are mostly acquired through a smartphone and then shared on Social Networks (SNs). On such a basis, this paper proposes an analysis for the source identification of a video uploaded on social networks, specifically, Twitter and Facebook. Furthermore, the paper evaluates different methods to build a reliable fingerprint and also introduces a novel method to generate a composite fingerprint by resorting to the use of PRNU noise. A tool to examine videos, oriented to forensic analysts, is also presented. Experimental results carried out on various videos, firstly uploaded and then downloaded from Facebook or Twitter, witness that the identification is still possible and under which conditions.

1. Introduction

Nowadays a huge amount of multimedia contents (images and videos) is generated in different ways with various devices and then uploaded on social networks (SNs). During the upload or once on-line, they are shared with other known users to be played or downloaded. At the time most of the SNs allow for the recording, through the use of a smartphone, and the uploading of a video clip at the same time. Facebook, Twitter and other SNs contain a huge number of videos and these contents constitute an interesting real-time source of information. In fact SNs could be of support during investigations which, always more, do an extensive use of social networks to reconstruct facts on the basis of the information contained within personal profiles (images and, in particular, videos) and associated with a specific account. Criminal activities like child pornography, fraud and terrorism are proliferating by misusing such digital contents.

Generally, these activities are done anonymously so it could be very useful to understand if a video posted by an unknown account used for illegal purposes it has been generated by the same video camera (smartphone) of another video uploaded on a known user account on a SN. In this way a connection can be established and this could help in addressing an on-going investigation and identifying possible suspects.

Uploading a video on a SN can severely reduce video quality by adding a layer of compression, sometimes resizing the video dimensions and cutting its length. So the question is: after such heavy processing is it still possible to determine if two videos come from the same video camera? The idea behind this work is to research a particular fingerprint that is able to achieve source identification in the case of such particular SNs videos.

The paper is organized as follows: Section 2 presents some previous works inherent to video source identification, while Section 3 describes how videos are managed on Twitter and Facebook. Section 4 introduces different modalities of PRNU estimation also proposing a new composite fingerprint. In Section 5 various experimental results are discussed to evaluate the performances of diverse kinds of fingerprints and in Section 6 a new specific tool for video forensic analysis is proposed. Finally Section 7 draws the conclusions and future works.

2. Related works

The main idea behind the approach of source identification is that each phase of the acquisition process leaves a sort of unique fingerprint on the digital content itself due to some intrinsic imperfections in the acquisition phase. In particular, the PRNU (Photo Response Non-Uniformity) noise is well known and used as fingerprint to identify a specific digital camera among a dataset of cameras [1]. The approach in [1] has also been extended to work with video camera identification and video forgery detection [2]. An adaptive weighting to improve the

* Corresponding author.
E-mail addresses: irene.amerini@unifi.it (I. Amerini), roberto.caldelli@unifi.it (R. Caldelli), andrea.delmastio@unifi.it (A.D. Mastio).

http://dx.doi.org/10.1016/j.image.2017.04.009
Received 23 December 2016; Received in revised form 22 April 2017; Accepted 23 April 2017
Available online 25 April 2017
0923-5965/

performance is proposed in [3] while Chen et al. in [4] try to identify digital camcorders by using the PRNU with various codecs and resolutions. In fact video cameras use CCD or CMOS chips as well as digital cameras and when the test video is long enough the obtained results are satisfactory. However, the task of source camera identification using videos is more challenging than the image counterpart, due to the degraded visual quality of videos and also to the static nature of video content. In particular, in [5], a study on compressed image and video is proposed stating that when the images (or video frames), from which the sensor fingerprint is estimated, are heavily lossy compressed an adjustment of the decision threshold is required to guarantee a certain false-alarm rate. Furthermore the technique presented by Lukas et al. [1] is applied to videos downloaded from YouTube [6] and for low resolution videos in [7]. Some experiments varying the codec, quality settings and recording resolution, are reported obtaining satisfactory results. In [8], the authors propose a method to identify streamed videos in wireless transmission; finally in [9] a different mechanism for estimating the reference PRNU is proposed finding that different video frame types (I and P) should have also different levels of reliability for PRNU estimation. An extended overview on video forensics which takes into account different issues concerning the matter is reported in [10].

3. Sharing videos on Facebook and Twitter

There are three ways to share videos on Twitter and Facebook: the user can record, edit and share videos from the SNs applications from iOS and Android smartphone, import videos from the device (smartphone and tablet) and finally upload videos through Twitter and Facebook web site. Twitter, in particular, supports MP4 and MOV video formats on mobile app and the user can upload videos up to 512 MB, that however, do not exceed 2 min and 20 s of length.[1] In Twitter the user can select a particular video clip to share, deleting a part of the video before tweeting it, by dragging and moving sideways. Facebook similarly support H.264 video in MOV or MP4 format and a recommended frame width no larger than 1280 pixels and however divisible by 16 pixels. Videos must be long less than 120 min and smaller, as file size, than 4 GB. In Table 1 characteristics of videos for upload compliance on Twitter and Facebook are summarized.

Obviously each video uploaded on Facebook and Twitter will need to be processed before other users can see it and the processing applied to the video is not known a priori. Once uploaded, it is possible to download video from Twitter by using some web services where you can copy the video link and choose the different download resolutions.[2] Concerning Facebook, instead, it is possible to save the video directly from the web browser at the maximum resolution provided by the SN according to the format of the uploaded video.

4. Fingerprint computation

The Photo Response Non-Uniformity (PRNU) noise is unique for each sensor, as demonstrated in [1] and it is generated by the imperfections due to the device construction. Usually PRNU noise is extracted from an image through a digital filtering operation and the fingerprint is obtained averaging multiple PRNUs obtained from images of the same digital camera. After that, the PRNU of the to-be-checked image is compared with the pre-computed PRNU fingerprints, belonging to a reference set and then it is assigned to a certain digital camera (if present within the reference set). In the particular scenario depicted by this paper a video is under analysis so, first of all, the video V is split in individual frames I_i $(i = 1 : N)$ where N is the total amount of frames in V. A wavelet denoising filter D [11] is used to filter out the scene

Table 1
Facebook and Twitter video options.

	Twitter	Facebook
Max upload length	2 min 20 s	120 min
Max upload size	512 MB	4 GB
Min upload res.	32 × 32	600 (width)
Max upload res.	1920 × 1200	1280 × 720 (recommended)
Max frame rate	40 fps	30 fps
Max bitrate	25 Mbps	–
Video format	MP4, MOV	MP4, MOV

content for each RGB color channel of the frame I_i leaving only the residual noise $n_i = I_i - D(I_i)$. Finally, the fingerprint FP is calculated for each color channel by averaging on a specified number of frames i, then converted to grey levels and finally post-processing operations are applied (e.g. Wiener filtering). The detection process to verify if a video V was taken with a video camera C, is performed using the normalized cross correlation computed between the fingerprint and the current test PRNU, that is $NCC = corr(FP, PRNU_{test})$, following the definition in [4]. In the video case, it is impossible to identify a digital camcorder from a single video frame, as occurred in the image case, because each frame is highly compressed by compression systems such as MPEG-x and so on. Therefore, it is necessary to estimate the PRNU of a test video on multiple frames as occurred for the fingerprint estimation.

So it is extremely important to understand the requirements to estimate a good fingerprint: how many frames are necessary and how to build the reference fingerprint. In this paper various possibilities for fingerprint estimation are examined in order to obtain a fingerprint for the reference dataset. In particular, it has been investigated the impact of the processing performed by social networks on PRNU estimation and also if different kinds of frames and their number induce a different reliability in PRNU estimation.

4.1. Classical fingerprint extraction

First of all, we have taken into consideration the classical technique for the fingerprint extraction as described in [4,5] evaluating different length of the chunk of frames (d_{ch}) within the video assumed for the estimation. In our scenario we have three kinds of different videos: the original video directly recorded by a smartphone and, after having upload it to SNs, the videos downloaded from Twitter and Facebook. A fingerprint for each of the three videos is generated respectively by using d_{ch} frames: FP_O, FP_{Tw}, FP_{Fb}. All of the three videos are obviously associated to the same smartphone.

4.2. Composite fingerprint extraction

Alternatively, a new approach called *composite fingerprint* is proposed where the reference pattern is built by using information coming from the original video and also from the videos downloaded from the SNs. A composite fingerprint should permit to take into account some changes on the PRNU noise introduced by the processing performed by the SNs onto the video. The mixed reference pattern FP_{comp} is obtained extracting the PRNU noise from chunks of frames of length d_{chO}, d_{chFb} and d_{chTw} respectively taken from the three available videos (O, Tw, Fb) combined as depicted in Fig. 1. Usually $d_{chO} = d_{chFb} = d_{chTw}$ but they could be different because of specific needs related to the application scenario. Then the PRNU is extracted from each frame through a digital filtering operation and finally the fingerprint is obtained by averaging on all the frames.

It is necessary to point out that only the original video (at least of d_{chO} length) need to be available for the production of the fingerprint because it can be uploaded on the various SNs and then downloaded to be mixed with the original version. So in a real application, the analyst does not need to have access to the three versions of the video necessary

[1] Such a limit was of 30 s and has been incremented on 21st June 2016 https://blog.twitter.com/2016/new-ways-to-tap-into-video-on-twitter.

[2] https://savedeo.com/sites/twitter.

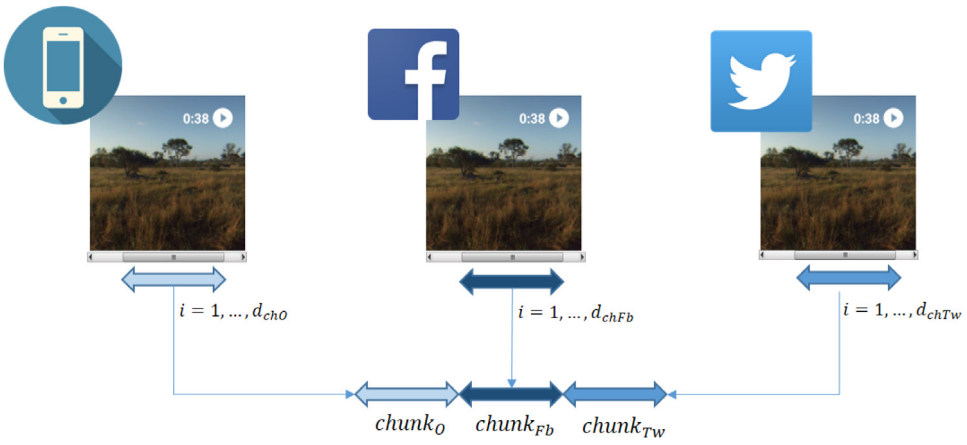

Fig. 1. The construction of the composite fingerprint by taking sub-parts (chunks) of the three different videos.

Table 2
List of the smartphones and features of the acquired videos.

Smartphone	Video res.	Format	Duration	Chunks
Samsung Galaxy S4	1280×720	MP4	4 m 1 s	7
Apple iPhone 5	1920×1080	M4V	3 m 10 s	5
LG Nexus 5	1920×1080	MP4	2 m 32 s	4
Nokia Lumia 830	1280×720	MP4	4 m 38 s	8
Samsung Galaxy S4 mini	1280×720	MP4	4 m	6

Table 3
$Av(PCE)$ on the evaluated FPs.

FP	Av (PCE)	S4	iPhone 5	Nexus 5	Lumia 830	S4 mini
FP_O	P_{Av}	780.23	10.55	14.46	185.39	13.65
	N_{Av}	0.15	−0.77	−0.46	3.33	−0.41
FP_{Tw}	P_{Av}	602.35	15.73	21.64	81.39	168.06
	N_{Av}	17.27	7.44	4.20	27.01	11.61
FP_{Fb}	P_{Av}	881.39	23.36	16.65	53.43	106.63
	N_{Av}	20.21	3.41	1.55	25.50	10.95
FP_{comp}	P_{Av}	1327.21	48.72	50.49	242.18	240.67
	N_{Av}	12.67	5.46	2.89	15.03	3.15
FP_{compI}	P_{Av}	176.62	9.01	5.63	36.63	21.09
	N_{Av}	2.63	4.09	4.27	6.26	2.48

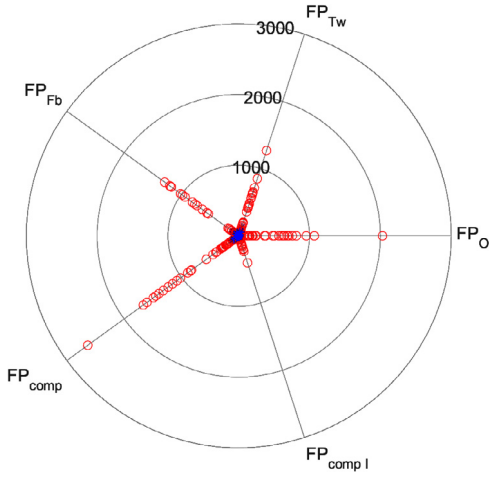

Fig. 2. The distribution of *PCE* values on the polar plot. Each straight line corresponds to a particular fingerprint computation of the considered five.

to compute the composite fingerprint. In this work, it has not been taken into account the case of obtaining the fingerprint by only resorting at videos coming from SNs that has been left to successive studies on more restrictive operative conditions.

4.3. I-frames composite fingerprint extraction

Finally another approach is evaluated i.e. to estimate the fingerprint only from I-frames (intra-coded frames) of the video. It is well known that I-frames are like conventional static image files and they do not require other video frames to be decoded. On the contrary P and B frames (inter-coded) contain motion-compensated difference information relative to previously decoded pictures and are more compressed then I-frames. For this reason I-frames could be more reliable than P-frames or B-frames for PRNU estimation [9]. We thus select a number n of I-frames within d_{ch} from a video V. The number of I-frames within a video depends on the GOP (Group of Pictures) size that is the distance, in terms of frames, between two I-frames, which varies from video to video. So a second version of the composite fingerprint called FP_{compI} is built up by using a variable number of frames. In this case the fingerprint is constructed by considering, as before, d_{chO} frames of the original video, but only I-frames contained within the chunk of length d_{chFb} and d_{chTw} in the case of Facebook and Twitter videos. In particular, the idea behind this choice is to assume that the I-frames are able to produce a more reliable fingerprint with respect to P and B ones and therefore they are more suitable to represent heavily processed videos like those coming from Twitter and Facebook.

5. Experimental results

In this section some of the different experimental tests that have been carried out are presented. First of all, the whole experimental set-up is introduced, subsequently the different kinds of fingerprint proposed in the previous section are compared and the achieved results are commented.

5.1. Set-up description

We selected 5 smartphones to produce various video files of different length at the default smartphone setting resolution (see Table 2). The videos contain scene with different contents as daylight outdoor scenes or indoor with poor illumination.

Each video clip has been uploaded on the two SNs under analysis, Facebook and Twitter, according to their resolution and length restrictions. When it is not possible to upload the entire video (in the Twitter case the upload limitation length is 2 min and 20 s as already evidenced in Table 1), the video is subdivided in different parts and

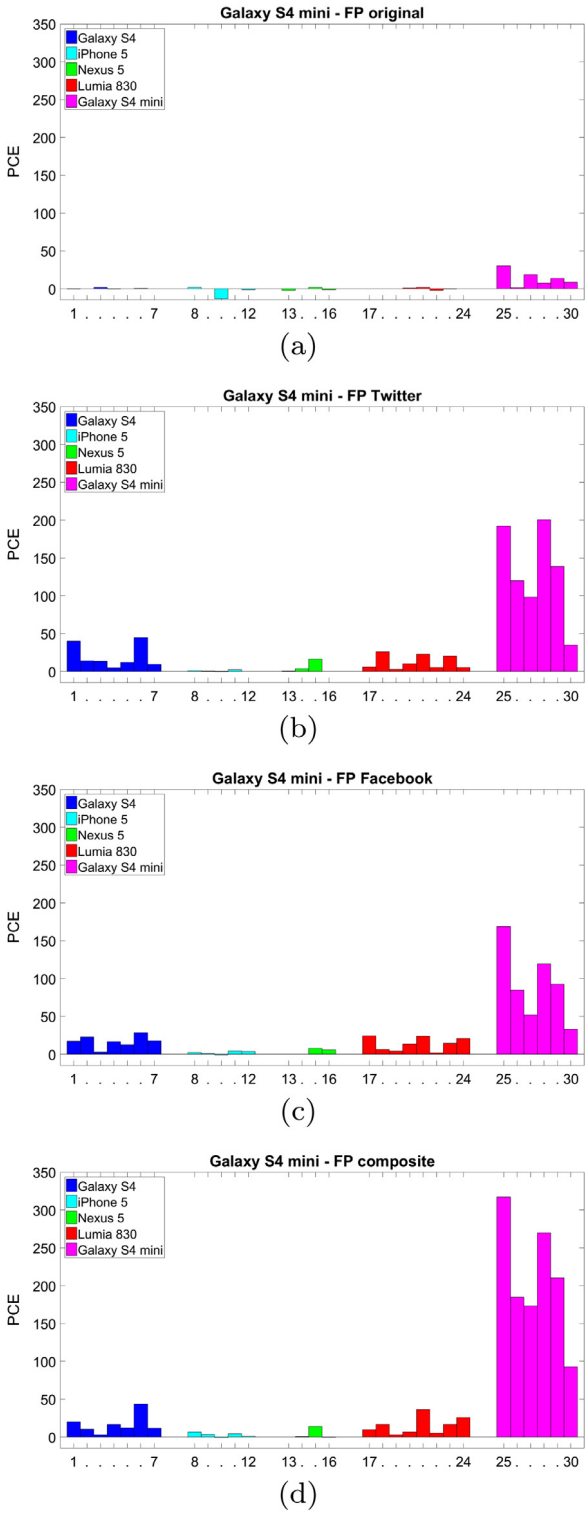

Fig. 3. PCE comparison among four different Galaxy S4 mini fingerprints FP_O, FP_{Fb}, FP_{Tw} and FP_{comp} on Twitter test videos (30 chunks).

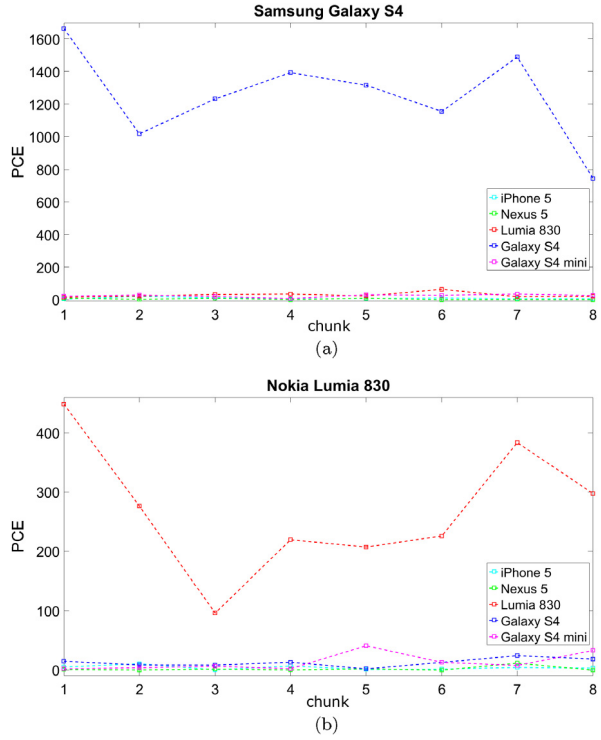

Fig. 4. PCE on a video test downloaded from Facebook vs the FP_{comp} dataset: Galaxy S4 (3 min 36 s, 8 chunks) (a) and Lumia 830 (3 min 36 s; 8 chunks) (b).

then recomposed after the download. All the video sequences have been downloaded at the resolution of 1280×720, MP4 format, so in the case of iPhone5 and LG Nexus 5 videos a resize is performed respect to the upload resolution (Table 2). The five proposed procedures for fingerprint extraction are taken into consideration: classical (Original (O)), Twitter (Tw) and Facebook (Fb)), Composite (Comp) and I-frames Composite (CompI). The related PRNUs are estimated according to what explained in Section 4.1. First of all, for each smartphone, fingerprints FP_O, FP_{Tw} and FP_{Fb} are calculated using d_{ch} frames from the original video recorded from the smartphone (for instance the first part of a video) and from the Facebook and Twitter videos as well. Though different values for d_{ch} have been analyzed, hereafter, for sake of conciseness, results are presented for $d_{ch} = 800$ frames. Furthermore, we construct the *composite fingerprint* FP_{comp} extracting the PRNU noise from frames combined from the three available videos (Original, Facebook and Twitter), as debated in Section 4.2 where $d_{chO} = d_{chFb} = d_{chTw} = 800$ for a total of 2400 frames. In particular, it is necessary to point out that the same original video composed by $d_{chO} = 800$ frames is the one uploaded both on Facebook and on Twitter. Finally, the FP_{compI} is constructed following the indications in Section 4.3 by using a variable number of frames depending on the number of I-frames available in d_{chFb} and d_{chTw} (27 I-frames per video chunk on average) while for the original video all the $d_{chO} = 800$ frames are considered as usual. In the following we evaluate the reliability of each considered fingerprint FP_O, FP_{Tw}, FP_{Fb}, FP_{comp}, FP_{compI} on different videos respect to those used for the fingerprint computation. Such video sequences are recorded, as before, from the same 5 smartphones and then downloaded from Facebook and Twitter after having been previously uploaded on it. In particular, the number of video parts (test chunks) evaluated for each video is evidenced in the last column of Table 2; for sake of coherence, each test chunk is itself composed by 800 frames (that corresponds to about 26.66 s of video with a frame rate of 30 fps). The goal of our analysis is to understand which fingerprint is the best choice for social networks video source identification. We judge the fingerprint reliability in terms of Peak-to-Correlation Energy (PCE) ratio that detects the presence of a peak in the *NCC*. The *NCC* is the cross correlation between the fingerprint itself and the PRNU of the video chunk under evaluation (see Section 4). The PCE value can be also negative.[3]

[3] http://dde.binghamton.edu/download/camera_fingerprint/.

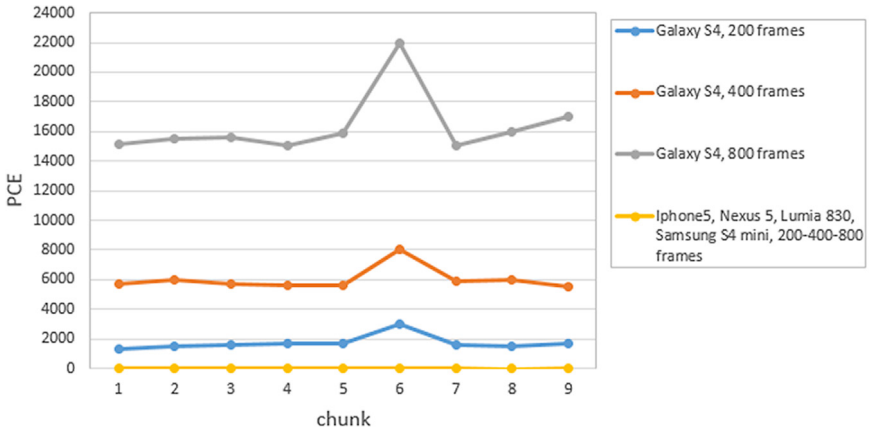

Fig. 5. PCE comparison on a Samsung Galaxy S4 video with all the fingerprints FP_{comp} in the dataset. The grey, orange and blue lines are related to FP_{comp} of the Samsung Galaxy S4 computed with 800, 400, 200 frames respectively. The correlation with the others FP_s (on average) are all collapsed in the yellow line.

Table 4
Comparison between FP_{comp} vs FP_O with $d_{ch} = 2400$. The $Av(PCE_{\mathbf{P}})$ is reported.

Smartphone	$Av(PCE_{\mathbf{P}})FP_{comp}$	$Av(PCE_{\mathbf{P}})FP_O, d_{ch} = 2400$
Galaxy S4	1327.21	918.81
iPhone 5	48.72	13.90
Nexus 5	50.49	50.12
Lumia 830	242.18	155.91
Galaxy S4 mini	240.67	146.63

5.2. Fingerprints evaluation

First of all, a complete overview on the five kinds of fingerprint configurations will be given to understand which one is the most convenient method to extract the PRNU fingerprint.

In Table 3, the average PCE values, $Av(PCE)$, on all the test chunks (30 chunks) is reported in terms of detection on Facebook and Twitter videos for all the evaluated kinds of fingerprints (FP_O, FP_{Tw}, FP_{Fb}, FP_{comp}, FP_{compI}). For example, in the column indicated with "S4" is reported the average PCE obtained when the FP_O (analogously for the other fingerprints) of the Galaxy S4 is correlated with all the chunks belonging to the Galaxy S4 itself (named P_{Av}); while with N_{Av} is intended the average PCE obtained correlating the Galaxy S4 FP_O with all the other smartphones test chunks (iPhone5, Nexus 5, Lumia 830, Galaxy S4 mini).

From the results obtained in Table 3, it is possible to point out that the composite fingerprint FP_{comp} performs quite well obtaining the higher P_{Av} values for all the smartphone fingerprints and getting N_{Av} small enough for the detection. The values of PCEs, P_{Av} and N_{Av}, for the cases FP_{Tw} and FP_{Fb} appear to be comparable with those obtained for FP_O fingerprint and do not seem to provide a significant improvement. It is necessary to point out that, in particular for the iPhone5 and Nexus 5, the P_{Av} values are quite low with respect to the others; this could be determined by the specific compression adopted within such devices, but, however, a certain degree of distinctiveness is still evidenced especially using the proposed composite fingerprint. Furthermore, in Fig. 2, the distribution of all the PCE values is reported ($PCE_{\mathbf{P}}$ in red representing the positive classes i.e. a smartphone is correlated with the respective fingerprint and $PCE_{\mathbf{N}}$ in blue for the negative classes). In particular, on each straight line a specific fingerprint computation is represented and the plot gives us an indication of the performance on each fingerprint. The *composite fingerprint* FP_{comp} again, demonstrates to have the higher distinctiveness among the other FP_s because the red points on the direction of FP_{comp} are more distant from the center with respect to the others. On the other side the blue points, that represents $PCE_{\mathbf{N}}$, are all close to the center as expected (in particular the negative PCE values are set to zero to improve the plot readability). It can be easily appreciated that the FP_{compI} seems the less effective among the fingerprints so it is omitted in the following presented evaluations.

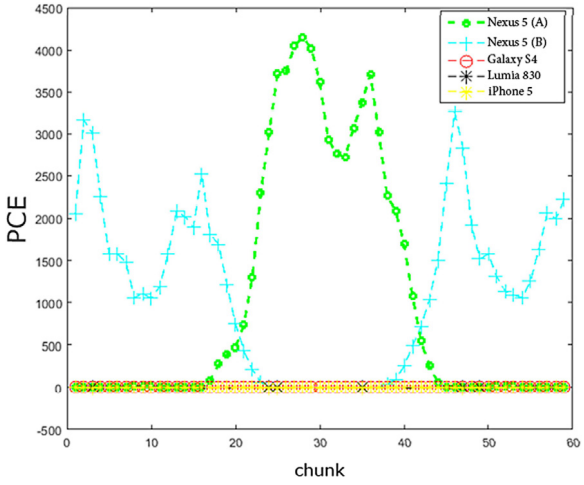

Fig. 6. PCE of a video assembled with a Nexus 5 (B), plus a Nexus 5 (A) and again a Nexus 5 (B) video fragments compared with the dataset of composite fingerprints (the LG Nexus 5 (B) is added to the dataset).

5.3. An in-depth analysis

In order to further verify the proposed fingerprint estimations we evaluate, for sake of clarity, a particular case. Let us consider the five Twitter test videos composed in total by 30 chunks coming from the 5 smartphones (from 1 to 7 chunks from the Galaxy S4, from 8 to 12 from the iPhone5 and so on). In Fig. 3 is reported the PCE values obtained correlating the FP_O, FP_{Tw}, FP_{Fb}, FP_{comp} fingerprints of the Galaxy S4-mini with all the 30 test chunks. It is possible to point out that the FP_{comp} is able to identify the correct chunks acquired by the Galaxy S4-mini smartphone and uploaded on Twitter more efficiently then the other FPs (see the purple columns in the histogram of Fig. 3(d); chunk from 25 to 30).

In Fig. 4, another case is reported, two unknown videos downloaded from Facebook are checked versus our FP_{comp} dataset, composed by 5 fingerprints associated to the 5 smartphones of Table 2. In Fig. 4(a), the video under test is correctly associated to the Galaxy S4 smartphone (which is correct according to the ground-truth) and in Fig. 4(b), the second video under evaluation has been identified as captured by a Lumia 830. A good distinctiveness is granted from the fact that the other correlations (with the other fingerprints) are around zero.

Another experiment has been performed to point out that the composite fingerprint FP_{comp} is suitable also with respect to a classical fingerprint FP_O extracted from the original video on a larger chunk

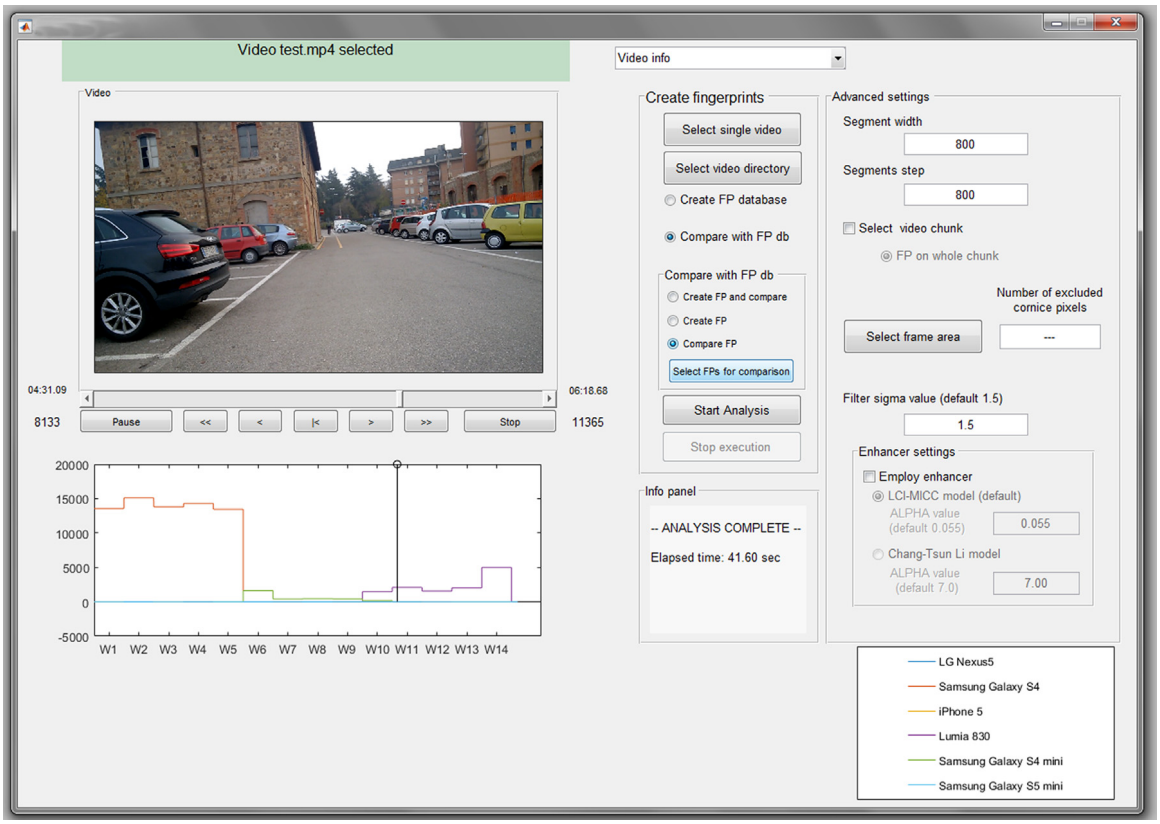

Fig. 7. The *Video Source Identification* tool.

of $d_{ch} = 2400$ frames (i.e. increasing by three the number of frames used for the estimation). This has been done to compare FP_{comp} which is practically calculated on 2400 frames, though it needs only 800 different ones (see Section 4), with FP_O when the same amount of frames are taken into consideration for the construction of the fingerprint. However it is worthy to underline that, in this case, a video clip with a superior time duration would be necessary to compute the FP_O: such a circumstance is not so easy to happen within a social network scenario. In Table 4 the average PCE, $Av(PCE_{\mathbf{P}})$, of the composite fingerprint FP_{comp} is compared with that of FP_O obtained with 2400 frames; results show that performances are still satisfactory.

Hereafter a further insight is given on the issue of the number of frames used to compute the fingerprint. In particular, the cases of 800, 400 and 200 frames have been considered. In Fig. 5 it is evidenced that employing a fingerprint with a greater number of frames is more suitable to evidence the distinction among fingerprints: the grey line which represents 800-frames fingerprint is more distant from the yellow one (other devices FPs) with respect to the orange (400-frames FP) and the blue (200-frames FP). So this shows that, to obtain a good trade-off between performances and number of frames, 800 frames constitutes a sufficient amount of pictures to achieve a reliable fingerprint.

Finally, a new experiment has been performed to check the behavior of the FP_{comp} fingerprint with respect to a post-processed video whose composition is unknown with the intent to simulate a possible social network case. In particular, to do this, we have taken two LG Nexus 5 smartphones (one already present in the dataset and a new one) and constructed a new video that is the composition of three video sequences: the first part, from chunk 1 to chunk 20, is coming from the LG Nexus 5 (named B), the second part (from chunk 21 to 43) from the LG Nexus 5 named (A) and finally, the third part (from 44 to 59) from the LG Nexus 5 (B) again. In Fig. 6 the result in terms of PCE is reported, in particular, it is evidenced that the two different LG Nexus 5 smartphones are well distinguished (cyan and green lines). This also proves that intra-model (devices with same brand and model) case can be managed by this kind of approach.

6. Tool description

In this section the tool to perform video source identification is described and the related GUI (Graphical User Interface) designed to support forensic analysts in their activity is shown in Fig. 7. The interface allows to select a single video file or multiple test videos; when the user works with a single video, information about resolution, frame rate, number of frames, duration, etc. are displayed in a specific drop-down menu (*Video info* at top-center of Fig. 7). In this tool different modalities of investigation are foreseen: when the forensic analyst has not any kind of prior information on a video origin a fully-automatic analysis it is necessary; on the other hand, if the operator it is interested in a specific part of a video a focused analysis on a sub-part has to be preferred. So the user is allowed to choose among different settings, made available by the tools, performing at default or at advanced level.

In the following are described the main options exploited by the tool:

Segment width: number of frames used to compute the fingerprint (default value is 800). When the length of the video is lower than this value, the whole video is considered for the fingerprint estimation; when the video is longer, a certain number of segments of such dimension are taken, resulting in several fingerprints for the same video;
Segments step: bias between the starting frame of one segment and the starting frame of the next one. Default is 800, that means that segments are contiguous; this is obviously true whenever *segment width* is equal to *segments step*. Varying the *segment step* value will result in overlapping parts of the video segments or inserting some gaps between subsequent segments of the video;
Selection of video chunk: an analyst can decide to select a sub-part (chunk) of the whole video, excluding, for example, too noisy clips. Two sliders appear under the video box, letting the user to select independently the start and the stop frames;

Selection of frame area: it permits to specify a frame area to be processed in order to ignore a cornice of pixels.

σ value and Enhancer: the value of σ is a parameter for the extraction of the fingerprint related to the filter used for PRNU estimation and it can be adjusted by the user or used in the default modality. Furthermore, two different *enhancers* [3,12], devised to improve the PRNU fingerprint, can be selected modifying their settings through another value named α.

Result presentation: a box in the bottom-left side of the GUI (see Fig. 7) shows the graphical representation of the results. In particular the *PCE* value, obtained for each video test segment correlated with each fingerprint in the database, is depicted. A different colored line is drawn for each element populating the dataset. When the video is played a marker slides on the depicted correlation graph: the analyst is thus able to check the exact correspondence between the visualized frame and the related fingerprint understanding which part of the video is recorded with a certain smartphone.

7. Conclusions and future works

This paper has proposed an analysis for the source identification of videos uploaded on different social networks, specifically, Twitter and Facebook. Five different kinds of fingerprint extraction methods have been evaluated and, in particular, a novel method to build a composite fingerprint to achieve better video source identification has been proposed. A tool useful for a forensic analyst has been introduced and adopted to carry out the experimental tests. Results obtained on various videos, firstly uploaded and then downloaded from Facebook or Twitter, have demonstrated that the device identification is still possible. Future works will be devoted to extend the experiments in an opener set scenario increasing the number of smartphones taken into account to realize a statistical analysis for a determination of a threshold. Furthermore, it could be interesting to investigate the case of Twitter videos directly uploaded on Facebook (and vice versa) making the identification even more difficult due to the increment of post-processing applied to the video and also evaluating different download resolutions for each social networks.

References

[1] M. Chen, J. Fridrich, M. Goljan, J. Lukas, Determining image origin and integrity using sensor noise, IEEE Trans. Inf. Forensics Secur. 3 (1) (2008) 74–90.

[2] N. Mondaini, R. Caldelli, A. Piva, M. Barni, V. Cappellini, Detection of malevolent changes in digital video for forensic applications, in: E.J. Delp, P.W. Wong (Eds.), SPIE Conference on Security, Steganography, and Watermarking of Multimedia Contents, Vol. 6505, 2007.

[3] C.T. Li, Source camera identification using enhanced sensor pattern noise, IEEE Trans. Inf. Forensics Secur. 5 (2) (2010) 280–287. http://dx.doi.org/10.1109/TIFS.2010.2046268.

[4] M. Chen, J. Fridrich, M. Goljan, J. Lukas, Source digital camcorder identification using sensor photo response non-uniformity, in: SPIE Conference on Security, Steganography, and Watermarking of Multimedia Contents, Vol. 6505, 2007, pp. 65051G–65051G–12. < http://dx.doi.org/10.1117/12.696519> .

[5] M. Goljan, M. Chen, P.C. Alfaro, J. Fridrich, Effect of compression on sensor-fingerprint based camera identification, in: Media Watermarking, Security, and Forensics 2016, IS&T Electronic Imaging 2016, 2016.

[6] W. van Houten, Z. Geradts, Source video camera identification for multiply compressed videos originating from Youtube, Digit. Investig. 6 (12) (2009) 48–60. http://dx.doi.org/10.1016/j.diin.2009.05.003. (http://www.sciencedirect.com/science/article/pii/S1742287609000310).

[7] D.-K. Hyun, C.-H. Choi, H.-K. Lee, Camcorder Identification for Heavily Compressed Low Resolution Videos, Springer, The Netherlands, Dordrecht, 2012, 695–701. http://dx.doi.org/10.1007/978-94-007-2792-2_68. (<http://dx.doi.org/10.1007/978-94-007-2792-2_68>).

[8] S. Chen, A. Pande, K. Zeng, P. Mohapatra, Live video forensics: source identification in lossy wireless networks, IEEE Trans. Inf. Forensics Secur. 10 (1) (2015) 28–39. http://dx.doi.org/10.1109/TIFS.2014.2362848.

[9] W.H. Chuang, H. Su, M. Wu, Exploring compression effects for improved source camera identification using strongly compressed video, in: 2011 18th IEEE International Conference on Image Processing, 2011, pp. 1953–1956. doi:10.1109/ICIP.2011.6115855.

[10] A.C. Kot, H. Cao, Image and Video Source Class Identification, Springer, New York, New York, NY, 2013, pp. 157–178. http://dx.doi.org/10.1007/978-1-4614-0757-7_5.

[11] M.K. Mihcak, I. Kozintsev, K. Ramchandran, Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising, in: Proc. of IEEE ICASSP, Phoenix, USA, 1999.

[12] R. Caldelli, I. Amerini, F. Picchioni, M. Innocenti, Fast image clustering of unknown source images, in: 2010 IEEE International Workshop on Information Forensics and Security, 2010, pp. 1–5. <http://dx.doi.org/10.1109/WIFS.2010.5711454>.