Can contact-free measurement of heartbeat signal be used in forensics?

Haque, Mohammad Ahsanul; Nasrollahi, Kamal; Moeslund, Thomas B.

Published in:
23rd European Signal Processing Conference (EUSIPCO), 2015

DOI (link to publication from Publisher):
10.1109/EUSIPCO.2015.7362487

Publication date:
2015

Document Version
Early version, also known as pre-print

CAN CONTACT-FREE MEASUREMENT OF HEARTBEAT SIGNAL BE USED IN FORENSICS?

Mohammad A. Haque, Kamal Nasrollahi, and Thomas B. Moeslund

Visual Analysis of People (VAP) Laboratory
Rundsburggade 14, 9000 Aalborg, Denmark
kn, mah, tbm@create.aau.dk

ABSTRACT

Biometrics and soft biometrics characteristics are of great importance in forensics applications for identifying criminals and law enforcement. Developing new biometrics and soft biometrics are therefore of interest of many applications, among them forensics. Heartbeat signals have been previously used as biometrics, but they have been measured using contact-based sensors. This paper extracts heartbeat signals, using a contact-free method by a simple webcam. The extracted signals in this case are not as precise as those that can be extracted using contact-based sensors. However, the contact-free extracted heartbeat signals are shown in this paper to have some potentials to be used as soft biometrics. Promising experimental results on a public database, have shown that utilizing these signals can improve the accuracy of spoofing detection in a face recognition system.

Index Terms— Forensics, biometrics, soft biometrics, heartbeat signals

1. INTRODUCTION

Forensic science deals with techniques used in criminal scene investigation for collecting information and its interpretation for the purpose of answering questions related to a crime in a court of law [1]. Such information is usually related to the human body or behavioral characteristics that can (or help to) reveal the identity of criminal(s). Such characteristics are extracted either from traces that are left in the crime scene, like DNA, handwritings, and fingerprints [2], or devices like cameras and microphones, if any, that have been recording visual and audio signals in the scene during the crime. Recorded visual signals as well as audio signals can be of great help as many different characteristics can be extracted from them, characteristics that can directly be used for identification (biometrics), or can help the identification process (soft biometrics).

Forensics investigations have long been utilizing such human characteristics. For example, the importance of DNA in [3], fingerprint in [4], facial images and its related soft biometrics in [5], [6], gait in [7] have been discussed for forensics applications. However, most of these biometric/soft biometric traits exhibit disadvantages in regards to accuracy, spoofing and/or unobtrusiveness. For example, fingerprint can be forged to breach the identification system, gait can be imitated, and facial image can be used in absence of the person. Thus, further investigations for new biometric traits are of interest of many applications. Human heartbeat signal is one of such emerging biometric traits.

Human heart is a muscular organ that works as a circulatory blood pump. When blood is pumped by the heart, some electrical and acoustic changes occur in and around the heart in the body, which is known as heartbeat signal [8]. Heartbeat signal can be obtained by Electrocardiogram (ECG) using electrical changes and Phonocardiogram (PCG) using acoustic changes. Both ECG and PCG heartbeat signals have already been used as biometrics for human identity verification in the literature. A review of such important ECG-based approaches can be obtained from [9]. On the other hand, a review of the important PCG-based identification methods can be found in [10]. The common drawback of all of the above mentioned ECG and PCG based methods for identity verification is the requirement of using obtrusive (contact-based) sensors for the acquisition of ECG or PCG signals from a subject. In another words, for obtaining heart signals using ECG and PCG the required sensors need to be directly installed on subject's body, which is obviously not always possible. Therefore, in this paper we look at contact-free measurement of the Heartbeat Signal from Facial Videos (HSFV) and investigate its distinctiveness potential. It is shown in this paper that such signals have distinctive features, however, their distinctiveness capability are not that high to use them as biometrics. Instead, it is shown that they can help improving the detection accuracy of a face spoofing detection algorithm in face recognition system, and hence can be used as a soft biometric in forensics applications, for example, with video-based face recognition algorithms.

The proposed system obtains HSFV signals from video sequences that are captured by a simple webcam, by tracing changes in the color of facial images that are caused by the heart pulses. Then, it extracts some distinctive features from these signals. Unlike ECG and PCG based heartbeat biomet-

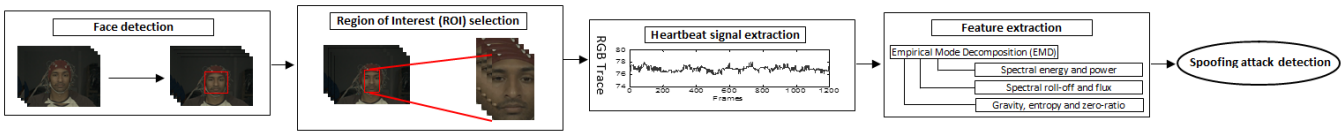


Fig. 1. The block diagram of the proposed system.

ric, the proposed HSFV soft biometric does not require any obtrusive sensor such as ECG electrode or PCG microphone. It is universal and permanent, obviously because every living human being has an active heart. It can be more secure than its traditional counterparts as it is difficult to be artificially generated, and can be easily combined with state-of-the-art face biometric without requiring any additional sensor.

The rest of this paper is organized as follows: the proposed system for detecting spoofing attacks in a face recognition system and its sub-blocks (including the heartbeat signals measurement and feature extraction) are explained in the next section. The experimental results are given in section 3. Finally the paper is concluded in section 4.

2. THE PROPOSED SYSTEM

The block diagram of the proposed system is shown in Fig. 1. Each of these sub-blocks of the system are described in the following subsections.

2.1. Facial video acquisition

The first step is capturing the facial video using a RGB camera, which is thoroughly investigated in the literature [11, 12]. As recent methods of facial video based heartbeat signal analysis utilized simple webcam for video capturing, we select a webcam based video capturing procedure.

2.2. Face detection

Face detection is accomplished by the well-known Haar-like features based Viola and Jones method of [13]. The face region is expressed by a rectangular bounding box in each video frames.

2.3. Region of Interest (ROI) selection

As the face area detected by the automatic face detection method comprises some of the surrounding areas of the face including face boundary, it is necessary to exclude the surrounding area to retain merely the area containing facial skin. This is accomplished by obtaining a Region of Interest (ROI) from the face by selecting 60% width of the face area detected by the automatic face detection method.

2.4. Heartbeat signal extraction

The heartbeat signal is extracted from the facial video by tracing color changes in RGB channels in the consecutive video frames. The average of the red, green and blue components of the whole ROI is recorded as the RGB traces of that frame. In order to obtain a heartbeat signal from a facial video, the statistical mean of these three RGB traces of each frame is calculated and recoded for each frame of the video. The resulting signal represents the heartbeat signal.

The heartbeat signal obtained from facial video by following the aforementioned approach is, however, noisy and imprecise. This is due to the effect of external lighting, voluntary head-motion, induced noise by the capturing system and the act of blood as a damper to the heart pumping pressure to be transferred from the middle of the chest (where the heart is located) to the face. Thus, we employ a denoising filter by detecting the peak in the extracted heart signal and discarding the out-lying RGB traces. The effect of the denoising operation on a noisy heartbeat signal obtained from RGB traces is depicted in Fig. 2. The signal is then passed through a Hodrick-Prescott filter [14] with a smoothing parameter (value = 2 in our case) in order to decompose it into trend and cyclic components. As heartbeat is a periodic vibration due to the heart pulse, we assume that the trend component comprises the noise in the heartbeat signal induced by voluntary head-motion. Thus, we obtain the denoised heartbeat signal from the cyclic component.

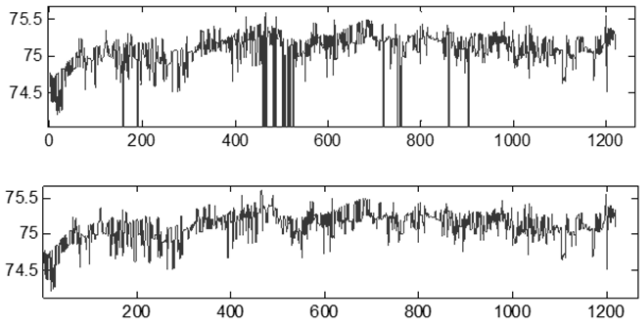


Fig. 2. An example of heartbeat signal before (top) and after (bottom) employing a denoising filter. On the x axis is the frame number and on the y axis is the RGB trace.

2.5. Feature extraction

The feature extraction from the denoised HSFV is accomplished by employing a decomposition method called Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) from [15]. The CEEMDAN decomposes the HSFB into a small number of modes called Intrinsic Mode Functions (IMFs). The number of IMFs can vary, but not less than 6 in this case. Thus, we considered first 6 IMFs. An example of first 6 IMFs for a HSFV in the case of a real face are shown in Fig. 3. We then calculated some spectral features of the original HSFV and each of the 6 IMFs for feature extraction. The extracted features are the statistical mean and variance of the spectral energy, power, low-energy, gravity, entropy, roll-off, flux, zero-ratio as stated in [16, 17]. As a whole, we extracted a feature vector of 112 elements for each facial video.

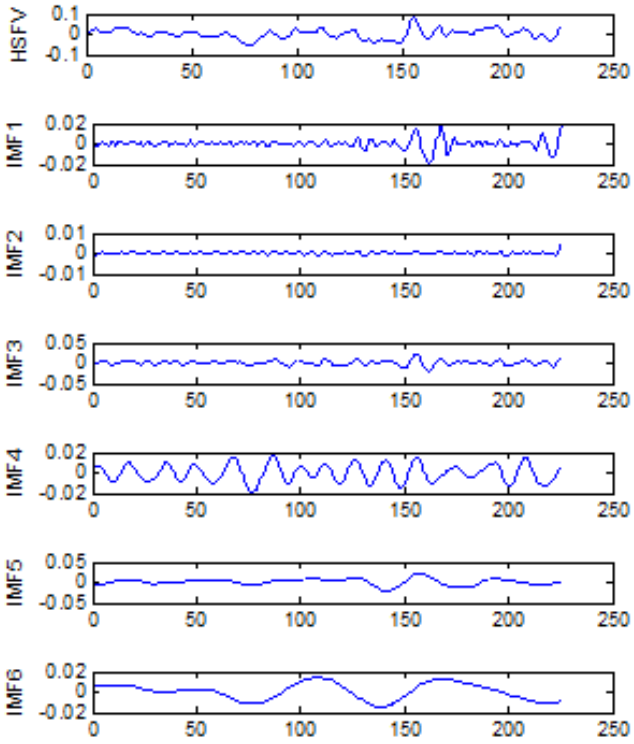


Fig. 3. An example of first 6 IMFs obtained for a HSFV of a real face. On the x and y axis are number of frames in the video and the amplitude of the signal, respectively.

2.6. Spoofing attack detection

We employ the Support Vector Machines of [18], with a tangent hyperbolic kernel function, as the classifier to discriminate between real facial video and spoofing attack.

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

3.1. Experimental environment

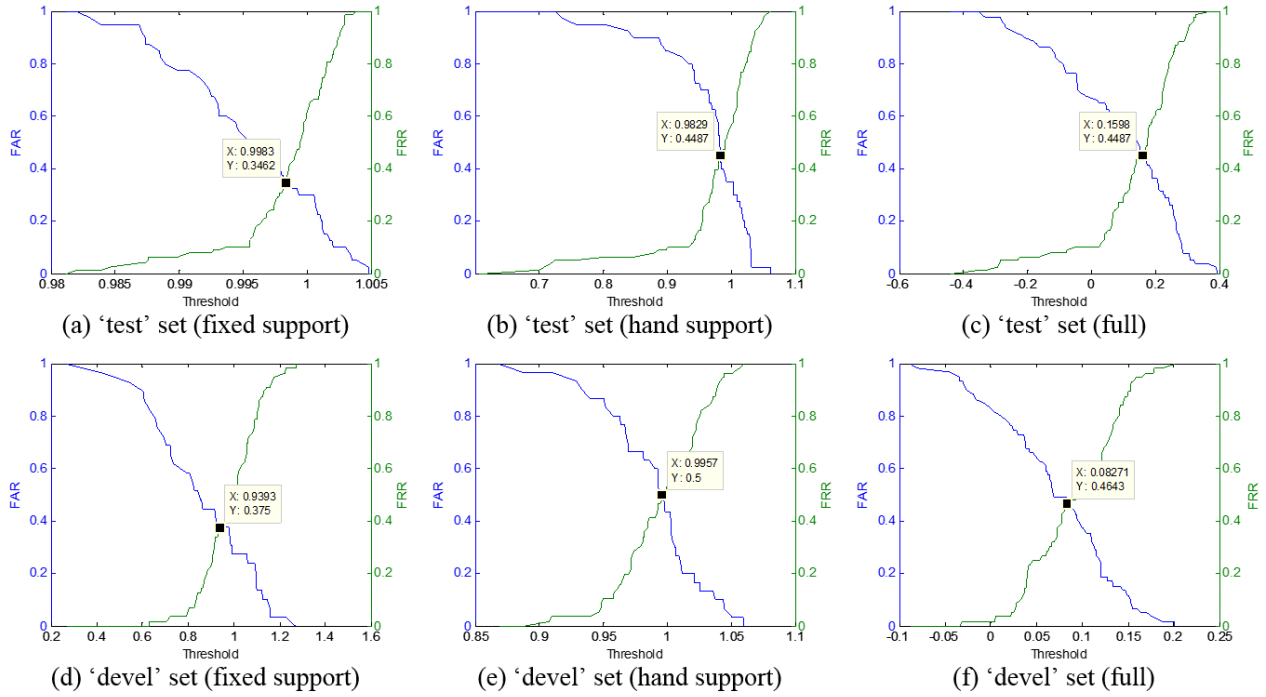
The performance of spoofing detection using the HSFV was evaluated in a system implemented in a combination of MATLAB and C++ environment by following the methodology of the previous section. We used the publicly available PRINT-ATTACK database [19] for spoofing attack detection. This database was collected by Anjos et al. and contains facial videos (each about 10 seconds long) captured by a simple webcam. The videos of the database are recorded in three scenarios. These are: video of real face, video of printed face held by operator's hand, video of printed face held by a fixed support. All these videos were then categorized into three sets: train, devel, and test. The details are given in Table 1. However, some of these videos are too dark to automatically detect face and extract heart signal. Thus, we discarded 2 videos from the train set and 4 videos from the devel set. The rest of the videos was used in the experiment.

Type	**"train"**	**"devel"**	**"test"**	**Total**
Real face	60	60	80	200
Printed face in hand	30	30	40	100
Printed face in a support	30	30	40	100
Total	120	120	160	400

Table 1. The Numbers of videos in different groups of the PRINT-ATTACK database [19].

3.2. Performance evaluation

A spoofing detection system exhibits two types of errors: accepting a spoofed face and rejecting a real face. First error is measured by False Acceptance Rate (FAR) and the second one is measured by False Rejection Rate (FRR). The performance can be depicted on a Receiver Operating Characteristics (ROC) graph where FAR and FRR are plotted against a threshold to determine the membership in true groups of real access and spoofed access. The ROC of the different combinations of datasets, after the training using train set, are shown in Fig. 4. From the results it is observed that the Equal Error Rates (EER), where FAR and FRR curves intersect, are different for different settings. When a printed face was shown in front of the camera by holding it in a fixed place the periodic variation in the heart signal of a real face can be discriminated with higher accuracy as shown in Fig. 4(a) and (d). But, when the printed face was held by hand, the hand shaking behavior affects the result. However, the results show that HSFV can retrieve the difference between real face and print attack moderately accurately.



(a) 'test' set (fixed support) (b) 'test' set (hand support) (c) 'test' set (full)

(d) 'devel' set (fixed support) (e) 'devel' set (hand support) (f) 'devel' set (full)

Fig. 4. ROC curves for employing the HSFV based spoofing detection on different combinations of the experimental datasets.

3.3. Performance comparison

We compare the performance of the HSFV with the baseline results for the test set of the PRINT-ATTACK database provided in [19]. The results are shown in Table 2.

Type	**"test"**
HSFV (fixed)	66,10
HSFV (hand)	66,10
HSFV (all)	61,39
Baseline (fixed)	**77,94**
Baseline (hand)	85,47
Baseline (all)	82,05
HSFV (fixed)+BG	75,64
HSFV (hand)+BG	**91,03**
HSFV (all)+BG	**85,90**

Table 2. Performance of HSFV in face spoofing detection in comparison to a baseline approach for PRINT-ATTACK database [19].

From the results it is observed that the HSFV alone cannot outperform the baseline approach. However, the experiment reveals that the HSFV is able to unveil some clues between real face and printed face shown to a biometric system, and can be a potential soft biometric to be used along with other features to achieve a doable result in face spoofing detection. This notion is shown in the last three rows of Table 2, where the features extracted from HSFV is fused with four background features (maximum, minimum, average and standard deviation of the signal obtained from the video frames background by following [19]), hereafter referred as BG. We employ a score-level fusion of probability estimates of classifier outputs obtained for HSFV and BG features. We observe that the fusion of HSFV and BG features improves the performance in face spoofing detection. One significant point to mention is that when BG features were fused with HSFV features in score-level, the spoofing detection system showed reduced performance than the baseline as indicated by the third-last row of Table 2. We believe that this is because of sensitivity of HSFV to the periodic noise induced by the camera capturing system.

3.4. Discussions about HSFV as a soft biometric for forensic investigations

Face recognition systems need to be robust against spoofing attacks. As printed face spoofing attack is very common in this regard, we investigate the potential of HSFV as a soft biometric for printed face spoofing attack detection in a face recognition system. Though from the results it is observed that the HSFV alone cannot provide very high accuracy, the experimental results show that the HSFV is able to unveil some clues between real face and printed face shown to a biometric system. When HSFV was fused with some other features (BG features in our experiment), the results were consid-

erably improved in most of the cases. Thus, we can infer that the HSFV carries some distinctive features and can be considered as a soft biometric. Hence, one could answer the question raised by the paper by a Yes answer, i.e., heartbeat signals extracted from facial images have some distinctive properties and might be useful for forensics applications.

4. CONCLUSIONS

This paper investigated the potential of the heartbeat signal from facial video as a new soft biometric. To do that, the distinctive properties that are carried in a heartbeat signal have been utilized to improve the accuracy of spoofing attack detection in a face recognition system. A description of the procedure of heartbeat signal extraction from facial video was provided and experimental results were generated by using a publicly available database for printed face spoofing attack detection in a face recognition system. The experimental results revealed that the contact-free measured heartbeat signal has the potential to be used as a soft biometric.

REFERENCES

[1] D. Meuwly and R. Veldhuis, "Forensic biometrics: From two communities to one discipline," in *Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG - Proceedings of the International Conference of the*, Sept 2012, pp. 1–12.

[2] A. Swaminathan, M. Wu, and K.J.R. Liu, "Digital image forensics via intrinsic fingerprints," *Information Forensics and Security, IEEE Transactions on*, vol. 3, no. 1, pp. 101–117, March 2008.

[3] D. Ehrlich, L. Carey, J. Chiou, S. Desmarais, S. El-Difrawy, L. Koutny, R. Lam, P. Matsudaira, B. Mckenna, L. Mitnik-Gankin, T. O'Neil, M. Novotny, A. Srivastava, P. Streechon, and W. Timp, "Mems-based systems for dna sequencing and forensics," in *Sensors, 2002. Proceedings of IEEE*, 2002, vol. 1, pp. 448–449 vol.1.

[4] W.S. Lin, S.K. Tjoa, H.V. Zhao, and K.J.R. Liu, "Digital image source coder forensics via intrinsic fingerprints," *Information Forensics and Security, IEEE Transactions on*, vol. 4, no. 3, pp. 460–475, Sept 2009.

[5] A.K. Jain, B. Klare, and U. Park, "Face matching and retrieval in forensics applications," *MultiMedia, IEEE*, vol. 19, no. 1, pp. 20–20, Jan 2012.

[6] J.E. Lee, A.K. Jain, and R. Jin, "Scars, marks and tattoos (smt): Soft biometric for suspect and victim identification," in *Biometrics Symposium, 2008. BSYM '08*, Sept 2008, pp. 1–8.

[7] H. Iwama, D. Muramatsu, Y. Makihara, and Y. Yagi, "Gait-based person-verification system for forensics," in *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on*, Sept 2012, pp. 113–120.

[8] B. Phibbs, *The Human Heart: A Basic Guide to Heart Disease*, M - Medicine Series. Lippincott Williams & Wilkins, 2007.

[9] M. Nawal and G.N. Purohit, "Ecg based human authentication: A review," *Int. J. Emerg. Eng. Res. Technol.*, vol. 2, no. 3, pp. 178–185, Jun 2014.

[10] F. Beritelli and A. Spadaccini, "Human identity verification based on heart sounds: Recent advances and future directions," *CoRR*, vol. abs/1105.4058, 2011.

[11] M.A. Haque, K. Nasrollahi, and T.B. Moeslund, "Real-time acquisition of high quality face sequences from an active pan-tilt-zoom camera," in *Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on*, Aug 2013, pp. 443–448.

[12] M.A. Haque, K. Nasrollahi, and T.B. Moeslund, "Constructing facial expression log from video sequences using face quality assessment," in *9th International Conference on Computer Vision Theory and Applications (VISAPP)*, 2014, pp. 1–8.

[13] Paul Viola and Michael J. Jones, "Robust real-time face detection," *Int. J. Comput. Vision*, vol. 57, no. 2, pp. 137–154, May 2004.

[14] R.J. Hodrick and E.C. Prescott, "Postwar u.s. business cycles: An empirical investigation," *Journal of Money, Credit and Banking*, vol. 29, no. 1, pp. 1–16, February 1997.

[15] M.E. Torres, M.A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, May 2011, pp. 4144–4147.

[16] M.A. Haque and J.M. Kim, "An analysis of content-based classification of audio signals using a fuzzy c-means algorithm," *Multimedia Tools and Applications*, vol. 63, no. 1, pp. 77–92, 2013.

[17] N.T.T. Nguyen, M.A. Haque, C.H. Kim, and J.M. Kim, "Audio segmentation and classification using a temporally weighted fuzzy c-means algorithm," in *Advances in Neural Networks ISNN 2011*, vol. 6676 of *Lecture Notes in Computer Science*, pp. 447–456. Springer Berlin Heidelberg, 2011.

[18] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sept. 1995.

[19] A. Anjos and S. Marcel, "Counter-measures to photo attacks in face recognition: A public database and a baseline," in *Biometrics (IJCB), 2011 International Joint Conference on*, Oct 2011, pp. 1–7.