Multimedia Tools and Applications (2021) 80:23313–23317
https://doi.org/10.1007/s11042-021-11108-3

GUEST EDITORIAL

Large-scale multimedia signal processing for security and digital forensics [SI 1163]

Ebroul Izquierdo[1] · Krishna Chandramouli[2] · Anthony T. S. Ho[3] · Hyoung Joong Kim[4]

Accepted: 25 May 2021/
Published online: 21 June 2021

It is widely acknowledged by the multimedia community that progress on new media applications has leapfrogged over the last decade. This significant trend can be traced back to the adoption of deep-learning technologies within main-stream scientific innovation. This trend has also led to the development of critical application including, among many others, generating text that describes images, automatic translation across multiple languages, automatic tagging of semantic objects in images and the ability to perform highly accurate face recognition and person detection. Since machine learning and more specifically deep-learning models aim to emulate the cognitive functions of humans, there is strong evidence that derived technologies and applications are often subjected to misuse. Instances of such technology misuse can be found in the pervasive propagation of misinformation, fake pictures and videos and fake news, where the validity of the information distributed is not properly authenticated. Despite efforts launched by platform owners such as Facebook, Google, Twitter and others to combat the spread of misinformation, the capabilities provided by automated multimedia signal processing technology to combat such challenges are often underexploited. Parallel and complementary to the scientific advancements in the field of machine learning and deep-learning, developments in the field of big-data systems have also enabled software application

✉ Ebroul Izquierdo
ebroul.izquierdo@qmul.ac.uk

Krishna Chandramouli
k.chandramouli@venaka.co.uk

Anthony T. S. Ho
a.ho@surrey.ac.uk

Hyoung Joong Kim
khj-@korea.ac.kr

[1] Multimedia and Vision Laboratory, School of Electronic Engineering and Computer Science, Queen Mary, University of London, London, UK

[2] Venaka Media Limited, London, UK

[3] University of Surrey, Guildford, UK

[4] Korea University, Seoul, Korea

and models to effectively handle and extract information large-volumes of data. Unfortunately, in many cases these two research fields continue developing in parallel, without exploitation of their synergies and complementarity. More importantly, such systems are not yet widely available to security and governmental organisations to experiment, validate and deploy as critical weapons to combat the threats originated by the widespread of misinformation.

Addressing the gap within the multimedia community on the nature of threats commonly faced by security organisations, this special issue presents a good sample of integrative research that bring together five thematic challenges into the field of security and digital forensics, namely (i) image authentication and verification; (ii) audio verification; (iii) biometric authentication; (iv) privacy-aware technologies and social implications; (v) criminal investigation for crimes against property, person and critical infrastructures. The total of 19 papers published in this special issue represent a broad call for research addressing the challenges outlined before.

The first ten papers summarise novel algorithms in handling image encryption. The first paper, by El-Khamy and Mohamed, introduces an efficient image encryption method which is inspired by DNA nitrogenous bases. The resultant image is a combination of DNA sequences that are transformed into four sub-images A, C, G, and T, DNA's bases. Chen's hyper-chaotic map is used to diffuse the resultant images, according to a control code. The four DNA images are then combined using a wavelet confusion algorithm to produce an encrypted image. Numerical simulation was used to examine the effectiveness of the encrypted image against different attacks. In the second paper, by Wang and Liu, the authors propose an application of chaotic Josephus scrambling and RNA computing in image encryption. The algorithm uses the classical 'scrambling-diffusion' process, and the pseudo-random sequences used in each stage are generated by the hyper-chaotic Lorenz system. During the scrambling phase, the randomness of the traditional Josephus traversal sequence is improved by using chaotic mapping, which makes the image scrambling effect better. Then during the diffusion phase, the grey value of the pixel is modified by using RNA modular operation and random substitution of RNA codons. The third paper, by Chidambaram et al., introduces a new hash algorithm for the RGB image is proposed that is able to validate the integrity and localise the tampered region within the image. In this paper authors adopt the approach of treating the whole image region as sensitive data for which one-way integrity verification code is generated block-wise to locate the areas of tampering. Integrity validation phase will compare the received digest and generated the digest from the received image. Blocks which are failed to pass in integrity validation will undergo the recovery process. The next contribution, by Zhang et al., proposes an effective and novel hybrid architecture, named Pixel-level Image Tampering Localization Architecture (PITLArc), which integrates the advantages of top-down detection-based methods and bottom-up segmentation-based methods. The authors present a typical fusion implementation of the proposed hybrid architecture on one outstanding detection-based method. Here, three methods are integrated resulting the contribution PITLArc to significantly improve their performance. A Dense Conditional Random Fields DenseCRFs-based post-processing method is introduced to further optimize the details of tampered regions. The fifth paper, by Abdel Raouf, proposes a new data hiding approach for image steganography based on the human visual properties using adaptive Least Significant Bits (LSB). Firstly, the author hypothesises that, the human eye has different sensitivity to RGB colour channels which permits different number of bits for every colour channel. Secondly, photos focus normally on their middle zone which permits hiding the secret message using a spiral way starting from the images' edges towards its centre. Both methods are used to enhance the visual appearance of

the stego image using the simple LSB replacement approach. This approach enables hiding bigger secret message with less real visual effect/distortion. The next contribution, by Muzaffer et al., proposes new copy-move forgery detection and localization technique independent from the characteristics of the forged regions. SIFT keypoints are obtained from CLAHE applied sub-images extracted from the input image by using RGB and L*a*b* color-spaces. Keypoint matching is realized on the sub-images and duplicated regions are determined to create a roughly marked image R. RANSAC is also applied in this stage and the generated homography matrix is used to construct transformed roughly marked image R ′. The method extracts DCT-based features from R and R ′ to localize exact borders of the tampered regions on the roughly determined areas by using a dynamic threshold. The proposed method has a new approach to determine the threshold dynamically. Tamper localization procedure also utilizes from morphological operations (chosen depending on the characteristic of the image) and Connected Component Labeling to determine exact forge boundaries. In their paper, Altay and Ulutaş present a robust Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD) based technique for copyright protection. The cover image is firstly decomposed into sub-bands by DWT. Low frequency sub-band is then divided into non-overlapping blocks. Blocks where watermark will be embedded are selected depending on their standard deviation values. Selected blocks are transformed to U, S, and V matrices by SVD. Watermark is embedded into the second row value in the first column of U component obtained by SVD. Embedding scaling factor is determined by Self-Adaptive Step Firefly Algorithm (SASFA) to balance robustness and transparency that are contradictory to each other. Firefly Algorithm (FA) is a simple, easy to implement, and flexible algorithm but it can pass over the global optimum or get trapped at local optimum. Therefore, SASFA, which constitutes the next step of each firefly based on its previous and present situations is used for global exploration of the solution space. Fibonacci-Lucas Transform (FLT) is applied to binary watermark to provide the security of watermarking scheme.

With the same topic (i), three papers addressing image-based image encryption were selected. Osorio and Moreno present the application of Brauer configuration algebras for encrypting multimedia data. The author contribution enables effective concealment of multiple secrets between a number of trusted parties. The proposal allows for the master shares to be encoded by sequences of integer numbers, thus saving storage space of the database servers involved in the protection process. According to the procedure, a master share is obtained via a digit sequencing algorithm (DSA), which define elements of a database as linear combinations of basic elements of a suitable Brauer configuration algebra (BCA). In the next paper, presented by Manikandan and Rengarajan, an algorithm suitable for encrypting DICOM and other types of images from medical domain is presented. RC6 cipher is used for encrypting the approximation coefficients (LL), obtained by applying the Haar wavelet transform on the plain image and combined with the redistributed (confused) detailed coefficients (LH, HL, HH). Generation of keys through governing equations of Combined Logistical Tent map, adds to the robustness of the algorithm against attacks. This algorithm works well for all types of images, including DICOM. The last paper in image encryption and watermarking is a review article presented by Li et al., survey typical reversible data hiding (RDH) algorithms that use PEE technology in combination: Conventional Prediction-error-expansion(C-PEE), Adaptive Prediction-error-expansion(A-PEE), Pairwise Prediction-error-expansion (P-PEE), Improve Pairwise Prediction-error-expansion(IP-PEE). A detailed experimental comparison results were presented that indicate the performance of these four different types of RDH algorithms. In the analysis of the results, authors propose a new performance measurement method based

on the PSNR and EC measurement, which is referred to as prediction error accuracy rate. The rate is determined by the ratio of the capacity that meets the embedding conditions to the total capacity of the image, combined with the measurement of this method, can better determine whether the performance of the algorithm.

This special issue contains two papers addressing the challenges in detecting tamper in audio sequences. In the first paper, Hemavathi and Kumaraswamy propose a countermeasure which can detect voice conversion spoofed speech based on artifact estimates obtained using source separation. The countermeasures extract the features from raw speech signals to differentiate between the natural and spoofed signals. Voice conversion algorithms mainly focus on transforming the spectral content of source to that of target. Hence, a lot of similarity is observed in target and spoofed speech. In the author contribution, instead of processing the signals directly, the speech is pre-processed using speech separation block where the artifacts introduced during the voice conversion is separated. Binary classifier based on CNN is built to classify the Time-Frequency representation of artifact estimate as natural or spoof signal. In the second paper, Verma and Khanna propose a system for speaker-independent cell-phone identification from recorded audio. This system is capable of dealing with test audio with different speech content and a different speaker compared to the training audio. Each recorded audio has the device fingerprint implicitly embedded in it, which encourages us to design a CNN-based system for learning the device-specific signatures directly from the magnitude of discrete Fourier transform of the audio. The paper contribution also addresses the scenario where the recorded audio is re-compressed due to efficient storage and network transmission requirements, which is a common phenomenon in this age of social media. The scenario of the cell-phone classification from the audio recordings in the presence of additive white Gaussian noise is addressed as well.

In the field of digital forensics two papers have been selected. In the first paper, by Agarwal and Bansal, the usefulness of pores (level 3 features) besides minutiae in latent fingerprint matching is examined. An algorithm based on Lindeberg's automatic scale selection method is proposed for pores extraction in latent fingerprints. The fusion of pores and minutiae at score level is used to re-rank the minutiae based latent matcher. The effectiveness of the proposed algorithm and pores utility are evaluated by observing and comparing the latent recognition accuracy obtained for minutiae matching and matching after fusion. Both minutiae and pores are automatically extracted in latent and reference fingerprints. In the second paper, by Wati et al., the authors contribution includes the use of Viola-Jones for face detection with Gabor Wavelet feature extraction and template matching for facial recognition. The face recognition algorithm from the authors has been evaluated in two real-world example scenarios namely (i) individual face recording and (ii) group recording, in which several faces were simultaneously captured.

In the fourth thematic cluster addressing privacy-aware technologies and social implications of biometric systems two additional papers are included. The first paper, by Climent-Perez and Florez-Revuelta, presents an RGB-only based visual privacy preservation filter, which capitalises on 'deep learning'-based segmentation and pose detectors. A background update scheme is presented, which limits leakage of sensitive information when detection fails. Dilation of the mask can further prevent information leakage, but a trade-off is necessary to correctly update background information. This is achieved via a specific study which is also presented. A comparative study is performed to determine the best configuration for privacy preservation. The second paper, by Binder et al., presents an overview of social implications as researched within the European project PERSONA. The author contributions include the

procedures needed for the assessment of their social, ethical, privacy and regulatory acceptance, particularly in view of the impact on both, the passengers crossing international territory and the border control authorities responsible for protecting the national borders. The paper presents a formal assessment of biometric technologies for real-world acceptance to cope with the increasing demand of global travellers crossing state borders.

The special issue closes with three selected papers within the broad thematic cluster of criminal investigation for crimes against property, person and critical infrastructures. In the first paper, Perez et al. introduce the research outcome from European research project MAGNETO. The paper summarises the modular approach adopted for the processing of information gathered from different information sources, and the extraction of knowledge to assist criminal investigation. The proposed platform provides novel technologies and efficient components for processing multimedia information in a scalable and distributed way, allowing Law Enforcement Agencies to make the analysis and a multidimensional visualization of criminal information in a single and secure point. In the second paper, Gomez-Silva et al. present a Deep Multi-Shot neural model for measuring the Degree of Appearance Similarity (MS-DoAS) between person observations. This model provides temporal consistency to the individuals' appearance representation and provides an affinity metric to perform frame-by-frame data association, allowing online tracking. The model has been deliberately trained to be able to manage the presence of previous identity switches and missed observations in the handled tracks. With that purpose, a novel data generation tool has been designed to create training tracklets that simulate such situations. The contribution aides the criminal investigation in identifying suspects fleeing the scene of crime in large crowded environments. The last paper, by Zhang and Kusrini, proposes the design and validation of a situation awareness component which is interfaced with the hardware component for controlling the focal length of the camera to protect critical infrastructures against UAV attacks. The continuous stream of media data obtained from the region of vulnerability is processed using the deep learning based object detector. The real-time computational efficiency has been validated during the pilot trials carried out within the European research project framework DEFENDER.

This Special Issue has assembled papers originating from well-known internationally reputed research institutions. The contributing authors were instrumental in the completion of the special issue and we would like to thank all of them. The anonymous referees also played a crucial role in the review and selection process ensuring the special issue includes only the submissions of the highest technical quality. Finally, we would like to thank the MTAP editorial team, especially, the Editor in Chief, Prof. B. Furht, for their help and very efficient handling of this special issue.