ICDP-2019

Abstracts/Notes

UNIVERSITY OF WESTMINSTER

The Institution of Engineering and Technology

sensors
an Open Access Journal by MDPI

VICTORIA
VIDEO ANALYSIS FOR INVESTIGATION OF CRIMINAL AND TERRORIST ACTIVITIES

BMVA

IEEE
Advancing Technology for Humanity

UKIVA
UK Industrial Vision Association

Day 1 **16th December**

14:00-18:00 Registration (115 New Cavendish)

14:15-14:30 **Welcome by ICDP General Chair and Local Chairs**

14:30-15:30 **Invited Talk: Bridging research in AI in academia and Industry. Dr Eduard Vazquez**, Research Technical Manager, AnyVision, UK. (Chair: Prof. Sergio A Velastin)

It's not exactly a secret that a dramatic change has occurred in the AI landscape. A change that has reinvented the relationship between academia and industry in the field, even drawing a question mark regarding the very same concept of AI. I'm going to explain my views on how the binomial academia-industry could work but also, and equally important, I will extend this relationship to a third concept, namely, real adoption of Computer Vision products in the real world. Research isn't happening any more just in the academia, but also in the industry. The borders between these two, previously well-separated worlds, are fading away. Industry is publishing papers in top conferences and academia is claiming benefits from the commercial applications. But such -whether good or bad- change cannot be understood without the incredible amount of investment that is being thrown into a plethora of new start-ups and, often times, questionable products.

Perhaps the two main questions are:

1. Is there a real, direct relation between the initial research occurring in a university and a final product with commercial success, in such manner that academia should claim benefits?
2. What is the effect that massive growth on investment is having in our field and the quality of future research and products?

In this talk, I'll try to shed some light into these two critical questions.

15:30-17:10 **Track 1**: Surveillance Systems and Solutions (Chair: Prof. Sergio A Velastin)

15:30-15:55 Behavior Detection as a Privacy-Enhancing Security Technology in Prison Cells, *Christopher Pramerdorfer (cogvis, Vienna, Austria), Martin Kampel (TU Wien, Vienna, Austria), and Reinhard Kreissl (VICESSE, Vienna, Austria)*

The everyday life of a prison is not only shaped by the objective of resocialisation but also, as a rule, by aspects of security, which have the highest priority. A particular challenge is the prompt detection of potentially dangerous behavior. In the face of scarce human resources and psychologically limited attention, the use of video live streams is not an optimal solution. An alternative are 3D sensors, which enable the automatic and robust detection of such behavior in detention rooms and other areas such as hospital wards or workshops, while respecting the privacy of the monitored people. In this paper we present our research on this matter, based on realistic data that was acquired in an Austrian prison over 3.5 months. We discuss the recording setup and resulting dataset, and present algorithms for detecting selected behaviors. The experimental results show that these behaviors can be detected reliably, demonstrating that automatic behavior analysis in 3D data is a promising means for supporting the security personnel.

15:55-16:20 Illegally Parked Vehicle Detection using Deep Learning and Key-point Tracking, *X. Gao, P. M. Birch, R. C. D. Young and C. R. Chatwin (University of Sussex, Brighton, UK)*

In this paper, we present a method for identifying and tracking illegally parked vehicles. This approach is based on deep learning for vehicles detection and hand crafted descriptors for the tracking which are designed to cope with occlusions. The tracking of the parked vehicle is achieved by key-point extraction of the detected vehicles and feature point matching. For each frame, a bounding box was generated to represent the vehicle and feature points extracted in that area. All parked vehicles have a unique ID which was generated by the Hungarian algorithm and Kalman filter, and the parked vehicle with the same ID was matched frame by frame. Based on this matching result, the stationary vehicles in the forbidden area can be tracked. Our approach tested efficiency and robustness on a public database and is shown to produce state of the art results.

16:20-16:45 Human Interaction Proofs (HIPs) Based on Multistable Images and Brauer Configuration Algebras (BCA), *M.A.O. Angarita (National University of Colombia), E. Izquierdo (Queen Mary University of London), and A.M. Canadas (National University of Colombia)*

Human Interactive Proofs (HIPs) are an important tool in e-Commerce, science and engineering, since they enable secure separation or identification of humans and computer bots. Although, many HIP techniques have been introduced and deployed in practical applications over the last two decades, the challenge of finding a waterproof method remains open. In this paper, we introduce a novel strategy to produce HIPs exploiting Brauer configuration algebras theory. The proposed approach is based on the generation of images of shapes that can be identified by humans but hard to be recognised by a computer program. Several experimental results are reported to demonstrate the robustness and feasibility of the proposed approach

16:45-17:10 Vitality Evaluation System for Street Architectures Based on Mobile OD Data, *Xianhan Zeng, Ning Wang, Zefei Gao, Zhicheng Liu, Junyan Yang and Qiao Wang (Southeast University, Nanjing, China)*

Urban vitality is closely related to the built environments such as street architectures in the city. We propose a vitality evaluation system which calculates the amount of human activities under the use of mobile phone signaling Origin-Destination (OD) data. Meanwhile, because of the strong correlation between urban safety and urban vitality, the potential security risks of street architectures of different categories at different moments in a day are able to be appraised by our method. Experiments on the database of the street view images and mobile OD data in Nanjing, China validate the practicality and accuracy of our method. Moreover, we study the distribution of Points of Interest (POI) nearby street architectures of different vitality. A series of heuristic conclusion which will help optimizing the planning of street facilities are drawn from our study.

17:10-18:30 Coffee break / Tech Award Demos

The IET's Vision and Imaging Network is giving an award to the best technology demo, Chair: Prof. Sergio A Velastin

Day 2 **17th December**

08:30-09:30 Registration at 115 New Cavendish Campus – University of Westminster

09:00-09:30 Continental Breakfast

09:30-09:45 **Welcome from Dr Peter Bonfield, Vice-Chancellor and President University of Westminster**

09:45-10:45 **Invited Talk: Body cameras: audio-visual learning and privacy-preserving analytics, Prof. Andrea Cavallaro,** Queen Mary University of London and Director of the Centre for Intelligent Sensing. (Chair: Dr Anastassia Angelopoulou)

The analysis of signals such as video, sound, and motion data from body cameras can help understand dynamic scenes, recognise interactions and classify physical activities. A body camera is indeed equipped with multiple sensors such as microphones and inertial measurement units, in addition to the imager. The analysis of these data is particularly challenging due to unconventional mounting and capturing conditions, motion blur, rapid changes in pose and various sources of noise. In this talk I will present the main challenges for learning, classification and processing signals from body cameras and show how multi-modal learning can be used to address these challenges. In particular, I will discuss applications such as person re-identification and action recognition, and present privacy-preserving classifiers for services that respect societal concerns arising from the use of these novel and pervasive sensing devices.

10:45-11:35 **Track 2: Robust Vision Algorithms** (Chair: Dr Anastasia Angelopoulou)

10:45-11:10 Robust Framework for Human Localization and Detection in Moving Train Carriage, *Muhammad Qasim Shafiq (University of Engineering and Technology Taxila, Pakistan), Saima Nazir (Fatima Jinnah Women University Rawalpindi, Pakistan), Muhammad Haroon Yousaf (University of Engineering and Technology Taxila, Pakistan), and Sergio A. Velastin (Queen Mary University of London, UK)*

Use of video surveillance cameras in public space is the recent solution to control vandalism acts and emergency incidents. Such type of incidents requires an urgent and an appropriate action by security personnel. Key problem for security personnel is to manage and monitor high volume of visual data. Since last decade, human detection for video surveillance systems is an emerging research area. There are several crucial factors that effect the performance of human/object detection, such as illumination changes, background clutter, dynamic background, occlusion, and camera orientation etc. In this paper, a hybrid approach is presented for the localization and detection of person inside a moving train with challenging environment. Our proposed framework contains two modules i.e. human localization and human detection. We proposed a hybrid approach using GMM background modeling for foreground extraction followed by head and face detection to be used as a clue for human detection. Along with head and face detection, Histogram of Oriented Gradient (HOG) feature representation is used for human localization. For detection, ensemble classifier outperforms SVM and KNN classifiers on BOSS Dataset (On Board Wireless Secure Video Surveillance) and achieved 90% accuracy.

11:10-11:35 Improving Text Recognition in Tor darknet with Rectification and Super-Resolution techniques, *Pablo Blanco-Medina, Eduardo Fidalgo , Enrique Alegre and Francisco Janez-Martin (Universidad de Leon, INCIBE -Spanish National Cybersecurity Institute, Spain)*

Text recognition can be used to retrieve textual information embedded in images. This task can be complex due to low-resolution and text orientation, which are problems commonly found in Tor darknet images. In this work, we evaluate the combination of three super-resolution algorithms, together with a rectification network, to enhance the performance of a text recognition algorithm. We evaluated these combinations in four state-of-the-art datasets, and in TOICO-1K, a Tor-based image dataset which was semi-automatically labelled for the task of Text Spotting in Tor darknet. We achieved the best performance increase in the ICDAR 2015 dataset, with an improvement of 3.77% when combining Residual Dense and the rectification networks. In TOICO-1K, we obtained a 3.41% of improvement when we combined Deep CNN and the rectification network. Our conclusion is that rectification performs slightly better than super-resolution when they are applied standalone, while their combination obtains the best results in the datasets evaluated.

11:35-12:00 Coffee Break

12:00-13:30 **The VICTORIA Project and Challenge**
VICTORIA (Video analysis for Investigation of Criminal and TerrORIst Activities) H2020 European Commission project. (Chair: Dr Anastasia Angelopoulou)

12:00-12:30 VICTORIA general overview, Xavier Mamy (IDEMIA)

12:30-13:30 VICTORIA live Demo, Ran Zhou (IDEMIA)

13:30-14:30 Lunch Break

14:30-15:30 **Invited Talk: Fake News Outbreak 2020: Can we Stop the Virus Spreading?, Dr. Antonios Michalas,** Department of Computing Sciences, Tampere University of Technology, Finland. (Chair: Prof. Dimitrios Makris)

Computer viruses have come a long way from the early days of personal computers when amateur hackers were designing simple software that could harm the proper functionality of a single machine. Nowadays, computer viruses have become more sophisticated and are able of exercising severe impacts aimed not only on large computer networks but social sphere of life as well (e.g. NHS WannaCry Attack). However, all viruses are driven by the very same principle: Self-replicating programs designed to spread itself from computer to computer. Nowadays, and with the intense use and spread of social media, many Internet users are facing the same challenge: Deciding whether to believe something on the Internet or not. Misinformation behaves like a virus. Fake news and false rumors reach more people, penetrate deeper into the social network, and can spread as fast as accurate stories. Even though misinformation, spin, lies and deceit have been around forever, in the past years it was easier to identify them since we were mainly reading news trough trusted and evaluated sources. However, making judgments about what to trust on the web can be really challenging. In this talk we will present a high-level overview of the problem of spreading fake news and its possible impact and will describe the base that needs to be built to successfully provide mechanisms that will reduce the spread of fake information.

15:30-17:10 **Track 3: Multicamera Surveillance** (Chair: Dr Epaminondas Kapetanios)

15:30-15:55 Multi-block fusion for vehicle re-identification, *Yunping Zhang and Krystian Mikolajczyk (Imperial College London, United Kingdom)*

The extensive coverage of surveillance camera networks has supported the ever-growing research of vehicle re-identification (re-ID) due to their significant applications in matching and tracking the vehicle-of-interest. The inherent challenging characteristics such as intra-class variance and inter-class similarity make such identification one of the most difficult tasks in computer vision. In this paper, we proposed a novel approach for vehicle re-id based on multi-block features. It implements the idea of information fusion from intermediate layers and multi-stage supervision into a fully convolutional neural network. We perform extensive experiments and analysis to demonstrate the effectiveness and superiority of our approach on two standard benchmarks.

15:55-16:20 Computer Vision meets Visual Analytics: Enabling 4D Crime Scene Investigation from Image and Video Data, *Thomas Pollok (Fraunhofer IOSB, Karlsruhe, Germany), Matthias Kraus (Universitat Konstanz, Konstanz, Germany), Chengchao Qu (Fraunhofer IOSB, Karlsruhe, Germany), Matthias Miller (Universitat Konstanz, Konstanz, Germany), Tobias Moritz (Fraunhofer IOSB, Karlsruhe, Germany), Timon Kilian , Daniel Keim and Wolfgang Jentner (Universitat Konstanz, Konstanz, Germany)*

In case of a crime or terrorist attack, nowadays much video footage is available from surveillance and mobile cameras recorded by witnesses. While immediate results can be crucial for the prevention of further incidents, the investigation of such events is typically very costly due to the human resources and time that are needed to process the mass data for an investigation. In this paper, we present an approach that creates a 4D reconstruction from mass data, which is a spatio-temporal reconstruction computed from all available images and video footage. The resulting 4D reconstruction gives investigators an intuitive overview of all camera locations and their viewing directions. It provides investigators the ability to view the original video or image footage at any specific point in time. Combined with an innovative 4D interface, our resulting 4D reconstruction enables investigators to view a crime scene in a way that is similar to watching a video where one can freely navigate in space and time. Furthermore, our approach augments the scene with automatic detections and their trajectories and enrich the crime scene with annotations serving as clues.

16:20-16:45 Domain Adversarial Training for Infrared-colour Person Re-Identification, *Nima Mohammadi Meshky, Sara Iodice and Krystian Mikolajczyk (Imperial College London, United Kingdom)*

Person re-identification (re-ID) is a very active area of research in computer vision, due to the role it plays in video surveillance. Currently, most methods only address the task of matching between color images. However, in poorly-lit environments, infrared images must be utilised which requires matching infrared to color images. In this paper, we propose a part-feature extraction network to better focus on subtle, unique signatures on the person which are visible across modalities. To train the model we propose a novel variant of the domain adversarial feature-learning framework. Through extensive experimentation, we show that our approach outperforms state-of-the-art methods.

16:45-17:10 Triplet Permutation Method for Deep Learning of Single-Shot Person Re-Identification, *M. J. Gomez-Silva, J.M. Armingol and A. de la Escalera (Universidad Carlos III de Madrid, Leganes, Madrid, Spain)*

Solving Single-Shot Person Re-Identification by training Deep Convolutional Neural Networks is a daunting challenge, due to the lack of training data, since only two images per person are available. This causes the overfitting of deep models, leading to degenerated performance. This paper formulates the Triplet Permutation method to generate multiple training sets, from a certain re-id dataset, by considering different constraints depending on the outcome that is sought to obtain. This is a novel strategy for feeding deep triplet networks to learn a single-shot re-identification model, which reduces overfitting. The improved performance of the model has been demonstrated over one of the most challenging re-identification datasets, PRID2011, proving the effectiveness of the method.

Day 3 18th December

08:30-09:30 Registration at 115 New Cavendish Campus – University of Westminster

09:00-09:30 Continental Breakfast

09:30-09:45 **Welcome by the IET's Visual and Imaging Network, Prof. Dimitrios Makris, IET**

09:45-11:50 **Track 4: Biometrics and Behaviour Analysis** (Chair: Prof. Dimitrios Makris)

09:45-10:10 Semi-automatic Progressive Enhancement for Latent Fingerprints, *Kittinuth Srisutheenon, Krisada Phromsuthirak, and Vutipong Areekul (Kasetsart University, Bangkok, Thailand)*

We propose a semi-automatic progressive enhancement method for latent fingerprints. This method requires three inputs; a latent image, a manual segmentation, and an initial block location. The method starts enhancing the initial block using a matched filter in the frequency domain, then the enhanced block is padded back on the input latent image. The padded image is fed back as an input image and the surrounding blocks of the enhanced block are progressively enhanced. The proposed method performs iterative enhancement and feedback until the enhanced blocks fill the entire segmentation. The proposed method is benchmarked against several state-of-the-art methods with the NIST SD27 database. The experimental results show that the proposed method outperforms other methods in terms of identification accuracy.

10:10-10:35 Spatial Signatures for EEG-based Biometric Person Recognition, *M. Al-Darkazali, Sanaul Hoque, and Farzin Deravi (University of Kent, Canterbury, UK)*

Biometric person recognition using EEG signals has received considerable attention in recent years. This paper proposes a new feature based on the co-activation of EEG sensors. A visual representation of this co-activation feature is used to illustrate the identity-bearing nature of the proposed feature. The DEAP database was used to evaluate the proposed feature which was presented in the form of a visual signature indicating the spatial correlations around the scalp of EEG signals for an individual. The results show a high identification accuracy irrespective of the emotional state of the data subjects.

10:35-11:00 Iris Image Recognition using Optimized Kohonen Self Organizing Neural Network, *J. Jenkin Winston, D. Jude Hemanth (Karunya Institute of Technology and Sciences, Coimbatore, India), A. Angelopoulou and E. Kapetanios (University of Westminster, London, UK)*

The pursuit to develop an effective people management system has widened over the years to manage the whopping increase in population. Any management system includes identification, verification and recognition stages. Iris recognition has become notable biometrics to support the management system due to its versatility and non-invasive approach. These systems help to identify the individual with the texture information distributed around the iris region. Many classification algorithms are available to help in iris recognition. But those are very sophisticated and require heavy computation. In this paper, an improved Kohonen self-organizing neural network (KSONN) is used to boost the performance of existing KSONN. This improvement is brought by the introduction of optimization technique into the learning phase of the KSONN. The proposed method shows improved accuracy of the recognition. Moreover, it also reduces the iterations required to train the network. From the experimental results, it is observed that the proposed method achieves a maximum accuracy of 98% in 85 iterations.

11:00-11:25 Biometric Presentation Attack Detection Using Stimulated Pupillary Movements, *Asad Ali, Sanaul Hoque, and Farzin Deravi (University of Kent, Canterbury, UK)*

Biometric systems can be subverted using presentation attack artefacts. This work presents a way to deal with the vulnerability to such spoofing attacks. In this work we propose the use of pupillary movements to detect such presentation attacks. The pupillary movements were stimulated by presentation of a moving visual challenge to ensure that some pupillary motion can be captured. Photo, 2D mask and 3D mask attack artefacts were evaluated based on data captured from 80 volunteers performing genuine attempts and spoofing attempts. The results indicate the effectiveness of the proposed pupillary movement feature to stop presentation attacks.

11:25-11:50 Feature Extraction Techniques for Human Emotion Identification from Face Images, *M. Kalpana Chowdary, D. Jude Hemanth (Karunya Institute of Technology and Sciences, Coimbatore, India), A. Angelopoulou and E. Kapetanios (University of Westminster, London, UK)*

Emotion recognition has been one of the stimulating issues over the years due to the irregularities in the complexity of models and unpredictability between expression categories. So many Emotion detection algorithms has developed in the last two decades and still facing problems in accuracy, complexity and real-world implementation. In this paper we propose two feature extraction techniques: Mouth region based feature extraction and Maximally Stable Extremal Regions (MSER) method. In Mouth based feature extraction method mouth area is calculated and based on that value the emotions are classified. In MSER method the features are extracted by using connecting components and then the extracted features are given to a simple ANN for classification.

11:50-12:10 Coffee Break

12:10-13:10 **Invited Talk: Ethics Standards for Autonomous and Intelligent Systems – building the foundation for trust, Dr Ansgar Koene,** University of Nottingham, United Kingdom. (Chair: Dr Epaminondas Kapetanios)

Automated image analysis has great potential for a wide range of policing and crime prevention applications ranging from the use of face recognition technology to rapidly scan for criminal suspects in CCTV footage, to real-time detection of violent behaviour patterns, analysis of satellite images for patterns associated with modern-day slavery and many more. The use of these technologies however requires high levels of public trust in the ethical oversight processes involved in their development and operation. As shown by the recent wave of legal challenges to the use of face recognition technology, this public trust cannot be taken for granted. One route towards establishing trust is through the use of standardized ethics-based development processes that provide clear evidence suitable for third-party verification. In this talk I will introduce the IEEE P70xx series of ethics-based Standards for Autonomous and Intelligent Systems, with an in-depth review of the IEEE P7003 Standard for Algorithmic Bias Considerations. By walking through a hypothetical example of using the IEEE P7003 Standard during development of a automated face recognition system it will be shown how the use of this Standard can contribute to providing a foundation for public trust in Autonomous and Intelligent Systems.

13:10-14:30 Lunch Break

14:30-15:45 **Track 5: Forensics and crime scene reconstruction (**Chair: Dr Sophie Triantaphillidou**)**

14:30-14:55 CPU vs GPU performance of deep learning based face detectors using resized images in forensic applications, *D. Chaves, E. Fidalgo, E. Alegre, F. Janez-Martin and J. Velasco-Mata (Universidad de Leon, Espana, INCIBE -Spanish National Cybersecurity Institute, Leon, Spain)*

Accurate and fast face detection is a crucial step in forensic applications such as surveillance, facial fugitives recognition, and child sexual abuse detection. Several deep-learning-based methods addressed the face detection problem with high accuracy but they require a large computation power and processing time. GPUs have been used to speed-up computations, however, nowadays there are available multiple generations of GPU architectures (e.g. Turing, Pascal, Keppler) making difficult to choose the most appropriate for face detection. In this work, we evaluated the speed-accuracy trade-off of three deep-learning-based face detectors in various CPUs/GPUs considering images reduced to different proportions as input. We successfully used this image resizing strategy in one previous work to improve the performance of face detection with Child Sexual Exploitation Material (CSEM). The results showed that the best speed-accuracy trade-off is achieved using the Pascal and Turing GPUs with images reduced to 50% of its original size.

14:55-15:20 impress: A Forensic Footwear Impression Dataset, *Manuel Keglevic (TU Wien, Austria), Silvia Wilhelm (Bundeskriminalamt, Federal Ministry of the Interior, Austria), and Robert Sablatnig (TU Wien, Austria)*

Footwear impressions are commonly found at crime scenes and are therefore a valuable source of evidence for criminal investigations. Forensic experts can show that a footwear impression was made by a specific shoe or impressions at different crime scenes were made by the same suspect by comparing individual characteristics of the impressions. However, this process is very time consuming and therefore automated solutions are desired. Yet, testing and training such methods requires datasets that on the one hand reflect real data from criminal cases and on the other hand provide ground truth information. To solve this, we created an acquisition line and captured footwear impressions of 300 different pairs of shoes under varying conditions with the help of the Austrian Police. In this work the creation of this dataset and the dataset itself is described in detail.

15:20-15:45 Drop impact dynamics of bloodstain on fabric, *Wilbur H. Galarion Jr., and Maricor N. Soriano (University of the Philippines, Diliman, Quezon City, Philippines)*

An investigation on the spreading of blood droplets in fabric, as well as finger formation around the rim of the spreading stain immediately after droplet impact was done. Blood droplets were release perpendicular to the fabric substrate, at varying impact velocities (0.44-4.4 m/s), and different substrate backing material (air, leather, aluminum). A camera was used to record the impact spreading, and formation of fingers. The maximum spreading diameter and the number of fingers formed scaled with the Reynolds and Weber number, respectively. Both increased with increasing impact velocity, but slightly decreased with harder backing material. It was also observed, that the number of fingers approaches a constant maximum value so that it remains constant even if impact velocity is further increased, deviating slightly from predictive models.

15:45-16:00 **Sensors Journal Best Conference Paper Prize Ceremony**

16:00-16:30 Panel Discussion

16:30-16:45 Concluding remarks from ICDP General Chair