White Paper on Industrial Applications of Computer Vision and Pattern Recognition

Giovanni Garibotto[1], Pierpaolo Murrieri[2], Alessandro Capra[3], Stefano De Muro[4], Ugo Petillo[4], Francesco Flammini[5], Mariana Esposito[5,6], Concetta Pragliola[5], Giuseppe Di Leo[7], Roald Lengu[5], Nadia Mazzino[5], Alfredo Paolillo[7], Michele D'Urso[2], Raffaele Vertucci[2], Fabio Narducci[8], Stefano Ricciardi[8], Andrea Casanova[9], Gianni Fenu[9], Marco De Mizio[10], Mario Savastano[11], Michele Di Capua[12], and Alessio Ferone[13]

[1] Independent Consultant
giovanni.garibotto@gmail.com
[2] Selex ES, Italy
[3] STMicroelectronics, Catania, Italy
[4] RFI, Rete Ferroviaria Italiana, Roma, Italy
[5] Ansaldo STS, Italy
[6] University of Naples, Federico II, Napoli, Italy
[7] Dept. of Industrial Engineering, University of Salerno, Italy
[8] Biometric and Image Proc. Lab., University of Salerno, Italy
[9] University of Cagliari, Italy
[10] CIRA Italian Aerospace Research Centre, Capua, Italy
[11] CNR, Napoli, Italy
[12] Unlimited Software srl, Naples, Italy
[13] Dept. of Science and Technology, University of Naples Parthenope, Italy

Abstract. The paper provides a summary of the contributions to the industrial session at ICIAP2013, describing a few practical applications of Video Analysis, in the Surveillance and Security field. The session has been organized to stimulate an open discussion within the scientific community of CVPR on new emerging research areas which deserve particular attention, and may contribute to the improvement of industrial applications in the near future.

1 Introduction

Computer Vision and Pattern Recognition play an ever increasing role in all the industrial applications because of the growing amount of imaging data and the need to support an autonomous intelligent management of them. Beside professional video recording and processing there is a continuous growing of digital imaging applications for individual and private users. Hence, the spread of applications of CVPR technology is also increasing, as well as the requirements for intelligence and sophisticated solutions. A short summary of some session lectures is briefly referred in this paper, to highlight the main innovation and research objectives experienced in some different fields.

A. Petrosino (Ed.): ICIAP 2013, Part II, LNCS 8157, pp. 721–730, 2013.
© Springer-Verlag Berlin Heidelberg 2013

The first part of the session is devoted to fundamental issues of Digital Imaging and Processing, considering multi-camera applications with distributed architectures and embedded solutions.

A second group of contributions are oriented to video analysis for security in the railway domain. Classical video surveillance applications and requirements analysis are described and discussed, including performance evaluation tools and methods. Quality control issues are also considered, like the analysis of the train profile, to achieve a high level of security for the railway service (both passengers and goods transportation).

Another topic of interest is the study of biometric solutions for homeland security and border (ports, airports and police investigations). The discussion is centered on the concrete opportunities for future development and achievements in this highly competitive domain, including normative and legislation constraints.

Finally we draw some conclusions to highlight the most relevant open issues, as perceived by the session speakers, and outline the main challenging objectives of applied research, looking for contribution and support by the whole scientific community in CVPR.

2 Context-Aware Imaging: The PANORAMA Project

The exponential increase of digital imaging in real-life applications has determined a shift from video recording and playback, towards highly sophisticated image processing and scene interpretation. This shift affects all the image processing chain, from *acquisition modules*, up to *video processing and reproduction*.

In most application domains there is a clear trend from "single-view imaging" to "multi-view imaging". In addition, acquisition devices need to become more *context-aware*, to better exploit all the properties of the current applications; they should also be aware of the existence of other cameras and sensors to achieve a full control and monitoring of the inspected scene.

The international project PANORAMA [1] is focused to the research and development of new video imaging technologies for a broad range of applications. Main objectives are:

- Ultra Wide Imaging, to achieve a wider monitoring and surveillance of the scene:
  - □ Integration of Multiple imaging sources and images from different angles
  - □ Multi modalities
  - □ Inter camera calibration
  - □ Algorithms to merge many views to a panorama or 3D representation
- Content Aware Imaging:
  - □ More intelligent image acquisition about which data to acquire (e.g. automatic ROI)
  - □ How to *optimally image* the material of interest by video content analysis.
- Context Aware Video Processing
  - □ Cameras become aware of higher-level user needs and the existence of other cameras

- ☐ Jointly optimized imaging in multi-camera systems.
- ☐ Combine multiple video streams in a more useful and drastically simplified representation.

Particular attention is devoted to real time processing of large amounts of image data with the development of new CMOS image sensor silicon supporting autonomous image acquisition. In all the application domains the user will be supported by a higher level of intelligence of the imaging system, so that she/he could concentrate more attention to the primary task objectives of the specific application.

The Panorama Consortium of research consists of 17 partners from 5 member countries, in 3 main application areas: Healthcare, Broadcasting services and Surveillance and Security. Italian partners are STMicroelectronics (one of the main semiconductor companies in the world) and University of Catania (Dep. of Mathematics and Computer Science).

Fig. 1. The PANORAMA project has been co-funded by grants from Belgium, Italy, France, the Netherlands, and the United Kingdom, and the ENIAC Joint Undertaking

3 Video Surveillance in the Railway Domain

The railway infrastructure is an open system, with physical assets which are mostly accessible at any time by anyone, like railway stations, rail yards, electrical stations, tunnels, viaducts, etc. As such it is naturally exposed to possible illegal actions like damages, robberies, vandalism, up to terroristic attacks, for many different reasons:

- Presence of a large number of passengers, often concentrated within a limited space, as a railway station;
- Large complexity and sparse geographical distribution of the whole system, with increased difficulty to guarantee a high protection level everywhere;
- Significant economical losses and relevant subjective impact (as perceived by the public opinion) by any possible criminal event or action happening in the railway transportation system.

For such reasons the Italian Railway Network (RFI), as managing authority of the National railway infrastructure, has identified the most sensible and critical sites of the railway system. Special interventions have been planned, and are in progress [2], to effectively increase the actual level of security of such sites, as well as to improve the feeling of the passengers and users (perceived security).

Video Surveillance technology represents one of the most important solutions to that problem, to provide a direct on-line monitoring of the inspected site. Moreover, they are frequently used in post-event analysis, to support police investigation after a crime or any other illegal action.

The strength and validity of Video Surveillance and Recording is highly improved by automatic processing (Intelligent Video Content Analysis VCA). Such Computer Vision technology is supposed to detect, in fully autonomous way, all abnormal events (within predefined categories) and produce an alarm message to help surveillance operators (acceptance/rejection of the alarm and activation of security interventions).

Some effort towards a standardization is in progress in the Transportation domain [3], aiming to define most common requirements for VCA functions, as well as basic criteria for performance evaluation Eight VCA alarms are considered as more representative in the Transportation field: Detection of Objects (and humans), Detection of unattended objects and baggage, Directional Analysis (direction not permitted), Removed Stationary Objects, Excessive velocity (object exceeding a specific speed), Loitering (long time presence of a person in a certain area), Crowding (space occupancy), People counting (people waiting in line or on a platform)

From the practical experience in the field, VCA has proved to be sufficiently mature and consolidated to address some visual tasks like:

- The presence of non-authorized persons in certain areas, that should be normally empty (and/or within pre-defined time-intervals)
- Virtual line crossing, like the yellow line on the railway platform, near the rail border,
- Face detection and blurring (ethical and privacy constraints)

There are other visual tasks which exhibit less accuracy and robustness in practical applications, especially when dealing with crowded scenes, like:

- Detection of abandoned baggage,
- Graffiti,
- Camera obstruction or tampering

The effectiveness of video analysis is always bound by the level of complexity of the scene. The railway environment is quite complex (context) and strongly influenced by variable operating conditions, which makes it difficult to compare the expected results from different industrial products of VCA from the market.

False-positive rate, FAR (False Alarm Rate), is a measure of the nuisance generated by a VCA system, (irrelevant alarm messages without any real threat); it is usually required to keep it very low (≤ 2 /day), to avoid overload the surveillance operator. On the other hand, also false-negative rate (real threats that are missed by the system, which is complementary to the POD as Probability of Detection), is a measure of reliability and must remain quite low (usually ≤ 5 %). The two requirements are in conflict, since a reduction of the first inevitably determines an increment of the second.

The development of a stable, objective and repeatable procedure for performance evaluation represents an important objective for all institutions and service departments in the transportation field, as well as in other domains.

In spite of the efforts in the recent years, there is not yet an agreement on a commonly accepted methodology, to identify the most appropriate KPI (Key Performance

Indicators) to be used in the experimental tests. In such procedure it is necessary to consider a number of factors, affecting system performance, like detection ranges, 3D shape and size of the objects, direction and speed of targets, objects' color and contrast with the background, environment light, etc.

Moreover, modern installations often consist in hundreds of cameras installed in the field, so that it is necessary to prove the correct coverage of the monitored space, and possibly provide the actual position (on the map layout) of the detected alarm event (beside its identification in the camera view).

It is worth to remark the importance to achieve VCA performance measure both on critical-alarm events (which are often artificial and actor-driven), as well as on "normal" routine operation. In these conditions the VCA system must prove its reliability, and robustness, not only to minimize FAR, but also to properly detect and classify some regular events (arrival and departure of a train, level of crowding, empty space, loading-unloading operations, etc.).

The use of video-recorded sequences in the different operating conditions is a common approach, widely used to compare both commercial products and research algorithms [4], with the availability of a large set of public data-bases (with ground-truth annotation). The best approach consists in the selection of a reference open technology to evaluate single significant algorithmic tasks (like object tracking, human and face detection, etc.). The OpenCV library [5] represents a natural (and highly qualified) candidate to achieve significant comparative results (against a commercial product).

The role of the scientific community of CVPR may become even more relevant for the achievement of a standard evaluation procedure for industrial applications, due to its independence position and the deep knowledge on the technology (current status and future potential).

4 Performance Evaluation of a Commercial VCA Solution

Automatic Video Content Analysis (VCA) is widely adopted to overcome the difficulty of real-time visual inspection of multiple video sources. However, in real environments where a high number of cameras are installed, the use of VCA may become critical in terms of false alarm rate, FAR. Major problems are due to physical phenomena like waiving trees, daylight change, camouflage, shadows, etc. To achieve a reliable VCA performance measure it is necessary to reproduce a testing environment as close as possible to the real infrastructure where the system is supposed to operate.

An experimental study has been carried out by the team of Ansaldo STS and the University of Naples [6] to evaluate the performance of a Commercial Off-The-Shelf (COTS) VCA system using different metrics both frame-based and object-based. The evaluation has been performed in a real on-board surveillance environment, i.e. on a vehicle used for rail-based mass-transit, by using some specific tools developed in Matlab. The evaluation methodology is based on a manually-generated Ground Truth (GT) and the analysis of the Algorithm Results (AR) automatically produced by the VCA software. The results obtained by the COTS software have been compared with the ones provided by an open-source system (i.e. i-SPY), to establish a reference metric for the analysis.

The alarm functions considered by the COTS VCA software are the following:

- **Cover:** it detects camera occlusion/blurring with paint-strokes or occluding objects.
- **Tamper:** it detects any move or relocation effects on the camera.
- **Stop:** it detects the presence of still objects in the scene.
- **Presence**: it detects objects/people moving in the scene.
- **Crowd:** it raises an alarm when the scene is overcrowded.

The main objective of the study has been robustness assessment of VCA algorithms to environmental variations as well as their overall reliability and time of response.

The evaluation has been concentrated on a few basic computer vision functions like object recognition, to verify that COTS solution features higher reliability (i.e. low number of false alarms). Another important measure has been the fragmentation rate of single blobs. Moreover, performance testing has been made also on high-level alarms associated to event recognition on board of the train-vehicle, to achieve a *black-box* performance measure in terms of both probability of detection and false alarm rate. An automatic tool developed in Matlab has been used to compute the selected performance indices quickly and reliably. The obtained results demonstrated a higher performance of the COTS system in all test conditions.

The proposed performance evaluation method is fairly general and applies equally well to other vehicles and testing environments. Moreover, the developed tool in Matlab can be easily extended to other relevant metrics (e.g. Track Matching Error, Latency of System Track, Track Completeness, etc.).

5 3D Shape Analysis in Railway Transportation Security

The aim of this study is a Video Monitoring System, developed by Ansaldo STS with the support of University of Salerno [7] to perform automatic defect detection on transiting rolling stocks. The main features and challenges of this research can be summarized as:

- use of line scan cameras acquiring high speed moving objects (train vehicles in transit), which present a different image projection system and several image quality problems compared to standard area scan cameras acquiring still objects;
- an uncontrolled illuminated environment, raising image acquisition quality problems;
- the definition of a pattern recognition and defect detection approach on rolling stock specific components.

The obtained results [6] are quite satisfactory and demonstrate the effectiveness of the proposed approach, with some additional remarks regarding future trends and possible evolution of the technology.

6 Advanced Sensors in Defense Systems and Logistics Applications

The ability to 'see' inside not accessible regions by using electromagnetic waves is an issue of interest in several applications. Hostage rescue, building surveillance, building clearance, and building search operations are a difficult challenge for military forces and law enforcement because of reduced situational awareness [8].

This contribution is aimed to refer some of the pattern recognition and image processing techniques exploited by Selex ES, to enhance performance and functionality of complex radar systems and to provide innovative and effective ways to perform augmented reality based on-field training and maintenance of mission critical systems.

The scenario complexity pushes towards new architectural and technological solutions. In this contribution a very compact multifunctional sensor able to localize and tracking humans inside critical areas is presented. The sensor combines innovative vital signs detection algorithms with human movement detection approaches. Typical applications of this sensor include:

- Police - Determine the presence and location of assailants or hostages in a building;
- Search & Rescue - Locate injured people inside buildings;
- Fire-fighters - Quickly determine whether people are trapped in a building.

The sensor exploits stepped frequency continuous wave radar technology and state-of-art detection algorithms enable the sensor to operate as a sensitive Doppler motion detector. Advanced packaging design results in the smallest, lightest through wall sensor available today. With immediate analysis and results, the proposed sensor provides first responders with critical information that may make the difference between life and death. In this contribution we will describe the developed architectural solution, and refer recent results confirming the actual possibility to detect and track humans beyond walls for homeland protection applications.

With regard to advanced training and maintenance of radar systems, the purpose of the research conducted by Selex ES is aimed to the development of an AR (Augmented Reality) platform specific to facilitate items localization in technical equipment by means of context-dependent visual aids, featuring a contactless interaction paradigm based on gesture recognition. One of the more compelling challenge faced by AR systems applied to industrial environments lies in highly accurate tracking and virtual-to-real co-registration. The approach adopted relies on redundant marker-based tracking, which provides a flexible and easily scalable solution to compute user's point-of-view and the associated operative volume. More specifically, we adopt a multi-marker technique to make the system more robust and fault-tolerant while keeping it simple and reliable in terms of the involved equipment. The goal is on the one hand to raise the probability that at least one single marker is correctly detected and recognized at any time, on the other hand to reduce the overall tracking error by means of a weighted average of each single marker's position/orientation. User-machine interaction is based on multiple finger detection and gesture recognition and exploits the same video stream used for user tracking to free the user from the need of any tangible I/O device. The solution developed exploits coloured fingertips tags to achieve a not instrumented two-hand finger based interface requiring only small rubber caps of different colours worn over user's fingertips.

Similarly, the objective of the CPILOS project [11] is to develop a technological platform, based on IT infrastructures and services, that can support critical processes, like secure tracking and tracing of goods in transport and logistics, using video surveillance facilities with radio-frequency identification (RFID) support.

7 Mobile Video Sensor in Surveillance Applications

Most Video Analysis applications are based on fixed camera systems, by exploiting background estimation over a long video sequence analysis. Anyway, it is worth to remind also the relevant sector of mobile CV applications, with on-board video sensors installed on a variety of mobile systems, like patrol cars (i.e. for Autonomous Number Plate Recognition, ANPR), PTZ units (human target detection and tracking) or unmanned vehicles for remote monitoring and inspection.

Contribution [9] addresses the topic of motion estimation for Unmanned Aerial Vehicles (UAV), in surveillance tasks, for both military and civil applications such as the remote monitoring of hostile area, coasts and boundaries surveillance, disaster management (earthquakes, flooding, fires, etc.). In such applications high definition images are needed. Furthermore, an UAV video coding system should meet both bandwidth (Behind Line-Of-Sight, BLOS, transmission) and computational (battery life) constraints.

The proposed motion estimation scheme makes use of global motion information provided by the onboard navigation system (using external position and orientation sensors). The observed scene is supposed to be approximated with a plane. Based on this simple hypothesis, a prediction of the motion field is obtained calculating the homography between the reference and the predicted frame. The obtained results are considered to be relevant in low-frame-rate video coding, for mobile devices, to achieve a significant reduction of computation and increasing battery life.

8 Biometrics in Security and Application Trends

On the basis of his long experience in the field [9], M. Savastano underlines the disillusion about the perspective of the market perceived in many different application domains, as a consequence of the economical crisis, still affecting the entire industrial world. Even with some distinctions, the crisis is global and consequently invests also the context of the biometric technologies which, may be more than others, are probably suffering the delicate transaction from the stage of experimental technological innovations to concrete commercial opportunity.

Anyway, a very promising area in biometrics is actually represented by the Automated Border Crossing (ABC) systems. They offer a service available to registered users or also to the holders of the new electronic passports for crossing automatically the borders and avoiding the queues which generally characterize the manual control of the travel documents. The systems, whose core is a biometric system which performs a comparison between a biometrics characteristic (fingerprint, iris, face,…) acquired by the traveler and a biometric reference, often stored in the travel document, are available in several international airports (through interactive kiosks) and the opinion of the experts is that the trend for ABC systems will be extremely positive in the coming years.

The deployment of ABC systems has been preceded by a series of preliminary experimental projects which have clearly defined the pros and cons of this kind of security approach. And, in Italy, the ABC systems are still in a very preliminary phase, with reasonable application and development opportunities, considering also the robust investment provided by the European Commission in this specific sector.

In this direction some applied research initiatives are likely to be successful with the arrangement of qualified partnership between industry and academic department. The role of a large scientific community like GIRPR can be extremely important to join together the required level of expertise.

9 Conclusions; Industrial Challenges for CVPR

The industrial session at ICIAP 2013 has been organized to represent a window over the world of ever growing applications of CVPR in the Security domain. Without any claim to represent an exhaustive list of such a wide range of applications it is intended to highlight some relevant technological objectives to be hopefully considered by the scientific community in a fruitful collaboration between advanced research and industry. Particular attention has been devoted to a better exploitation of a-priori knowledge and contextual information, to improve the level of scene understanding and interpretation (Context-Aware Imaging). Multi-camera video analysis and processing has been also identified as a fundamental goal of applied research, as well as edge-processing capabilities. Human target detection (by using extended biometric features) and re-identification within a network of cameras, as well as behavior understanding, are also primary research objectives of the CVPR technology.

The intelligent analysis of video stream (both live and recorded) through VCA has been widely discussed in the session with particular attention to the development (and possibly standardization) of Performance Evaluation methods to be accepted by the industrial community. It is a common opinion that the CVPR scientific community of GIRPR can play an important role in this field, by moving beyond the classical evaluation methodology, which has been mainly constrained to quantitative measures of algorithmic results in the academic research. It is necessary to extend such methodology to consider the VCA system as a whole, including high level recognition functions and the availability of intelligent tools for self-awareness and self-calibration and improve the robustness of the solution to variable environmental and operational conditions. One contribution of the session has been devoted to investigate the use of other imaging sensors, beside traditional video cameras, in Security applications. In particular, interesting results have been referred regarding the possibility to detect human beings beyond walls in homeland security, by using a special radar sensor configuration to operate as a sensitive Doppler motion detector. Finally, some experimental results have been referred on the use of an Augmented Reality AR system, applied to industrial environment to support remote assistance in advanced training and maintenance operations.

References

1. Capra, A.: PANORAMA, Ultra Wide Context Aware Imaging, http://www.panorama-project.eu
2. De Muro, S., Borzani, M.: La Sicurezza corre sul binario. Dossier VideoSorveglianza. Safety & Security, 22–24 (2010)
3. Video Content Analysis (VCA) Recommended Practice for Transportation Applications, APTA Draft, TSWG1 – VCA RP Working document (2010), http://www.apta.com
4. Baumann, A.: A Review and Comparison of Measures for Automatic Video Surveillance Systems. EURASIP Journal on Image and Video Processing 2008, 30 (2008), http://jivp.eurasipjournals.com/content/pdf/1687-5281-2008-824726.pdf
5. OpenCV, Open Source Computer Vision library, http://www.opencv.org
6. Esposito, M., Flammini, F., Pragliola, C.: Application of Video Analytics for Surveillance on-Board Train. In: ICIAP (2013)
7. Di Leo, G., Lengu, R., Mazzino, N., Paolillo, A.: Pattern recognition for defect detection in uncontrolled environment railway applications. In: Petrosino, A. (ed.) ICIAP 2013, Part II. LNCS, vol. 8157, pp. 753–757. Springer, Heidelberg (2013)
8. Ascione, M., Buonanno, A., D'Urso, M.: A method based on passive acoustic sensors for detection of vital signs in closed structures. In: Proc. of the IEEE Int. Conf. on Instrumentation and Measurement Technology I2MTC, Graz (2012)
9. Angelino, C.V., Cicala, L., De Mizio, M., Leoncini, P., Baccaglini, E., Gavelli, M., Raimondo, N., Scopigno, R.: H.264 sensor aided video encoder for UAV BLOS missions. In: Petrosino, A. (ed.) ICIAP 2013, Part II. LNCS, vol. 8157, pp. 749–752. Springer, Heidelberg (2013)
10. Savastano, M.: Cross-jurisdictional aspects of biometrics. ID World. In: 11th Annual World Summit, Frankfurt (D) (2012)
11. Carullo, M., Cavaliere, G., De Prisco, A., Di Capua, M., Petrosino, A., Padovano, D., Nave, G., Ruggeri, D.: Stock control through video surveillance in logistics. In: Petrosino, A. (ed.) ICIAP 2013, Part II. LNCS, vol. 8157, pp. 740–748. Springer, Heidelberg (2013)