Submitted to ELECTRONICS LETTERS

MPEG-4-BASED ERROR RESILIENT CODING WITH CLOSED-LOOP INTERACTION AMONG PERIPHERAL AND CENTRALIZED PROCESSING MODULES

Claudio Sacchi (1), Nicola Conci (1), Luigi Prisciandaro (2), and Francesco De Natale (1)

Abstract – In this letter[1], we are going to propose a novel MPEG-4-based video coding whose resilience is implemented by exploiting Region-of-Interest (RoI) scalability together with a closed-loop interaction between a central processing system, devoted to object detection and tracking, and the video coding embedded at sensor level. Results in terms of processing reliability, achieved in a formal way, can prove the usefulness of the proposed approach in the context of scene monitoring applications.

1. Introduction – Scene monitoring is one of the most emerging applications of video processing. Scene monitoring is essential in video-based surveillance, domotics, automated road traffic management, ambient intelligence, etc. In the typical system configuration, multiple cameras send video contents related to single or multiple scenes to a central hub. The central hub collects and processes such contents in order to generate some outputs useful to report the current situation of the scene. In such a framework, basic image processing tasks to be fulfilled are detection and tracking of foreground objects present in the scene. Effective object detection and tracking requires the real-time transmission of high-quality image sequences over unreliable and band-limited channels. This generally imposes the adoption of lossy coding techniques and prevents information retransmission in case of errors. Under such conditions, the received video contents may be affected by channel errors able at involving frame losses and/or heavy artifacts tampering wide image areas. For this reason, some kind of countermeasure is necessary in order to avoid the failure of the processing chain. The problem of recovering channel errors in scene monitoring applications is not trivial. The main objective of error recovery is not to maximize the PSNR (as in broadcasting applications), but to reduce and possibly eliminate in real-time the hugest noisy artifacts that could compromise the correct working of the image processing modules.

A feasible solution for post-processing error recovery in video-based surveillance systems has been proposed in [1] for JPEG-coded image sequences. In this letter, we are proposing a novel error resilient video coding technique based on MPEG-4 standard, targeted to scene monitoring applications. The error resilience is obtained by means of RoI scalability features typical of MPEG-4 coding [2] jointly exploited

---

[1] This work has been partially supported by MIUR in the framework of the project: “Study and Development of Innovative Smart Systems for Highly Reconfigurable Mobile Networks” Project COFIN 2005099984_001.

with a closed-loop interaction among peripheral sensor level and centralized processing level. At our best knowledge, this is the main innovation point proposed in the present work. RoI scalability has been designed in order to allow users to manually select some interesting regions of the displayed image. In our analysis, RoI scalability is employed by the image processing chain in order to automatically select some regions of crucial interest for object detection and tracking tasks (i.e. image areas containing foreground elements). The mechanism of error resilience is detailed in Section 2. In Section 3 some experimental results will be shown. Finally, letter conclusions are drawn in Section 4.

2. MPEG-4 coding with interactive error resilience – The core of our idea is shown in the diagram of Fig.1. At peripheral level, we have a digital sensor provided with a basic degree of processing capabilities. In particular, the MPEG-4 coding is embedded in the sensor. The proposed coding technique is targeted to video processing systems with a central hub devoted to collect and process visual information coming from multiple sensors. In the present letter, we have considered the use of RoI scalability to improve the resilience of the system against channel errors. The system shown in Fig.1 is based on the wireless transmission of MPEG-4 coded image sequence with real-time UDP transport protocol and RTP packetization [3]. Under such conditions, wrong packets are lost packets and heavy artifacts can be produced in the decoded frames. Serious problems arise when such artifacts tamper frame areas containing objects to be detected, tracked and recognized. In such a situation RoI scalability can help to repair such damages and to improve system performances in terms of correct object detection. The error resilience mechanism can be summarized as follows:

i) The embedded MPEG-4 encoder begins to encode frames without any kind of scalability. In parallel, a RoI database is created at sensor level and physically stored in a buffer of finite dimensions. The database content is initialized by a set of encoded RoIs roughly estimated by the “intelligent” sensor using a simple background-frame subtraction. The database is indexed by the position of the RoI in the image plane;

ii) The central processing system starts to work on the decoded sequence. It detects and tracks foreground objects in the scene. The image areas corresponding to foreground objects are labeled by the system as “RoIs”. The information about the 2D coordinates of top-left and bottom-right points of selected RoIs is then sent to the peripheral encoder through the feedback channel;

iii) The peripheral encoder starts to encode the RoIs on the basis of the information received by the central hub with the best quality allowed. In such a way the RoI database is dynamically updated. The position of the RoI in the current frame is then compared with the same one in the previous frame and the RoI area is incremented in order to embrace the moving object in the following frame. Such a mechanism of RoI enlargement is based on the utilization of a Kalman filter and has been adopted in order to face possible drifts of the video-based object tracker;

iv) The central processing system detects artifacts in the scene due to channel errors (the mechanism of error detection is similar to the one adopted in [1]). If these artifacts are produced in background areas of the scene, a simple concealment based on zero-motion replacement [4] is adopted (this is the reason why a buffer is used also by the central hub). Otherwise, if the detected artifact tampers an area of the image previously labeled as RoI, the remote processing issues a RoI transmission request to the peripheral encoder and sends it by means of the feedback channel. The RoI transmission request contains the information about the RoI position in the image plane. In such a way, the wanted RoI is retrieved from the database, packetized and transmitted through the forward channel;

v) The central processing level receives the requested RoI and produces an output sequence with the damaged areas repaired. The codec process can restart by i).

3. Experimental results – We tested the proposed resilient MPEG-4 coding in the framework of a domotic application. The experimental setup consists of a peripheral sensor emulated by an analog video camera connected to a laptop by a frame grabber. The video coding system is based on the XVID release of the MPEG-4 standard. The central processing system, consisting of another laptop, is located inside a monitored house. In this kind of setup, potential delays occurring in the selective RoI transmission described at the point iv) of Section 2 should not be very critical. In fact, the peripheral sensor and the central hub are physically very close. Moreover, the monitored scene is sufficiently "slow-motion". The bi-directional wireless connection has been implemented by means of an IEEE 802.11b radio link. We have considered a UDP/RTP transmission of an MPEG-4 video sequence at a rate of 50Kb/s. The tested image sequence consists of 130 frames at 384x288 pixels per frame. As the available bandwidth of IEEE 802.11b connections ranges from 1Mb/s to 11Mb/s, the small bandwidth overhead due to feedback channel (few kb/s) can be easily supported. In the present dealing, we adopted the performance evaluation methodology detailed in [5] and based on the tracing of Receiver Operating Characteristics (ROC) curves. ROC curves are the most effective tools for assessing video processing systems targeted to object detection, as they map one against another two critical performance indicators like false detection probability (P_{fals}) and correct detection probability (P_{det}). Visual results shown in Fig.2 can provide "at a glance" an idea of the quality improvement provided by the proposed error resilient coding even in presence of heavy channel errors (packet error rate PER=0.5). The left frame is obtained by the transmission of the MPEG-4 coded sequence without any error resilience mechanism. The right frame is the result of the replacement of the damaged RoI with the corresponding one transmitted “on request” by the peripheral sensor. However, the most impressive results are shown in Fig.3 where ROC curves are drawn. The first curve is the ideal ROC, related to the error-free transmission of the image sequence. The ROC curve related to the noisy transmission (PER=0.5) adopting the proposed error resilience mechanism is very close to the ideal ROC. This means that the foreground objects are correctly detected. On the contrary, the ROC curve related to the MPEG-4 transmission obtained without any kind of error resilience is very far from ideal performances.

4. Conclusion – A novel methodology for scalable and resilient image coding based on MPEG-4 layer has been studied targeted to video processing systems for object detection and tracking. The proposed coding algorithm can profitably exploit RoI scalability and a closed-loop interaction between peripheral sensors and the remote processing hub in order to improve the quality of the received frames in those areas of specific interest for the application context. Results showed a dramatic performance improvement concerning both visual quality and, above all, image processing reliability.

References

[1] F. Granelli, C. Sacchi, C. Regazzoni, “Real-time error correction algorithm for noise-corrupted JPEG bit-streams”, *Electronics Letters*, Vol. 35, No. 14, 8 July 1999, pp. 1147-1148.

[2] N. Brady, “MPEG-4 Standardized Methods for Compression of Arbitrarily Shaped Video Objects”, *IEEE Trans on Circ. and Sys. for Video Tech.*, Vol. 9, No. 8, Dec. 1999, pp. 1170-1189.

[3] W. Stallings, *“Data and Computer Communications”*, (4th Ed.), Macmillan: 2000.

[4] Y. Wang, Q.F. Zhu “Error Control and Concealment for Video Communication: A Review”, *Proceedings of the IEEE*, Vol. 86, No. 5, May 1998, pp. 974-997.

[5] F. Oberti, A. Teschioni, C. Regazzoni, “ROC curves for performance evaluation of video sequences processing systems for surveillance applications”, *Proc. of ICIP 1999 Conf.*, Vol. 2, pp. 949-953.

Authors’ affiliations

(1) University of Trento, Department of Information and Telecommunication Technology (DIT), Via Sommarive 14 I-38050, Trento, Italy, e-mail: sacchi@dit.unitn.it (2) Create-Net International Research Centre, Via Solteri 38, I-38100, Trento (Italy), e-mail: luigi.prisciandaro@create-net.it

Figure captions

Figure 1. Diagram of the error-resilient MPEG-4 coding and transmission system

Figure 2. Object detection obtained by MPEG-4 coding without error resilience (left frame) and obtained by using MPEG-4 coding with closed-loop error resilience (PER=0.5).

Figure 3. Assessment of video-processing performance obtained by means of ROC curves (PER=0.5).

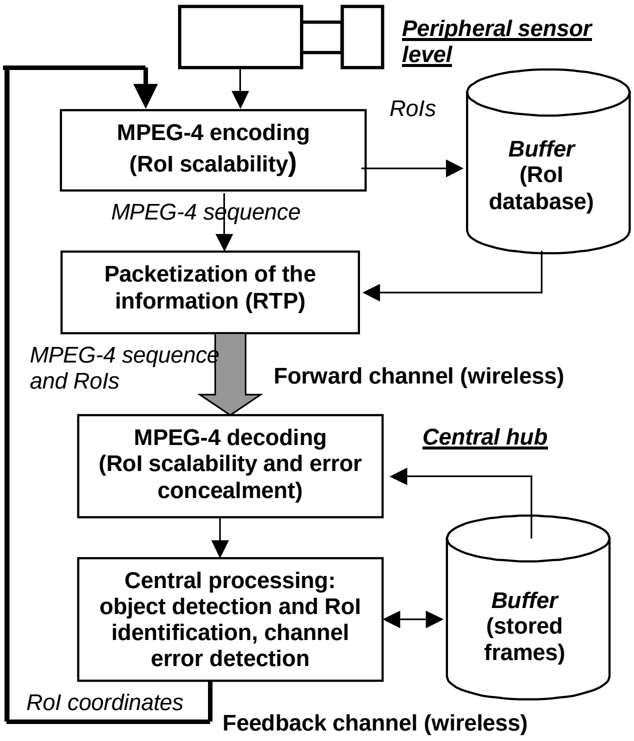

Figure 1

Figure 2

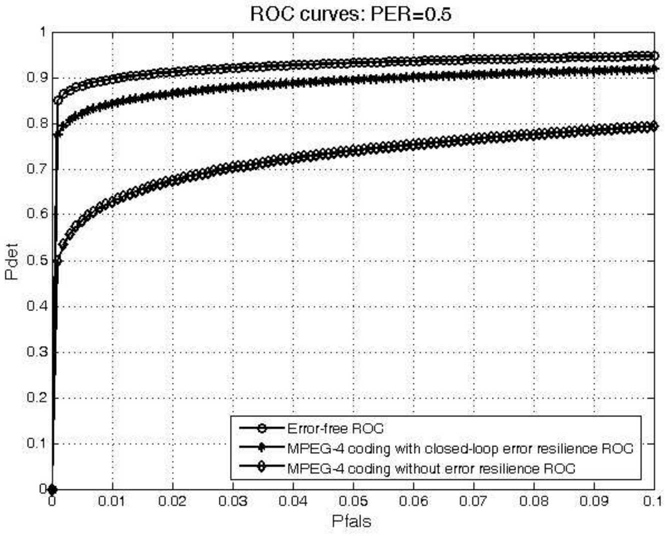

Figure 3