Multiple Priority Region-of-Interest H.264 Video Compression Using Constraint Variable Bitrate Control

Chung-Ming Huang, *IEEE Senior Member* and Chung-Wei Lin, *IEEE Student Member*
Laboratory of Multimedia Mobile Networking
Department of Computer Science and Information Engineering
National Cheng Kung University, Tainan, Taiwan, R.O.C.
Correspondence: huangcm@locust.csie.ncku.edu.tw

***Abstract*—With the growth of digital video surveillance markets and requirements of high-quality surveillance data, an efficient video compression technique that is suitable for video surveillance applications and is compatible with emerging coding standards, e.g., H.264, is required urgently. This paper presents the method and the corresponding system using region-of-interest H.264 video coding based on constraint variable bitrate control for multiple moving priority regions. The proposed Multiple Moving region-of-interest (RoI) macroblock decision and an RoI-based Constraint Variable Bitrate Control called MM-RoI-CVBC are able to identify each MB coding priority and increase PSNR (recognition rate) of automatically or maually detected RoI. In comparison with H.264 JM fixed-quantization parameter coding, H.264 JM constant bitrate control (CBR) and other RoI coding algorithms, the proposed MM-RoI-CVBC method can obviously enhance the quality of RoI with less coding complexity in subject to the target bitrate.**

***Index Terms*—region-of-interest coding, H.264, video surveillance, bitrate control.**

1. INTRODUCTION

With the growth of digital video surveillance markets and requirements of high-quality surveillance data, traditional video compression techniques are not enough to resolve several emerging challenges, e.g., higher coding complexity and lower recognition rate on human eyes' region-of-interest (RoI). Thus, a novel video compression technique that is compatible with the emerging coding standard, e.g., H.264/AVC [1], is required to improve the quality of RoI with less coding complexity. Although RoI-based video compression was well researched for other video coding standards, it is still interesting for related H.264 video surveillance applications because of newly-defined profiles/tools in H.264 and multiple objects moving in the surveillance environment.

The major concept of RoI-based video coding is to enhance the visual quality of human eyes' region-of-interest (RoI) and sacrifice the quality of non-RoI by dropping more high-frequency coefficients. The procedures of RoI-based video coding can be mainly classified into two aspects: (i) pre-object segmentation and tracking and (ii) post-bitrate control.

*This research is supported by the National Science Council of the Republic of China, Taiwan under the contract number NSC 97-2219-E-006-002, and Intel Microelectronics Asia Ltd, Taiwan Branch.

In pre-object segmentation and tracking, Hu. et al. described the standard procedure, including environment modeling, motion segmentation, object classification and motion tracking [2]. In the aspect of environment modeling, Wang. et al. adjusted camera parameters to acquire best video qualities based on external information on brightness, contrast, and location [3]. Then, regional motion segmentation and frame differencing were adopted to separate different objects' motion from a static background [4]. In the calibrated global space, segmented objects can be tracked using techniques of color histograms or a gradient-based foreground detection [5]. Besides, FMO (Flexible Macroblock Ordering) is an emerging tool specified in H.264 baseline and extension profiles. FMO is able to group distinct object-MBs into distinct slices and related information is recorded in a MB assignment map (MBAmap). Then, distinct slice groups can be coded or transmitted with distinct coding/transmission priorities to achieve goals of RoI-based video coding or error-resilience channel coding [6]. However, FMO is not supported by the main profile and too complicated to encode/decode in real time for the video surveillance application.

In post-bit rate control, well-known rate control methods, e.g., H.263 TMN [7], MPEG-4 VM [8] and H.264 JM, play major roles to allocate proper bit budgets and adjust the quantization parameter (QP) to minimize visual distortion in subject to limited bandwidth resources. However, these methods of traditional rate control consider that priorities of all MBs are the same with each other. This assumption is not always correct for surveillance applications because sometimes RoI has less difference signals (complexity) but need more bit budgets. For surveillance applications, multiple object-based video coding was further proposed in MPEG-4 (ISO/IEC 14496), which is able to compress separated objects, composite them using a scene descriptor and represent them in distinct spatiotemporal cyberspace. Among these methods of multiple-object video coding, to increase PSNR (recognition rate) of individual RoI, past RoI-based bitrate control attempted to allocate distinct number of MB bit-budgets to multiple-level priority MBs (objects) based on remaining bit budgets and coding complexity, i.e., higher priority, more bit-budgets [9][10]. However, these improved RoI-based bitrate control methods need (i) preload and preanalyze later n frames or (ii) a complicated quadratic R-Q (rate-quantization) function to update QP and its parameters. Thus, it is also not suitable for the real-time surveillance application because of much computing time-consumption.

In this paper, we propose a method and system of region-of-interest H.264 video coding using constraint variable bitrate control for multiple moving priority regions, called MM-RoI-CVBC. MM-RoI-CVBC contains two technical contributions: (i) multiple moving region-of-interest MB decision and (ii) an RoI-based constraint variable bitrate control algorithm. According to object-tracking results, the proposed multiple moving region-of-interest MB decision is able to identify each MB priority. Then, MBs can be classified into three levels- Priority 3: Background, Priority 2: RoI-contour extension and Priority 1: RoI itself. Distinct prioritylevel MBs can be assigned distinct QPs. According to the characteristic of the human visual model, QP-delta values, Δ_1 and Δ_2, can be determined, e.g., a QP-set $\{QP_{ROI}, QP_{ROI}+\Delta_1, QP_{ROI}+\Delta_2\}$ for RoI, RoI-contour extension and Background, respectively. QP_{ROI} can be initialized properly based on the target bitrate and frame spatiotemporal resolutions. Since frame complexity (MAD) and RoI are different in each frame, the output bitrate will vary frame-by-frame. Thus, the proposed MM-RoI-CVBC can be regarded as the extension of variable bitrate control. To further control total output bitrate, we propose the constraint variable bitrate control algorithm to adjust QP_{ROI} and avoid the encoded bits exceeding or being below the target bitrate.

The rest of this paper is organized as follows. Section 2 introduces Yuriy's object tracking algorithm and proposed multiple moving region-of-interest MB decision. Section 3 introduces the proposed RoI-based constraint variable bitrate control algorithm, including QP-delta selection, frame-level QP_{ROI} initialization and QP_{ROI} adjustment. Section 4 exhibits experiment results. Section 5 has concluding remarks.

2. The Proposed Multiple Moving RoI MB Decision in MM-RoI-CVBC

In this Section, we will integrate Yuriy's object tracking algorithm in MM-RoI-CVBC and introduce our proposed multiple moving region-of-interest MB decision based on tracking results.

In the aspect of object tracking, Chesnokov Yuriy's object tracking algorithm published in CodeProject is modified for satisfying our system requirements firstly [11]. The way of Yuriy's object tracking is to compare a pre-defined background with incoming captured images and draw each object's boundary (called blobs) based on their motion vectors and RGB differences. Fig. 1 shows tracking results of the *Hall* standard video using Yuriy's algorithm, in which multiple tracking objects' blobs can be recognized and marked. In addition to automatic object-tracking, RoI can be also user-defined manually, e.g., RoI made by the upper left corner point (called the calibrated point), the frame height is 162 (pixel) and the frame width is 226 (pixel), which is shown in Fig. 2.

According to detected object's information, each MB coding priority can be classified into three levels: 3- Background, 2- RoI-contour extension and 1- RoI itself. In the proposed

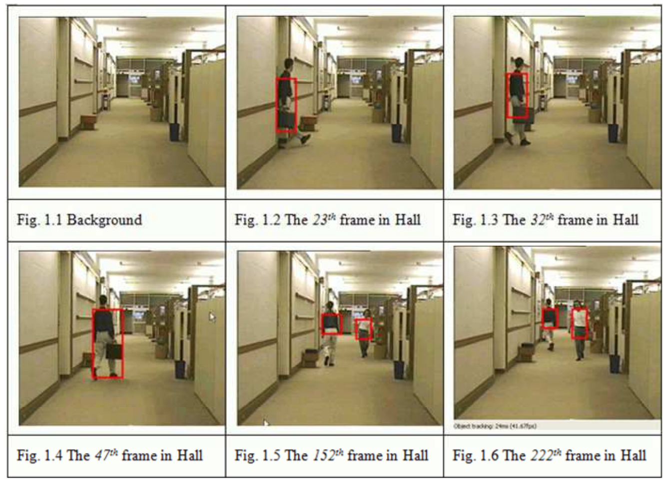

Fig. 1. The tracking results of the *Hall* standard video using Chesnokov Yuriy's object tracking algorithm.

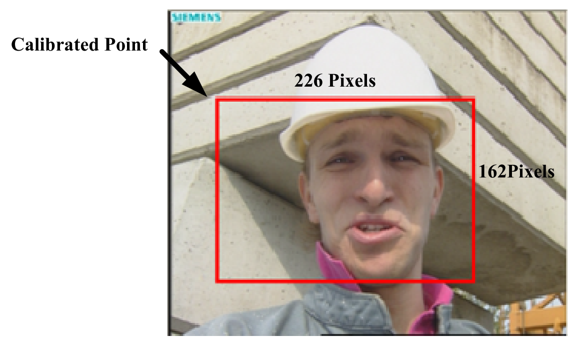

Fig. 2. The manually user-defined RoI in the CIF Foreman video.

multiple moving RoI MB decision, three major issues we concern are as follows:

1) How to determine each MB coding priority efficiently at given information about blobs' positions and sizes, and record them in an RoI mapping table.
2) How to determine each MB coding priority when multiple blobs overlap.
3) How to determine each MB coding priority when object tracking fails, e.g., no objects are detected when external illumination changes.

To resolve the aforementioned technical problems, the proposed multiple moving region-of-interest MB decision is proposed and is depicted in Fig. 3. Six major steps are as follows:

Step 1. Set the priority of all MBs as 3 (Background).

Step 2. If no object is detected or object tracking fails, the RoI mapping table is the same as the previous one temporally and go back to Step 1; otherwise, go to the next recursive multiple moving region-of-interest MB decision.

Step 3. Find the index of MBs located within each detected object from the coordination (x,y) to ($x + w$,$y + h$), in which (x,y) is the calibrated position, w is the object width and h is the object height. Then, the priority of these MBs was assigned Priority 1 (RoI).

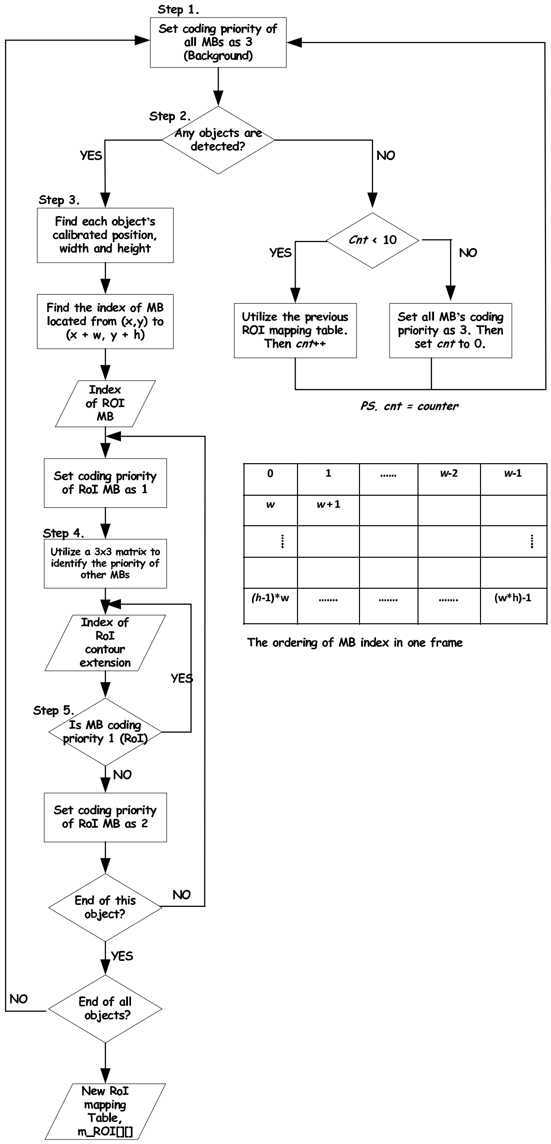

Fig. 3. The flowchart of the proposed multiple moving RoI MB decision in MM-RoI-CVBC.

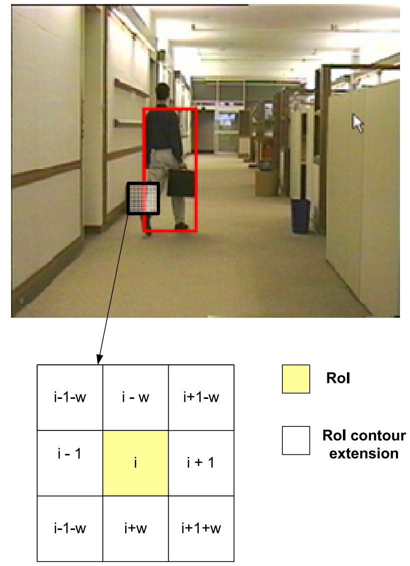

Fig. 4. The 3x3 matrix for identifying the priority of other MBs (yellow area: RoI and others: RoI contour extension).

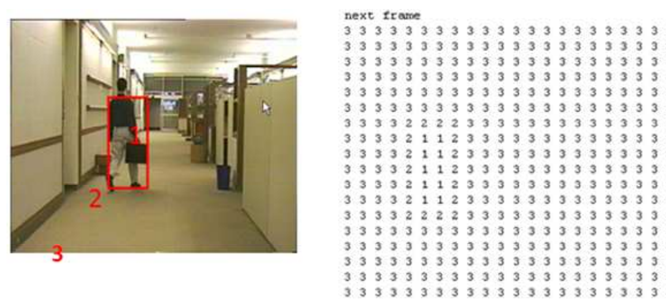

Fig. 5. The results of the proposed multiple moving RoI MB decision in MM-RoI-CVBC.

Step 4. Utilize a 3x3 matrix to determine MB indexes of RoI, RoI-contour extension and its coding priority. Referring to Fig.4, the central cell of this matrix belongs to the detected object's contour and the other eight surrounding matrix cells belong to the RoI-contour extension.

Step 5. Check the priority of RoI-contour extension MBs when multiple objects overlap. If the priority has been assigned 1 (RoI), it means that this MB falls into other object's inner area; otherwise, its' priority is assigned 2 (RoI-contour extension).

Step 6. Repeat Steps 2-5 until all detected objects are processed. The final RoI mapping table is presented as m_ROI[i][j], in which i means the MB index, m_ROI[i][0] is the index of the detected object, which the i^{th} MB belongs to, and m_ROI[i][1] is the priority of the i^{th} MB.

Fig. 5 shows the results of the proposed multiple moving RoI MB decision using MM-RoI-CVBC, including the definition of MB coding priority and the generated RoI mapping table.

3. The RoI-based Constraint Variable Bitrate Control Algorithm

In Section 2, the proposed multiple moving RoI MB decision is able to determine the coding priority of multiple moving MBs and generate the RoI mapping table. According to the RoI mapping table, RoI, RoI-contour extension and Background are assigned a QP set $\{QP_{ROI}, QP_{ROI}+\Delta_1, QP_{ROI}+\Delta_2\}$, in which $0 \leq \Delta_1 \leq \Delta_2$. It is because "larger QP, more distortion" in the traditional rate-distortion model. Three research issues we concern are as follows:

1) The QP-delta (Δ_1 and Δ_2) selection: Let us consider the following situations: (i) The smaller QP-delta values may result in slight blocking effects because the PSNR difference of adjacent MBs is smaller. But the quality of RoI is similar to that of non-RoI, the effect of RoI video coding is not obvious and less bitrate can be saved. (ii) The larger QP-delta value is able to increase PSNR of RoI and save more bit-rates by dropping high

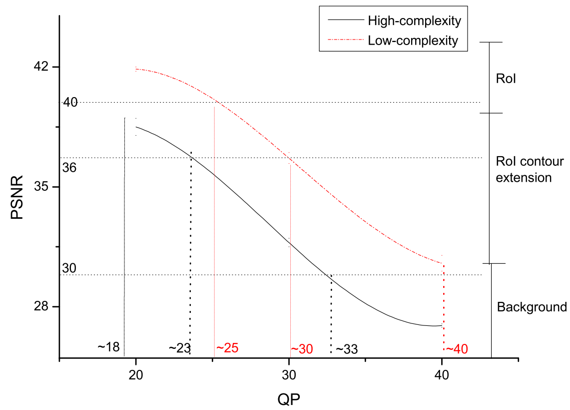

Fig. 6. The result of the Distortion-Quantization curve.

coefficients of the background. How to determine Δ_1 and Δ_2 in consideration of the characteristic of human visual model and RoI-based video coding is our first research issue to be addressed.

2) The initial QP_{ROI} decision: How to determine the proper initial QP_{ROI} in consideration of the characteristic of captured videos.

3) The dynamic parameter QP_{ROI} adjustment: In our proposed MM-RoI-CVBC, since each frame has distinct complexity and all MB QPs are determined based on the importance of distinct regions in advance, the numbers of encoded bitrate of each frame are different with each other. It belongs to variable bitrate video coding. Thus, the third research issue is how to adjust the parameter QP_{ROI} to make the encoded bitrate approach the target bitrate in subject to the storage capacity or bandwidth resources in the surveillance environment.

In the first issue, the relationship of Distortion-Quantization is modeled based on many gathered standard video sequences at the beginning. Considering the characteristic of the human perception, PSNR ($\geq$30dB) is usually regarded as the minimum acceptable video quality (Level3), and PSNR ($\geq$40dB) is regarded as the excellent quality (Level1) [12]. Without loss of generality, the average PSNR (34dB-36dB) is regarded as the middle video quality (Level2). Thus, in our work, video qualities can be classified into three levels for RoI-based video coding: Level3 for Background, Level2 for RoI contour extension and Level1 for RoI. Fig. 6 depicts the result of Distortion-Quantization curves, in which the dot curve is drawn for higher-complexity data and the solid curve is drawn for lower-complexity data. From the observation of Fig. 6, proper QPs regrading distinct quality levels should be $\{25,30,40\}$ and $\{18,24,33\}$ for distint complexity data. Thus, reasonable QP-delta values Δ_1 and Δ_2 are 5 and 15, respectively, in the QP set $\{QP_{ROI}, QP_{ROI}+\Delta_1, QP_{ROI}+\Delta_2\}$. It satisfies the condition of QP-delta [-25,26] in H.264. Besides, QP_{ROI} is within [1,31] because all H.264 quantization parameters should be limited within [1,51].

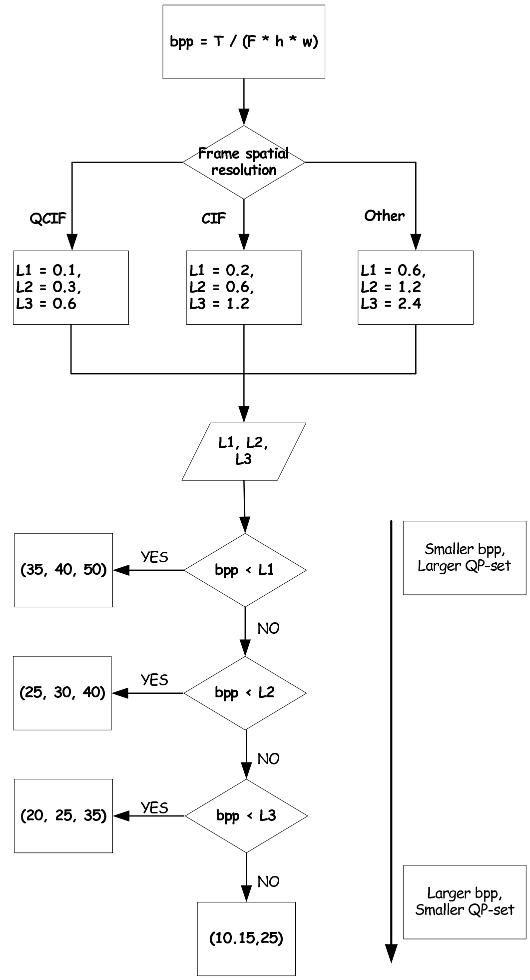

Fig. 7. The procedure of the initial QP-set decision.

In the second issue, a proper initial QP_{ROI} can be determined based on the target bitrate and frame spatiotemporal resolutions. A parameter bpp (bits per pixel) is given by

$$bpp = T/(F \cdot h \cdot w) \tag{1}$$

where T is the target bitrate, F is the frame rate, h is the height and w is the width.

Fig. 7 shows the procedure of the initial QP-set decision, in which $L1$, $L2$ and $L3$ are three thresholds depending on the frame spatial resolution. Through a great number of off-line experiments, one of four QP-sets, e.g., (35,40,50), (25,30,40), (20,25,35) and (10,15,25), are selected when bpp and thresholds $L1/L2/L3$ are known.

In the third issue, the quantization parameter of RoI should be updated frame-by-frame based on object-tracking results, buffer fullness and the difference of encoded and target bitrates. Fig. 8 shows the dynamic parameter QP_{RoI} adjustment, which is listed as follows:

Step 1. If no object is detected, a uniform QP-set (40,40,40) is assigned to quantize all MBs, and then discard the following control procedure.

Step 2. Calculate current buffer fullness B by $B_t = B_{t-1} - R_t + T/F$, where t is the time index and R_t is the

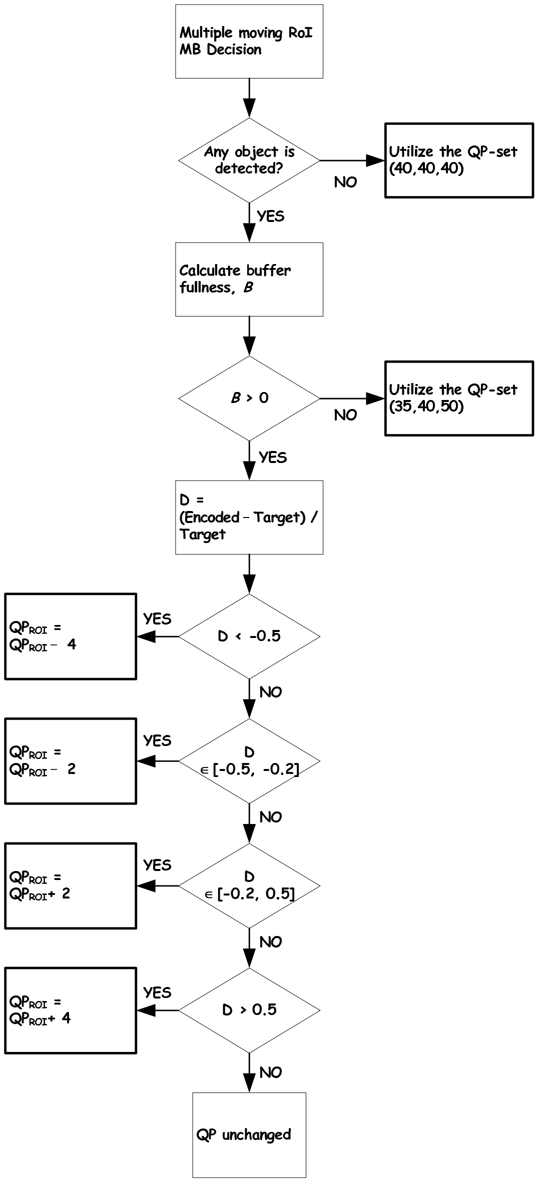

Fig. 8. The dynamic parameter QP_{RoI} adjustment in proposed MM-RoI-CVBC.

encoded bitrate. If the crisis of buffer overflow occurs when $B \geq 0.8 \cdot B_{max}$, the largest QP-set (35,40,50) will be assigned; otherwise, go to Step3.

Step 3. Calculate the percentage (D) of the difference of the encoded bitrate and target bitrate.

Step 4. QP_{ROI} can be adjusted step-wisely based on the range of D, e.g., QP_{ROI} -4 if D $\leq$ -0.5, -2 if D $\in$(-0.5,-0.2], unchanged if $D \in$ (-0.2,+0.2], QP_{ROI} +2 if D $\in$(0.2,0.5], QP_{ROI} +4 if D $\geq$0.5. That is, QP_{ROI} should increase when the previously encoded bitrate is larger than the target bitrate; otherwise, QP_{ROI} should decrease.

4. Experimental Results

In our experiments, we mainly compare the performance of the proposed MM-RoI-CVBC with (1) those of fixed-QP and constant bitrate control (CBR) in original H.264 JM13.2, (2) Sun's method [9] and (3) Sivanantharasa's method [6]. Both Sun's and Sivanantharasa's methods are improved based on the original H.264 JM13.2 constant bitrate control . The comparison items include PSNRs of RoI/RoI extension/Background, the error of the encoded bitrate and the target bitrate, and the computing time-consumption of bitrate control. The used video sequences are Hall and Foreman. Table I tabulates major H.264 encoding parameters, e.g., the number of encoded frames, the initial QP for I/P slices, RDO (Rate Distortion Optimization) and RatControlEnable.

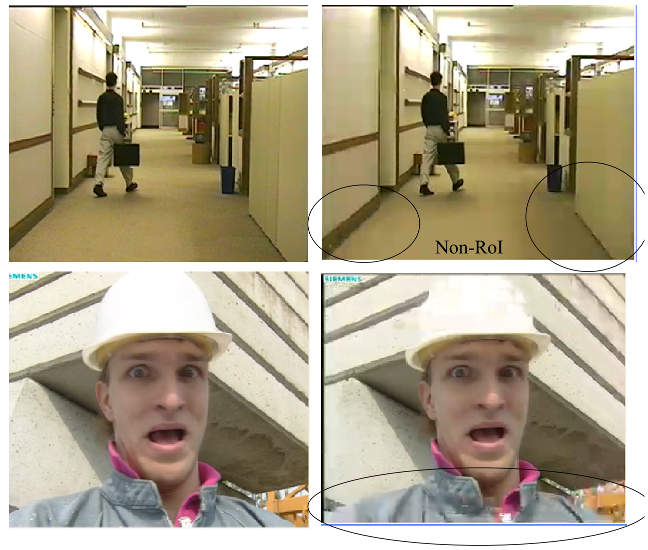

Fig. 9. The results of RoI images encoded by MM-RoI-CVBC (Upper: Hall and Bot.: Foreman).

At the beginning, the video $Hall$ is encoded by fixed QPs, and then the output bitrates 1826Kbps and 206Kbps are regarded as the high bitrate and low bitrate in other coding algorithms. From these experiment results tabulated in Table II, the following conclusions are made:

1. In our proposed MM-RoI-CVBC, the quality of RoI is better than that of non-RoI obviously, which is depicted in Fig. 9, e.g., floor regions (Hall) and bottom regions (Foreman) are blurred.

2. Comparing the quality of encoded images, the quality of RoI is possibly better or worse than that of non-RoI in JM13.2, Sun's and Sivanantharasa's methods. The proposed MM-RoI-CVBC can obtain better PSNR of RoI about 2-8 dB gain than other RoI-based video coding algorithms, but the quality of non-RoI will be sacrificed.

3. Besides, the error of encoded bitrate of MM-RoI-CVBC is only about 3% because our proposed MM-RoI-CVBC is able to control bit budgets effectively. The result of MM-RoI-CVBC is similar to those of traditional JM13.2 CBR, Sun's and Sivanantharasa's methods.

4. Comparing total time-consumption of the rate control part, our proposed MM-RoI-CVBC spends less coding time per frame (0.4ms) than JM13.2 (6ms), Sun's (9ms) and Sivanantharasa's (6ms). It is because that other coding algo-

TABLE I
MAJOR H.264 ENCODING PARAMETERS. (E: ENABLE, D: DISABLE)

	JM FixedQP	JM CBR	MM-RoI-CVBC	Sun [9]	Sivanantharasa [6]
FrameToBeEncoded	180 (Foreman) and 300 (Hall)				
QPISlice/QPPSlice	20/30/40	30	30	30	30
PSliceSearch16x16	E	E	E	E	E
RDO	D	D	D	D	D
RateControlEnable	D	E	D	D	D

TABLE II
COMPARISONS OF THE EXPERIMENT OF HALL.

Methods	Fixed QP			JM 13.2		Proposed		Sun [9]		Sivanantharasa [6]	
Parameters	20	30	40	1826	206	1826	206	1826	206	1826	206
PSNR (dB), Avg→	41.76	36.36	30.08	41.60	36.45	40.79	34.08	41.60	36.28	41.41	36.46
RoI→	-	-	-	41.58	34.35	49.00	38.56	41.57	34.51	41.23	34.44
RoI contour extension →	-	-	-	41.38	34.89	45.06	36.08	41.39	35.09	40.94	35.04
non-RoI →	-	-	-	41.66	37.15	39.99	33.56	41.66	37.02	41.53	37.16
Error rate(%)	-	-	-	-2.3	-1.5	-3.6	+2.8	-2.3	-1.0	-2.5	-1.3
Output bitrate (Kbps)	1826	206	40	1783	203	1759	211	1783	204	1780	203
Total time consumption for rate control (ms/frame)	0.083	0.085	0.085	6.54	6.14	0.41	0.38	10.47	8.96	5.99	5.87

rithms contain the complicated rate-distortion model and more procedures for bit-allocation and QP-decision additionally.

5. CONCLUSION

This paper has proposed the method and the corresponding system using region-of-interest H.264 video coding based on constraint variable bitrate control for multiple moving priority regions, called MM-RoI-CVBC. The proposed multiple moving region-of-interest (RoI) macroblock decision is able to efficiently determine each MB coding priority based on multiple automatically or manually detected moving objects. The special cases of multiple objects overlap and tracking failure are considered. In addition, the proposed RoI-based constraint variable bitrate control is able to increase PSNR (recognition rate) of RoI and avoid buffer overflow/underflow. In comparison with H.264 JM fixed-quantization parameter (QP) coding, constant bitrate control (CBR) and other RoI coding algorithms, the proposed MM-RoI-CVBC can enhance the quality of RoI regions about 2-8dB with less coding complexity in subject to constraint bit budgets.

REFERENCES

[1] ISO/IEC JTC1, Information Technology- Coding of Audio-Visual Objects- Part 10: Advanced Video Coding, ISO/IEC FDIS 14496-10, 2003.

[2] W. Hu, T. Tan, L. Wang and S. Maybank, "A survey on visual surveillance of object motion and behaviors", *IEEE Transactions on System, Man and Cybernetics- Part C: Applications and Reviews*, VOL. 34, No. 3, pp.334-352, August 2004.

[3] J. Wang, W.Q. Yan, M.S. Kankanhalli, J. Ramesh and M.J.T. Reinders, "Adaptive monitoring for video surveillance", *Proceedings of IEEE International Conference on Information, Communications and Signal Processing*, pp.1139-1143, 2004.

[4] L. Li, Q. Zeng, Y. Jiang and H. Xia, "Spatio-temporal motion segmentation and tracking under realistic condition", *Proceedings of IEEE International Symposium on Systems and Control in Aerospace and Astronautics*, pp.229-232, 2006.

[5] L. Lee, R. Romano and G. Stein, "Monitoring activities from multiple video streams: establishing a common coordinate frame", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, VOL. 22, NO. 8, pp.758-767, August 2000.

[6] P. Sivanantharasa, W.A.C. Fernando and H. Kodikara Arachchi, "Region of Interest Video Coding with Flexible Macroblock Ordering", *Proceedings of IEEE International Conference on Image Processing*, pp. 53-56, 2006.

[7] ITU-T SG16 Video Coding Experts Group, "Video Codec Test Model, Near-Term, Version 8 (TMN8)", September, 1997.

[8] ISO/IEC JTC1/SC29/WG11, "MPEG-4 Video Verification Model v18.0, Coding of Moving Pictures and Audio N3908", January 2001.

[9] Y. Sun, I. Ahmad, D. Li and Y.Q. Zhang, "Region-Based Rate Control and Bit Allocation for Wireless Video Transmission", *IEEE Transactions on Multimedia*, VOL. 8, NO. 1, pp.1-10, February 2006.

[10] D. Agrafiotis, D.R. Bull, N. Canagarajah and N. Kamnoonwatana, "Multiple Priority Region of Interest Coding with H.264", *Proceedings of IEEE International Conference on Image Processing*, pp. 53-56, 2006.

[11] [On-line] http://www.codeproject.com/KB/audio-video/object_tracker.aspx

[12] M.C. Lin, L.R. Dung and P.K. Weng, "An ultra-low-power image compressor for capsule endoscope", *Biomed Eng.*, VOL. 5, 2006.