A FAST AND PERFORMANCE-MAINTAINED TRANSCODING METHOD BASED ON BACKGROUND MODELING FOR SURVEILLANCE VIDEO

Mingchao Geng[1], Xianguo Zhang[1], Yonghong Tian[1], Luhong Liang[2], Tiejun Huang[1]
[1]Institute of Digital Media, Peking University,
Beijing, 100871, P. R. China
[2]Institute of Computing Technology, Chinese Academy of Sciences,
Beijing, 100190, P. R. China
{mcgeng, xgzhang, yhtian, lhliang, tjhuang}@jdl.ac.cn

***Abstract*—Low-complexity and high-performance surveillance video transcoding methods play an important role for a wide range of surveillance video transmission and storage applications. Towards this end, the special characteristics of surveillance video should be utilized for transcoding. In this paper, we propose a fast and performance-maintained transcoding method. This method firstly divides macroblocks (MBs) into foreground MBs, foreground border MBs and background MBs. Statistics show that the three categories have different distributions of prediction modes, motion vectors and reference frames. Following this, we adopt different transcoding strategies in terms of removing the redundant prediction modes, narrowing motion search range and reducing reference frames. In particular, we propose an algorithm to exploit the decoded motion vector to adaptively calculate motion search range. Experimental results show that, compared with the recent background modeling based full-decoding-full-encoding, our transcoding method saves more than 93% time with ignorable quality loss.**

Keywords- surveillance video transcoding; background modeling; motion search range; full-decoding-full-encoding; decrease complexity

I. INTRODUCTION

Video surveillance systems are widely used for safety and communication applications. With an exponential increase of networked and widely deployed high-definition cameras, how to transcode surveillance video to adapt various application settings is increasingly becoming an important issue. In a video teleconferencing system, for example, low-complexity, high-efficient and bit-rate scaling transcoding techniques can be helpful for transferring video data among different teleconferencing clients (e.g., mobile devices) under various networks.

Generally, the most efficient bit-rate scaling video transcoding approach is the traditional full-decoding-full-encoding (FDFE). However, due to the complexity problem mostly caused by motion estimation (ME) and mode decision (MD), FDFE is not feasibly used in practical transcoding systems. To decrease the ME complexity, several fast transcoding methods using motion vector refinement were proposed by [1-4], with comparable transcoding performance to FDFE. Meanwhile, methods for saving MD complexity [5-7] are also widely investigated in the past years. A zero-block decision based scheme was introduced by Wu et al. [5], where the zero-block decision scheme is used to skip impossible inter and intra prediction modes, consequently leading to 93% saving of computation time on average. Nevertheless, none of the above video transcoding techniques focus on surveillance video. When they are applied to surveillance video, the transcoding efficiency has much room for improvement because the special characteristics of surveillance video are not utilized.

Intuitively, a reasonable transcoding solution for high-efficient surveillance video is to transcode the foreground objects and background separately. We denote it as object-oriented transcoding throughout this paper. Following this idea, object-oriented methods in [8-9] were proposed to divide an input frame into foreground and background regions, and then transcode background with low quality. However, object-oriented methods often need complex object segmentation and usually focus on subjectively defined "foreground objects." For surveillance transcoding, the subjective measurement is a debatable problem, especially considering various security requirements.

To realize high-efficient, bit-rate scaling and objective-quality-measured surveillance video transcoding, a background modeling based transcoding method (referred to as BGT here) was introduced by Zhang et al. [10]. BGT uses the generated background frame as an additional long-term reference frame in transcoding. The experimental results showed that BGT saves nearly half of the bit-rate while maintaining the PSNR performance compared with the traditional FDFE. Nevertheless, BGT still follows the FDFE framework and its complexity is too high since no methods are specially designed to save the time cost of ME and MD. To further eliminate redundant computation and maintain the high performance in background modeling based transcoding, this paper proposes a fast and performance-maintained surveillance video transcoding algorithm based on background modeling. Because different blocks in one frame of surveillance video might have different motion characteristics, we propose to firstly classify blocks according to the modeled background frame, and then transcode them with corresponding strategies.

[1] Dr. Yonghong Tian is the corresponding author.

978-0-7695-4711-4/12 $26.00 © 2012 IEEE
DOI 10.1109/ICME.2012.65

Our fast and performance-maintained background modeling based transcoding method (namely FP-BGT) works as follows: Firstly, a background frame is generated from the original decoded frames by a low-complexity way. After that, the modeled background frame is used to divide the MBs in current frame into foreground MBs (FMs), foreground border MBs (FBMs) and background MBs (BMs) with simple threshold judgment. Thirdly, different ME and MD strategies in transcoding will be respectively used for the three categories of MBs. Candidate reference frames are selectively skipped due to low correlation, and prediction modes are regularly forbidden according to mode distribution and decoded mode type. In particular, an algorithm suitable for different categories is proposed to exploit the relationship between decoded motion vector and predicted motion vector in encoder to adaptively reduce motion search range. Overall, our method can significantly decrease the redundant computation, meanwhile maintaining the gain that BGT has achieved. Experimental results show that on average FP-BGT can save 96.98%/96.98% on CIF and 96.71%/96.72% on SD transcoding time over traditional FDFE/BGT respectively. Meanwhile, the transcoding quality is slightly worse than BGT by not more than 0.1dB, but obtains 1.26dB/1.15dB PSNR gain on CIF/SD compared with traditional FDFE.

The rest of this paper is organized as follows. Sec. 2 analyses the distribution of prediction modes, motion vectors and reference frames. Sec. 3 presents the proposed method. Sec. 4 shows the experimental results and Sec. 5 concludes this paper.

II. PROBLEM ANALYSIS

The number of reference frames, the motion search range and prediction modes in ME and MD are three key factors to decrease the transcoding complexity. For different motion characteristics, BMs generally tends to select larger block size prediction modes in MD, background prediction and nearby motion search position in ME. Above tendency, nevertheless, is not followed by FMs and FBMs. Therefore, we firstly make an experimental analysis of their distributions in BGT on the three kinds of MBs and sub-macroblocks (SMBs), including BMs, FMs, and FBMs. This analysis can enlighten us how to reduce the redundant transcoding complexity from the three aspects. Note that the analysis results are obtained from an experiment of transcoding four CIF surveillance video of Crossroad, Overbridge, Snowgate, Snowroad (as shown in Fig. 1) in BGT, with the bit-rate at about 512kbps.

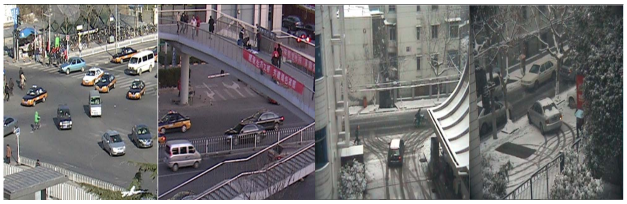

Figure 1. Examples of CIF video frames in the experiments

A. Analysis of Reference Frames Used in Transcoding

To clearly and objectively present selected reference frame in different categories, the reference frame number is set to 5 during analysis of the experiments, where the modeled background frame is treated as the long-term reference frame. The selected times of each reference frame in MBs is respectively recorded to calculate the percentage among BMs, FMs and FBMs. This analysis can indicate the most and lowest used reference frame in surveillance transcoding, which points out redundant reference frames to diverse categories. As Fig. 2 shows, the first reference frame is necessary for all the MBs, and the long-term reference frame also takes up a large percent for BMs and FBMs. Furthermore, the first, second and the long-term reference frames together take up about 83% in BMs or FBMs; while in FMs, the first two reference frames takes up about 86%. Therefore, the first two and the long-term reference frames are necessary for transcoding BMs and FBMs, whereas only the first two reference frames are indispensable to FMs.

B. Analysis of the Proper Motion Search Range

To avoid performance loss in video coding and transcoding, motion search range for each MB should be larger than the "Real MVD". Here the so-called Real MVD means the difference between the predicted motion vector (PMV) from the neighboring MBs and the best matched motion vector. Therefore, the distribution of the Real MVDs can indicate the proper size of motion search range. Fig. 3 shows the distribution of the Real MVDs for BMs, FMs and FBMs. It is shown that, 99% of the Real MVDs are less than 1 integer pixel in BMs, so the transcoding integer motion search range can be set to 1. For FMs and FBMs, the ratio of larger Real MVDs cannot be neglected, because the transcoding efficiency is more easily influenced by these larger Real MVDs rather than smaller Real MVDs. From the statistics, we can see that, in surveillance video transcoding, motion vector in BMs is close to the predicted motion vector. In FMs and FBMs, nevertheless, the predicted motion vector might be quite different from the decoded MV, so the motion search range should be narrowed according to the difference between them.

C. Analysis of Prediction Modes in Transcoding

It is clearly that the used intra- and inter-prediction modes are entirely different among BMs, FMs and FBMs. Used prediction modes in MBs are recorded to get the proportion of each prediction mode, then barely selected mode type in transcoding can be revealed. The analysis of prediction modes enlighten us how to refine candidate mode types, which can reduce complexity in MD and ME. As Fig. 4 shows, SKIP and inter 16×16 prediction modes are selected almost 96% in BMs. Therefore, the small-size mode is forbidden in BMs. For FMs and FBMs, however, 8×8, 8×4, 4×8 and 4×4 prediction modes take over 10% and 39%, so we treat one small-size mode as candidate mode when the decoded MB use it. Another interesting finding is that the I16MB intra prediction in FBMs is barely used, because FBMs contain both foreground and background pixels. These distributions indicate that, the used modes in surveillance video transcoding are sharply varied in background and foreground.

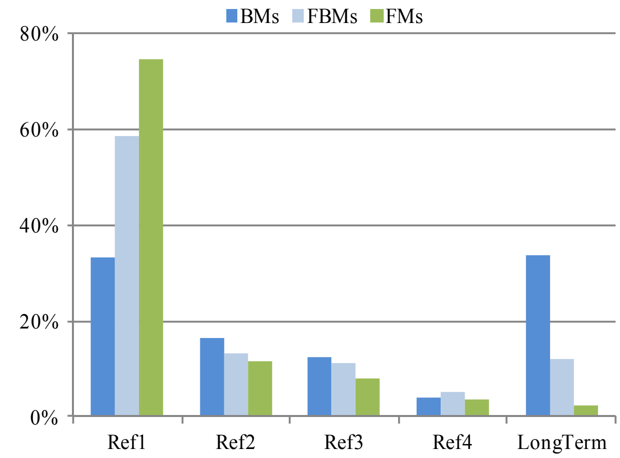

Figure 2. The distribution of selected reference frames

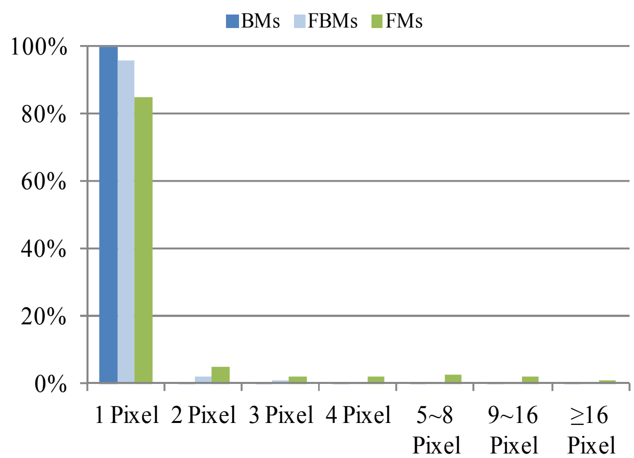

Figure 3. The distribution of the Real MVDs

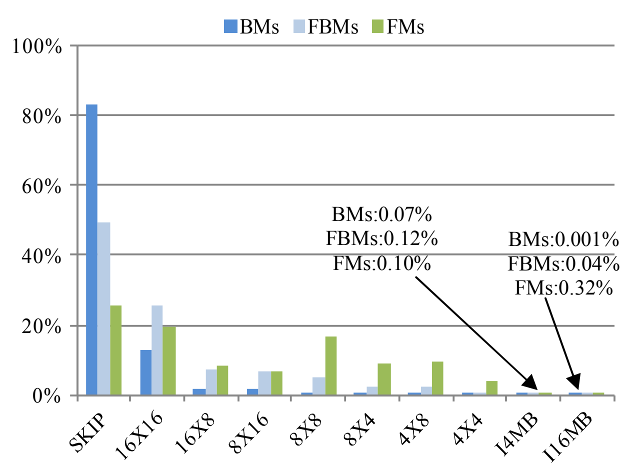

Figure 4. The distribution of used intra- and inter-prediction modes

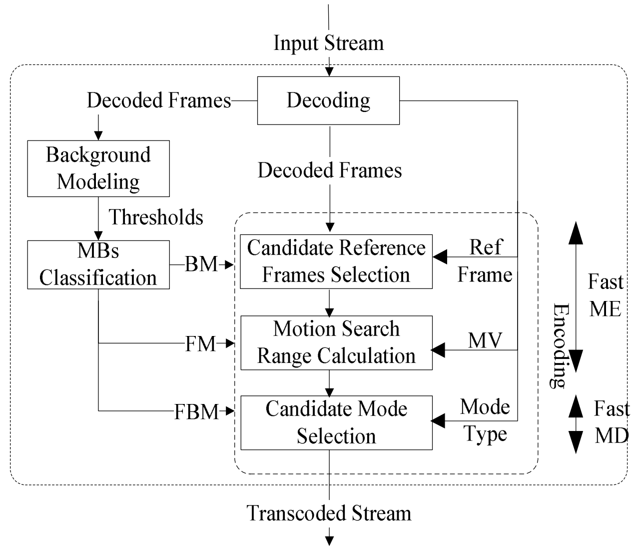

Figure 5. The framework of the proposed method

III. THE PROPOSED METHOD

A. *The Framework*

As shown in Fig. 5, FP-BGT consists of the following modules: the Background Modeling module that is used to generate the background frame, the MB Classification module that classifies MBs into BMs, FMs and FBMs, the Candidate References Frame Selection module that aims at reducing the number of reference frames, the Motion Search Range Calculation module that is used to reduce motion search range, and the Candidate Mode Selection module that refines the mode decision procedure. In this framework, the fast transcoding process includes the following steps:

Firstly, a background frame is generated from the original decoded frames. Then the modeled background frame is used to classify MBs into FMs, FBMs and BMs.

After that, BMs, FMs and FBMs are respectively processed as follows:

1) Candidate Reference Frames Selection: To reduce the number of reference frames, only most frequent reference frames are used for each MB category.

2) Motion Search Range Calculation: Different motion search ranges are calculated according the MBs category, so as to further decrease ME time.

3) Candidate Mode Selection: Mode refinement is used to decrease the complexity in MD. Unnecessary mode types are forbidden in this step to reduce ME and MD complexity.

According to the distribution of reference frames, MVDs and used prediction modes analyzed in Sec. 2, FMs should be transcoded in a meticulous way such that relatively complex transcoding strategies are employed for FMs. In contrast, FBMs are located between BMs and FMs, and then moderate-complexity ME and MD methods will be used for fast transcoding. At last, BMs should be treated by large simplified transcoding method, because BMs are always static in surveillance video.

B. Background Modeling and MB Classification

To a transcoder, added techniques should introduce low complexity and achieve acceptable accuracy. Therefore, in this paper, the modeled background frame is generated by a low-complexity Running Gaussian Average method [11]. To classify MBs kind, each MB is divided into sixteen 4×4 sub-blocks, and each 4×4 sub-block will be judged as “foreground” or “background” by comparing the residual value with the thresholds. Let C_i and BG_i denote one pixel value in MB and the corresponding value in background frame, respectively, β be the threshold for each pixel which is used as a constant value, and then the judgment of each sub-block can be expressed as the follow decision rule:

$$\sum_{i=1}^{16}\left(\left(C_i - BG_i\right) - \beta\right) > 0 \qquad (1)$$

A 4×4 sub-block is marked as “foreground” if the residual value exceeds the threshold value β. Then a MB is categorized into BMs, FBMs or FMs by the “foreground” number of 4×4 sub-blocks. One MB will be classified as FM if the number is greater than 8, and classified as BM if none of 4×4 sub-blocks is “foreground”. Others are FBMs.

C. Reference Frame Refinement

In reference frame buffer, to BMs, the long-term and nearby reference frame contribute most to transcoding efficiency, while other reference frames take low correlation with background. To the contrary, FMs and FBMs contain strong correlation among these reference frames. According to correlation and analysis of three categories, the simplified candidate reference frame pool is shown in Table 1.

TABLE I. CANDIDATE REFERENCE FRAME

Three categories	BMs	FMs	FBMs
Candidate reference frame	First, Long-term	First, Second, Long-term, Reference frame from decoded MB	

For BMs, only two reference frames are added to candidate pool; while more than three reference frames are employed in FMs and FBMs. Due to the decrease of candidate reference frame in BMs, the redundant computation in ME is obviously reduced. On the other hand, transcoding efficiency is retained by moderate simplification in FMs and FBMs.

D. Modified Motion Estimation

In this paper, the *PMVDs* are the difference value between the motion vector from the decoder (MV_{dec}) and the PMV in the encoder. Actually, in the decoder, one MB often contains multiple MV_{dec} for different SMBs. Thus PMV will be compared with each MV_{dec} to find the maximum *PMVD* value ($PMVD_{max}$). One $PMVD_i$ vector is calculated by

$$PMVD_i(X,Y) = \left(PMV_i(X) - MV_{dec}(X), PMV_i(Y) - MV_{dec}(Y)\right) \qquad (2)$$

Given all these *PMVDs*, the $PMVD_{max}$ vector will be recorded and used in Motion Search Range Calculation module. Equation 3 and 4 show value of $PMVD_{max}$ in X- and Y-offset:

$$PMVD_{max}(X) = MAX(PMVD_0(X), PMVD_1(X), \ldots.) \qquad (3)$$

$$PMVD_{max}(Y) = MAX(PMVD_0(Y), PMVD_1(Y), \ldots.) \qquad (4)$$

$PMVD_{max}$ indicates the max distance between MV_{dec} and PMV, which objectively enlighten us a proper motion search range. According to the Real MVDs distribution analysis and $PMVD_{max}$ value, the adaptive motion search range calculation algorithm is presented in Fig. 6. R_{org} is the original motion search range, and R_{mod} means modified range. $PMVD_{max}$(i) is equal to $PMVD_{max}$(X) or $PMVD_{max}$(Y).

Input value:
R_{org}: the original input motion search range;
$PMVD_{max}$(i): the maximum value of *PMVD*(X) or *PMVD*(Y);
d_1: the extra search range for FBMs;
d_2: the extra search range for FMs;
Init value:
The output value R_{mod} is initialized to R_{org} at first.
Calculation procedure:
For each SMB, ensuring its classification first, then
A. To BMs, R_{mod} is set to 1.
B. To FMs and FBMs:
If the SMB is a FBMs, set *Flag to 0*; Else set *Flag* to 1.
if ($PMVD_{max}$ (i)== 0)
$R_{mod} = 1 + 1 \times Flag$;
else if ($PMVD_{max}$ (i)== 1)
$R_{mod} = d_1 + d_2 \times Flag$;
else if ($PMVD_{max}$ (i)<= R_{org} - d_1 - d_2)
$R_{mod} = PMVD_{max}(i) + d_1 + d_2 \times Flag$;
else
$R_{mod} = R_{org}$;
Output value: R_{mod}

Figure 6. Adaptive motion search range calculation

Fig. 6 shows following search range calculation principles: For BMs, motion search range is set to 1 for motionless feature, and it introduces ignorable performance loss. To FMs and FBMs, nevertheless, motion search range narrowing should be methodic to avoid large transcoding quality decrease. It is obviously that motion search range in FMs should be larger than in FBMs, for high motion characteristic in FMs. Therefore, motion search range is calculated as follows: If $PMVD_{max}$(i) is equal or less than 1, that means the SMB is low moving, then motion search range can be confined in small range, such as between 1 and 4. When $PMVD_{max}$(i) is larger than 1 and less than R_{org}-(d1+d2), motion search range of SMB should be adaptively varying with $PMVD_{max}$(i) to avoid large quality loss and reduce redundant computation as more as possible. So motion search range is set to ($PMVD_{max}$(i)+d1+d2) on FMs and ($PMVD_{max}$(i)+d1)on FBMs in this case.

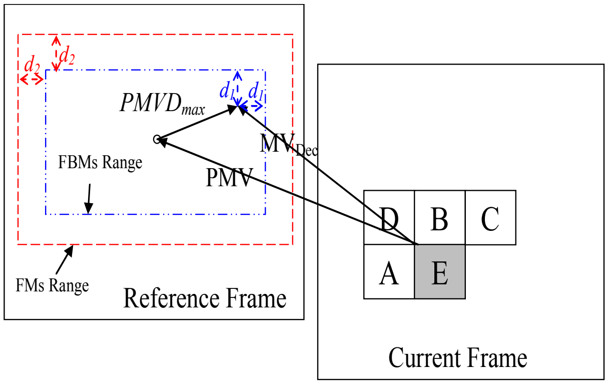

Figure 7. PMVD and modified motion search range

Besides, through experiments, we found that low quality increase will be contributed when d1 and d2 exceed 2, then both of d1 and d2 are set to 2 in this paper. At last, for severe moving SMBs, motion search range is equal to original range. Fig. 7 shows an example for *PMVDs* and finally modified motion search range in FMs and FBMs.

The motion search procedure occupies the great mass of ME computation, and Motion Search Range Calculation module methodically narrow search range of diverse categories. More important, FMs and FBMs which impact large to transcoding efficiency are carefully treated by flexible search range calculation. The facts above decide that this module significantly reduce redundant computation with a slight quality loss.

E. Mode Decision Refinement

To static region, the large size prediction modes will be mostly selected, which means smaller prediction modes can be forbidden in BMs. But the smaller mode should be employed to FMs and FBMs, for size of moving objects generally being smaller than a MB size.

The final mode decision refinement method is clearly listed in Table 2, where S denotes the lowest size of decode mode is equal or greater than 8×8 block size. In this paper, the candidate prediction mode pool contains three levels, and each level achieves various degree of mode size. As shown in Table 1, only SKIP and 16×16 modes are allowed in BMs, regardless of other factors. To FMs and FBMs, 8×4, 4×8, 4×4 modes are optional, and the lowest size of decode mode decides which one is added to candidate pool. I16MB mode is discarded in FBMs, for unsuitable to FBMs area containing background and foreground.

TABLE II. MODE DECISION REFINEMENT METHOD

Decode Mode	FMs	FBMs	BMs
S	level1, I16MB	level2	16×16, SKIP
8×4	level2, I16MB		
4×8			
4×4	level3, I16MB	level3	
I4MB			

*S: Decode Mode Size={ SKIP,16×16,16×8,8×16,8×8, I16MB},
*level1={SKIP,16×16,16×8,8×16,8×8,I4MB},
*level2={SKIP,16×16,16×8,8×16,8×8,I4MB,8×4,4×8},
*level3={SKIP,16×16,16×8,8×16,8×8,I4MB,8×4,4×8,4×4}

Figure 8. Examples of SD video frames

IV. EXPERIMENTAL RESULTS

A. Experimental Setup

To verify proposed method, transcoding efficiency comparison and time saving proportion of FP-BGT compared with FDFE and BGT respectively are included in our experiments. FDFE is the traditional cascade decoder-encoder scheme without extra techniques, and BGT is the background modeling based transcoding method [10].

FDFE, BGT and the proposed FP-BGT are implemented on H.264 reference software JM17.2, and the common testing parameter setups are listed in Table 3. Surveillance streams are transcoded by Baseline Profile of H.264/AVC and only the first frame is set to IDR frame type in this paper.

The long-time surveillance videos [12] captured by static cameras (shown in Fig. 1 and Fig. 8) with difference motion characteristics and foreground proportion are well to evaluate FP-BGT. The experimental results are listed in Table 4, including the efficiency and time saving of FP-BGT respectively compared with FDFE and BGT.

TABLE III. PARAMETERS CONFIGURATION

Parameter	Value	Parameter	Value
Porfile	Baseline	Used MODE	ALL
Rate Control	Disable	Framerate	25
Entropy Coding	UVLC	Frame Structure	IPPP
Search Range	32	IntraPeriod	0
RD Optimization	High	SAD Method	Hadamard
Motion Search	Fast Full	Reference Num	5

TABLE IV. PERFORMANCE COMPARISON

FP-BGT vs	PSNR(△dB)		Bitrate(△%)		TimeSave(△%)	
	FDFE	**BGT**	**FDFE**	**BGT**	**FDFE**	**BGT**
CIF(352×288)						
Crossroad	1.23	-0.11	-27.77	3.21	96.50	96.50
Overbridge	0.82	-0.10	-24.04	3.90	95.73	95.72
Snowgate	1.90	-0.09	-60.63	5.66	98.06	98.05
Snowroad	1.10	-0.10	-40.49	5.64	97.65	97.65
Avg.	**1.26**	**-0.10**	**-38.23**	**4.60**	**96.98**	**96.98**
SD(720×576)						
Bank	1.34	-0.09	-55.93	5.02	98.15	98.16
Crossroad	1.11	-0.09	-26.25	3.16	97.13	97.13
Office	0.42	-0.08	-17.06	3.59	93.84	93.86
Overbridge	1.74	-0.05	-62.75	3.13	97.72	97.71
Avg.	**1.15**	**-0.08**	**-40.50**	**3.72**	**96.71**	**96.72**

B. Transcoding Efficiency and Complexity Analysis

As shown in Table 4, compared to FDFE, FP-BGT achieves an average PSNR gain of 1.26dB on CIF and 1.15dB on SD surveillance stream, with totally 96.98% and 96.71% time saved meanwhile. Due to the modeled background frame technique, FP-BGT realizes high transcoding performance on surveillance video, and the computation is sharply reduced by proposed fast algorithm. Besides, compared with BGT, FP-BGT achieves slight PSNR gain loss: averagely 0.10dB quality loss comes up on CIF stream and 0.08dB loss on SD stream. The classification strategy and distinguishing transcoding procedure help to retain high efficiency. On the base of that, redundant complexity is greatly decreased: almost 96.98% and 96.72% transcoding total time is saved on CIF and SD surveillance stream.

Experimental results indicate that FP-BGT is proved to be a fast and performance-maintained transcoding method. The removing redundant prediction mode, adaptively narrowing motion search range and reducing candidate reference frame techniques sharply increase the transcoding speed. Besides, the classification strategy mainly reserves the quality originating from modeled background frame.

C. Additional Experiments

Additionally, we respectively analyze the contribution ratio of proposed algorithms in Table 5. The proposed motion search range results in significant total-time saving, because it reduces 85% of the search points on average, as shown in Table 6. Besides, Candidate Reference Frame Selection module achieves more than 31% transcoding time saving and ensures 0.036dB quality loss, and Candidate Mode Selection module averagely decrease 13% time.

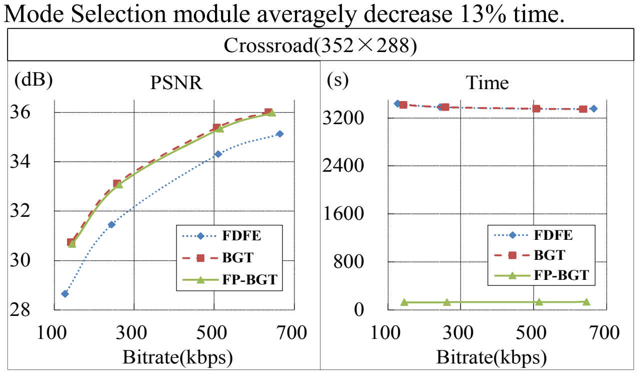

Figure 9. RD-curves and time saving example

TABLE V. CONTRIBUTION DISTRIBUTION IN PROPOSED METHOD (%)

Seq.	Crossroad	Overbridge	Snowroad	Snowate	**average**
Search Range	93.91	93.02	94.77	94.90	**94.15**
Mode Reduced	12.29	10.69	16.73	14.00	**13.43**
Ref.Refinement	34.05	31.17	47.89	42.08	**38.80**

TABLE VI. SEARCH POINTS REDUCTION (%)

Crossroad	Overbridge	Snowroad	Snowate	**average**
84.88	84.17	85.67	85.93	**85.16**

V. CONCLUSION

In this paper, we propose a fast surveillance video transcoding method based on background modeling. One significant advantage of our transcoding method is that it can achieve high efficiency and significantly decrease redundant computation by using category-specific transcoding strategies. That is, different transcoding procedures are used for different categories of MBs. Moreover, an algorithm is also developed to exploit the decoded motion vector to adaptively calculate motion search range. Results show that compared with FDFE, the proposed method averagely maintains 1.26dB/1.15dB PSNR gain on CIF and SD surveillance stream, with more than 93% time saved. For future work, we will concentrate on accurate classification strategy and effective surveillance video analysis technology.

ACKNOWLEDGMENT

This work is partially supported by grants from National Basic Research Program of China under contract No. 2009CB320906, the Chinese National Natural Science Foundation under contract No. 61035001 and No. 61072095, and Fok Ying Dong Education Foundation under contract No. 122008.

REFERENCES

[1] J. Youn, M.T. Sun, and C.W. Lin, “Motion vector refinement for high performance transcoding,” IEEE Trans. Multimedia, vol. 1, pp. 30-40, Mar. 1999.

[2] Y. Shin, N. Son, N. D. Toan, and G. Lee, “Low-complexity heterogeneous video transcoding by motion vector clustering,” in Information Science and Applications (ICISA), 2010 International Conference on, pp. 1–6, April 2010.

[3] K.-T. Fung and W.-C. Siu, “Low complexity H.263 to H.264 video transcoding using motion vector decomposition,” presented at the IEEE Int. Symp. Circuits Systems, May 2005.

[4] H. Kalva, P. Kunselmann “Dynamic motion estimation for transcoding P frames in H.264 to MPEG-2 transcoders,” IEEE T. Consum. Electr., vol. 54, 2008.

[5] C. Wu, and Y. Lin, “Efficient inter/intra mode decision for H.264/AVC inter frame transcoding” Source: Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP 2009), 3697-700, 2009

[6] P. Zhang, Q.-M. Huang, and W. Gao, “Key techniques of bit-rate reduction for H.264 streams,” in Proc. Pacific-Rim Conf. Multimedia (PCM), 2004, pp. 985–992.

[7] X. Lu, A. M. Tourapis, P. Yin, and J. Boyce, “Fast mode decision and motion estimation for H.264 with a focus on MPEG-2/H.264 transcoding,” in Pro. 2005 IEEE Int. Symp. Circuits Syst., May 200.

[8] A. Vetro, H. Sun, and Y. Wang, “Object-based transcoding for scalable quality of service,” presented at the IEEE Int. Symp. Circuits and Systems (ISCAS’2000), Geneva, Switzerland, May 28–31, 2000.

[9] T. Hata, N. Kuwahara, T. Nozawa, and et al. “Surveillance system with object-aware video transcoder,” in International Workshop on Multimedia Signal Processing. Shangai, China (2005)

[10] X. Zhang, L. Liang, Q. Huang, and et al. “A background model based method for transcoding surveillance videos captured by stationary camera,” PCS, 2010.

[11] M. Piccardi, “Background subtraction techniques: a review,” in Proc. IEEE Int. Conf. Systems, Man, Cybernetics, 2004.

[12] ftp://124.207.250.92/public/seqs/video/Surveillance/ (by AVS member).