EFFICIENT MODE DECISION FOR NOISY VIDEO TRANSCODING

Anna Zhang[1,2,3], *Zhen Han*[1,2,3*], *Zhongyuan Wang*[1,2,3]

1. State Key Lab of Software Engineering, School of Computer,
Wuhan University, Wuhan, 430072, China
2. National Engineering Research Center for Multimedia Software, School of Computer,
Wuhan University, Wuhan, 430072, China
3. Hubei Province Key Laboratory of Multimedia and Network Communication Engineering, Wuhan
University, Wuhan, 430079, China

ABSTRACT

In many practical transcoding applications, such as video surveillance, the source videos are often contaminated by noise. The presence of noise not only results in poor compression efficiency and visual quality, but also imposes an adverse effect on the performance of subsequent video analysis tasks. Thereby it is very necessary to denoise the video. In this paper, we propose an efficient mode decision method for H.264 noisy video transcoding through analysing the effect of noise on mode decision. In the algorithm, we use the information available from previously decoded MBs to decide which modes can be overpassed with little loss to the rate-distortion performance. Experimental results show that our method saves the computational complexity nearly 65%, without noticeable rate-distortion(R-D) loss in comparison with the "Full-coding" (cascaded transcoding) method.

***Index Terms*—**Video transcoding, Noisy video, Mode decision, Video denosing.

1. INTRODUCTION

Video transcoding is a core technology for providing universal multimedia access by converting one compressed video to another. Over the past years, many approaches on transcoding including bit-rate transcoding, spatial/temporal transcoding, error-resilient transcoding, standards transcoding have been proposed.

In all of these approaches, the input compressed videos to the transcoder are noiseless. However, in many practical transcoding applications, the source videos are often contaminated by noise during acquisition, storage and transmission. Directly transcoding a video without considering noise could deteriorate its visual quality due to magnifying noise. Thereby for noisy video transcoding applications, such as video surveillance, mobile video communication, low-complexity video denoising techniques should be adopted in conventional transcoder to transcode noisy videos to de-noised ones.

Current video surveillance and mobile video communication systems with lower video bitrates and resolutions usually adopt common video codes such as H.264/AVC to compress captured videos. In order to reduce the complexity of mode decision in H.264/AVC, many fast mode decision algorithms for transcoding based on the correlation between macro block modes in the input and output videos have been proposed.

Several fast mode decision methods for MPEG-2 to H.264 transcoding have been proposed in [1-3]. The method proposed in [1] uses directional features gathered from the MPEG-2 video to predict the intra modes in the H.264 video. The cascaded algorithm proposed in [2] utilized the DCT coefficients to determine the intra mode size directly and transform the 8×8 DCT coefficients to 4×4 integer transform coefficients to reduce the complexity. The paper [3] proposes a mode skipping rule of intra prediction mode decision based on the relationship between DCT energy trend in the decoded MB of MPEG-2 bitstream and the intra prediction modes of H.264/AVC. A Fast inter mode decision algorithm for P frames in MPEG-2 to H.264/AVC transcoding was proposed in [4]. The algorithm presented in [5] realize macro block type and motion vector information, proposing a zero-block inter/intra mode decision scheme for MPEG-2 to H.264 transcoding. The literature [6-7] present a fast macro block mode decision algorithm based on machine learning techniques for MPEG-2 to H.264 P-frame transcoding.

Other mode decision methods for H.264 transcoding are proposed in [8-11]. An efficient inter/intra mode decision for H.264/AVC frame-skipping transcoding was proposed in [8]. The literature [9] proposes a direct inter-mode selection algorithm for P-frames in fast homogeneous H.264/AVC bit-rate reduction transcoding. The literature [10] makes use of several directional measures that are based not only on the pixels in an image block but also on the neighboring pixels. Two types of mode mapping algo-

The research was supported by the National Natural Science Foundation of China (61671332), the Natural Science Fund of Hubei Province (2015CFB406,2016CFB573), the Applied Basic Research Program of Wuhan City (2016010101010025).* corresponding author.

rithms for H.264/AVC to HEVC transcoding are proposed in [11]. In the first solution, a single H.264/AVC coding parameter is used to determine the outgoing HEVC partitions using dynamic thresholding. The second solution uses linear discriminant functions to map the incoming H.264/AVC coding parameters to the outgoing HEVC partitions.

Compared with mode decision in traditional video transcoder, mode decision in noisy video transcoder is more complicated because of weaker correlation between the input noisy video and output denoised video. The methods above are not suitable for noisy video transcoding. Therefore it is very important to develop specially-designed mode decision method for noisy video transcoding applications.

In this paper, we propose a novel mode decision method for H.264 to H.264 noisy video transcoding through analysing the effect of noise on mode decision. The basic mode decision rules for noisy video transcoding are proposed firstly and then the specific inter and intra mode decision methods are presented respectively. The inter/intra mode decision method predict the basic MB mode type choice from three candidates {intra, inter, skip}. The inter sub-mode decision selects the block sizes for H.264/AVC motion compensation from a set of seven block sizes (i.e., P16×16, P16×8, P8×16, P8×8, P8×4, P4×8, and P4×4) for inter-coded MBs. The intra sub-mode decision selects the block sizes from {I16×16, I8×8, and I4×4} and the prediction directions for intra-coded MBs.

This paper is organized as follows. Section 2 analyses the effect of noise on mode decision and the MB mode correlation between the input noisy video and output denoised videos. Section 3 proposes the mode decision method. In section 4, the experimental results are given. Section 5 makes a conclusion.

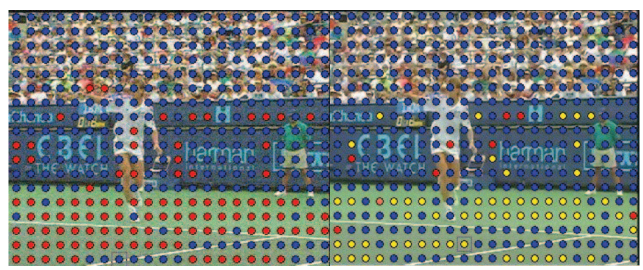
2. PROPOSED CODING SCHEME

The best mode of one MB depends on its spatial and temporal correlation. If the spatial correlation is stronger than the temporal correlation, the best mode is intra mode. Otherwise, the best mode is inter mode. The intra/inter prediction errors of one MB can be used to approximately estimate its spatial/temporal correlation. The smaller the intra/inter prediction errors are, the stronger the spatial/temporal correlation is. Therefore, the effect of noise on the intra/inter prediction errors should be considered for intra/inter mode decision in noisy video transcoder.

Let ΔE_{intra} and ΔE_{inter} represent the increase of intra and inter prediction errors of one MB caused by noise separately. Let $N_{current}(x,y)$, $N_{ref,intra}(x,y)$ and $N_{ref,inter}(x,y)$ represent the additive noise located in (x,y) in current MB, intra prediction reference block and inter prediction reference block respectively. Because the additive noise is irrelevant to original MB pixels, ΔE_{intra} can be approximately estimated according to the difference of $N_{current}(x,y)$ and $N_{ref,intra}(x,y)$. The bigger the difference is, the bigger ΔE_{intra} is. Similarly, ΔE_{inter} can be approximately estimated according to the difference of $N_{current}(x,y)$ and $N_{ref,inter}(x,y)$.

The inter prediction reference block is the noisy block in reference frame. The intra prediction reference block is more complicated than the inter prediction reference block. H.264 supports different intra prediction modes: DC prediction mode, vertical prediction mode, horizontal prediction and diagonal prediction modes. For DC prediction and diagonal prediction modes, the pixels in intra prediction reference block are the weighted average of pixels of adjacent blocks. For example, DC prediction mode uses only one reference value to predict the entire block and this value is obtained by averaging the spatially neighbour pixels of the block.

The inter prediction reference block is the noisy block in reference frame, so the level of $N_{ref,inter}(x,y)$ is the same as $N_{current}(x,y)$. For the intra prediction reference block, the weighted average can be regarded as the low pass filter for noisy block, so the level of $N_{ref,intra}(x,y)$ is more likely smaller than that of $N_{ref,inter}(x,y)$ and ΔE_{intra} is smaller than ΔE_{inter}. As shown in Fig. 1, red, blue and yellow blocks mean intra-coded MB, inter-coded MB and SKIP-coded MB separately. This means that intra-coded MB in noisy video might be inter-coded in denoised video but inter-coded MB in noisy video might not be intra-coded in denoised video.



(a) **(b)**

Fig.1 The 32th frame of Stenfan sequence size of 352×288: (a) the noisy sequence; (b) the denoised sequence.

Whether one intra-coded MB in noisy video should be inter-coded in denoised video or not depends on the motion activity of this MB. If one intra-coded MB in noisy video belongs to motion active region and its temporal correlation is rather weak, it should be still intra-coded in denoised video.

3. PROPOSED MODE DECISION METHOD

According to above discussion, the basic mode decision rules for noisy video transcoding can be summarized as:

1. If one MB is inter-coded in noisy video, it should be inter-coded in denoised video.

2. If one MB is intra-coded in noisy video and belongs to motion active region, it should be intra-coded in denoised video; if it belongs to motion non-active region, it should be inter-coded in denoised video. More details see section 3.2.

3.1. Inter mode decision

According to the first point of the basic mode decision rules, inter mode should be mapped to inter mode in noisy video transcoder. For inter-inter mode mapping, the consistency of motion vectors should be considered. Weaker the consistency of motion vectors is, smaller the block-size is. Owing to randomness of noise, the consistency of motion vectors in noisy video is weaker than that in denoised video. This means that smaller block-size MB in noisy video might be bigger block-size in denoised video but bigger block-size MB in noisy video might not be smaller block-size in denoised video. Therefore block-size of MB in noisy video is mapped to same and bigger block-size. For example, P16×8 is mapped to {P16×16, P16×8, SKIP}.Fig.2 shows the specific mapping rules for inter modes.

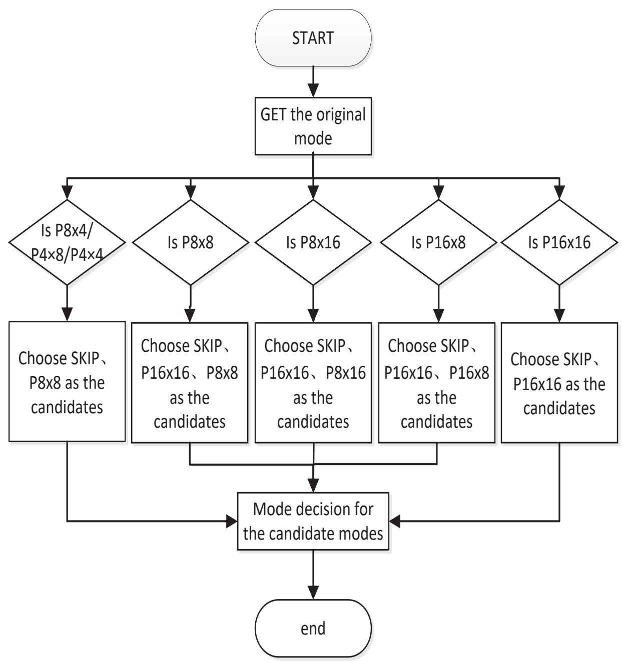


Fig.2 Flowchart of inter mode decision method

3.2. Intra mode decision

According to the second point of the basic mode decision rules, intra mode should be mapped to inter or intra mode in noisy video transcoder according to motion activity of MB. One simple way to estimate the motion active region is to employ the binary differences between adjacent denoised frames. We take the 16×16 pixels into consideration to determine whether one MB belongs to the motion active region. If the percentage of pixels in the motion region is more than threshold T_1, we consider that the MB belongs to the motion active region and intra mode should be mapped to intra mode. If the percentage is less than threshold T_2, we consider that the MB belongs to the motion non-active region intra mode should be converted to inter mode. Otherwise, we consider that the MB belongs to the region whose motion activity could be determined and intra/inter modes should all be chose as the candidate modes.

For intra-intra mode mapping, the consistency of prediction directions should be considered. Weaker the consistency of prediction directions is, smaller the block-size is. Similar to Inter-Inter mode decision, the consistency of prediction directions in noisy video is weaker. Therefore block-size of MB in noisy video is also mapped to same and bigger block-size. For example, I8×8 is mapped to {I16×16, I8×8}.

Table 1. Intra mode decision method

Original mode	$P > T_1$	$P < T_2$		Other
		P=0	Other	
I4×4	I4×4 I8×8 I16×16	SKIP P16×16	P4×4 P8×8 P16×16	P4×4 P8×8 P16×16 I4×4 I8×8 I16×16
I8×8	I8×8 I16×16	SKIP P16×16	P8×8 P16×16	P8×8 P16×16 I8×8 I16×16
I16×16	I16×16	SKIP P16×16	P16×16	P16×16 I16×16

For intra-inter mode conversion, intra-coded MB in noisy video belongs to motion non-active region. If the percentage of pixels in the motion region is 0, we consider that the MB belongs to the static region and it should be mapped to {P16×16, SKIP}. Otherwise, other inter modes except SKIP mode are considered. For example, I8×8 is mapped to {P8×8, P16×16}.

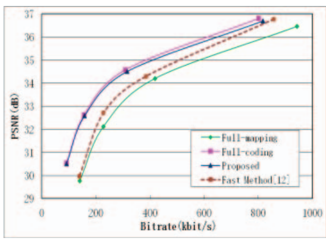
The specific mapping strategy is shown in Table 1. The percentage of pixels in the motion region is marked by P.

4. EXPERIMENTAL RESULTS

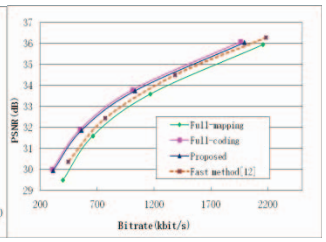
In our experiment, a noisy video transcoder was implemented using the version 18.4 of reference JVT software. The length of GOP is 30 and iframe rate is 30 frames per second. QP is set 24 to 44 with a step of 4. Other parameters are all default settings. Several video sequences with additive Gaussian noise are used as test sequences. We set the noise variance of Y to 10. The video sequence is denoised by mean filter after decoding in transcoder. The thresholds {T_1, T_2} are set to {50%, 40%}.

We chose the cascaded transcoding method marked by "Full-coding" , the mode-reusing method marked by "Full-mapping" as reference methods and a fast encoding method marked by "Fast method"[12] for comparisons because of no previous similar work. Our algorithm is marked by "Proposed". Among them, "Full-coding" method gives the best R-D performance and thus used for the benchmark, while "Full-mapping" method represents the fastest speed without extra computation on mode decisions.

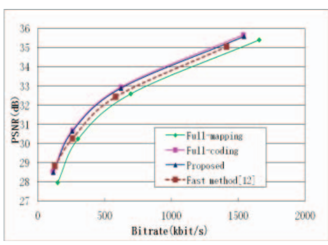
Experimental results of the PSNR on the different bit rates curves are depicted in Fig. 3. As depicted in Fig.3, our algorithm has a significantly better R-D performance than "Full-mapping" method and performs similarly as "Full-coding" method.



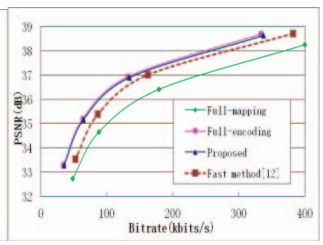
(a)Foreman



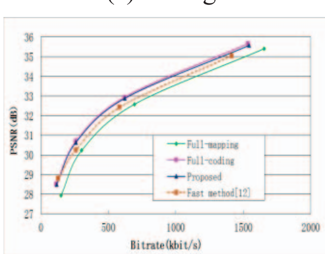
(b) Football



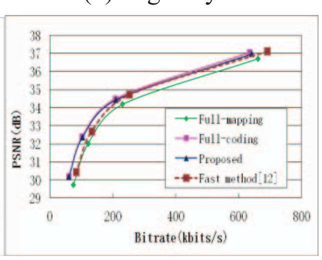
(c) Coastguard



(d) Highway



(e) Hall



(f) News

Fig.3 Experimental results on rate distortion cures.

Experimental results of the speed are shown in Table 2. ∆Time indicates the percentage of the total transcoding time change compared with "Full-coding" method. Their positive values mean increment while negative values mean decrement. From Table 2, we can see that our algorithm gives the excellent results. The running time of "Full-mapping" method is reduced by about 75%, which is the fastest without extra computations on mode decisions. The running time of "Proposed" method is reduced by about 65%, which is significantly faster than "Full-coding" method.

Table 2. Speed performance

Video sequence	Method	∆Time
Foreman	Full-mapping	-78.24%
	Fast method[12]	-42.52%
	Proposed	-62.97%
Football	Full-mapping	-79.35%
	Fast method[12]	-49.16%
	Proposed	-63.82%
Coastguard	Full-mapping	-79.03%
	Fast method[12]	-46.49%
	Proposed	-60.61%
Highway	Full-mapping	-79.42%
	Fast method[12]	-47.33%
	Proposed	-60.01%
Hall	Full-mapping	-80.01%
	Fast method[12]	-58.43%
	Proposed	-70.74%
News	Full-mapping	-80.61%
	Fast method[12]	-63.05%
	Proposed	-70.41%

5. CONCLUSION

In this paper, we propose a novel mode decision method for H.264 noisy video transcoding through analysing the effect of noise on mode decision. Compared with the "Full-coding" method, our algorithm achieves almost same R-D performance, meanwhile reduces the transcoding time by about 65%. Compared with the "Full-mapping" method without extra computations on mode decisions, at the same bit rates, our algorithm achieves significantly better R-D performance. This is very important for noisy video transcoding applications, such as video surveillance, mobile video communication.

6. REFERENCES

[1] Tang Q, Nasiopoulos P. Efficient motion re-estimation with rate-distortion optimization for MPEG-2 to H. 264/AVC transcoding[J]. Circuits and Systems for Video Technology, IEEE Transactions on, 2010, 20(2): 262-274.

[2] Wei G, Yang G, Zhang Y, et al. Fast MPEG-2 to H.264 Transcoding for Intra Frame in Transform Domain[J]. International Journal of Digital Content Technology & Its Applications, 2013, 7:606-613.

[3] Liu X, Yoo K Y, Kim S W. Low complexity intra prediction algorithm for MPEG-2 to H. 264/AVC transcoder[J]. Consumer Electronics, IEEE Transactions on, 2010, 56(2): 987-994.

[4] Wang P, Zhang X, Huang H. Fast Inter Mode Decision Algorithm Based on Mode Mapping and Decision Tree for P Frames in MPEG-2 to H.264/AVC Transcoding[J]. Journal of Signal Processing Systems, 2013, 79(1):33-43.

[5] Su C J, Lin Y. Zero-block inter/intra mode decision for MPEG-2 to H. 264/AVC inter P-frame transcoding[J]. IET Image processing, 2010, 4(6): 494-504.

[6] Fernández-Escribano G, Kalva H, Cuenca P, et al. A fast MB mode decision algorithm for MPEG-2 to H. 264 P-frame transcoding[J]. Circuits and Systems for Video Technology, IEEE Transactions on, 2008, 18(2): 172-185.

[7] Fernández-Escribano G, Kalva H, Martínez J L, et al. An MPEG-2 to H. 264 video transcoder in the baseline profile[J]. Circuits and Systems for Video Technology, IEEE Transactions on, 2010, 20(5): 763-768.

[8] Jiang J H, Lin Y. Efficient mode decision for H.264/AVC frame-skipping transcoding[J]. Electronics Letters, 2011, 47(11):650-651.

[9] Song B, Liu H, Qin H, et al. A Direct Inter-Mode Selection Algorithm for P-Frames in Fast H.264/AVC Transcoding[J]. Ieice Transactions on Communications, 2012, E95.B(6):2101-2108.

[10] Kim H, Lim S, Yu S. Fast intra-mode decision method for luma and chroma blocks for H.264/AVC[J]. Multimedia Tools & Applications, 2015, 74(13):4641-4653.

[11] Peixoto E, Shanableh T, Izquierdo E. H. 264/AVC to HEVC video transcoder based on dynamic thresholding and content modeling[J]. Circuits and Systems for Video Technology, IEEE Transactions on, 2014, 24(1): 99-112.

[12] Lee J Y, Park H W. A Fast Mode Decision Method Based on Motion Cost and Intra Prediction Cost for H.264/AVC[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2012, 22(3):393-402.