WYNER-ZIV TO H.264 VIDEO TRANSCODER FOR MOBILE TELEPHONY

J. L. Martinez[1], G. Fernandez-Escribano[1], H. Kalva[2], W.A.C. Fernando[3], A. Garrido[1]
[1]Instituto de Investigación en Informática de Albacete. Universidad de Castilla-La Mancha, Albacete. Spain
{joseluismm,gerardo,antonio}@dsi.uclm.es
[2]Department of Computer Science and Engineering. Florida Atlantic University, Boca Raton, FL. USA. hari@cse.fau.edu
[3]Center Communications Systems Research, University of Surrey, GU2 7XH, United Kingdom. W.Fernando@surrey.ac.uk

***Abstract*— Wyner-Ziv (WZ) coding has been receiving attention from the research community as it offers a more symmetric complexity in video communication by reducing the complexity of video encoding. This encoding paradigm, however, increases the complexity of video decoding. This paper presents a novel video communication architecture that exploits the low complexity WZ encoding without being burdened by high complexity WZ decoding. At the core of the proposed architecture is a WZ to H.264 transcoder that converts WZ video to H.264 to reduce the amount of resources necessary for decoding. This architecture enables low complexity encoding and decoding on mobile devices. Using very simple WZ encoders and H.264 decoders, a communication system will only need very low complexity devices at both the sender and receiver sides. Such low complexity communication devices are highly desirable for practical consumer applications. The paper also presents the low complexity algorithms for WZ to H.264 transcoding. The time reduction of the improved WZ/H.264 transcoder offers a very acceptable 86% complexity reduction with negligible Rate-Distortion loss.**

I. INTRODUCTION

In mobile to mobile video communications, both the transmitter and receiver devices may not have necessary computing power, resources or complexity constraints to perform complex video algorithms (both coding and decoding). Traditional video codecs typically have highly complex encoders and quite less complex decoders. Today the H.264/AVC [1] standard is being widely used due to its better performance. The solution adopted in these mobile to mobile applications is to sacrifice quality and reduce the encoding complexity by using only the lower complexity encoding tools available n h.264; for example, intra 4x4 is typically not used. On the other hand, Wyner-Ziv (WZ) [2] coding, a particular case of *Distributed Video Coding* (DVC) [3] has received great interest from multimedia research community. This new coding paradigm, with promising results, offers low complexity encoders using more complex decoders, i.e., a reversal of the asymmetry in complexity compared to the traditional video coding approach.

To make WZ based solutions practical, this paper proposes a WZ to H.264 transcoder for low complexity video communication systems. The proposed system includes WZ encoders at the transmitter side while the receiver uses the H.264/AVC decoders there by reducing the complexity on both sides. The system also includes a transcoder device that converts the WZ video from the sender to H.264 video for lower complexity decoding at the receiver. Basically, both sending and receiving devices shift their complexity to the network resulting in less complex user devices. The addition of a transcoder in the network thus enables the low complexity end user devices. This scenario is depicted in Figure 1. A basic transcoder performs full WZ decoding procedure on the video signal from the transmitter and then encodes it to H.264. The transcoder thus has to handle two complex processes: WZ decoding and H.264 encoding. This paper presents algorithms for decreasing the complexity of this transcoder by reducing the complexity of the H.264 encoding stage. The key contributions of this paper are: 1) communication system to exploit the low complexity encoding in WZ coding, 2) low complexity MB mode decisions in WZ to H.264 transcoding, 3) identification of research issues in making practical use of WZ coding.

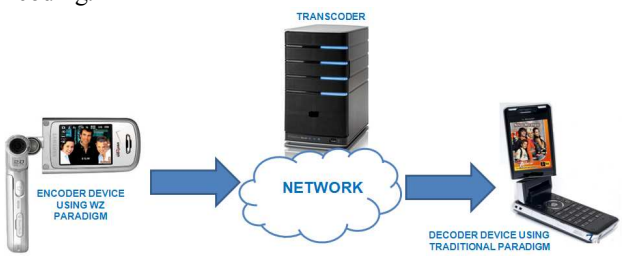

Figure 1. Multimedia Communication System using a WZ/H.264 transcoder

II. WZ/H.264 VIDEO TRANSCODER

Transcoding algorithms are not new in video research community [4] but, as far as we know, there are no WZ to H.264 transcoding approaches in the literature. The WZ to H.264 transcoding solution has a highly desirable feature – a low complexity encoder which is essential for mobile video communication. In order to achieve real time mobile-to-mobile video communications the transcoder device located in the network should provide real time performance while minimizing computational resources required. To achieve this complexity reduction, the proposed system reduces the complexity of the H.264 encoding task using some information

This work has been jointly supported by the Spanish MEC and European Commission FEDER funds under grants "Consolider Ingenio-2010 CSD2006-00046" and "TIN2006-15516-C04-02", and by JCCM funds under grant "PAI06-0106" and presented was developed as a part of VISNET II, a European Network of Excellence (http://www.visnet-noe.org), funded under the European Commission IST FP6 programme.

978-1-4244-2559-4/09/$25.00 ©2009 IEEE

extracted in the WZ decoding stage. In H.264 the Macro Block Code Decision and Motion Estimation are the more complex tasks that should be accelerated.

A. Inter Prediction: Motion Estimation

Motion Estimation is the most complex task done in traditional video codecs; it refers to the process of acquiring the motion vectors which correspond to a similar block in previously coded frames. In WZ coding similar motion estimation is done at the decoder to generate side information using the past and future frames. Using these motion vectors generated in WZ decoding stage can focus the search range and search windows form this previous calculated motion vectors and reduce and /or localize the motion search.

B. Inter Prediction: Mode Decision

H.264 supports motion compensation block sizes ranging from 16x16 to 4x4 with many options between the two; this method is known as tree structured motion compensation. The Mode Coded Decision is done based on the sum of the absolute differences between the current block and previous / past ones inside the search window and the length of the motion vectors.

In this paper, we focus our attention on a part of the inter-frame prediction: the macro block partition mode decision, one of the most computationally intensive tasks involved in the H.264/AVC encoding process. The proposed approach was developed based on the insights from our work on MPEG-2 to H.264 transcoding [5]. In order to accurately determine mode decision without the complex cost based approach, a decision tree using Machine Learning tools was obtained using statistical information (mean and variance) from the *residual side information* (difference between the estimated side information and the decoded frame) and the motion vectors generated in the side information generation. The key idea behind this approach is to determine fast MB coding mode decisions using easily computable features derived from WZ coding stage.

III. RESULTS

In order to evaluate our approach, the proposed transcoder has been implemented in the H.264/AVC reference software 13.2 version [1]. The training stage was done using 10 frames of flower and garden sequence in QCIF format, therefore, this sequence has been excluded in the performance evaluation. In the simulation, the Group of Pictures (GOP) is 2 where the key frames in the WZ coding were directly transcoded to I frames in H.264. Therefore the accelerated approach was implemented for P frame transcoding. Four different sequences with different motion / textures characteristic were used in the simulation and the Rate Distortion (RD) performance is shown in the Figure 2. As Figure 2 shows the RD performance is practically the same between our approach and the reference transcoder (full decoding followed by full encoding). In order to give more details about the RD performance, these RD differences have also been calculated using *Bjontegaard's* method [6]. These detailed differences are shown in Table 1. Table 2 shows the % of time reduction of our proposed transcoder per sequence and QP compared to the reference one; as you can see the average on time reduction is around 86%.

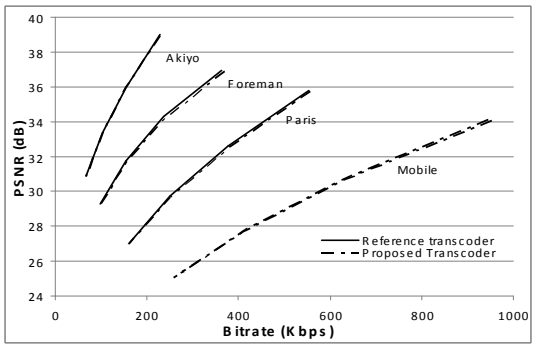

Figure 2. RD performance for QCIF sequences

Table 1. % Time Reduction of the proposed transcoder

% of Time Reduction (ΔTime) over 150 frames (five seconds) for QCIF sequences, within the range [28,40] for the Quantification Parameter (QP)																
Sequence	Akiyo				Foreman				Mobile				Paris			
QP	28	32	36	40	28	32	36	40	28	32	36	40	28	32	36	40
ΔTime	85.80	85.69	84.96	82.18	81.40	86.15	84.95	84.79	75.71	87.84	87.99	87.40	87.28	89.04	87.05	85.98

Table 2. ΔPSNR and ΔTBitrate for QCIF sequences

Sequence	Akiyo	Foreman	Mobile	Paris
ΔPSNR	-0,001	-0,017	-0,008	-0,006
ΔBitrate	0,000	0,065	0,035	0,021

III. CONCLUSIONS

This paper proposes a WZ to H.264 video transcoder as a solution for mobile-to-mobile video communications. We show that transcoders can be used to exploit the low complexity encoding inherent in WZ coding. The proposed solution eliminates the need for complex H.264 encoding on end user devices for application such as video telephony. This is a first step in order to achieve real time communication with low complexity at the both sender and receiver devices which is a highly desirable feature in consumer devices. These first approach results show an extremely closer RD performance with an 86% complexity reduction.

REFERENCES

[1] ISO/IEC International Standard 14496-10:2003, "Information Technology – Coding of Audio - Visual Objects – Part 10: Advanced Video Coding".

[2] Wyner and J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder". IEEE Trans.on Inform. Theory, vol. IT-22, pp. 1–10, Jan. 1976.

[3] B. Girod, A. Aaron, S. Rane, D.R. Monedero, "Distributed Video Coding". Proc. of IEEE Special Issue on Advances in Video Coding and Delivery, Vol.93, No. 1, pp.1-12, 2005.

[4] Anthony Vetro; Charilaos Christopoulos; Huifang Sun, "Video transcoding architectures and techniques: An overview", IEEE Signal Processing Magazine, vol. 20, pp. 18–29, Mar. 2003.

[5] G. Fernandez-Escribano, H. Kalva, P. Cuenca, L. Orozco-Barbosa, and Antonio Garrido, "A Fast MB Mode Decision Algorithm for MPEG-2 to H.264 P-Frame Transcoding," IEEE Transactions on Circuits and Systems for Video Technology,Vol. 18, Issue 2, Feb, 2008. Pp 172-185.

[6] Bjontegaard G (2001) Calculation of average PSNR differences between RD-Curves. Presented at the 13th VCEG-M33 Meeting, Austin, TX, April