D-Cast: DSC based Soft Mobile Video Broadcast

Xiaopeng Fan∗
Harbin Institute of Technology
92 West Dazhi Street
Harbin, China
fxp@hit.edu.cn

Feng Wu
Microsoft Research Asia
5 Danling Street
Beijing, China
fengwu@microsoft.com

Debin Zhao
Harbin Institute of Technology
92 West Dazhi Street
Harbin, China
dbzha@hit.edu.cn

ABSTRACT

Mobile video broadcasting is a popular application of wireless network. However, the traditional layer based approaches have limited supports to the accommodation of users with diverse channel conditions. The newly emerged Softcast approach provides smooth multicast performance but is not very efficient in inter frame compression. In this work, we propose a new video multicast approach: D-cast. Instead of using conventional close loop prediction (CLP), D-cast is based on distributed source coding (DSC) theory. This helps D-cast to avoid error propagation but still achieve high compression efficiency in inter frame coding. D-cast outperforms softcast 3-5dB in video PSNR while maintaining the similar graceful degradation feature as softcast. In addition, D-cast is efficient in not only narrow band channel but also wide band channel.

General Terms

Video multicast, hieratical modulation, scalable video coding

Keywords

Distributed source coding, Motion compensation, Syndrome coding, Coset coding, Softcast, D-cast

1. INTRODUCTION

Wireless video broadcast aims to transmit video signal simultaneously to multiple users. The main challenge is the difference of the users in the channel conditions and hence the video quality they can expect. A typical wireless video broadcast scheme based on the DVB-T standard[1] accommodates diverse users by combining a layered transmission scheme[15][9] and scalable video coding (SVC) scheme [17][14]. SVC encodes the video signal into one base layer and multiple enhancement layers. The base layer contains the minimum information bits to decode the video, while the enhancement layer(s) contains additional information bits to refine the video. The state-of-art layered transmission scheme is hierarchical modulation (HM)[13]. HM superimposes the two layer bits in one wireless symbol and allow the user to decode BL only or all bits according to their own channel condition. Therefore, with SVC and HM, low SNR users can receive rough video signal while high SNR users can receive high quality video signal. However, this typical scheme has following three disadvantages: Firstly, to achieve scalability generally decreases the compression efficiency in video coding. Secondly, layered transmission in PHY layer, such as HM, reduces the SNR of base layer. Thirdly, such scheme only provides limited choices of BL and EL rates, e.g. DVB-T standard specifies 3 BL rates and 5 EL rates. This creates cliff effects in video quality as opposed to continuously changing channel condition.

Recently, a novel approach called Softcast has been proposed for wireless video broadcasting[7]. Softcast transmits the linear transform of the video signal directly in analog channel without quantization, FEC and modulation. It naturally supports broadcasting to users with different SNR, since the channel noise is directly transformed into reconstruction noise of the video. It also supports progressive video transmission by changing the number of the transformed coefficients. However, Softcast exploits intra frame redundancy only and thus is not very efficient in the aspect of video signal compression. Also, to apply traditional inter frame coding in Softcast is invalid due to the inter frame error propagation, i.e. all the errors in a previous frame copies to the following frame and the video quality decreases consecutively. In a recent improved version of Softcast, the utilization of 3D-DCT partially enables inter frame compression[6]. However, without motion compensation the inter frame redundancy is still not fully exploited in this updated version. Another disadvantage of Softcast is the inefficiency when the channel bandwidth is greater than the video bandwidth (pixels/s).

In this paper, we propose a wireless video multicast approach called 'D-Cast'. D-cast is, to the best of our knowledge, the first distributed source coding (DSC) [11] based throughout solution for wireless video broadcasting. Traditional residue coding is not efficient for video multicast system such as softcast, due to the unpredictable reconstruction error and the error drifting effects. However, DSC is efficient and drifting-free when the predicted signal is only known at
the decoder. In this work, we apply DSC techniques in video multicast to achieve high efficient inter frame compression and meanwhile avoid error drifting. The proposed D-cast approach achieves efficient and smooth multicast performance by comprehensively utilizing linear transform, coset coding, syndrome coding, decoder motion estimation techniques. Instead of transmitting (the linear transform of) the video signal itself, D-cast transmits the coset code [10] of the video signal by raw OFDM. This significantly reduces the variance of the signal but the receiver can still decode the signal with the help of the inter frame prediction. To avoid wrong decision of coset decoding, we also transmit a small amount of LDPC code as syndrome [8] to protect the signal. In experiments, the proposed approach achieves significant gain over Softcast. Furthermore, the proposed approach supports progressive image/video transmission and is highly efficient in both narrow and wide band channel.

The rest of the paper is organized as follows: Section 2 is a brief review of the related works. Section 3 introduces the proposed framework. Section 4&5 details the proposed encoder and decoder respectively. Section 6 explains efficient implementation of D-cast in narrow and wide band channel. Section 7 presents experimental results and Section 7 concludes the paper.

2. RELATED WORKS

Softcast is a simple but comprehensive design for wireless video multicast, covering the functionality of video compression, channel coding and PHY layer transmission in one scheme. At the server side, Softcast consists of three steps: transform, power allocation and wittening. Transform removes the spatial redundancy of a video frame. Power allocation minimizes the total distortion by optimally scaling the transform coefficients. Wittening transforms the coefficients by Hadamard matrix to create packets with equal average power and equal importance. At the user side, instead of doing direct inverse of what server does, Softcast uses an LLSE estimator as the opposite operation of power allocation and witting. All the steps are linear operations thus the channel noise is directly transformed into reconstruction noise of the video. Therefore, Softcast has the graceful degradation property when channel condition becomes worse. In addition, by skipping low importance coefficients Softcast can also efficiently broadcast video in narrow band channels.

However, Softcast exploits intra frame redundancy only and thus is not very efficient in the aspect of video signal compression. Also, to apply traditional inter frame coding in Softcast is invalid due to the inter frame error propagation, i.e. all the errors in a previous frame copies to the following frame and the video quality decreases consecutively. In a recent improved version of Softcast, the utilization of 3D-DCT partially enables inter frame compression[6]. However, without motion compensation the inter frame redundancy is still not fully exploited in this updated version.

Another disadvantage of Softcast is the inefficiency when the channel bandwidth is greater than the video bandwidth (pixels/s). In this situation, since all the DCT coefficients have been transmitted once, softcast has to retransmit them and reduce the noise by averaging the multiple versions of DCT coefficients. This has been shown to be not efficient when bandwidth becomes large.

The main difficulty to enable inter frame coding in softcast is the error propagation problem. Typical inter frame coding schemes utilize close loop prediction (CLP), i.e. encode the motion compensated difference between a video frame and its previous frame. However, in Softcast, the noise of each frame will add to its following frames if we only transmit their difference. This will successively reduce the reconstruction quality frame after frame. The main reason to cause this problem is that the encoder cannot exactly know the decoder reconstruction frame.

To compress a source with its prediction only available at decoder is a typical problem in distributed source coding(DSC). As shown in Fig. 2, X is the source representing the current video frame, S is its side information representing the predicted frame. The theoretical foundations of DSC, the Slepian-Wolf theorem[16] and the Wyner-Ziv theorem[18], presents an important but somehow surprising conclusion that, a source X can be efficiently compressed with its predictor S only available at the decoder. Practically, DSC employs coset coding[12] or syndrome coding[4][8]. Accompanied by the advances of the practical solutions, DSC has found considerable usage in video compression[5].

3. FRAMEWORK OF D-CAST

The proposed D-cast approach comprehensively utilizes coset coding, syndrome coding, decoder motion estimation and linear transform to remove both intra frame redundancy and inter frame redundancy in linear coding based wireless video multicast system.

Fig. 3 depicts the framework of D-Cast, which is a two layers system. D-cast first transforms the original image into DCT domain. Then it applies two successive coset coding module to get, for each DCT coefficients, several most important bits (MSBs), several least important bits (LSB) and the remaining middle bits. The middle bits and the LSBs are compressed and transmitted by two separate layers respectively, while the MSBs are discarded since they can be recovered at the decoder.

In the lower layer, the transmission of the LSBs is through power allocation and wittening as softcast does. At the receiver, the LLSE decoder first recovers the LSBs (with noise). Meanwhile, the prediction of the current frame is through motion compensated extrapolation (MCE), and the predicted image is transformed into DCT domain. Then, with the LSBs and the predictors, the transform coefficients of the current frame are recovered by coset decoding. Since both the LSBs and the predictors may have noise inside them, the decoded coefficients may suffer errors. The correction of these errors are the main task of the upper layer.

In the upper layer, the middle bits are compressed through syndrome coding at very low bit rate. The syndrome bits are protected by FEC at 1/2 rate and modulated by BPSK. This guarantees that all the users can receive the syndrome bits. At the receiver, with demodulation and channel decoding we get the LLR of each syndrome bits. These syndrome bits help to correct the errors of the (noisy) middle bits from the lower layer. Then, with the LSBs, the corrected middle bits and the predicted coefficients, D-cast restores the original coefficients by coset decoding.

Both the modulated syndrome bits and the soft coded LSBs are transmitted over OFDM channel.

Fig. 4 shows the basic idea of D-cast. The MSBs have high correlation with the decoder prediction, thus they can be guessed at the decoder and need no transmission. The
Figure 1: Softcast

Figure 2: Compression of a source $X$ when its side information $S$ is only available at the decoder

Figure 3: Diagram of D-cast
LSBs has little correlation with the decoder prediction, thus they need to be transmitted. Middle bitplanes can also be estimated at the decoder, but the estimation may not be fully correct. Thus middle bitplanes needs some syndrome bits to protect.

4. D-CAST ENCODER

This section presents the details of the proposed D-cast encoder.

4.1 Coset encoding

Coset encoding is a typical technique used in DSC. With coset code, a source can be efficiently compressed when its predictor only available at the decoder. The coset code achieves efficient compression by throwing away partial information of the source signal. Nevertheless, the code is still decodable as long as the predictor is available at the decoder, because the lost partial information can be recovered with the help of the predictor.

![Figure 5: Coset code example](image)

Fig.5 illustrates the basic idea of the coset code. Consider the compression of an integer random variable $X$ which is uniformly distributed in $[0,255]$. Without prediction, it would take 8 bits to represent $X$. Suppose $Y$ is another integer random variable and $X - Y$ is uniformly distributed in $[-4,3]$. $Y$ serves as the side information (i.e. the predictor at the decoder). If $Y$ is available at both encoder and decoder, we can convey $X$ by sending the residue $X - Y$. Since $X - Y$ is uniformly distributed in $[-4,3]$, it would take only 3 bits to represent $X$. If $Y$ is only available at the decoder, we can convey $X$ using the following algorithm (Fig.5) [19].

We send the remainder of $X$ modulo 8, which is 3, to the decoder. This modulo operation actually partitions the space of all possible inputs into 8 different groups or say cosets. The remainder 3 can be considered as the index of the coset that $X$ belongs to. Since the coset index is uniformly distributed in $[0,7]$, it takes 3 bits to express. The decoder receives the coset index and observes the side information $Y$. The coset index suggests many possible candidates of $X$, and each is congruent to 3 modulo 8. However, it can be verified that the decoder can always find out $X$ by selecting the candidate which is closest to $Y$. In the above example, we can always communicate $X$ using 3 bits with and without $Y$ at the encoder, as long as $Y$ is available at the decoder.

The proposed D-cast uses a similar coset code as the one in the above example, except that our coset takes real value input and real value output. The proposed approach consists of two successive coset encoders. The first one divides each transform coefficient $X$ by a step $q_1$ and get the remainder $R$

$$R = X - \left\lfloor \frac{X}{q_1} \right\rfloor + \frac{1}{2} q_1$$

This is somehow equivalent to throwing away some MSBs of $X$. The second coset encoder divides $R$ by a smaller step $q_2$ to get the remainder $L$

$$L = R - \left\lfloor \frac{R}{q_2} \right\rfloor + \frac{1}{2} q_2$$

In some sense $L$ represents LSBs although $L$ is real value. In D-cast, we let

$$q_1 = 2^N q_2$$

where $N$ is an integer. In the second coset encoder, we also calculate

$$M = (\frac{R - L}{q_2}) \mod (2^N)$$

as the middle bits of each coefficient.

Fig.6 shows the DCT coefficients before and after coset encoding. The left is the original DCT coefficient while the right is the $L$ value. Obviously the magnitude of some large coefficients becomes much smaller after coset coding. This saves the power to transmit the signal or equivalently means increased SNR under constrained power.

In the following subsection we will present how to further encode and transmit the $L$ and $M$.

4.2 Power allocation and Wittening

The coset data, i.e. the $L$ value of each DCT coefficients are encoded using several linear operations. First the DCT coefficients are divided into 64 subbands and for each subband $i$ we calculate the variance $\sigma_L^2(i)$ of the $L$ values. Then all $L$ are scaled for optimal power allocation in terms of minimizing the distortion[6],

$$\tilde{L}_i = g_i L_i$$

$$g_i = \left( \frac{\sigma_L^{-1}(i)}{P} \right)^{1/2} \sum \sigma_L(i)$$

After this optimal scaling, the variances of the $\tilde{L}$ values of each subbands are still different. To redistribute energy, the $\tilde{L}$ values from different subbands are combined together to form vectors and the new vectors are transformed by Hadamard matrix. This creates packets with equal energy and equal importance[6].
4.3 Syndrome encoding

The syndrome encoder is for distributed source compression (DSC)[8]. It takes each bitplanes of all the $M$ as input and produces both syndrome bits and the systematic bits (i.e. the original bitplanes of $M$). It forwards the syndrome bits to the following FEC module but discard the systematic bits. It uses a very high bit rate LDPC code thus the bitplanes are compressed at very low bit rate.

4.4 Packaging and transmission

Both the coset values (after wittening) and the syndrome bits are transmitted to PHY layer. These values and bits are packaged into multiple packets. Each packet consists of a sequence of syndrome bits and a sequence of coset values. In PHY layer, the syndrome bits are coded by 1/2 rate FEC and BPSK modulation. This guarantees in high probability the lossy transmission of the syndrome bits. The coset values are directly mapped into transmitted signal. This direct mapping is by modifying the existing 802.11 PHY layer to allow raw data to bypass the FEC and QAM[6].

5. D-CAST DECODER

The D-cast decoder is composed of three parts: the decoder motion estimation module predicts the video frame by MCE; the lower layer decoder reconstructs a noisy version of the current frame via coset decoding; and the higher layer decoder refines the frame by syndrome decoding.

5.1 Decoder Motion Estimation

D-cast performs motion estimation (ME) at decoder, which is the so-called motion compensated extrapolation (MCE). Before decoding the frame $n$, the decoder first estimates the motion vectors between the reconstructed frame $n - 1$ and frame $n - 2$, and then smooths the motion field by median filter[3]. If frame $n - 2$ is not available, all the MVs of frame $n - 1$ are set to zero. After that, the decoder extrapolates the motion vectors to frame $n$ as shown in Fig. 7. The extrapolation is based on the assumption of translational motion. Inversely extending each motion vector of frame $n - 1$ reaches an intersection point in the frame $n$. Each block in frame $n$ finds its nearest intersection point and gets the MV corresponding to that point. Then, with the MVs and the frame $n - 1$, the decoder predicts the frame $n$ by motion compensation. Finally, the predicted frame $n$ is transformed into DCT domain.

5.2 LLSE at decoder

The Linear Least Square Estimator (LLSE) is to reconstruct the coset value $\hat{L}$ with minimum distortion. The received signal of the lower layer can be written as:

$$Y = HGL + N$$

where $H$ is Hadamard matrix, $G$ is the diagonal matrix for power allocation, and $N$ is the channel noise. Let $C = HG$, then the LLSE of $L$ is

$$\hat{L} = \Sigma_L C^T (C \Sigma_L C^T + \Sigma_N)^{-1} Y$$

where $\Sigma_L = \text{diag}(E(\text{LL}^T))$ and $\Sigma_N = E(\text{NN}^T)$.

5.3 Coset decoding

With the decoded coset value $\hat{L}$ and the side information $S$ (i.e. the predicted DCT coefficients), the lower layer of the receiver reconstructs the DCT coefficients by coset decoding. According to (1)-(3), we have

$$L = X - \left[ \frac{X}{q_2} + \frac{1}{2} \right] q_2$$

Receiving the coset value $\hat{L}$ at the decoder, there are multiple possible reconstructions of $X$ forming a coset $C$.

$$C = \{ \hat{L}, \hat{L} \pm q_2, \hat{L} \pm 2q_2, \hat{L} \pm 3q_2, \ldots \}$$

D-cast selects in $C$ the one nearest to the side information $S$ as the reconstruction of the DCT coefficient.

$$\hat{X} = \arg \min_{c \in C} |c - S|$$
Note that when channel SNR is 5dB and estimate according to the channel SNR. In D-cast, we assume the lowest noise in the side information. Then the coset decoding is correct, i.e. the distance between the original signal and the coset value is a scalar value sent to decoder together with those syndrome bits.

Fig.8 illustrates the principle of the coset coding in constellation view. The coset code vertically generates multiple vertical lines and horizontal lines which forms many intersection points. The original signal is quantized to its nearest intersection point and the quantization residue is the coset vector we will transmit. The coset vector is much shorter than the original signal, which means compression. At the decoder, the coset data indicates multiple possible locations of the original signal (the cross point in the figure). The decoder reconstructs the signal by selecting from all possible locations the one nearest to the predicted signal.

5.4 Syndrome decoding

The purpose of the syndrome decoding is to correct the errors of the coset decoding result.

The syndrome decoding in D-cast includes two steps: FEC decoding and DSC decoding. Both perform belief propagation (BP).

FEC decoder is a soft-in-soft-out one. It gets the log likelihood ratio (LLR) information from PHY layer as input and forwards the output LLR to the DSC decoder as the soft information.

The DSC decoder gets soft information of the syndrome bits from FEC decoder and estimates the soft information of the systematic bits from the coset decoder outputs. Performing BP on these soft information produces the soft information of the encoder input, i.e. the middle bitplanes of original signal. Then making hard decision corrects the errors of the coset decoding result. Combining the corrected middle bits and the coset value reconstructs the coset value. Finally a new coset decoding reconstructs the signal again with the coset value and the side information.

The estimation of the soft information of the systematic is as follows: The distribution of the original signal given the predicted signal (i.e. the side information S) is assumed to be Lapacian as most DVC approaches does[2]. The probability of a particular bit in the M to be 0 or 1 is calculated by accumulation on all possible ranges.

Fig.9 illustrates the functionality of the syndrome coding in constellation view. The coset value suggests multiple possible locations of original signal. The decoder finds the one nearest to the predicted signal as the reconstructed signal. However, the predicted signal may suffer noise when SNR is low. In that situation, the coset decoding may make wrong decisions. The syndrome decoding result indicates the value of each bit in middle bitplanes. This value as a constraint reduces the set of the possible locations. Since the possible locations becomes apart from each other, the decoder can find the original signal (approximately) even if the predicted signal is noisy.

5.5 Refinement stage

As mentioned before, D-cast performs decoder ME. After reconstructing the current frame, D-cast performs ME between current frame and previous frame. This step has two functionalities. The first is to estimate the MVs for future usage in MCE. The second is to refine the predicted signal and the reconstructed signal.

This refined predicted signal is based on ME, and is much better than the MCE based predicted signal. Thus we let D-cast to perform in pixel domain a minimum mean square error (MMSE) estimation:

\[ x^* = \alpha \hat{s} + (1 - \alpha) \hat{x} \]  
\[ \alpha = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_{\hat{x}}^2} \]  

where \( \sigma_s^2 \) is the variance of the original reconstruction noise, and \( \sigma_{\hat{x}}^2 \) is the variance of the prediction noise. The prediction noise variance is estimated by the SAD in ME, while the reconstruction noise variance is estimated by the SNR of the channel.

6. VIDEO MULTICAST IN NARROW OR WIDE BAND CHANNELS

The aforementioned approach is for the situation that the channel bandwidth (symbols/s) is equal to the video bandwidth (pixels/s) approximately. In this section we will discuss how to efficiently transmit the video over narrow or wide band channels.

6.1 Video multicast in narrow band channel

When the channel bandwidth is smaller than the video bandwidth, D-cast reduces the number of transmitted symbols by discards some DCT bands. As mentioned before, D-cast divided all the coefficients into 64 subbands. D-cast encoder discards high frequency bands in a reverse zigzag order. This is mainly for better visual quality since human visual system is less sensitive to high frequency components than low frequency components. At the decoder, these discarded bands directly copy signal from the corresponding bands of the predicted frame at the decoder. Thus the distortion of these discarded bands are exactly the power of prediction error.

6.2 Video multicast in wide band channel

The softcast approach becomes less efficient when the channel bandwidth is greater than the video bandwidth. In this situation, softcast utilizes the bandwidth by retransmission. It averages the multiple samples and hence reduces the
Figure 8: Coset decoding in D-cast

Figure 9: Syndrome decoding in D-cast
noise power. This only provides about 3dB gain in PSNR, each time the bandwidth doubled.

In contrast, D-cast is efficient in wide band channel. D-cast utilizes the bandwidth by retransmitting coset values $L$ with successively decreased coset step size $q_2$. At beginning, D-cast generates coset value $L$ with large step size $q_2$. Then, if the bandwidth is not fully utilized, D-cast generates new coset value $L'$ with smaller step size $q_2$. D-cast generates and transmits new coset values until the bandwidth is fully occupied.

At the decoder, D-cast first decodes the frame as presented in the previous section. Then receiving the new coset value $L'$ and utilizing reconstructed frame $X'$ as better side information, D-cast reconstructs a refined version $X''$ by coset decoding. In this way, D-cast keep refining the reconstructed frame until all coset values utilized.

The basic principle of this approach is to consider the reconstructed frame as a new but better side information and re-encode the current frame to get higher PSNR. Since the step size $q_2$ decreases successively, the coset value $L$ also becomes smaller and smaller. To transmit these smaller $L$ with constrained power means reduced noise power or increased SNR.

In D-cast the $q_2$ decreases exponentially with respect to the bandwidth. Thus the video PSNR increases linearly with the bandwidth $B$. This compares favorably with the softcast where the PSNR increases linearly with $\log(B)$.

7. EXPERIMENTS

In experiments, we evaluate the performance of the proposed D-cast in video multicast and progressive image transmission.

We compare D-cast with Softcast[7][6]. The video packets are transmitted to OFDM. The OFDM signal is transmitted over AWGN channel. The receiver passes the signal to the OFDM module to perform CFO corrections, channel estimation and correction, and phase tracking. Then it inverts the operations of the transmitter and forwards the soft information to video decoding layer.

7.1 Video multicast

The first test is video multicast to users with diverse SNR. The channel bandwidth is equal to the video bandwidth (i.e. the number of video pixels per second). Note that both softcast and D-cast have to transmit some additional information (e.g. the meta data in softcast, or the syndrome bits in D-cast). In our implementation, we let $q_1 = 2q_2$ such that $M$ only contains 1 bitplane. The syndrome encoder uses a 0.99 rate LDPC code. With FEC, the bandwidth occupation of syndrome bits are about 1.97% of the total bandwidth. In our test, the bandwidth occupied by these additional information is compensated by discarding some DCT bands as explained in section 6. The video test sequences are 'foreman_qcif.yuv' and 'bus_qcif.yuv'. The video frame rate is 30Hz. The GOP structure is 'IPPP...'. The results are given in Fig.10. The 3D-DCT based softcast [6] is about 2.5dB better than the 2D-DCT based softcast[7]. Our D-cast is 3.5dB better than the 2D-DCT based softcast. The visual quality comparison is given in Fig.11. Obviously D-cast has better visual quality than both 2D-DCT based softcast and 3D-DCT based softcast.

7.2 Progressive image transmission

The second test is progressive image transmission. When the channel bandwidth is wider than the video bandwidth, both D-cast and softcast supports to improve the video quality by transmitting additional information. Softcast simply retransmits the coded video signal hence reduce the power of noise. D-cast transmits coset value with successively decreased coset step size. In this test D-cast transmits only coset values but no syndrome bits. With successively decreased coset step size, the coset code reduces the reconstruction errors and improves the image PSNR. The results are given in Fig.12. D-cast is much better than softcast in terms of image quality. More importantly, in D-cast the PSNR increases linearly with the bandwidth $B$, which compares favorably with softcast where the PSNR increases linearly with $\log(B)$.

8. CONCLUSIONS

In this paper we propose a DSC based video multicast approach: D-cast. D-cast mainly utilizes linear transform to encode and transmit video signal, thus performs gracefully in video multicast. D-cast applies DSC principle into video multicast and benefits from motion compensation while avoiding the error propagation. D-cast comprehensively utilizes linear transform, coset coding and syndrome coding, and achieves better performance than existing softcast approach. Furthermore, D-cast supports progressive image transmission efficiently in wide band channel.

9. ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundations of China: 60736043 and the Major State Basic Research Development Program of China (973 Program 2009CB320905).

10. REFERENCES


Figure 10: Performance comparison

Figure 11: Visual quality comparison, the 5th frame of foreman_qcif.yuv, AWGN channel, SNR=7dB


