# WAVECAST: WAVELET BASED WIRELESS VIDEO BROADCAST USING LOSSY TRANSMISSION

*Xiaopeng Fan*[1]*, Ruiqin Xiong*[2]*, Feng Wu*[3] *and Debin Zhao*[1]

[1]School of Computer Science, Harbin Institute of Technology, China
[2] School of EECS, Peking University, Beijing, China
[3] Microsoft Research Asia, Beijing, China
E-Mail: fxp@hit.edu.cn

## ABSTRACT

Wireless video broadcasting is a popular application of mobile network. However, the traditional 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: WaveCast. Different from softcast, WaveCast utilizes motion compensated temporal filter (MCTF) to exploit inter frame redundancy, and utilizes conventional framework to transmit motion information such that the MVs can be reconstructed losslessly. Meanwhile, WaveCast transmits the transform coefficients in lossy mode and performs gracefully in multicast. In experiments, WaveCast outperforms softcast 2dB in video PSNR at low channel SNR, and outperforms H.264 based framework up to 8dB in broadcast.

***Index Terms***— 3D wavelet, wireless video broadcast, softcast, DCast, WaveCast

## 1. INTRODUCTION

Wireless video broadcasting is a popular application aiming to transmit video signal simultaneously to multiple users of possibly different channel conditions. The main challenge is the difficulty to fully utilize each user's channel capacity and provides each user best video quality under his channel condition. The conventional digital video broadcasting (DVB) [1] framework based on H.264 [2] and 802.11 [3] can hardly accommodate diverse users in broadcast due to the stair effect: The server should encode the video source at one rate. However, if some users' channel capacity turn out to be less than this source rate, then their reconstruction quality will degrade sharply; while if some users' channel capacity are more than the source rate, their reconstruction quality will not improve accordingly. A typical approach to mitigate this stair effect is scalable video coding (SVC) [4][5]. With SVC, users of low channel capacity can receive rough video signal while users of high channel capacity can receive high quality video signal. However, SVC decreases the compression efficiency of video signal. In addition, SVC can not entirely mitigate the stair effect, but divides one 'big stair' into two or three 'small stairs'.

The conventional framework utilizes lossy compression and lossless transmission, and thus can hardly accommodate diverse users with different channel capacity. Recently, a novel all-in-one framework named as Softcast, with lossless compression and lossy transmission has been proposed for wireless video broadcasting [6, 7]. Different from conventional frameworks, Softcast simply compresses the video by discrete cosine transform (DCT) and transmits the coefficients directly in lossy mode. This makes the reconstruction quality depends only on the receiver's channel SNR. A high SNR user can automatically get high quality video while a low SNR user can also decode the video. In wireless broadcasting application, Softcast achieves significant gain over the conventional DVB framework [7].

One disadvantage of Softcast is the inefficiency in exploiting the inter frame redundancy of video signal. This problem is partially solved by the utilization of 3D-DCT in a recent improved version of Softcast[7]. However, without motion compensation, the inter frame redundancy is still not fully exploited in the updated version. In [8, 9], motion compensation and lossy transmission are combined together in a distributed video coding [10] framework. However, to the best of our knowledge, there is no solution till now to enable motion compensation and lossy transmission together in conventional video coding framework.

In this paper, we propose a wireless visual communication framework called WaveCast based on lossy transmission and 3D wavelet transform[11, 12, 13]. WaveCast utilizes motion compensated temporal filter (MCTF) [14, 15] to exploit inter frame redundancy. This helps to improve the coding efficiency owning to the aligned motion. WaveCast transmits the motion information by conventional communication

This work was done in Microsoft Research Asia (MSRA) where Dr. Xiaopeng Fan was a visiting researcher in 2011-2012.

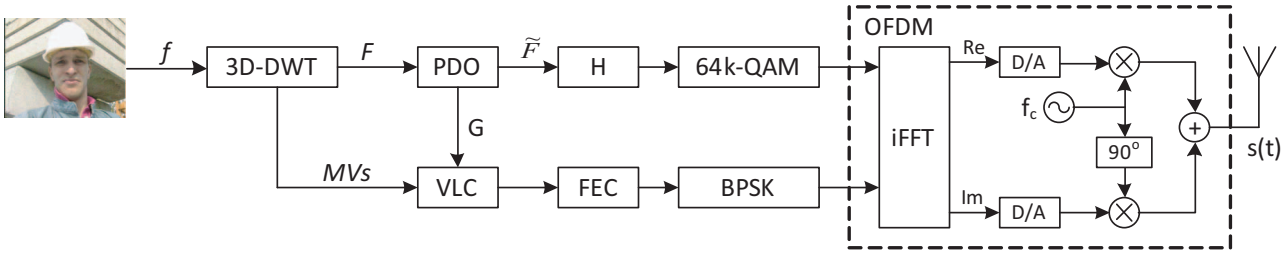


**Fig. 1**. WaveCast server

framework, such that the transmission of MVs are lossless. Different from conventional framework, WaveCast uses lossless compression and lossy transmission. The server sends full information of video to the channel, while how much information can be received by the client(s) depends on the channel condition. This makes the system much more robust to the fluctuation of the channel SNR compared with conventional visual communication systems. In broadcast, the framework can utilizes each client's channel as much as possible, and let each client to get as much visual information as its channel condition allows. In experiments, the proposed approach achieves significant gain over H.264 based conventional framework as well as Softcast.

The rest of the paper is organized as follows: Section 2 introduces the proposed framework. Section 3 presents experimental results and Section 4 concludes the paper.

## 2. PROPOSED WAVECAST

The proposed WaveCast framework is a wireless visual communication framework with jointly designed source coding, channel coding and physical layer. WaveCast utilizes 3D wavelet transform to compress the video signal. Compared with the 3D DCT transform in Softcast, 3D wavelet is more efficient to remove temporal redundancy owning to aligned motion. WaveCast communicates the motion information by traditional communication method, such that the MV transmission is lossless. WaveCast transmits the wavelet coefficients in lossy mode as Softcast does. In broadcast, WaveCast achieves similar graceful performance as Softcast.

Fig.1 depicts the framework of WaveCast server. Firstly, WaveCast transforms the input video signal by 3D wavelet. The transmission of the wavelet coefficients are in lossy mode similar to Softcast. The wavelet coefficients are scaled for power distortion optimization (PDO) and then transformed by Hadamard as precoding for broadcasting channel. The resulting coefficients are mapped to complex symbols by a very dense constellation (64K-QAM): each coefficient is quantized into 8-bit integer number and every two integers compose one complex number of 64k possible values. The meta data, including the MVs and the scaling factor of PDO, are coded by using conventional scheme consisting of variable length coding (VLC), forward error correction (FEC) and binary phase-shift keying (BPSK) mapping. Finally, the modulated symbols are passed into OFDM module undergoing iFFT and D/A conversion, and the analog signals are then modulated with carrier waves to generate transmitted signal.

The client side of WaveCast is depicted in Fig. 4. The OFDM module receive the signal and reconstruct the modulated complex symbols of both the scaled coefficients and the meta data. The meta data are demodulated and decoded first to get the MVs and the scaling factor. Then the scaled coefficients are reconstructed by inverse 64k-QAM and inverse Hadamard. The inverse 64k-QAM here do nothing but decoupling each complex value back into two real values. Each real value here is actually the 8-bits integer number plus channel noise. After inverse Hadamard transform, the coefficients are estimated by MMSE. At last, with the MVs, the coefficients are inversely transformed by 3D wavelet, to generate the final reconstruction.

### 2.1. MCTF based 3D Wavelet Transform

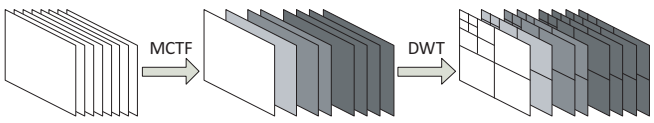


**Fig. 2**. 3D wavelet transform

As shown in Fig. 2, the 3D wavelet transform consists of two steps: motion compensated temporal filter (MCTF) and 2D spatial wavelet transform.

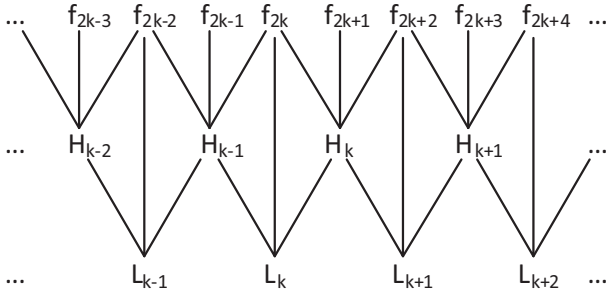


**Fig. 3**. The lifting structure of MCTF

The first step of 3D wavelet is to apply the MCTF recursively on the video signal. The MCTF used in this work is the Barbell-lifting based temporal wavelet filters[15]. Fig. 3 shows the lifting structure of MCTF. MCTF decomposes the input video sequence $f$ into high pass frames $H$ and low pass frames $L$. Let $n$ be the number of input video frames, and $f_i$ be (a vector representing) the $i^{th}$ frame. For each odd frame $f_{2k+1}$, $k \in \{0, 1, \cdots, n/2-1\}$, a bidirectional motion estimation (ME) is performed. Let function $\mathcal{B}_{2k\rightarrow 2k+1}(\cdot)$ and $\mathcal{B}_{2k+2\rightarrow 2k+1}(\cdot)$ represent forward prediction and backward prediction respectively. The high pass frames are generated by

$$H_k = \alpha_0 \mathcal{B}_{2k\rightarrow 2k+1}(f_{2k}) + f_{2k+1} + \alpha_1 \mathcal{B}_{2k+2\rightarrow 2k+1}(f_{2k+2}) \quad (1)$$

where $\alpha_0 = \alpha_1 = -0.5$. And the low pass frames are generated by

$$L_k = \beta_0 \mathcal{B}_{2k-1\rightarrow 2k}(H_{k-1}) + f_{2k} + \beta_1 \mathcal{B}_{2k+1\rightarrow 2k}(H_k) \quad (2)$$

where $\beta_0 = \beta_1 = 0.25$, and function $\mathcal{B}_{2k-1\rightarrow 2k}(\cdot)$ and $\mathcal{B}_{2k+1\rightarrow 2k}(\cdot)$ represent the inverse process of motion compensation. Taking (1) into (2), it is clear that the MCTF here forms a (5,3) wavelet transform in temporal direction.

The MCTF is recursively applied on the low pass frames to generate temporal subbands. After each level of MCTF, the video frames are updated by

$$H_{n/2+k} \leftarrow H_k, \quad f_k \leftarrow L_k, \quad k = 0, 1, \cdots, n/2-1 \quad (3)$$

The new frames $f_k$ ($k \in \{0, 1, \cdots, n/2-1\}$) are actually the low pass frames of last level of MCTF. Then the MCTF is applied on the new frames $f_k$ ($k \in \{0, 1, \cdots, n/2-1\}$) to generate new low pass frames and high pass frames. The encoder recursively applies several levels of MCTF, and generate finally a low pass frame $L_0$ and several high pass frames $H_1, H_2, \cdots, H_{n-1}$.

After MCTF, the low pass and high pass frames are transformed by 2D wavelet $\mathcal{W}(\cdot)$ to get wavelet coefficients $F$.

$$F = \{\mathcal{W}(L_0), \mathcal{W}(H_1), \mathcal{W}(H_2), \cdots, \mathcal{W}(H_{n-1})\} \quad (4)$$

### 2.2. Power Distortion Optimization

The power distortion optimization (PDO) is to minimize the reconstruction distortion under power constraint. PDO includes two steps. The first step is to normalize the wavelet coefficients such that equal coefficients in different subband will have equal importance to the reconstruction. In our WaveCast, 3D wavelet transform includes 4 levels of recursive MCTF decomposition and 3 levels of spatial wavelet transform. From (1) and (2), the temporal wavelet is a (5,3) wavelet with filters $[-1, 2, -1]/2$ and $[-1, 2, 6, 2, -1]/8$. Thus the norm of the low pass filter (LPF) and the high pass filter (HPF) are $a = \sqrt{3/2}$ and $b = \sqrt{23/32}$ respectively. For 4 levels of recursive MCTF decomposition, the transformed coefficients belonging to 5 different temporal frequency are normalized by the factors $b^4, ab^3, ab^2, ab, a$ respectively. For spatial transform, a (9,7) wavelet is used, and the norm of LPF and HPF are $c = 1.14$ and $d = 0.89$ respectively. For 3 levels of spatial transform, the transformed coefficients belonging to 4 different spatial frequency are normalized by the factors $d^3, cd^2, cd, c$ respectively.

The second step is to scale the coefficients such that the total power is optimally allocated to each subband in the transmission. The wavelet coefficients $F$ (after normalization) are divided into several subbands and for each subband $i$ we calculate the variance $\sigma^2_{F(i)}$. We compare these variance with a threshold. If the variance is smaller than the threshold, then the corresponding subband is skipped. All the rest coefficients are scaled for optimal power allocation between different subbands. Let $P_{coeff}$ be the total power to transmit the coefficients, and $G_i$ be the scaling factor of subband $F_i$. According to [6], the optimal power allocation of the coefficients $F$ is

$$\tilde{F}_i = G_i F_i, \quad G_i = \left( \frac{\sigma^{-1}_{F(i)} P_{coeff}}{\sum \sigma_{F(i)}} \right)^{1/2} \quad (5)$$

### 2.3. Packaging and Modulation

The transmission of the wavelet coefficients $\tilde{F}$ are in lossy mode. After power allocation, the variances of each subband are still different. To redistribute energy, the coefficients from different subbands are combined together to form several new vectors and each new vector has similar norm. Then the new vectors are transformed by Hadamard matrix,

$$D = H\tilde{F}. \quad (6)$$

Then the transformed coefficients $D$ are randomly grouped together. This creates packets with equal energy and equal importance. The coefficients of each packet is then directly mapped to complex symbols by a very dense constellation, 64K-QAM. Each coefficient is quantized into 8-bit integer number and every two integers compose one complex symbol (of 64k possible values). Let $\mathcal{Q}_{8bits}(\cdot)$ be the quantization function. The 64K-QAM can be expressed as

$$\vec{D}_i = \mathcal{Q}_{8bits}(D_{2i}) + j\mathcal{Q}_{8bits}(D_{2i+1}). \quad (7)$$

An inverse FFT is computed on each packet of symbols, giving a set of complex time-domain samples. These samples are then quadrature-mixed to passband in the standard way. The real and imaginary components are first converted to the analogue domain using D/A converters; the analogue signals are then used to modulate cosine and sine waves at the carrier frequency, $f_c$, respectively. These signals are then summed to give the transmission signal, $s(t)$.

$$s(t) = \mathsf{Re}\{\mathsf{iFFT}\{\vec{D}\}e^{2\pi j f_c t}\} \quad (8)$$

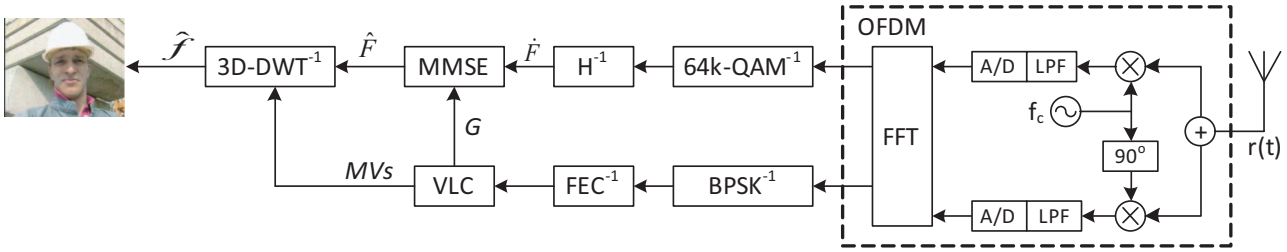


**Fig. 4**. WaveCast client

### 2.4. Transmission of Motion Information

The transmission of motion information and the scaling factors $G$ are through traditional communication scheme consisting of entropy coding, channel coding and modulation. The scaling factors $G$ are quantized by an 8-bits scalar quantizer and coded by fix length coding. A predictor for each MV is generated based on the neighboring blocks' MVs. Then the prediction residues together with the mode information are compressed by universal variable length coding (UVLC). The UVLC table is as follows.

| Input | Output |
|---|---|
| 0 | 1 |
| $\pm 1$ | 01x |
| $\pm 2, \pm 3$ | 001xx |
| $\pm 4, \pm 5, \pm 6, \pm 7$ | 0001xxx |
| $\cdots$ | $\cdots$ |

**Table 1**. UVLC for MV coding

The compressed bitstream is then further coded by forward error correction (FEC) code. The FEC code is the one used in 802.11 standard. It is a 1/2 convolutional code with generator polynomials $\{133, 171\}$ as shown in Fig.5.

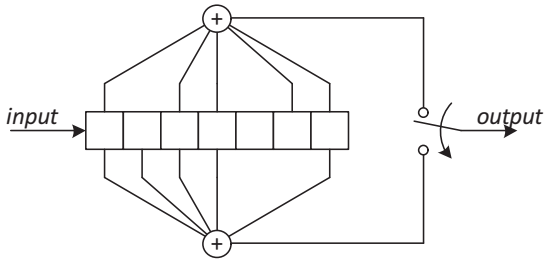


**Fig. 5**. The 1/2 convolutional code with generator polynomials $\{133, 171\}$

The coded bits are then mapped to complex symbols by BPSK constellation. Finally, the complex symbols are passed into OFDM module for transmission.

### 2.5. Receiver

The receiver gets signal $r(t)$ from the channel and reconstruct the video signal. The OFDM module receives the signal and reconstruct the modulated complex symbols of both the scaled coefficients and the meta data. The received signal $r(t)$ is quadrature-mixed down to baseband using cosine and sine waves at the carrier frequency. After applying the low-pass filters, the OFDM module samples and converts the baseband signals to digital numbers, and use a forward FFT to convert them back to the frequency domain.

The frequency domain signal after FFT includes transform coefficients and meta data. The meta data are decoded first to get the MVs and the scaling factor. The soft information of the original bitstream is estimated by a soft BPSK detector. Then the receiver uses Viterbi algorithm to correct the errors in the bitstream, and then decodes the bitstream to get MVs and scaling factors $G$.

The scaled coefficients are reconstructed by inverse 64k-QAM and inverse Hadamard. The inverse 64k-QAM here just splits each complex symbol back into two real values. Each real value here is actually the 8-bits integer number plus channel noise. After inverse Hadamard transform, the coefficients are estimated by MMSE. Let $\dot{F}$ be the coefficients after inverse Hadamard transform. $\dot{F}$ can be written as:

$$\dot{F} = \tilde{F} + N \tag{9}$$

where $N$ is the equivalent channel noise after inverse Hadamard transform. Under the assumption that the channel noise is whiten Gaussian, $N$ is also whiten Gaussian. The MMSE reconstruction of original $F$ is

$$\hat{F}_i = \frac{\sigma^2_{F(i)}}{\sigma^2_{F(i)} G_i^2 + \sigma_N^2} \dot{F}_i \tag{10}$$

where $\sigma_N^2$ are the noise variance.

At last, with the MVs, the wavelet coefficients $\hat{F}$ are inversely transformed by 3D wavelet, to generate the final reconstruction $\hat{f}$. The inverse transform consists of 2D inverse wavelet transform and inverse MCTF.

$$\hat{f} = \mathsf{MCTF}^{-1}\{\mathcal{W}^{-1}\{\hat{F}\}\} \tag{11}$$

## 3. EXPERIMENTS

In experiments, WaveCast is compared with both Softcast[6, 7] and conventional framework. We have implemented two versions of Softcast based on 2D-DCT and 3D-DCT respectively, i.e. Softcast2D[6] and Softcast3D[7]. The conventional framework is based on H.264 and 802.11. We use JM14.2 software as H.264 codec and ues 'baseline' profile. We implement standard 802.11 PHY layer with FEC and QAM modulations. The H.264 coded video data is packed into RTP packets of length 1200 bytes. We insert into each RTP packet a 32-bits CRC, and then encode each packet separately by FEC code. We generate the 1/2 convolutional code with polynomials $\{133, 171\}$ as FEC code following the 802.11 standard. The FEC coded bits are mapped to the channel symbols by QAM modulation, including BPSK, QPSK, 16QAM and 64QAM. All the modulated signals are transmitted over simulated OFDM channels. The channel decoding is by soft Viterbi algorithm. After that, the decoder performs CRC check for each RTP packet and forward those error-free packets to H.264 decoder. The H.264 decoder is able to tolerant a small percentage of RTP packet loss, by utilizing error concealment. In our test, we have config the H.264 decoder to use the most complex error concealment method in JM14.2, the motion copy one, to get best reconstruction quality.

To evaluate average performance of each framework, we create a monochrome 512-frame test video sequence, called 'all_seq', by combining the first 32 frames of the 16 CIF sequences ( 'akiyo', 'bus', 'coastguard', 'crew', 'flower', 'football', 'foreman', 'harbour', 'husky', 'ice', 'news', 'soccer', 'stefan', 'tempete', 'tennis', and 'waterfall'). For all framework, the GOP length is 32 frames. For both WaveCast and H.264, the search range of ME is $32 \times 32$ and the MV precision is $1/4$ pixel. In ME, WaveCast uses only $16 \times 16$ block size, while H.264 uses all the 7 block size from $4 \times 4$ to $16 \times 16$.

All the framework utilizes equal bandwidth, 1.15MHz. For Softcast and WaveCast, this constraint is satisfied by skipping those subbands with lowest variance. Note that WaveCast needs to transmit MVs, which occupies averagely 0.12MHz of bandwidth. Thus WaveCast skipped slightly more subbands than Softcast. The conventional framework utilizes four recommended combination of channel coding and modulation from 802.11a. We calculate the corresponding bit-rates respectively according to the bandwidth, and set the bit-rates constraint to the H.264 encoder for rate control.

The first experiment assumes that only the decoder knows the channel SNR. The experimental result is given in Fig.6. According to the result, all the four combinations of conventional framework suffer very serious stair effect. For example, the combination 'H.264,1/2FEC,16QAM' performs good when channel SNR is between 13dB to 14dB, but is not good when channel SNR is out of this range. When the channel SNR becomes more than 14dB, the reconstruction quality

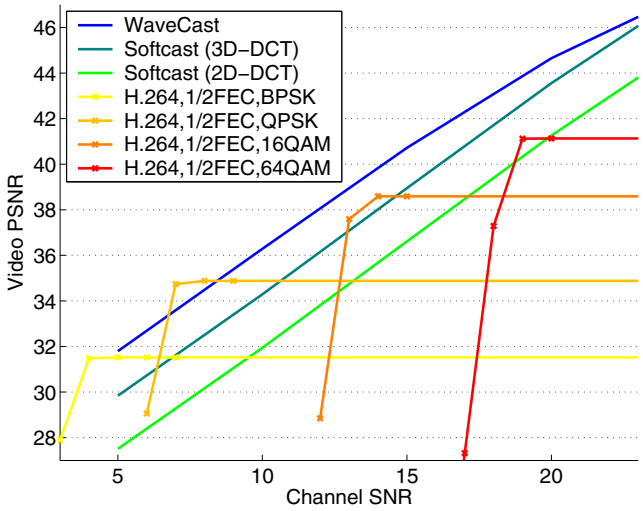


**Fig. 6**. Multicast performance comparison

does not increases. When the channel SNR becomes 12dB, the reconstruction quality drops very quickly. When the channel SNR becomes even lower, the video decoder cannot work since almost all received RTP packets has bit error. In contrast, the three all-in-one frameworks does not suffer the stair effect. When the channel SNR increases, the reconstruction PSNR increases accordingly, and vice versa. WaveCast is the best one among the three all-in-one frameworks. When the PSNR is between 30dB and 40dB, WaveCast is constantly 2dB and 4dB better than Softcast3D and Softcast2D respectively. When the PSNR of WaveCast becomes more than 40dB, WaveCast gains less over Softcast, possibly due to skipping more subbands than Softcast.

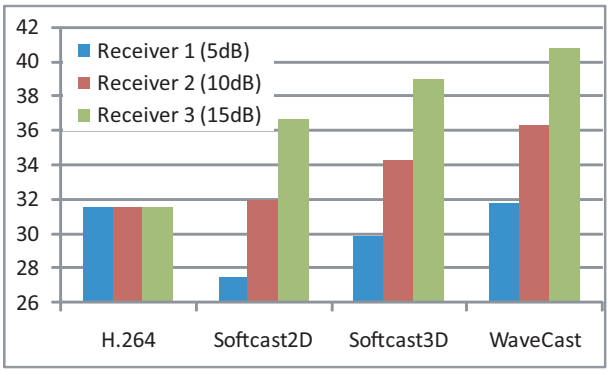


**Fig. 7**. Multicast to three receivers.

We then let all the frameworks to serve a group of three receivers with diverse channel SNR. The SNR for each receiver is 5dB, 10dB and 15dB respectively. The test result is given in Fig.7. In conventional H.264 based framework, the server has to choose low speed combination of FEC and QAM of 802.11, because otherwise the receiver of 5dB will

not be able to decode the video. In light of this, although the other two receivers have better channel condition, they will also only receive low speed 802.11 signal and reconstruction low quality video. In Softcast and WaveCast, the server can accommodate all the receivers simultaneously. Using WaveCast, the user of 5dB can get similar reconstruction quality as using H.264 based conventional framework. Meanwhile, the users of 10dB and 15dB get 4dB and 8dB better reconstruction quality respectively by using WaveCast than conventional framework.

## 4. CONCLUSION

In this paper, we propose a wireless video broadcasting framework called WaveCast, based on 3D wavelet transform. WaveCast takes advantage of both conventional framework and the recently emerging Softcast framework. Similar to Softcast, WaveCast transmits the transform coefficients in lossy mode and performs gracefully in multicast. Meanwhile, WaveCast utilizes MCTF as temporal transform such that the motion is aligned, and utilizes conventional framework to transmit motion information such that the MVs can be reconstructed losslessly. In experiments, WaveCast outperforms Softcast 2dB in average at low channel SNR, and outperforms H.264 based framework up to 8dB in multicast.

## 5. ACKNOWLEDGEMENT

This work was supported in part by the Major State Basic Research Development Program of China (973 Program 2009CB320905), the Program for New Century Excellent Talents in University (NCET) of China (NCET-11-0797), the National Science Foundation of China (NSFC) under grants 61100095 and the Fundamental Research Funds for the Central Universities (Grant No. HIT.BRETIII.201221).

## 6. REFERENCES

[1] "Digital Video Broadcasting (DVB)," Website, 2009, `http://www.etsi.org/deliver/etsi_en/300700_300799/300744/01.06.01_60/en_300744v010601p.pdf`.

[2] T. Wiegand, G. Sullivan, G. Bjintegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 560–576, 2003.

[3] IEEE 802.11 Working Group et al., "IEEE 802.11-2007: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE 802.11 LAN Standards 2007*, 2007.

[4] Feng Wu, Shipeng Li, and Ya-Qin Zhang, "A framework for efficient progressive fine granularity scalable video coding," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 11, no. 3, pp. 332 –344, mar 2001.

[5] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the h.264/avc standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 9, pp. 1103 –1120, sept. 2007.

[6] Szymon Jakubczak, Hariharan Rahul, and Dina Katabi, "One-Size-Fits-All Wireless Video," in *Proc. Eighth ACM SIGCOMM HotNets Workshop*, New York City, NY, October 2009.

[7] Szymon Jakubczak and Dina Katabi, "A cross-layer design for scalable mobile video," in *Proceedings of the 17th annual international conference on Mobile computing and networking*, New York, NY, USA, 2011, MobiCom '11, pp. 289–300, ACM.

[8] Xiaopeng Fan, Feng Wu, and Debin Zhao, "D-Cast: DSC based Soft Mobile Video Broadcast," in *ACM International Conference on Mobile and Ubiquitous Multimedia (MUM)*, Beijing, China, December 2011.

[9] Xiaopeng Fan, Feng Wu, Debin Zhao, Oscar C. Au, and Wen Gao, "Distributed soft video broadcast (DCAST) with explicit motion," in *Data Compression Conference (DCC)*, Snowbird, UT, USA, April 2012.

[10] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," in *Proc. IEEE*, 2005, vol. 93, pp. 71–83.

[11] D. Taubman and A. Zakhor, "Multirate 3-d subband coding of video," *IEEE Transactions on Image Processing*, vol. 3, no. 5, pp. 572–588, 1994.

[12] J.R. Ohm, "Three-dimensional subband coding with motion compensation," *IEEE Transactions on Image Processing*, vol. 3, no. 5, pp. 559–571, 1994.

[13] S.J. Choi and J.W. Woods, "Motion-compensated 3-d subband coding of video," *IEEE Transactions on Image processing*, vol. 8, no. 2, pp. 155–167, 1999.

[14] A. Secker and D. Taubman, "Lifting-based invertible motion adaptive transform (limat) framework for highly scalable video compression," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1530–1542, 2003.

[15] R. Xiong, J. Xu, F. Wu, and S. Li, "Barbell-lifting based 3-d wavelet coding scheme," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1256–1269, 2007.