

Analisis Kinerja *Rate-Distortion Codec Wyner-Ziv Video Coding* Berbasis *Discrete Wavelet Transform*

I Wayan Shandyasa

Magister Teknik Elektro
Program Pasca Sarjana Universitas Udayana
Denpasar, Indonesia
shandyasa@gmail.com

I Made Oka Widyantara

Lab. Sistem Komunikasi Jurusan Teknik Elektro
Fakultas Teknik Universitas Udayana
Denpasar, Indonesia
oka.widyantara@unud.ac.id

Abstrak—Pengkodean video *Wyner-Ziv* (WZVC) merupakan salah satu paradigma baru dalam pengkodean video, dimana korelasi temporal sumber di eksploitasi disisi *decoder*. *Encoder* video hanya mengirimkan informasi seminimal mungkin, dan selanjutnya *decoder* video akan membuat skema pembangkitan *side information* (SI) menggunakan *frame-frame* sebelumnya yang sudah didekodekan. Pembangkitan SI di *decoder* adalah persoalan utama untuk menghasilkan kinerja *rate-distortion* (RD) yang mendekati kinerja RD pengkodean video konvensional seperti MPEG dan H.264. Paper ini akan mereview penerapan teknik *discrete wavelet transform* (DWT) pada *codec* WZVC. Dengan sifat hirarki berlapis, DWT dapat membuat eksploitasi spasial dari orde rendah sampai orde tinggi. Dengan kemampuan ini, *codec* WZVC dapat menerapkan DWT untuk mengeksploitasi hirarki lapisan orde tinggi untuk pembangkitan SI di *decoder*. Implementasi DWT pada *codec* DWT dapat diklasifikasikan menjadi (i) penerapan langsung DWT sebagai teknik transformasi dan (ii) menggunakan hirarki lapisan orde rendah sebagai *auxiliary information* (AI) untuk meningkatkan akurasi pembangkitan SI di *decoder*. Studi literatur menunjukkan bahwa implementasi DWT pada *codec* WZVC mampu menghasilkan kinerja RD yang lebih baik dibandingkan dengan H.264 pada pengkodean Intra. Hal ini memungkinkan *encoder* video dapat dirancang berkompleksitas rendah.

Kata kunci—WZVC; DWT; *Rate-Distortion*; *Side Information*; *Auxiliary Information*;

I. PENDAHULUAN

Teknik pendistribusian video saat ini banyak dilakukan berbasis *client-server*, contohnya seperti pada aplikasi sistem monitoring video. Pada aplikasi ini, sisi *client* memiliki sumber daya terbatas dibandingkan dengan sisi *server*. Dengan skema ini, *client* akan memiliki keterbatasan dalam memproses data video untuk dikirimkan secara *real-time* ke *server*. Teknik pengkodean video yang dikembangkan saat ini seperti MPEG dan H.264 berbasis pengkodean prediksi, dimana encoding video pada sisi *client* sangat kompleks dibandingkan proses decoding di sisi *server*. Maka penerapan teknik pengkodean ini tidak efisien diterapkan pada aplikasi monitoring video berbasis *client-server*.

Untuk memberikan solusi terhadap aplikasi dimana encoding video harus berkompleksitas rendah, saat ini, sebuah teknik pengkodean video yang dikenal sebagai WZVC [1]. WZVC menurunkan kompleksitas dengan memindahkan

prosedur estimasi dan kompensasi gerak ke *decoder* (*server*). Skema pengkodeannya adalah *decoder* harus membangkitkan sebuah *frame side information* (SI) menggunakan beberapa *frame referensi* yang tersedia di *decoder*. *Frame SI* adalah sebuah *frame prediksi* untuk sebuah *frame* yang dikirimkan oleh *encoder*. Semakin baik *frame SI*, maka makin sedikit bit yang harus dikirimkan oleh *encoder*. Realisasi *codec* WZVC didasarkan pada dua teori informasi yaitu teorema Slepian-Wolf [2] dan Wyner-Ziv [3]

Dengan fokus pada proses pembangkitan SI, beberapa *codec* WZVC menerapkan teknik *Discrete Wavelet Transform* (DWT) [4]-[8]. Dengan struktur hirarki yang dimiliki oleh DWT, beberapa skema dalam membangun sebuah *codec* WZVC bisa menggunakan DWT untuk mendapatkan mapping korelasi spasial dalam *frame* WZ. Dari mapping ini akan diperoleh layer yang paling penting untuk dikirimkan ke *decoder* sedangkan korelasi spasial orde rendah akan dibuang [7]. Di sisi *decoder*, layer orde tinggi akan diprediksi menggunakan mapping korelasi dari *frame SI* yang akan tersedia di *decoder*. Skema lain penerapan DWT pada *codec* WZVC adalah menggunakan layer orde rendah dari DWT sebagai *auxiliary information* (AI) untuk pembangkitan *frame SI* di *decoder*. Kombinasi AI dan metode *motion-compensated frame interpolation/extrapolation* (MCFI/E) diharapkan mampu menghasilkan *frame SI* yang lebih akurat [8].

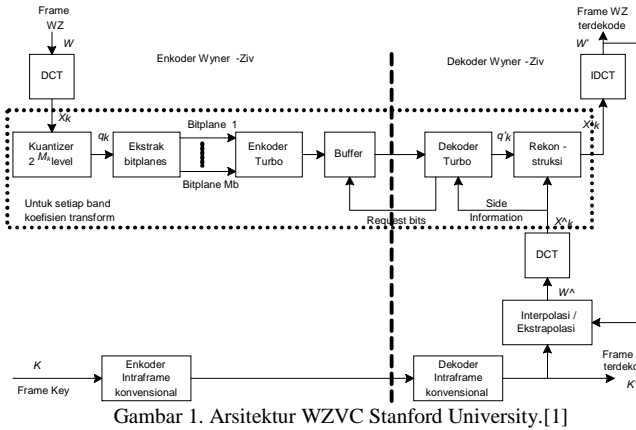
Paper ini akan mengevaluasi implementasi DWT pada *codec* WZVC dalam bentuk perbandingan kinerja *rate-distortion* (RD). Hasil evaluasi dapat dijadikan sebagai dasar pengembangan desain *codec* WZVC.

Selanjutnya paper ini diorganisasikan sebagai berikut; Bab II menjelaskan mengenai teori dasar penggunaan WZVC, Bab III menjelaskan mengenai *codec* WZVC dengan DWT dan AI, Bab IV membahas tentang analisis kinerja RD *codec* WZVC berbasis DWT, dan akhirnya Bab V adalah kesimpulan dari paper ini.

II. STATE OF THE ART WZVC

Pengkodean video WZ pertama kali dikembangkan di Stanford University tahun 2002 untuk domain piksel dan dikembangkan dalam domain transform[1]. Solusi pengkodean ini dikenal dengan *Wyner-Ziv Video Coding* (WZVC) Stanford. Arsitektur utama WZVC Stanford

dikarakteristikkan oleh pengkodean Slepian-Wolf berbasis frame dengan Turbo kode sebagai pengkodean kanal [2]-[3], dan sebuah kanal umpan balik untuk membentuk kontrol laju di decoder[1]



Gambar 1. Arsitektur WZVC Stanford University.[1]

Seperti ditunjukkan pada Gambar 1, pada sisi encoder WZVC pertama akan mengklasifikasi frame menjadi frame WZ (dikodekan WZ) dan frame Key (dikodekan dengan pengkodean video konvensional, H.264). Sebuah transformasi DCT diterapkan untuk eksploitasi spasial frame WZ. Koefisien transformasi selanjutnya dikuantisasi seragam, dan indeks-koefisien selanjutnya dikodekan Turbo dimana bitstream yang dihasilkan dikirimkan ke decoder secara bertahap sesuai permintaan decoder. Kanal *feedback* digunakan oleh decoder untuk meminta *bitstream* tambahan ke encoder.

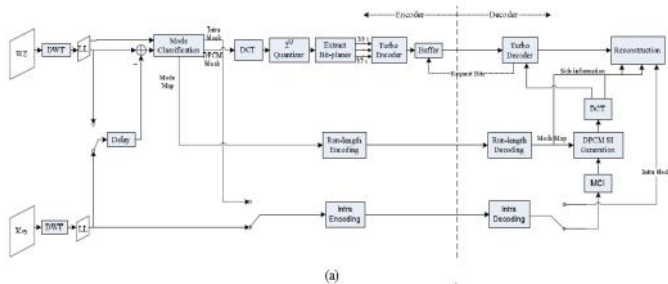
Pada sisi decoder, decoder membuat SI untuk setiap frame WZ dengan membentuk *motion-compensated frame interpolation/extrapolation* (MCFI/E) menggunakan frame-frame terdekat yang sudah didekodekan. SI untuk setiap frame WZ merupakan sebuah estimasi (versi noise) dari frame WZ asli. Semakin baik kualitas estimasi, maka makin kecil jumlah error decoder Turbo yang harus diperbaiki sehingga semakin sedikit jumlah laju bit yang dibutuhkan.

III. IMPLEMENTASI DWT PADA CODEC WZVC

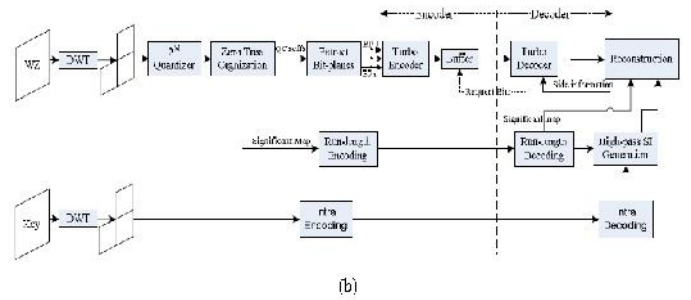
Berdasarkan codec WZVC Stanford, beberapa codec WZVC domain DWT telah dikembangkan

3.1 WZVC domain DWT

DWT diimplementasikan sebagai solusi alternatif menggantikan transformasi DCT pada codec WZVC Stanford.



(a)



(b)

Gambar 2. (a) Diagram low-pass subband (b) Diagram high-pass subband WZ coding.[6]

A. Enkoding WZVC Domain DWT

Untuk dapat mendukung *one level scalability* setiap frame dari input video dipecahkan kedalam *low-low pass* (LL) *subband* untuk mengurangi layer resolusi dan *three high-pass subband*. *Low-pass subband* dan *high-pass subband* adalah frame Key, seluruhnya di encodekan dengan DCT berbasis *intra block coding*. Prosedur pengkodean frame WZ dapat dijelaskan sebagai berikut:

- *Low-pass Subband WZ Coding*

Subband low-pass dari WZ frame diencodekan dengan menggunakan pengkodean DPCM adaptif WZ dan pengkodean blok intra. Seperti ditunjukkan pada Gambar 2.a, koefisien-koefisien *subband low-pass* dipartisi ke blok-blok. Setiap blok diklasifikasikan secara adaptif menjadi mode intra dan mode mode DPCM. Sinyal residu DPCM adalah perbedaan antara blok-blok dalam subband low-pass frame sekarang dan blok-blok *co-located* di subband low-pass frame referensi. Mode klasifikasi didasarkan pada perbandingan mean squared error (MSE) residu DPCM dengan varian blok sekarang, yaitu

$$\text{mode} = \begin{cases} \text{intra} & \text{if } \text{VAR}(X_{\text{intra}}) < \alpha \text{MSE}(X_d) \\ \text{DPCM} & \text{else} \end{cases} \quad (1)$$

- *High-pass Subband WZ Coding*

Untuk mendapatkan statistik korelasi *high-order* dalam domain wavelet, metode *zero-tree* digunakan untuk mencari koefisien *subband high-pass*. Seperti ditunjukkan pada Gambar 2.b, prosedur pengkodeannya adalah (1). untuk subband berbeda, diterapkan kuantisasi scalar seragam dengan M-level berbeda. Koefisien-koefisien kuantisasi diurutkan sesuai struktur *zero-tree* dan peta signifikansi dapat diperoleh. Selanjutnya, peta signifikansi di encodekan *run-length* dan *bitplanes* subband high-pass terkuantisasi diencodekan secara sekuensial dengan encoder Turbo.

B. Dekoding WZVC Domain DWT

Frame kunci didekodekan dengan decoding frame Intra. Terdapat dua strategi dekodig frame WZ yaitu menurunkan resolusi dan full resolusi.

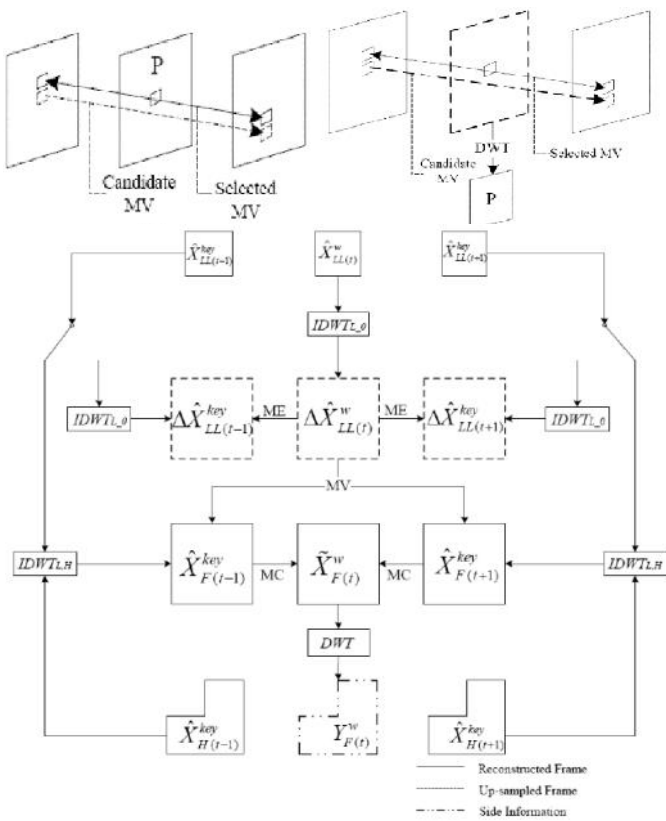
- Dekoding WZ dengan menurunkan resolusi

Seperti ditunjukkan pada Gambar 1a, pertama kali, frame prediksi Y (SI) untuk resolusi yang diturunkan dibangkitkan

dengan MCI (Gambar 3a). Jika decoder tidak memperoleh akses ke subband high-pass, prediksi Y pada saat t diinterpolasi dengan frame resolusi rendah yang telah didekodekan pada saat $t-1$ dan $t+1$. Selanjutnya, mode peta frame WZ sekarang didekodekan dengan decoder *run-length*. Sesuai dengan mode peta, bagian Y_D dari Y diekstrak dan residu DPCM Y_D diperoleh dari $Y_D = Y - X^{ref}$, dimana X^{ref} adalah perbedaan frame. Selanjutnya Y_D di transformasi DCT dan digunakan sebagai SI dalam domain DCT.

• *Full Resolution WZ Decoding*

Untuk decoding WZ full resolution terbagi lagi menjadi 2 yaitu *Low-pass subbands decoding* dan *High-pass subbands decoding*. Dimana WZ decoding untuk *low-pass* dan *high-pass* sama dengan penjelasan sebelumnya yang menjadi pembedanya adalah bagaimana cara mencari atau menentukan *Side Information*nya. Untuk dapat menentukan prediksi dari frame dapat dilihat dari gambar 3(b)



Gambar 3. (a) SI untuk reduced-resolution WZ decoding (b) SI untuk full resolution WZ decoding.[6]

Gambar 4. Perbaikan generasi SI untuk high-pass subband [6]

Untuk Gambar 3(b) pertama frame prediksi dari *full-resolution* diinterpolasi dan direkonstruksi ulang oleh frame kunci, kemudian DWT disisipkan kedalam Y frame dan untuk *low-pass subband* dari Y dipergunakan dalam frame prediksi dari *reduced-resolution* Y dengan menggunakan informasi dari frekuensi spasial subband yang lebih tinggi maka layer akan terkompensasi. Untuk *high-pass subband* WZ decoding dijelaskan pada gambar 3(b). Dengan menggunakan Mapping berarti ketika frekuensi tinggi untuk subband SI di atur kembali

berdasarkan *zero-tree* didalam mapping. *High-pass subband* telah di decode. SI yang telah dihaluskan diusulkan untuk *high-pass subband WZ decoding*. [6] *motion information* dari *reduced-resolution* layer dan *full-resolution* layer memiliki korelasi yang tinggi. Keduanya memiliki korelasi inter-band antara DWT subband. Di dalam skema korelasi inter-band pada frame di eksploitasi odengan menggunakan korelasi temporal dari *reduced-resolution* yang tidak memiliki sampel. Perbaikan dari SI untuk 2 level DWT dapat dilihat pada gambar 5. Dimana *low-pass subband* (LL) mengalami *reduced-resolution* layer.

3.2 Codec WZVC dengan DWT sebagai *Information Auxiliary*

Skema codec WZVC ini diajukan oleh [8]. Seperti ditunjukkan pada Gambar 5, frame Key dari urutan video dikompres dengan menggunakan *codec intra-frame konvensional*. Frame WZ di enkodekan dengan *Codec WZ spatial-aided low-delay*. Detail dari codec ini dapat dijelaskan sebagai berikut :

• *Pengkodean auxiliary information spasial*

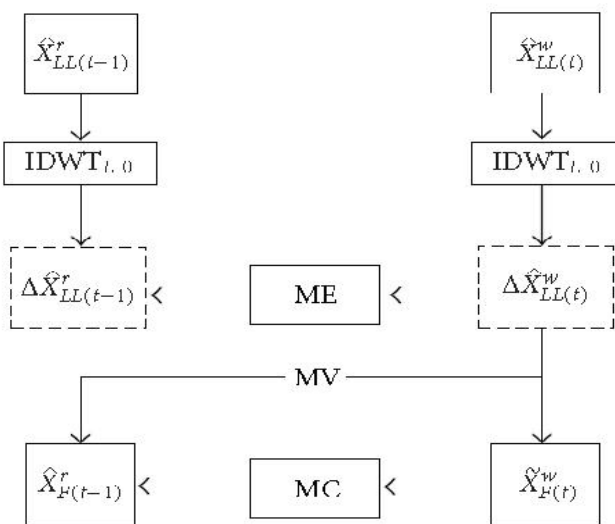
DWT diadopsi sebagai perangkat untuk menghasilkan informasi auxiliary. Di encoder, untuk setiap frame WZ, DWT-2D level satu dengan filter biorthogonal 9/7 diterapkan untuk didekomposisi frame original dan *subband low-low pass* (LL) dari frame sekarang yang digunakan sebagai *auxiliary information* spasial. Hasilnya adalah resolusi AI adalah setengah dari frame aslinya. Untuk menurunkan pengulangan temporal, DPCM dibentuk antara subband LL terdekat untuk encode subband LL. Pada pengkodean DPCM, residu dihitung sebagai perbedaan antara subband LL sekarang dan frame referensi sebelumnya. Kemudian residu di transformasi DCT dan dikuantisasi, Akhirnya koefisien-koefisien kuantisasi diencodekan dengan H.264/AVC. Jika frame referensi adalah frame Jey, maka subband LL dari frame intra full resolusi yang telah direkonstruksi harus diperoleh dengan DWT untuk membentuk frame referensi dalam pengkodean DPCM.

• *SA-MCE-Based Side Information Generation.*

Gerak kompensasi ekstrapolasi adalah metode umum dalam skema *low-delay WZ coding*. Untuk metode MCE, seperti yang ditunjukan pada paper [1], gerak antara frame decoder pada saat t_1 dan waktu t_2 diperkirakan dan *motion estimation* digunakan untuk mengekstrapolasi SI pada waktu t , namun karena tidak adanya informasi dari frame, metode MCE sangat tidak efektif, oleh karena itu, *auxiliary informationspatial* dibantu metode MCE diadopsi dalam paper[8]. Usulan skema SA-MCE generasi SI digambarkan seperti gambar 5. Prosedur rinci dari skema tersebut dalam rangka memperoleh *motion information* untuk *motion compensation* pada *high resolution* AI pada resolusi rendah perlu diunsample pertama, selanjutnya pencarian *motion* dapat dilakukan pada saat upsample resolusi rendah (LL), atau upsample frame-resolusi rendah saat ini dan sebelumnya. Frame resolusi tinggi direkonstruksikan (L-H). Karena kurangnya subband high-pass, yang upsample resolusi rendah

frame mengalami artefak, seperti *blending*, aliasing, dan ubin. Resolusi rendah dalam frame (L) dapat mengganggu pencocokan ketika blok dibandingkan dengan blok kerangka acuan berkualitas tinggi (H). Sebelumnya upsample frame-resolusi rendah memiliki artefak yang sama, sehingga efek artefak bisa dibatalkan oleh artefak serupa. Oleh karena itu melakukan DWT dan IDWT untuk mendapatkan frame upsample dari band LL, bahkan dalam kasus frame sebelumnya menjadi *frame key*. Invers dari DWT transform IDWT_{L,0}. Digunakan untuk upsample subband LL. Dari hasil analisis [8] hasil numeric untuk tingkat distorsi, dapat ditemukan bahwa ketika kualitas AI ditingkatkan, kinerja WZVC dapat ditingkatkan, Oleh karena itu lebih banyak bit dialokasikan untuk AI *coding* daripada frame WZ *coding* yang menginduksi kualitas DPCM kode LL-band menjadi tinggi.

Gambar 5. Metode SA-MCE [8]



Reconstructed frame

--- Up-sampled frame

IV. HASIL DAN PEMBAHASAN

Pada bagian ini akan dipaparkan kinerja RD dari codec WZVC domain DWT dan codec WZVC dengan DWT sebagai IA.

Dalam paper yang membahas mengenai penggunaan domain DWT didistribusikan dalam skema pengkodean video yang mendukung skalabilitas spasial. Pengurangan resolusi layer dikodekan oleh adaptif DPCM WZ dan modus Intra skema pengkodean. Koefisien high-pass subband yang organized oleh *zero-tree* dan kode oleh PCM WZ encoder. Dalam resolusi penuh WZ decoding, efek aliasing kompensasi SI metode generasi diadopsi untuk mengurangi lapisan-resolusi. Selain itu, skema generasi SI halus untuk subband *high-pass* decoding diusulkan. Skema yang diusulkan mengurangi MC efek aliasing dalam domain wavelet dan kinerja Skema yang diusulkan menjanjikan perbandingan DCT domain WZ coding sementara dengan dukungan skalabilitas spasial yang lebih baik.

Dalam paperyang membahas mengenai skema *spatial-aided low-delay WZ coding*. Dalam skema ini, *subband low-pass* dari frame WZ dihasilkan oleh DWT digunakan sebagai tambahan spasial informasi dan dikodekan oleh DPCM. Pada dekoder, Informasi tambahan spasial diterjemahkan terlebih dahulu. dengan melakukan *motion estimation* pada upsampled spasial tambahan informasi, MV lebih akurat diperoleh dengan membandingkan MCE berbasis SI generasi. Peningkatan ini memungkinkan kita untuk menerapkan efisiensi *high-low delay WZ coding*. Dalam studi lebih lanjut, analisis yang lebih umum akan dianggap pada skala penuh saja. The low-pass subband dikodekan dan ditransmisikan sebagai informasi tambahan. The high-pass subband dapat dikodekan secara independen oleh *spatial-aided low-delay WZVC*. Dalam hal ini, semua dampak yang dibawa oleh penipisan, interpolasi berikutnya, dan *simple-coarse* kuantisasi dapat dianggap pada skala penuh secara lebih umum. Selain itu, untuk sepenuhnya mengeksplorasi karakteristik yang diusulkan SA - WZVC.

Berdasarkan penjelasan diatas hasil kinerja RD codec WZVC domain DWT difungsikan untuk meningkatkan akurasi pembangkitan frame SI yang nantinya dipergunakan sebagai acuan dalam proses encoder maupun decoder sebuah video yang akan di transmisikan dalam sebuah system aplikasi *client-server* yang memiliki keterbatasan sumber daya pada *client* dengan menggunakan DWT pada WZVC sebagai solusi untuk mengeksplorasi korelasi spasial pada frame WZ. Dengan struktur hierarki yang dimiliki oleh DWT kita mendapatkan mapping korelasi spasial dalam frame WZ dan memperoleh layer yang paling penting untuk dikirimkan ke decoder, disisi decoder layer orde tinggi akan diprediksi menggunakan mapping korelasi dari frame SI yang sudah di DWT dan frame SI diperoleh dengan skema satu atau lebih frame referensi yang tersedia di decoder. Dengan menggunakan layer orde rendah dari DWT sebagai informasi tambahan atau AI yang digunakan oleh decoder sebagai informasi pendukung pada proses pembangkitan SI.

KESIMPULAN

Paper ini telah memaparkan model implementasi DWT pada codec WZVC untuk meningkatkan akurasi pembangkitan frame SI. Dari dua model implementasi DWT, telah ditunjukkan bahwa Analisis dari kinerja Rate-Distortion bergantung pada model implementasi DWT pada codec WZVC. DWT diterapkan pada WZVC sebagai solusi untuk mengeksplorasi korelasi spasial pada frame WZ. Dengan struktur hierarki yang dimiliki oleh DWT kita mendapatkan mapping korelasi spasial dalam frame WZ dan memperoleh layer yang paling penting untuk dikirimkan ke decoder, disisi decoder layer orde tinggi akan diprediksi menggunakan mapping korelasi dari frame SI yang sudah di DWT dan frame SI diperoleh dengan skema satu atau lebih frame referensi yang tersedia di decoder. Dengan menggunakan layer orde rendah dari DWT sebagai informasi tambahan atau AI yang digunakan oleh decoder sebagai informasi pendukung pada proses pembangkitan SI.

REFERENSI

- [1] Aaron, A., Rane, S., Setton, E. dan Girod, B., "Transform Domain Wyner-Ziv Codec for Video", *Proceedings of SPIE in Visual Communications and Image Processing*, San Jose, California, USA, Vol. 5308, hal. 520-528, Jan.2004
- [2] Slepian, D. dan Wolf, J.K., "Noiseless Coding of Correlated Information Sources", *IEEE Transaction Information Theory*, Vol. IT-19, No.4, hal. 471- 480, Jul.1973.
- [3] A. D. Wyner and J.Ziv, "The rate-distortion function for source coding with side information at the decoder" *IEEE Trans. On Information Theory*, vol. IT-22, no. 1-10, Jan. 1976.
- [4] Xun Gou, Yan Lu, Feng Wu, Wen Gao, "Distributed Video Coding Using Wavelet" *IEEE TransISCAS* 2006, pp.5427-5430.
- [5] R.Benardini, R. Rinaldo, P.Zontone, A.VItali, "Performance Evaluation Of Distributed Video Coding Schemes" *IEEE Trans ICASSP* 2008, pp.709-712.
- [6] Bo Wu, Xiangyang Ji, Debin Zhao, Wen Gao, "Wavelet Based Distributed Video Coding With Spatial Scalability", *IEEE Trans* 2008, pp.3458-3461.
- [7] James E. Fowler, Marco Tagliaschi, Beatrice Perquet-Popescu, "Wavelet-Based Distributed Source Coding Of Video", diakses di internet pada tanggal 18 oktober 2013: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.101.5160&rep=rep1&type=pdf>.
- [8] Bo Wu, Xiangyang Ji, Debin Zhao, Wen Gao, "Spatial-Aided Low-Delay Wyner-Ziv Video Coding", *Research Article EURASIP Journal on Image and Video Processing* 2009 11 pages.

Halaman Ini Sengaja Dikosongkan