

Evaluasi Kinerja *Rate-Distortion* Dengan Penerapan Teknik Pembangkitan *Side Information* Pada Codec *Wyner-Ziv Video Coding*

Anak Agung Ayu Putri Ardyanti
Magister Teknik Elektro
Universitas Udayana
Denpasar, Indonesia
putri.ardiyanti@unud.ac.id

I Made Oka Widyantara
Lab. Sistem Telekomunikasi, Jurusan
Teknik Elektro
Universitas Udayana
Denpasar, Indonesia
oka.widyantara@unud.ac.id

Linawati
Jurusan Teknik Elektro
Universitas Udayana
Denpasar, Indonesia
linawati@unud.ac.id

Abstrak—*Wyner-Ziv Video Coding (WZVC)* adalah paradigma baru dalam pengkodean video, dimana, eksploitasi korelasi temporal video dilakukan sepenuhnya di dekoder. Sasarannya adalah menurunkan kompleksitas enkoder, dengan menggeser teknik estimasi dan kompensasi gerak ke sisi dekoder. Skema umum WZVC adalah mendekodekan frame dari enkoder, dengan membangkitkan *side information (SI)* di dekoder. Semakin akurat frame SI yang dihasilkan untuk sebuah frame yang dikirimkan oleh enkoder, maka kinerja *rate-distortion (RD)* codec WZVC akan semakin baik. Saat ini telah dikembangkan dua codec WZVC dengan metode pembangkitan SI berbeda yaitu codec DISCOVER dan codec WZVC-EM. Codec DISCOVER menerapkan metode *bidirectional motion-compensated frame interpolation (B-MCFI)*, dimana sebuah frame SI diperoleh dengan estimasi dan kompensasi gerak dua frame referensi yang tersedia di dekoder. Sedangkan codec WZVC-EM menerapkan metode pembelajaran medan gerak secara iteratif berbasis pada algoritma Expectation-Maximization (EM). Metode ini akan memperbaharui medan gerak secara iteratif untuk meningkatkan akurasi soft SI. Paper ini akan mengevaluasi kinerja RD kedua codec tersebut pada struktur *group of picture (GOP)* berbeda. Sasarannya adalah untuk memperoleh kerangka kerja codec WZVC yang efisien untuk dikembangkan lebih lanjut. Hasil evaluasi menunjukkan bahwa kinerja RD codec WZVC-EM lebih baik untuk ukuran GOP lebih besar dari dua.

Kata Kunci—*Side Information, Wyner-Ziv Video Coding, Codec DISCOVER, Codec WZVC-EM, Estimasi dan kompensasi gerak.*

I. PENDAHULUAN

Perkembangan digital khususnya bidang video saat ini mengalami kemajuan yang sangat pesat mulai aplikasi *video on demand* sampai *tv streaming*. Pengkodean video konvensional seperti H.26x dan MPEG.x berbasis pengkodean prediktif dimana enkoder bekerja lebih tinggi daripada dekoder. Metode ini sangat cocok untuk media penyimpanan digital dan *video streaming* karena sumber aplikasi dikodekan

sekali di enkoder dan didekodekan beberapa kali. Namun dalam aplikasinya metode ini tidak cocok untuk pengkodean video yang diterapkan seperti pada *wireless video surveillance, wireless PC camera*, dimana seluruh aplikasi seperti ini menginginkan kinerja enkoder menjadi ringan. Munculah teknik pengkodean video berdasarkan teori informasi Slepian-Wolf (SW) [1] dan Wyner-Ziv (WZ) [2] yaitu *Distributed Video Coding* [3] yang sering disebut juga dengan DVC [1] [2]. Teorema ini menawarkan solusi dimana kinerja enkoder menjadi lebih rendah dengan memindahkan efisiensi medan gerak ke dekoder. Pengkodean sumber terdistribusi inilah yang menjadi dasar dari pengembangan *Wyner-Ziv Video Coding (WZVC)*. WZVC menerapkan teorema WZ dengan pembangkitan SI di dekoder. Dengan skema ini, enkoder hanya mengirimkan informasi yang terbatas yang tentang sebuah frame untuk menurunkan kapasitas bit transmisi. Selanjutnya, dekoder akan mendekodekan frame video dengan membangkitkan SI, yaitu skema estimasi dan kompensasi gerak menggunakan hanya frame-frame yang telah didekodekan sebelumnya di dekoder sebagai frame referensi. Semakin akurat SI yang dihasilkan, maka semakin baik kinerja *rate-distortion (RD)* codec WZVC.

Sebagai usaha untuk merealisasikan sebuah codec WZVC praktis beberapa kelompok peneliti telah mengajukan berapa rancangan codec WZVC. Codec pertama diusulkan [3][4][5] dan dikenal sebagai codec (*coder/decoder*) WZVC Stanford. Codec ini menerapkan teknik *motion compensation interpolation/extrapolation (MCI/E)* untuk membangkitkan SI, menggunakan dua frame referensi di dekoder. Codec ini tidak sepenuhnya praktis, karena menggunakan frame referensi yang disediakan langsung di dekoder.

Untuk meningkatkan kinerja RD codec WZVC, beberapa codec WZVC telah diusulkan dalam literature seperti codec DISCOVER [6] dan WZVC-EM [7]. Kedua codec menggunakan kerangka kerja codec WZVC Stanford, tetapi dengan metode pembangkitan SI yang berbeda. Codec DISCOVER menerapkan metode *Bidirectional Motion-Compensated Frame Interpolation (B-MCFI)*, dimana sebuah frame SI diperoleh dengan estimasi dan kompensasi gerak dua frame referensi yang tersedia di dekoder. Sedangkan codec WZVC-EM menerapkan metode

pembelajaran medan gerak secara iteratif berbasis pada algoritma Expectation-Maximization (EM). Metode ini akan memperbaharui medan gerak secara iteratif untuk meningkatkan akurasi soft SI.

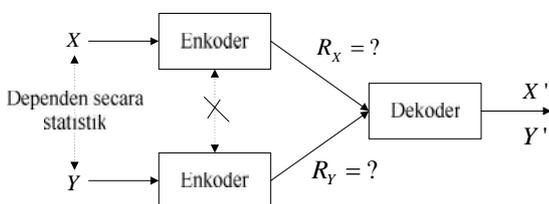
Perbedaan metode SI yang digunakan untuk mendesain codec WZVC menunjukkan bahwa parameter utama untuk meningkatkan kinerja codec WZVC adalah bagaimana mendapatkan sebuah metode pembangkitan SI yang efisien. Oleh karena itu, paper ini akan mengevaluasi lebih lanjut penerapan metode pembangkitan SI ini dengan membandingkan kinerja RD kedua codec pada ukuran GOP yang berbeda. Sampai saat ini belum ada literatur yang mengkaji evaluasi kinerja RD kedua codec tersebut. Sasarannya adalah akan diperoleh sebuah kerangka kerja codec WZVC yang efisien untuk dikembangkan lebih lanjut.

Selanjutnya, paper ini diorganisasikan sebagai berikut: Bab 2 menjelaskan tentang *state of the art* codec WZVC, Bab 3 menjelaskan tentang pembangkitan SI pada codec WZVC DISCOVER dan WZVC-EM, Bab 4 membahas tentang Kinerja RD kedua codec tersebut, dan akhirnya Bab 5 adalah kesimpulan dari paper ini.

II. STATE OF THE ART WZVC

2.1 Pengkodean Sumber Terdistribusi

Teori pengkodean sumber terdistribusi ditujukan untuk pengkodean dua atau lebih sekuen random dependen secara terdistribusi. Setiap sumber diencodekan secara independen dan bitstreamnya dikirim ke dekoder tunggal untuk decoding secara gabungan. Untuk mengkompres dua sumber dependen secara statistik, laju encoding minimum R yang dibutuhkan untuk memulihkan kedua sumber tanpa kesalahan di dekoder gabungan, oleh teori informasi Shannon rekonstruksi *lossless* dari X dan Y dinyatakan dengan $R_X \geq H(X)$. Dimana teorema Informasi menegaskan bahwa ketika R memenuhi $R \geq R_X + R_Y \geq H(X,Y)$, X dan Y dapat direkonstruksi dengan akurat di dekoder, selama kedua sumber tersebut saling berkorelasi dan laju bit total R lebih besar dari entropy gabungan $H(X,Y)$. Sebuah pertanyaan muncul apakah memungkinkan untuk memulihkan dua sekuen dependen ini dengan probabilitas kesalahan rekonstruksi kecil menggunakan laju bit $R_X < H(X)$ atau $R_Y < H(Y)$.



Gambar 1 Skema enkoding independen dan dekoding gabungan untuk dua sumber [1]

Slepian dan Wolf [1] telah mempelajari persoalan pada Gambar 1 dan ini merupakan studi pertama tentang Pengkodean Sumber Terdistribusi. Kombinasi laju bit yang

mungkin dari R_X dan R_Y untuk merekonstruksi X dan Y dengan probabilitas error kecil, dinyatakan dengan:

$$R_X \geq H(X|Y) \tag{1}$$

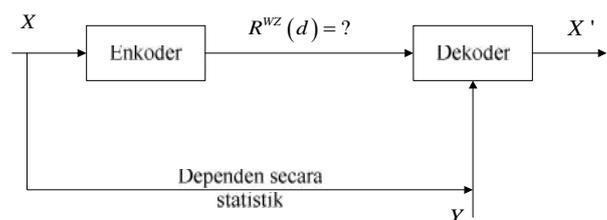
$$R_Y \geq H(Y|X) \tag{2}$$

$$R_X + R_Y \geq H(X,Y) \tag{3}$$

Konsep pengkodean Wyner-Ziv (WZ) [1] sangat cocok diterapkan pada aplikasi pengkodean video. Dalam literatur pengkodean video dikenal sebagai *Wyner-Ziv Video Coding* (WZVC). Dalam kasus ini, beberapa frame sekuen video diencodekan Wyner-Ziv, sedangkan frame sisa diencodekan menggunakan standar pengkodean citra/video tradisional seperti JPEG atau H.264. Dekoder, dengan menggunakan frame-frame yang diencodekan secara tradisional kemudian membangkitkan estimasi frame sekarang melalui teknik interpolasi/ekstrapolasi frame. Frame ini yang kemudian digunakan dalam proses dekoding frame WZ sekarang (diencodekan WZ).

Teorema SW [1] dan WZ [2] yang sudah sangat dikenal, dalam Teori Informasi menegaskan bahwa adalah mungkin untuk mengkompres dua sinyal dependen secara statistik dengan cara terdistribusi (enkoding terpisah, dekoding gabungan) menggunakan laju bit yang serupa seperti yang digunakan dalam sistem di mana sinyal diencodekan dan didecodekan secara bersama-sama (gabungan), seperti pada pengkodean sumber tradisional (gambar 1).

Wyner dan Ziv [2] telah mempelajari kasus khusus pada pengkodean SW yang berkaitan dengan laju bit di koordinat $(H(X|Y), H(Y))$, di Gambar 2. Kasus khusus ini berkaitan dengan pengkodean sumber sekuen X dengan mempertimbangkan sekuen Y , dikenal sebagai *side information* (SI), tersedia di dekoder (Gambar 2) dalam banyak literatur dikenal sebagai kompresi lossy dengan dekoder SI. Disebut kompresi lossy karena Wyner dan Ziv mempertimbangkan distorsi (D) rata-rata yang dapat diterima, diantara sekuen X dan versi dekodenya (X').

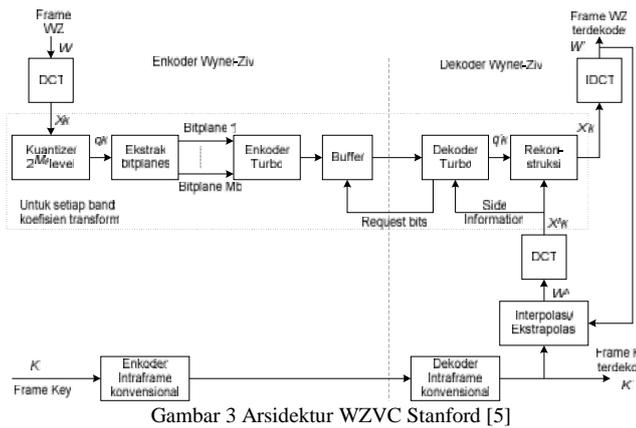


Gambar 2 Kompresi lossy dengan side information [2]

Pada Gambar 2 dapat dilihat dimana X adalah sekuen yang akan diencodekan sehingga disebut sebagai informasi utama, dan Y disebut *side information* adalah dianggap tersedia di dekoder.

2.2 Codec WZVC Domain Transform

Berdasarkan pengkodean sumber terdistribusi, sebuah codec WZVC praktis domain transform pertama kali dikembangkan di Stanford University [5]. Solusi pengkodean ini dikenal dengan *Wyner-Ziv video Coding* (WZVC) Stanford.



Gambar 3 Arsitektur WZVC Stanford [5]

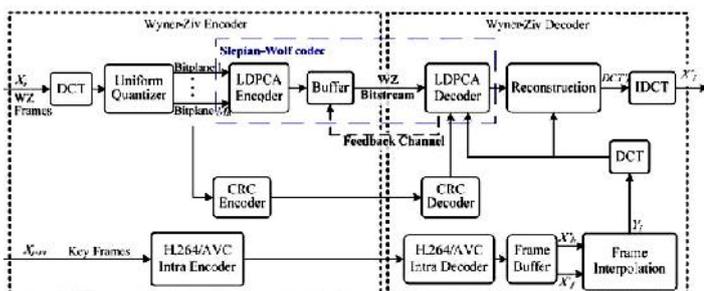
Seperti ditunjukkan pada Gambar 3, codec WZVC Stanford pada prosedur encoding frame WZ di domain transform dilakukan dengan mentransformasi DCT dan kuantisasi frame WZ. Teknik pengkodean kanal Turbo code diterapkan untuk pembangkitan bit parity dari bitplane indeks kuantisasi. Bit parity ini nantinya akan dikirim ke dekoder secara bertahap untuk memperbaiki error dalam SI. Dimana control laju dibentuk di dekoder dengan saluran umpan balik, dekoder meminta kembali bit parity jika prosedur dekoding tidak berhasil mengestimasi frame WZ.

III. PEMBANGKITAN SIDE INFORMATION PADA CODEC WZVC

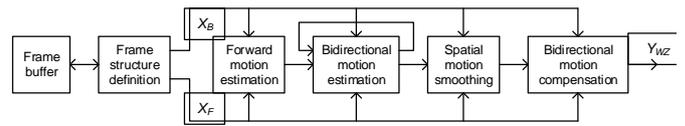
Sebagai usaha meningkatkan kinerja RD codec WZVC Stanford, beberapa desain codec WZVC baru telah diusulkan, dengan fokus utama perbaikan pada metode pembangkitan SI. codec DISCOVER [6] menerapkan teknik B-MCFI dan WZVC-EM [7]. metode pembelajaran tanpa supervisi medan gerak berbasis pada algoritma EM. Detail algoritma pembangkitan SI di kedua codec tersebut dijeaskan pada subbab berikut.

3.1 WZVC DISCOVER

Pada codec WZVC DISCOVER, pembangkitan SI menggunakan metode *bidirectional motion-compensated frame interpolation* (B-MCFI) berbasis blok yang dikembangkan oleh [8]. Metode menggunakan dua frame referensi yaitu satu frame temporal sebelumnya (X_B) dan frame temporal berikutnya (X_F).



Gambar 4 Arsitektur codec WZVC DISCOVER [9]



Gambar 5 Kerangka kerja interpolasi frame untuk pembuatan SI [10]

Kerangka kerja interpolasi frame untuk pembangkitan SI seperti yang ditunjukkan pada gambar diatas dapat dijelaskan sebagai berikut [10]:

• Algoritma *Forward Motion Estimation*

Algoritma pencocokan blok digunakan untuk mengestimasi medan gerak antara frame X_B dan frame X_F dimana jumlah vektor gerak, kedua frame X_B dan X_F akan di *low pass filter* (LPF) terlebih dahulu sebelum dicocokkan. Algoritma *Forward Motion Estimation* (FME) menggunakan kriteria *matching mean absolute difference* (MAD) berdasarkan pada *cost function* (CF) minimal dan dinyatakan sebagai [10]:

$$(v_x, v_y) = \arg \min_{d_x, d_y} CF(d_x, d_y) \quad ; d_x = [-M, M]; d_y = [-M, M] \quad (4)$$

$$MAD(d_x, d_y) = \frac{1}{N} \sum_{(x,y) \in B} |X_F(x, y) - X_B(x + d_x, y + d_y)| \quad (5)$$

• *Bidirectional Motion Estimation* (BME)

Algoritma BME diterapkan untuk memperbaiki vektor-vektor gerak pada blok sekarang di frame interpolasi Y sebelumnya dan berhubungan dengan blok-blok tetangga terdekatnya yang telah diperoleh pada prosedur FME dan dirumuskan sebagai berikut[10]:

$$(x_1, y_1) = (x, y_i) + MV(B_i) \quad (6)$$

$$(x_2, y_2) = (x, y_i) - MV(B_i) \quad (7)$$

• *Spatial Motion Smoothing*

Algoritma *spatial motion smoothing* diterapkan untuk menurunkan jumlah vektor gerak yang salah dengan membandingkan motion vektor yang salah dengan medan gerak yang sebenarnya pada codec DISCOVER menggunakan *weighted median vector filter* dengan formula sebagai berikut[10]:

$$\sum_{j=1}^N w_j \|x_{wvmf} - x_j\|_L \leq \sum_{j=1}^N w_j \|x_i - x_j\|_L \quad (8)$$

Pemilihan bobot dibentuk selaras dengan error prediksi, dinyatakan sebagai:

$$w_j = \frac{MSE(x_c, B)}{MSE(x_j, B)} \quad (9)$$

Mean square error (MSE) menyatakan *matching* yang telah berhasil antara blok sekarang B di frame X_F dan blok di frame X_B yang dikompensasi dengan vektor x_c dan x_j . Pada area sekitar hidung terdapat kesalahan interpolasi medan gerak, karena vektor-vektor gerak hanya meminimalkan MSE antara frame-frame Key tetapi tidak menyatakan medan gerak yang sebenarnya. Kesalahan interpolasi medan gerak pada Gambar selanjutnya dihaluskan dengan *weighted vector median filter* menggunakan vektor-vektor gerak tetangga sebagai faktor bobot.



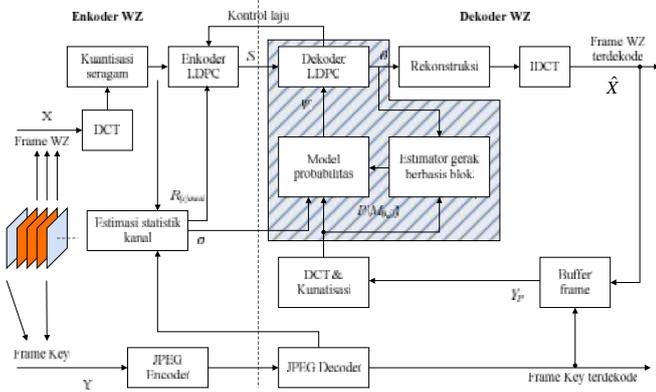
Gambar 6 Frame #7 dari sekuen video Foreman (a) dengan (b) tanpa *spatial motion smoothing* [11]

• *Bidirectional Motion Compensation*

Setelah vektor-vektor akhir telah diperoleh, frame interpolasi Y dibentuk dengan bidirectional motion compensation. Diasumsikan bahwa interval waktu antara frame Key XB dengan frame interpolasi Y adalah sama dengan interval waktu antara frame interpolasi Y dengan frame Key XF, sehingga setiap frame Key memiliki bobot yang sama (1/2), ketika estimasi gerak dibentuk .

3.2 Codec WZVC-EM

Varodayan dkk [12] mengajukan desain codec WZVC dengan skema pembangkitan SI menggunakan metode pembelajaran medan gerak tanpa supervisi berbasis algoritma EM, dan dikenal sebagai codec WZVC-EM. Seperti ditunjukkan pada Gambar 7, untuk mengkodekan frame WZ, dekoder membentuk skema pembelajaran gerak tanpa supervisi (*unsupervised*) menggunakan *bitstream* frame WZ secara langsung dan satu frame SI yang sudah didekodekan sebelumnya. Medan gerak yang dihasilkan selanjutnya digunakan untuk membangkitkan soft SI. Detail kerangka kerja codec WZVC-EM dapat dijelaskan sebagai berikut :



Gambar 7 Arsitektur codec WZVC berbasis EM domain transform [12]

• Pada enkoder, frame *Key* (*Y*) di kodekan dan ditransmisikan dengan pengkodean konvensional JPEG. Enkoding frame WZ (*X*) dilakukan secara independen dari frame *Y*. Enkoder Wyner-Ziv mentransformasi frame *X* ke dalam koefisien *DCT* berbasis blok $k \times k$ (X^T), dimana *k* adalah juga ukuran pencocokan blok pada estimator gerak. Koefisien-koefisien transform (X^T) kemudian dikuantisasi menjadi indek-indek kuantisasi (X^{TQ}), menggunakan tabel kuantisasi rekomendasi JPEG. Indek-indek koefisien terkuantisasi X^{TQ} dikomunikasikan ke dekoder menggunakan LDPC code [12].

• Pada sisi dekoder, piksel-piksel dari tidak digunakan secara langsung karena sistem *lossless* bekerja dalam domain transform. Oleh karenanya, di *DCT* berbasis blok $k \times k$, yang menghitung transform seluruh blok $k \times k$ di seluruh pergeseran piksel. Dengan cara ini, koefisien-koefisien transform dari setiap kandidat gerak dari tersedia di : estimator gerak dan model probabilitas, untuk memperoleh *soft* SI (*E*). Akhirnya, indek-indek koefisien kuantisasi dari *X*, dipulihkan dari sistem *lossless*, direkonstruksi di pusat *bin-bin* kuantisasinya (*centroid reconstruction*) dan ditransformasi balik untuk menghasilkan frame rekonstruksi \hat{X} .

Pada gambar 7, bagian yang diarsir adalah bagian penerapan algoritma EM dalam codec WZVC-EM dan secara detail dijelaskan sebagai berikut:

• Model

Menurut [7] : model distribusi probabilitas posterior dekoder dari sumber *X* berbasis pada parameter *soft estimate* dekoder LDPC θ untuk menghitung distribusi probabilitas posteriori dari *M* adalah:

$$P_{app} \{M\} \equiv P\{M | Y, S; \theta\} \propto P\{M\} P\{Y, S | M; \theta\} \tag{12}$$

dengan step kedua hukum Bayes. Bentuk ini menegaskan sebuah bentuk solusi iteratif EM.

• Algoritma E-step

E-step memperbaharui distribusi *M* yang telah diestimasi dengan mengacu parameter θ . Secara sederhana estimasi *M* dilakukan di vektor gerak berbasis blok $M_{(u,v)}$ dan *S* diabaikan. Pada ukuran blok *k*, setiap blok $k \times k$ dari $(t-1)$ dibandingkan dengan setiap *Y*, pada blok *co-located* dan sekitarnya pada range pencarian gerak tetap. Untuk sebuah blok (u,v) dengan piksel kiri atas dilokasikan di (u,v) , distribusi pada pergeseran $M_{(u,v)}$ diperbaharui dan dinormalisasikan [7]:

$$P_{app}^{(t)} \{M_{u,v}\} := P_{app}^{(t-1)} \{M_{u,v}\} P\{\hat{Y}_{(u,v)+M_{u,v}} | M_{u,v}; \theta_{u,v}^{(t-1)}\} \tag{13}$$

dimana $(u,v)+M_{u,v}$ adalah blok $k \times k$ dengan piksel kiri atas di $((u,v)+M_{u,v})$ dan $P\{\hat{Y}_{(u,v)+M_{u,v}} | M_{u,v}; \theta_{u,v}^{(t-1)}\}$ adalah probabilitas pengamatan $(u,v)+M_{u,v}$ dimana vektor $M_{u,v}$ dari $X_{u,v}$ (diparameterkan oleh $\theta_{u,v}^{(t-1)}$).

Secara umum, probabilitas bahwa pencampuran SI memiliki nilai di pixel (i,j) adalah:

$$\begin{aligned} \mathbb{E}^{(t)}(i, j, \mathcal{S}) &= \sum_m P_{app}^{(t)} \{M_{i,j} = m\} P\{X(i, j) = \mathcal{S} | M_{i,j} = m, \hat{Y}\} \\ &= \sum_m P_{app}^{(t)} \{M_{i,j} = m\} P_z \{\mathcal{S} - \hat{Y}_m(i, j)\} \end{aligned} \tag{14}$$

dimana $p_z(z)$ adalah *probability mass function* (pmf) dari noise additive independen *Z*, dan \hat{Y}_m adalah frame rekonstruksi sebelumnya terkompensasi pada *m* konfigurasi gerak.

Persamaan 14 menunjukkan bahwa penjumlahan bobot meliputi seluruh versi pergeseran dari \hat{Y}_m . Selama seluruh distribusi $M_{(i,j)}$ diperoleh, maka operasi pencampuran

memungkinkan seluruh kandidat dari proses pergeseran berkontribusi secara partial untuk soft SI, (i, j) .

Model blok kandidat ini membatasi ketergantungan statistik (soft matching) antara $x[i]$ dan $(u, v)[i, m_i]$ dengan parameter Laplacian (\dagger), sehingga :

$$V_l = \frac{1}{1 + 1 + \dots + 1} \quad (15)$$

dimana,

$$1 = \begin{cases} \exp\left(-\frac{l^2}{2}\right), & \text{jika } 0 \leq l \leq 2^{m-1} \\ \exp\left(-\frac{(2^m - l)^2}{2}\right), & \text{jika } 0 \leq l \leq 2^m \end{cases}$$

(16)

• Algoritma M-step

Pada M-step, dekoder LDPC memperbaharui menggunakan yang telah dibangkitkan menggunakan metode *joint bitplane LDPC decoding* [13].

$$\begin{aligned} \hat{u}^{(t)} &:= \arg \max_{\Theta} P\{\hat{Y}, S, \Theta\} \\ &= \arg \max_{\Theta} \sum P_{app}^{(t)}\{M = m\} P\{\hat{Y}, S | M = m, \Theta\} \end{aligned}$$

(17)

Sebelum dinormalisasi pada $\omega \in \{0, \dots, 2^d - 1\}$, $^{(t)}$ dihitung sebagai:

$$\hat{u}^{(t)}(i, j, \hat{S}) := \mathbb{E}^{(t)}(i, j, \hat{S}) \prod_{g=1}^d (r_g^{(t)})^{1_{[S_g=1]}} (1 - r_g^{(t)})^{1_{[S_g=0]}} \quad (18)$$

dimana r_g merupakan bit ke- g dalam *Gray mapping* dari nilai luminance dan $1_{[\cdot]}$ merupakan fungsi indikator. M-step juga membangkitkan *hard estimate* \hat{x} dengan mengambil satu nilai yang paling mungkin untuk setiap piksel yang sesuai dengan .

$$\hat{X}(i, j) = \arg \max_{\hat{S}} \hat{u}^{(t)}(i, j, \hat{S}) \quad (19)$$

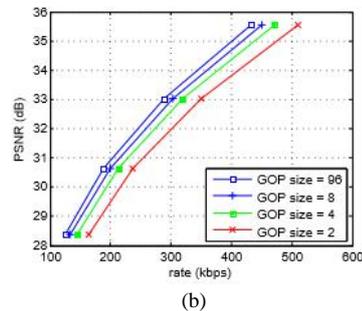
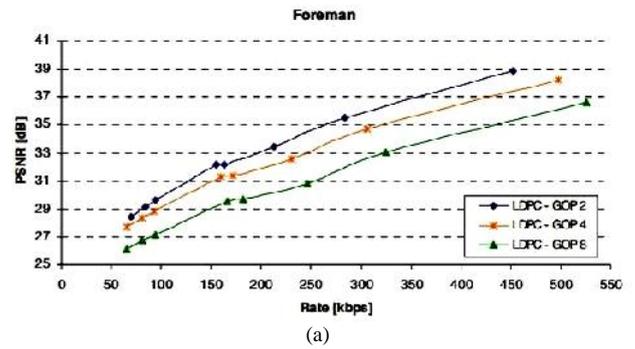
dengan iterasi melalui M-step dan E-step, dekoder LDPC meminta bit-bit syndrome tambahan jika estimasi tidak konvergen. Algoritma berhenti ketika *hard estimate* \hat{x} menghasilkan syndrome yang identik dengan S .

IV. HASIL DAN PEMBAHASAN

Pada bagian ini akan dibahas evaluasi kinerja RD codec WZVC DISCOVER dan WZVC-EM, masing-masing dengan menerapkan metode pembangkitan SI yang berbeda dimana telah dipaparkan pada bab sebelumnya. Pembahas kinerja RD dilihat dari hasil kedua codec yaitu codec dilihat dari hasil GOP 2, 4 dan 8 dari video Foreman (sequen video ini dijadikan video percobaan oleh [8] dan [10]). Video Foreman sendiri merupakan contoh video dengan gerak cepat.

Secara umum codec WZVC DISCOVER pada gambar 8 (a) menunjukkan penurunan kinerja RD terlihat dari penurunan grafik dengan menggunakan GOP ukuran 2, 4 dan 8. Hal ini menunjukkan bahwa teknik B-MCFI masih belum dapat berjalan dengan baik ketika GOP lebih besar dari 2 dimana korelasi temporal dengan pergerakan cepat untuk Foreman belum mampu diestimasi. Jarak temporal frame

referensi berubah jauh sehingga sulit untuk diperkirakan frame selanjutnya yang terdapat di frame WZ.



Gambar 8 Performa RD GOP 2 Foreman : (a) WZVC DISCOVER [8] dan (b) WZVC-EM [10]

Gambar 8 (b) WZVC-EM menunjukkan bahwa terjadi peningkatan grafik pada GOP ukuran lebih dari 2, pada hal ini peningkatan ini disebabkan oleh medan gerak yang terus diperbaharui (interaktif) dan motion field yang diperbaiki terus menerus. Teknik pembelajaran tanpa supervisi medan gerak ini melakukan perbaikan estimasi medan gerak secara kontinyu sehingga frame SI kompensasi (soft SI) mampu mendekati frame WZ. Kondisi ini menyebabkan codec WZVC-EM mampu meningkatkan kinerja RD terutama pada ukuran GOP lebih besar dibandingkan codec WZVC DISCOVER.

V. KESIMPULAN

Paper ini telah memaparkan dan menjelaskan skema berbeda pada pembangkitan soft SI yang digunakan pada setiap codec WZVC, mulai dari metode *Motion-Compesated Frame Interpolation/ Extrapolation* (MCFI/E) pada codec WZVC Stanford, *Bidirectional Motion-Compesated Frame Interpolation* (B-MCFI) pada codec DISCOVER, sampai skema pembelajaran tanpa supervisi medan gerak pada codec WZVC-EM. Teknik pembelajaran tanpa supervisi medan gerak melakukan perbaikan estimasi medan gerak secara terus-menerus sehingga frame SI kompensasi (soft SI) mampu mendekati frame WZ. Kondisi ini menyebabkan codec WZVC-EM mampu meningkatkan kinerja RD terutama pada ukuran GOP lebih besar dibandingkan codec WZVC Discover.

Semakin akurat Soft SI semakin identik dengan frame WZ, dalam paper ini dapat diketahui bahwa pada codec

WZVC DISCOVER dengan menggunakan metode *Bidirectional Motion-Compensated Frame Interpolation* (B-MCFI) hanya mampu bekerja baik pada GOP 2, hal ini menyebabkan enkoder bekerja lebih banyak dibandingkan Codec WZVC-EM. Dari hasil ini, pengembangan codec WZVC sebaiknya diarahkan pada sistem pembelajaran di SI yang nantinya akan mampu memperbaiki medan gerak awal untuk menghasilkan medan gerak baru untuk menambah kinerja RD dan mengurangi beban di enkoder.

REFERENSI

- [1] 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.
- [2] Wyner, A.D. dan Ziv, J., "The Rate-Distortion Function for Source Coding with Side Information at the Dekoder", *IEEE Transaction. Information Theory*, Vol. IT-22, No. 1, hal. 1–10, Jan.1976.
- [3] Girod, B., Aaron, A., Rane, S., dan Monedero, D.R., "Distributed Video Coding", *Proceedings of IEEE, Spscial Issue on Advances in Video Coding and Delivery*, Vol. 93, No. 1, hal. 71-83, Jan. 2005.
- [4] Aaron, A., Zhang, R., dan Girod, B., "Wyner-Ziv Coding of Motion Video", *Asilomar Conference on Signals, Systems and Computer*, Pacific Grove, USA, Vol. 1, hal. 240-244, 3-6 Nov. 2002
- [5] 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.
- [6] Artigas, X., Ascenso, J., Dalai, M., Klomp, S., Kubasov, D. dan Ouaret, M., "The DISCOVER Codec: Architecture, Techniques and Evaluation", *Proceedings of Picture Coding Symposium*, Lisbon, Portugal, 2007.
- [7] Varodayan, D., "Adaptive Distributed Source Coding", Desertasi Ph.D. Stanford University, 2010.
- [8] Kubasov, D., Nayak, J. dan Guillemot, C., "Optimal Reconstruction in Wyner-Ziv Video Coding with Multiple Side Information", *Proceedings of IEEE International Workshop Multimedia Signal Processing*, Chania, Greece, hal. 183-186, 2007.
- [9] Pereira, F., Ascenso, J., Brites, C., "Studying the GOP Size Impact on the Performance of Feedback Channel-based Wyner-Ziv Video Codec", *IEEE Pacific Rim Symposium on Image Video and Technology, Santiago, Chile, December 2007*.
- [10] Ascenso, J., Brites, C. dan Pereira, F. "Content Adaptive Wyner-Ziv Video Coding Driven by Motion Activity", *Proceedings of IEEE International Conference Image Processing*, Atlanta, Georgia, 2006.
- [11] Ascenso, J., Brites, C. dan Pereira, F. (2005), "Improving Frame Interpolation with Spatial Motion Smoothing for Pixel Domain Distributed Video Coding", *Proceedings of EURASIP Conference Speech Image Processing, Multimedia Communication Services*, Smolenice, Slovak Republic, 2005.
- [12] Varodayan, D., Chen, D., Flierl, M. dan Girod, B., "Wyner-Ziv Coding of Video with Unsupervised Motion Vector Learning", *EURASIP Signal Processing: Image Communication Journal, Special Issue on Distributed Video Coding*, Vol. 23, No. 5, hal. 369-378, Jun.2008.
- [13] Varodayan, D., Mavlankar, A., Flierl, M. dan Girod, B. (2006), "Distributed Coding of Random Dot Stereogram with Unsupervised Learning of Disparity", *Proceedings of IEEE International Workshop Multimedia Signal Processing*, hal. 5-8