

TEKNIK ESTIMASI GERAK PENCARIAN PENUH DENGAN AKURASI SETENGAH PIKSEL UNTUK *FRAME RATE UP CONVERSION VIDEO*

P. A. Satya Prabhawa¹⁾, I M. O. Widyantara²⁾, G. Sukadarmika³⁾

^{1,2,3)}Lab. Sistem Telekomunikasi, Jurusan Teknik Elektro Fakultas Teknik Universitas Udayana
Email: ary.satya07@gmail.com, oka.widyantara@unud.ac.id, sukadarmika@ee.unud.ac.id

ABSTRAK

Saat ini Teknologi video digital banyak digunakan pada aplikasi hiburan, contohnya adalah TV Digital dengan format HD. Dengan frame rate tinggi, pengkodean video akan menghasilkan laju bit lebih tinggi yaitu sampai 15 – 30 fps. Permasalahannya adalah kapasitas saluran transmisi memiliki kapasitas terbatas. Solusinya adalah menurunkan laju bit dengan menurunkan jumlah frame video ke penerima. Skema ini dikenal dengan Frame Rate Up-Conversion (FRUC) video, dimana frame yang di encoder akan direkonstruksi kembali di decoder dengan membangkitkan frame intermediate (FI). FI dibangkitkan dengan teknik Motion Compensation Interpolation (MCI). Terkait dengan metode FRUC, penelitian ini mengajukan skema MCI unidirectional dengan pencarian gerak akurasi setengah piksel. Pada skema ini, sebuah motion vector (MV) kandidat akan dicari di frame referensi, proses estimasi gerak dilakukan dengan menambah piksel sisipan diantara piksel eksisting. Sasarannya adalah meningkatkan akurasi MV kandidat. Hasil simulasi menunjukkan bahwa metode yang diajukan lebih baik sampai sebesar masing – masing 3,21 dB dan 3,11 dB pada wilayah pencarian 7 dan 15 piksel dibandingkan dengan metode frame repetition untuk sekuen video foreman dan hall monitor.

Kata Kunci : frame rate, decoding, FRUC, motion estimation, unidirectional, frame intermediate.

1. PENDAHULUAN

Saat ini Teknologi video digital banyak digunakan pada aplikasi hiburan, contohnya adalah TV Digital dengan format HD. Dimana gambar video yang ditampilkan memiliki frame rate tinggi yaitu sebesar 15 – 30 fps. Pada sisi pengkodean video digital dengan frame rate tinggi menghasilkan laju bit tinggi. Pada sisi lain, rekonstruksi video yang baik dengan frame rate tinggi menjadi pertimbangan utama dalam mendesain layanan video. Salah satu cara yang dapat diterapkan untuk mendesain layanan video yaitu mengkodekan data video untuk menurunkan jumlah frame yang dikirim. Sedangkan pada sisi penerima (*decoding*), data video direkonstruksi dengan prosedur penyisipan (*Frame Data Intermediate*). Sasarannya adalah untuk meningkatkan laju frame pada penerima. Teknik ini sering disebut FRUC Video.

Teknik FRUC ini menggunakan metode (MCI) dua arah yaitu *Forward* – MCI dan *Backward* – MCI. Metode ini selanjutnya dikenal sebagai *Unidirectional* – MCI. Pada Metode ini akan diperoleh dua frame kandidat *forward* dan *backward*, kemudian mengkompensasi *frame* tersebut menjadi *frame* final. Teknik ini akan diterapkan pada pembangkitan FI pada teknik FRUC. Teknik FRUC ini menggunakan standar teknik motion estimation (ME).

Skema ME dilakukan dengan mencari motion vector sebuah blok frame eksisting di

blok dalam frame referensi [1]. Permasalahannya adalah penggunaan satu frame referensi akan membatasi informasi yang di eksploitasi hanya pada frame tersebut. Jika pergerakan objek yang di estimasi tidak linier, maka motion vector kandidat yang dihasilkan belum mampu secara akurat mewakili blok yang di estimasi.

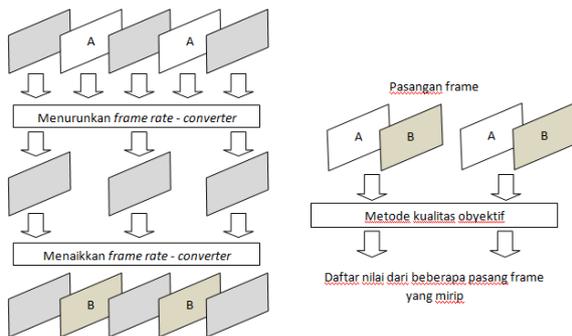
Dua langkah pencarian setengah piksel tidak banyak mempengaruhi lokasi blok kandidat setengah piksel. Itu hanya bisa menguji sekitar lokasi seleksi blok setengah piksel pada langkah pertama (pencarian akurasi integer piksel) [2]. Metode akurasi setengah piksel adalah menambahkan atau menyisipkan sebuah frame baru diantara frame referensi dan frame sekarang. Dimana penyisipannya tiap setengah piksel dari pencarian satu piksel, yang artinya setelah disisipkan piksel baru sebelumnya dengan menggunakan metode pencarian satu piksel, kemudian disisipkan lagi setengah piksel. Maka waktu yang dibutuhkan lama tetapi kualitas hasilnya akan sangat baik dibandingkan dengan metode pencarian satu piksel.

Metode Estimasi Gerak Pencarian Penuh dengan akurasi setengah piksel dilakukan dengan cara memperoleh nilai intensitas setengah piksel (*half-pel intensity*) dari nilai averaging posisi integer piksel yang berdekatan [3].

Oleh karena itu, paper ini akan menggunakan skema ME untuk FRUC dengan dua frame referensi. Skema ini dikenal dengan ME *unidirectional*. Untuk meningkatkan akurasi ME lebih lanjut, paper ini akan menerapkan metode ME dengan pencarian penuh akurasi setengah piksel.

2. KAJIAN PUSTAKA
2.1 Frame Rate Up – Conversion (FRUC) Video

Gambar 2.1 mengilustrasikan proses dari mekanisme FRUC. Pada sisi *encoder* jumlah frame yang dikirim akan diturunkan. Sedangkan pada sisi penerima (*decoding*), data video direkonstruksi dengan prosedur penyisipan FI. Sasarannya adalah untuk meningkatkan laju frame pada penerima. Untuk mengetahui kualitas hasil frame estimasi yang akan disisipkan pada *decoder* maka dapat digunakan metode pengukuran kualitas objektif dengan membandingkan frame asli dengan FI [4].



Gambar 2.1 Mekanisme FRUC [4]

Secara umum algoritma FRUC terbagi menjadi dua kategori yakni: (i) tidak mempertimbangkan informasi dari pergerakan objek yakni dengan metode *Frame Repetition* dan *Frame Averaging*. (ii) mempertimbangkan informasi dari pergerakan objek dengan teknik MCI. Teknik ini menyandarkan pada suatu proses yang disebut dengan ME.

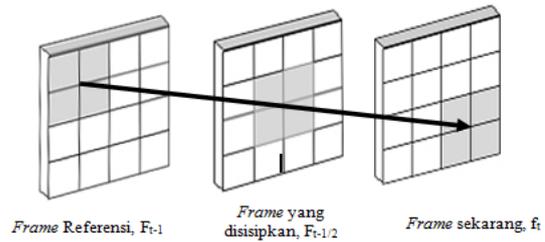
2.2 Metode Motion Compensation Interpolation (MCI)

Dalam kompensasi gerak *Frame Rate Up – Conversion* (MC – FRUC), *frame* yang hilang diinterpolasi ke informasi vektor gerak dihitung diantara *frame* referensi dan *frame* baru [5]. Teknik MCI yang digunakan pada algoritma FRUC yakni sebagai berikut:

a. Estimasi Blok Unidirectional

Pada FI, blok diisi data dari pergerakan blok yang menjadi referensi

sebanyak MV yang diperoleh [6]. Pada gambar 2.2 dibawah ini mengilustrasikan metode *Unidirectional* pada arah maju.



Gambar 2.2 Metode Unidirectional Pada Arah Maju [1]

Dari 2 buah *frame* yang telah didekodekan yakni *frame* f_{t-1} dan f_{t+1} maka untuk mencari *frame intermediate* sebagai *frame* interpolasi f_t . MV terbaik untuk FI adalah nilai rata-rata terkecil dari pengurangan sisi maju dan mundur [1].

2.3 Estimasi Gerak

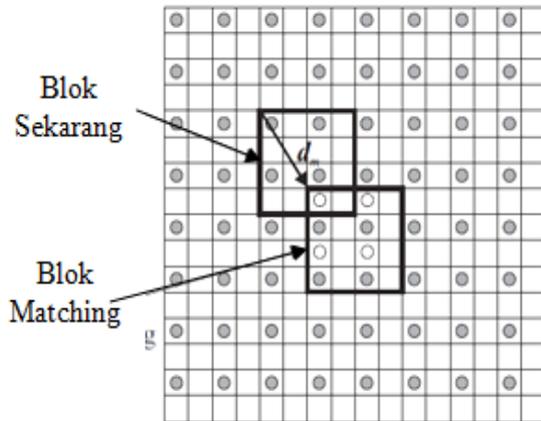
Estimasi gerakan merupakan proses untuk menentukan pergerakan obyek-obyek pada video sekuensial. Pergerakan obyek tersebut dikenal dengan istilah MV. Metode yang sangat populer dan banyak digunakan adalah pendekatan berbasis area yang dikenal dengan *Block Matching*. Mekanisme *Block Matching* dilakukan dengan skema pencarian gerak dalam ukuran window tertentu. Estimasi Gerak membuat model base *frame* baru pada data yang tersedia dalam satu atau lebih *frame* encoded sebelumnya yaitu *frame* referensi [4]. Dibawah ini akan dijelaskan tentang bagian dari estimasi gerak yaitu metode pencarian penuh dan metode pencarian setengah piksel (*half – pel accuracy*).

2.3.1 Pencarian Penuh Akurasi Satu Piksel

Metode pencarian penuh ikut dalam pencarian blok referensi terbaik terhadap semua kemungkinan lokasi blok dalam *search windows* yang ditentukan. Contoh hasilnya yaitu mencari *match* kemungkinan terbaik dan memberikan PSNR terbesar diantara semua algoritma *block matching* [7]. *Macroblock* dibandingkan dengan semua posisi *matching* yang tepat dalam *search window* dan mencari MAD, dan semua nilai MAD yang dibutuhkan dibandingkan kemudian mendapatkan posisi dimana nilai MAD minimum untuk posisi *matching* terbaik.

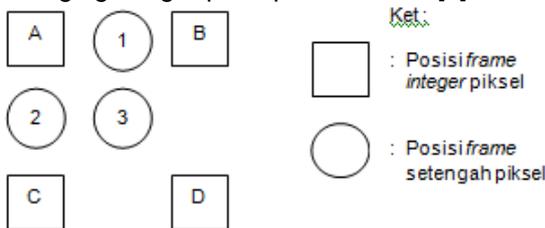
2.3.2 Metode Akurasi Setengah Piksel (ME – ½ P)

Metode ME – ½ P menyisipkan frame intermediate diantara frame referensi dan frame sekarang tiap setengah piksel. Permasalahannya dengan penggunaan *stepsize* pecahan adalah tidak adanya titik-titik sampel dalam *tracked frame* yang bersesuaian dengan titik-titik sample di *anchor frame*. Untuk merealisasikan sebuah *step size* dengan 1/d piksel, *tracked frame* diinterpolasi dengan faktor d. Sebagai contoh, untuk d = 2 [5], skema estimasi gerak ditunjukkan pada Gambar 2.3



Gambar 2.3 Pencocokan block dengan metode ME – ½ P [7]

Gambar 2.4 menggambarkan posisi berdekatan integer piksel yang diberi label dari A sampai D. Posisi setengah piksel diberi label dari a sampai c. Nilai intensitas setengah piksel a diperoleh dari posisi *averaging integer* A dan B. Untuk nilai ke b, posisi A ke C yang digunakan. Nilai ke c diperoleh dari nilai *averaging integer* pada posisi A ke D [8].



Gambar 2.4 Model Metode Akurasi Setengah Piksel [8]

Dari penjelasan diatas, pencarian penuh ME – ½ P akan memperbanyak jumlah blok kandidat sampai dua kali lipat. Metode pencarian penuh ikut dalam pencarian blok referensi terbaik terhadap semua kemungkinan lokasi blok dalam *search windows* yang ditentukan.

2.4 Metode Pengukuran Kualitas Video

Kualitas *frame* hasil estimasi di dekoder (FI) pada algoritma FRUC di penelitian ini menggunakan parameter *Peak to Peak Signal to Noise Ratio* (PSNR). Distorsi frame dihitung dari PSNR, yaitu:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \dots \dots \dots (2.1)$$

dimana, untuk format frame video adalah QCIF (176 × 144) piksel, MSE dinyatakan sebagai:

$$MSE = \frac{1}{176 \times 144} \sum_{i=0}^{175} \sum_{j=0}^{143} \{X(i, j) - \hat{X}(i, j)\}^2 \dots \dots \dots (2.2)$$

dengan $X(i, j)$ adalah koordinat piksel pada frame video asli, dan $\hat{X}(i, j)$ adalah koordinat piksel pada FI.

3. METODOLOGI PENELITIAN

3.1. Data Penelitian

Penelitian akan diperoleh dengan menggunakan kondisi-kondisi test seperti dibawah ini, yaitu :

Foreman dan Hall. Sekuen-sekuen ini menunjukkan tipe berbeda dalam konten adegan, yaitu dari pergerakan adegan tinggi (Foreman) dan pergerakan adegan rendah (Hall). Jumlah *frame* untuk setiap sekuen : Foreman = 400 *frame* dan Hall = 300 *frame*.

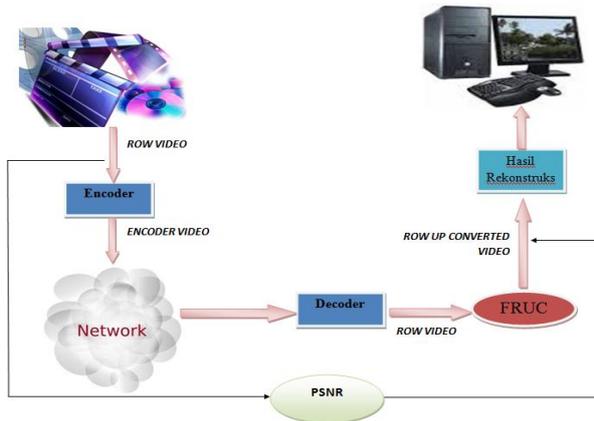


Gambar 3.1 Contoh Frame Dari Sekuen Video (a) Foreman dan (b) Hall

Format seluruh sekuen video adalah QCIF (176 x 144 piksel) dengan laju frame 30 fps. Ukuran blok: 8 x 8 piksel dengan range pencarian gerak adalah ± 50 piksel arah *vertikal* dan *horizontal*.

3.2. Gambaran Umum Penelitian

Sebuah rancangan estimasi gerak *Unidirectional* dengan menerapkan metode pencarian metode pencarian penuh akurasi estimasi gerak setengah piksel yang digunakan untuk memperoleh MV sebagai informasi untuk menyisipkan *frame intermediate* ke dalam urutan video asli di *dekoder*.



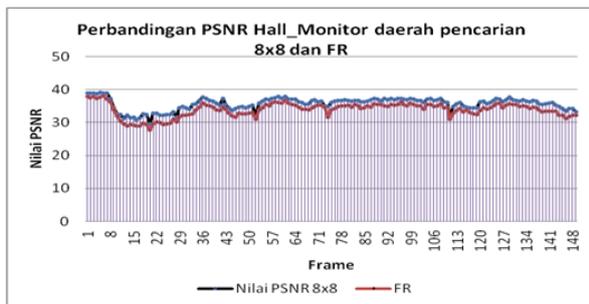
Gambar 3.2 Desain Suatu Video di *capture* Dari Sebuah Sumber Kemudian Ditampilkan Kembali Dengan Mekanisme FRUC

4. HASIL DAN PEMBAHASAN

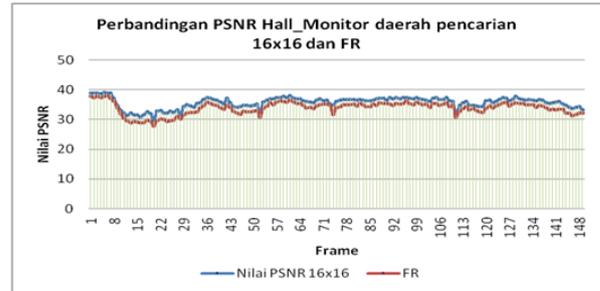
4.1 Analisis Trace PSNR

Dapat dilihat kinerja kedua algoritma FRUC dalam grafik *per-frame* nilai PSNR berikut.

Dari Gambar 4.1 dapat dilihat untuk nilai tertinggi dan terendah dari inputan video hall monitor dengan algoritma FRUC *frame repetition* menghasilkan nilai PSNR sebesar 38,06 dB dan 27,81 dB pada FI ke-11 untuk PSNR tertinggi dan ke-40 untuk PSNR terendah serta memiliki nilai PSNR rata – rata sebesar 33,84 dB. Pada algoritma ME – ½ P memiliki nilai PSNR rata – rata sebesar 35,66 dB serta nilai tertinggi dan terendah pada skenario makroblok 8 x 8 piksel dan diameter pencarian 7, yakni sebesar 39,18 dB dan 29,61 dB. Nilai tersebut berada pada FI ke-9 untuk PSNR tertinggi dan ke-39 untuk PSNR terendah. Sedangkan nilai tertinggi dan terendah pada skenario makroblok 8 x 8 piksel dan diameter pencarian 15 piksel, yakni sebesar 39,19 dB dan 29,56 dB. Nilai tersebut berada pada FI ke-9 untuk PSNR tertinggi dan ke-39 untuk PSNR terendah serta memiliki nilai PSNR rata – rata sebesar 35,66 dB.



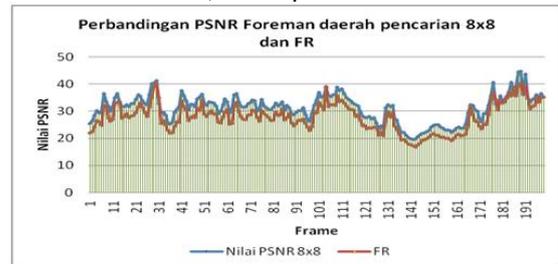
(a)



(b)

Gambar 4.1 Perbandingan nilai PSNR video Hall Monitor dengan makroblok 8 x 8 (a) diameter pencarian 7 dan (b) diameter pencarian 15

Dari Gambar 4.2 dapat dilihat untuk inputan video foreman menghasilkan nilai PSNR yang lebih variatif. Nilai PSNR tertinggi dan terendah yang dihasilkan algoritma FRUC *frame repetition* sebesar 40,42 dB dan 16,84 dB pada FI ke-377 untuk PSNR tertinggi dan ke-285 untuk PSNR terendah serta memiliki nilai PSNR rata – rata sebesar 27,86 dB. Sedangkan algoritma FRUC ME – ½ P memiliki nilai PSNR rata – rata sebesar 31,07 dB serta nilai PSNR tertinggi dan terendah sebesar 44,63 dB pada FI ke-377 dan 19,68 dB pada FI ke-283 untuk skenario dengan makroblok 8x8 dengan daerah pencarian 7 piksel dan pada algoritma FRUC dengan makroblok 8x8 dengan daerah pencarian 15 piksel memiliki nilai PSNR rata – rata sebesar 30,97 dB serta menghasilkan nilai PSNR tertinggi dan terendah sebesar 44,63 dB pada FI ke-377 dan 19,04 dB pada FI ke-283.



(a)



(b)

Gambar 4.2 Perbandingan nilai PSNR video Foreman dengan makroblok 8 x 8 (a) diameter pencarian 7 dan (b) diameter pencarian 15

Untuk inputan video foreman menghasilkan nilai PSNR yang lebih variatif. Nilai PSNR tertinggi dan terendah yang

dihasilkan algoritma FRUC *frame repetition* sebesar 40,42 dB dan 16,84 dB pada FI ke-377 untuk PSNR tertinggi dan ke-285 untuk PSNR terendah serta memiliki nilai PSNR rata – rata sebesar 27,86 dB. Sedangkan algoritma FRUC ME – ½ P memiliki nilai PSNR rata – rata sebesar 31,07 dB serta nilai PSNR tertinggi dan terendah sebesar 44,63 dB pada FI ke-377 dan 19,68 dB pada FI ke-283 untuk skenario dengan makroblok 8x8 dengan daerah pencarian 7 piksel dan pada algoritma FRUC dengan makroblok 8x8 dengan daerah pencarian 15 piksel memiliki nilai PSNR rata – rata sebesar 30,97 dB serta menghasilkan nilai PSNR tertinggi dan terendah sebesar 44,63 dB pada FI ke-377 dan 19,04 dB pada FI ke-283.

Jika dilihat perbedaan nilai PSNR tertinggi dan terendah antara algoritma pembandingan dan yang diusulkan, algoritma yang diusulkan yaitu metode ME – ½ P memiliki nilai PSNR yang tinggi dibandingkan dengan algoritma *frame repetition*. Sehingga, dari kedua data algoritma di atas dapat kita buat selisih nilai PSNR kedua algoritma tersebut dalam tabel berikut.

Tabel 4.1 Selisih nilai PSNR rata-rata, tertinggi dan terendah algoritma FRUC ME – ½ P dengan algoritma FRUC *frame repetition* Foreman dan Hall Monitor.

Sekuen	Parameter	Makroblok 8x8 dengan daerah pencarian 7 piksel	Makroblok 8x8 dengan daerah pencarian 15 piksel
Foreman	Average	1,82 dB	1,82 dB
	Max	1,12 dB	1,13 dB
	Min	1,81 dB	1,75 dB
Hall Monitor	Average	3,21 dB	3,11 dB
	Max	6,56 dB	4,21 dB
	Min	2,84 dB	2,20 dB

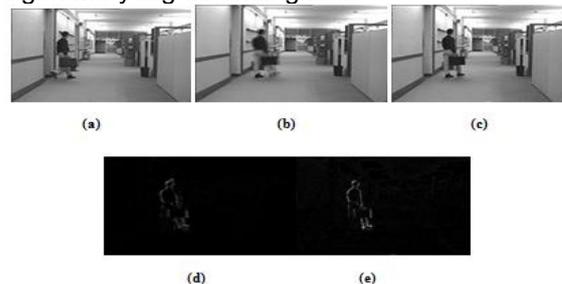
Dari Tabel 4.1 untuk kedua algoritma baik algoritma FRUC pembandingan maupun algoritma FRUC ME – ½ P dapat dilihat bahwa algoritma ME – ½ P lebih tinggi dibandingkan dengan algoritma pembandingan yaitu *frame repetition*. Hal ini dilihat dari selisih nilai pada Tabel 4.1 menunjukkan nilai PSNR rata-rata untuk metode algoritma FRUC ME – ½ P lebih tinggi 1,82 dB dari pada algoritma *Frame Repetition* (FR) pada video *hall monitor* untuk skenario makroblok 8x8 dengan daerah pencarian 7 piksel dan pada daerah pencarian 15 piksel. Sedangkan pada video *Foreman* untuk skenario makroblok 8x8 dengan daerah pencarian 7 piksel antara algoritma FRUC ME – ½ P dengan algoritma *frame repetition* memiliki selisih nilai PSNR rata – rata sebesar 3,21 dB dan pada daerah pencarian 15 piksel memiliki selisih sebesar 3,11 dB.

Jadi dapat disimpulkan bahwa algoritma FRUC yang berdasarkan pada MCI lebih baik dari algoritma FRUC yang tidak berdasarkan MCI. Walaupun kompleksitas dari teknik yang berdasarkan pada MCI ini lebih tinggi dibandingkan FR namun menghasilkan kualitas *frame* yang lebih baik ditinjau dari nilai PSNR.

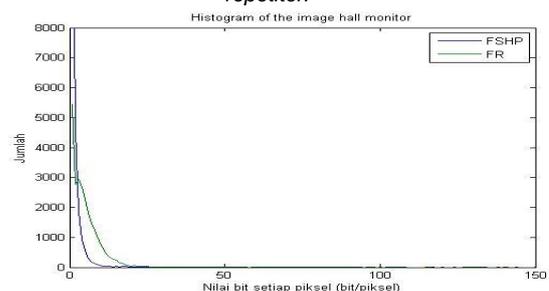
4.2 Analisis Rekonstruksi *Frame Intermediate*

4.2.1 Inputan Hall Monitor

Dari Gambar 4.3 dapat dibandingkan secara visual bahwa algoritma yang diusulkan lebih baik dalam merekonstruksi *frame intermediate*. Pergerakan objek pada video inputan ini terdapat pada saat seseorang dalam gambar tersebut bergerak keluar. Algoritma yang diusulkan mampu mendekati aslinya dalam merekonstruksi pergerakan tersebut. Dalam gambar dibawah ini akan diperlihatkan grafik perbandingan kedua algoritma yang dibandingkan.



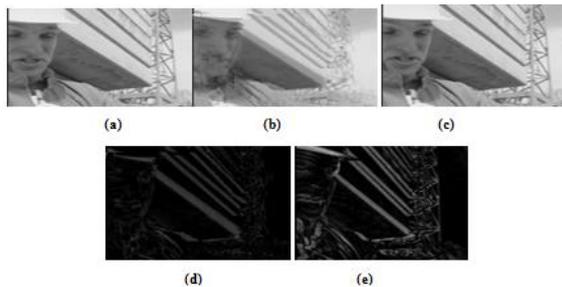
Gambar 4.3 *frame intermediate* dari inputan hall monitor dengan membandingkan (a) *frame* asli (b) *unidirectional full search half pixel accuracy* (c) *frame repetition* (d) *error prediction image* dari *unidirectional full search half pixel accuracy* (e) *error prediction image* dari *frame repetition*



Gambar 4.4 Grafik perbandingan *error prediction image* pada *frame intermediate* dari inputan hall monitor.

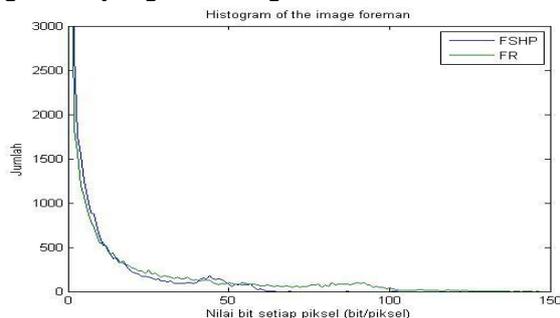
Dari Gambar 4.4 di atas dapat disimpulkan bahwa secara visualisasi algoritma FRUC yang diusulkan lebih baik daripada algoritma FRUC pembandingan untuk inputan hall monitor karena nilai bit per piksel algoritma FRUC yang diusulkan lebih banyak yang mendekati nol.

4.2.2 Inputan Foreman



Gambar 4.5 frame intermediate dari inputan foreman dengan membandingkan (a) frame asli (b) unidirectional full search half pixel accuracy (c) frame repetition (d) error prediction image dari unidirectional full search half pixel accuracy (e) error prediction image dari frame repetition

Dari Gambar 4.5, dapat dicermati bahwa *error prediction image* dari algoritma FRUC perbandingan lebih banyak menghasilkan error daripada algoritma FRUC yang diusulkan. Pada gambar di atas pergerakan seorang menunjukkan latar belakangnya. Algoritma FRUC yang diusulkan lebih baik dalam merekonstruksi gambar mendekati gambar asli daripada algoritma perbandingan. Dalam gambar dibawah ini akan diperlihatkan grafik perbandingan kedua algoritma yang dibandingkan.



Gambar 4.6 Grafik perbandingan *error prediction image* pada frame intermediate dari inputan hall monitor.

Dari Gambar 4.6 di atas, dapat disimpulkan bahwa *error image* kedua algoritma ini memiliki jumlah kesalahan yang tidak memiliki rentang yang jauh namun algoritma yang diusulkan masih lebih baik.

5. Kesimpulan

Berdasarkan pengujian dan analisis yang telah dilakukan terhadap algoritma FRUC *Unidirectional Full Search Half Pixel Motion Estimation* dan *Frame Repetition* pada pengkodean deretan *frame* video, maka dapat disimpulkan :

Bahwa dengan menggunakan algoritma FRUC metode unidirectional dua arah yaitu forward dan backward kualitas hasil algoritma FRUC estimasi gerak pencarian penuh dengan akurasi setengah piksel lebih

baik dibandingkan dengan algoritma perbandingan yaitu *frame repetition*. Itu dapat dilihat dari selisih dari kedua algoritma tersebut masing – masing 3,21 dB dan 3,11 dB pada wilayah pencarian 7 dan 15 piksel dibandingkan dengan metode *frame repetition* untuk sekuen video foreman dan hall monitor.

REFERENSI

- [1] Jostell, J. dan Isberg, A.2012.“Frame Rate Up-Conversion Of Real-Time High-Definition Remote Surveillance Video”. Master's Thesis, Universitas Teknologi Chalmers, Swedia.
- [2] Mahdavi-Nasab H. and Kasaei S. 2008. “New Half-Pixel Accuracy Motion Estimation Algorithms for Low Bitrate Video Communications”. Scientia Iranica
- [3] Widya Y. dan Oka W. 2013. “Penerapan Teknik Fast Motion Estimation Pada Algoritma Frame Rate Up – Conversion Video”. Prosiding Conference on Smart – Green Technology in Electrical and Information Systems Bali, 14 – 15 Nov.
- [4] Hegde, G, Cyril Prasanna Raj P, P.R.Vaya. 2009. “Implementation of Systolic Array Architecture for Full Search Block Matching Algorithm on FPGA”. European Journal of Scientific Research
- [5] S.-J. Kang, K.-R. Cho, dan Y. H. Kim. Nov. 2007 “Motion Compensated Frame Rate Up-Conversion Using Extended Bilateral Motion Estimation,” IEEE Trans. Consumer Electronics, vol. 53, no. 4, pp. 1759–1767.
- [6] L. Zhang, C. Wang, W. Zhang, and Y. P. Tan, “Frame Rate UpConversion With Edge-Weighted Motion Estimation and Trilateral Interpolation”. 2010. IEEE Trans. Circuits Syst. Video Technol.,vol. 20, no. 6, pp. 886-893, June.
- [7] Barjatya, A. 2004. “Block Matching Algorithms For Motion Estimation”. IEEE Transactions on Circuits and System for Video Technology _____.
- [8] Flierl, M, Bernd Girod. 2007. “Half-Pel Accurate Motion-Compensated Orthogonal Video Transforms”. IEEE DCC. Max Planck Center for Visual Computing and Communication Stanford University