

## TEKNIK ESTIMASI DAN KOMPENSASI GERAK PADA VIDEO CODING FGS (Fine Granularity Scalability)

Oleh :

**Agus Purwadi , Afrizal Yuhaneff , Lince Markis**

Staf Pengajar Jurusan Teknik Elektro Politeknik Negeri Padang

E-mail : aguspurwadi99@yahoo.com

### ABSTRACT

*Motion estimation is a process to determine the movement of an object on video sequential. The movement of objects is known as motion vector. A motion vector indicates a shift point between the current frame with the reference frame. Of motion vector is obtained, it would seem that the movement of the dots between the observed frame. In this study using the algorithm block matching SAD (Sum of Absolute Different), the search process is done per pixel. To determine the quality of the movement of objects in each frame interpolation is obtained by calculating the PSNR value. PSNR values range from 35 to 40 dB. From the research conducted using the 90 frame interpolation obtained PSNR value decreases.*

*Keywords : motion estimation compensation, FGS (Fine Granularity Scalability)*

### PENDAHULUAN

Pada pengkodean video FGS (*Fine Granularity Scalability*) yang merupakan perbaikan dari MPEG-4 menerapkan pengkodean prediksi interframe didasarkan pada teknik estimasi dan kompensasi gerak. Pada teknik ini, hanya frame-frame error prediksi yang dikodekan, yaitu perbedaan antara frame-frame asli dengan frame-frame hasil proses rekonstruksi. Maka setiap macroblock pada frame asli akan di ambil prediksinya dari frame sebelumnya yang telah direkonstruksi. Informasi gerak dari sebuah macroblok dalam frame rekonstruksi dinyatakan oleh sebuah vektor perpindahan dua dimensi atau vektor gerak. Pada pengkodean video FGS (*Fine Granularity Scalability*), proses *block matching* akan dihasilkan satu vektor gerak untuk setiap proses kompensasi gerak. Vektor-vektor gerak

yang dihasilkan oleh metode ini hanya dapat referensi pixel-pixel didalam area frame. Oleh karenanya, macroblock macroblock di tepian frame, tidak dapat diprediksi dengan baik. Optimalisasi teknik estimasi dan kompensasi gerak untuk mempertinggi kualitas pengkodean prediksi interframe dapat dilakukan dengan meningkatkan area pencarian vektor gerak untuk mengkompensasi pergerakan obyek video di tepian frame.

### METODOLOGI

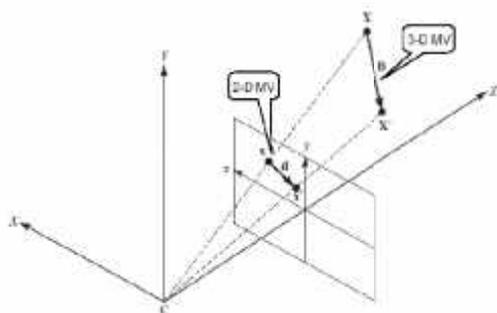
#### Estimasi Gerak dan Kompensasi

*Motion* dapat dikategorikan dalam model 2 dimensi atau model 3 dimensi. Dua dimensi *motion estimation* merupakan bagian yang penting pada sistem video processing. Dua dimensi *motion estimation* sering kali sebagai syarat langkah awal untuk struktur 3-D

dan *motion estimation*. Tiga dimensi *motion estimation* digunakan untuk menggambarkan *motion* (gerakan) pada obyek dalam ruang 3 dimensi, sebagai contoh dalam aplikasi *computer vision* untuk *tracking* obyek dalam *video coding* dan *motion compensation*.

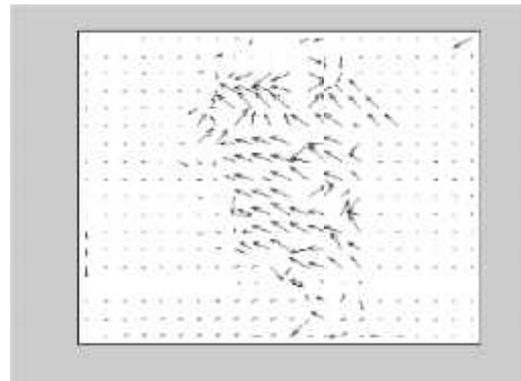
Dalam beberapa aplikasi, kita dapat mempresentasikan *motion* dengan model dua dimensi. Gambar 2. diperlihatkan proyeksi pada perpindahan obyek pada perencanaan *image* dan hubungan antara *motion* 3-D dan *motion* 2-D.

Ketika titik pada obyek berpindah dari  $X = x, y, z$  pada waktu  $t_1$  ke  $X' = (x + dx, y + dy, z + dz)$  pada waktu  $t_2 = t_1 + dt$ , ini adalah proyeksi *image* berpindah dari  $x = (x, y)$  ke  $x' = (x', y') = (x + dx, y + dy)$ . Tiga dimensi *motion vector* pada  $X$  adalah tiga dimensi pemindahan  $d(X, t_1, t_2) = (X' - X)$  dan 2-D *motion vector* pada  $x$  adalah dua dimensi perpindahan  $d(x, t_1, t_2) = (x' - x)$ .



Gambar 1. Proyeksi Perpindahan Obyek

Bidang *motion* dari  $t_1$  ke  $t_2$  adalah representasi dari  $d(x, t_1, t_2)$  pada semua posisi *image*  $x$  pada waktu  $t_1$ . Kita dapat melukis bidang *motion* dari grafik vektor, seperti diperlihatkan pada gambar 2., di mana arah dan magnitudo pada tiap-tiap tanda panah dalam arah dan magnitudo pada *motion vector* dalam *pixel* di mana tanda panah diletakkan pada arah asal.



Gambar 2. Dua Dimensi Bidang *Motion*

Bidang *motion* sangat berhubungan dengan *zoom* pada kamera dan rotasi pada kamera seperti diperlihatkan pada gambar 2.

Global *motion* dapat diestimasi menggunakan model perbedaan seperti translasi, transformasi geometri, turunan *mapping* atau proyeksi *mapping*. Satu atau lebih model yang digunakan adalah 6 parameter model turunan. Untuk dimasukkan dalam global bidang *motion*, 6 parameter turunan model *motion* yang dipergunakan dalam simulasi adalah:

$$x' = a_0 + a_1 \times x + a_2 \times y \quad \dots\dots (1)$$

$$y' = a_3 + a_4 \times x + a_5 \times y \quad \dots\dots(2)$$

di mana  $[x, y]$  adalah posisi *pixel* dalam *image* asal,  $[x', y']$  adalah posisi hubungan *pixel* dan *image* referensi, dan  $[a_0, a_1, a_2, a_3, a_4, a_5]$  adalah global parameter *motion*

**PERANCANGAN SISTEM**

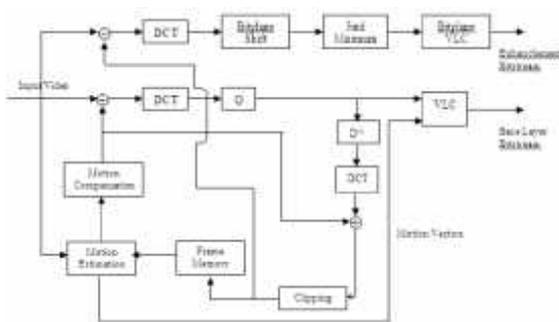
**FGS (*Fine Granularity Scalability*)**

FGS (*Fine Granularity Scalability*) merupakan perbaikan (amandemen) pada streaming video MPEG-4, adalah untuk membuat respon pada pertumbuhan yang diinginkan pada standart video coding dalam kanal atau juga dalam internet, dengan range yang lebar pada bit rate dan range yang lebar pada variasi bit ratenya. Dengan base-

layer coding yang merupakan bentuk amandemennya, yaitu merupakan bentuk yang sederhana untuk mendapatkan ketelitian dalam tingkat ketelitian coding yang tinggi dengan tingkat kompleksitas yang rendah.

Secara mendasar perencanaan dari kompresi video dapat diklasifikasikan ke dalam dua pendekatan yaitu: *scalable* dan *non-scalable* video coding [8]. Secara sederhana dapat diperlihatkan untuk enkoder dan dekoder dalam intra mode (unit video coding yang dipresentasikan tanpa referensi kode data sebelumnya) dan hanya digunakan DCT (discrete cosine transform).

Pada enkoder video non-scalable dibangkitkan satu kompresi urutan bit. Sebaliknya pada enkoder video scalable kompresi baris video di bagi dalam berbagai sub stream, kompresi sub stream yang pertama dinamakan *base substream*, yang mana dapat dikodekan berdiri sendiri dan memberikan kualitas visual yang kasar. Sedang kompresi substream kedua adalah *enhancement substream*, yang mana dapat dikodekan bersama-sama dengan base substream dan akan mendapatkan kualitas visual yang lebih bagus. Kualitas skalabilitas, ukuran image atau frame rate dinamakan SNR (Signal to noise ratio) untuk spatial atau temporal.



Gambar 3. Struktur Enkoder video coding FGS

### Algoritma Estimasi dan Kompensasi Gerak

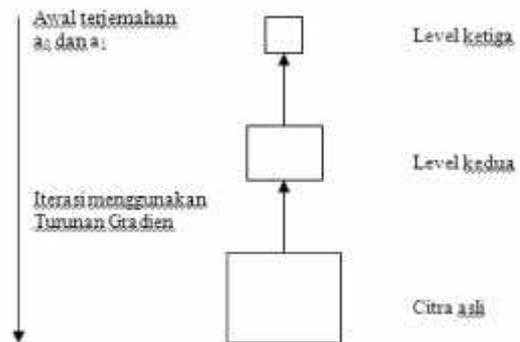
Estimasi gerakan merupakan bagian yang sangat penting dalam sistem video processing. Di dalam penelitian ini akan mengevaluasi pendekatan estimasi gerakan didasarkan pada posisi gradien dan implementasi hirarkis.

Dalam algoritma ini, kita mengusulkan suatu metoda yang efisien untuk estimasi gerakan dari urutan citra. Metoda yang umum adalah dapat mengakomodasi berbagai model gerakan. Teknik estimasi gerakan ini dirancang untuk memperkecil *sum of squared differences* (SSD) antara *current frame* dan kompensasi gerakan frame sebelumnya.

$$SSD = \sum_{x,y \in R} (I(x,y) - I'(x',y'))^2 \dots\dots(3)$$

$$SAD = \sum_{x,y \in R} (I(x,y) - I'(x',y')) \dots\dots(4)$$

Algoritma ini terdiri dari tiga langkah, pertama sebuah tiga level *low pass* piramid citra dibangun untuk meningkatkan pemusatan dan mengurangi komputasional yang komplek. *Low-Pass filters* dan *down sampling* digunakan untuk membangun ini piramida. Teknik ini dinamakan *Three-level Hierarchical Implementation*.



Gambar 4. Hirarki Estimasi

Kemudian suatu translasi awal diperkirakan dengan full-pixel ketepatan ada di puncak piramida. Inisial estimasi ini diperoleh oleh suatu teknik *matching*, yang mana memperkecil SAD dengan menggunakan modifikasi suatu pencarian n-step. Di dalam tiga *stage*, suatu posisi gradien dieksekusi pada masing-masing tingkatan piramida yang di mulai dari inisial translasi pada level kasar. Gerakan parameter dihitung dengan penyederhanakan SSD itu. Karena tergantung pada SSD pada parameter gerakan adalah *non linear*, kita menggunakan prosedur iterasi sebagai berikut :

$$a^{(t+1)} = a^{(t)} + H^{(-1)}b \quad \dots\dots\dots (5)$$

Di mana:

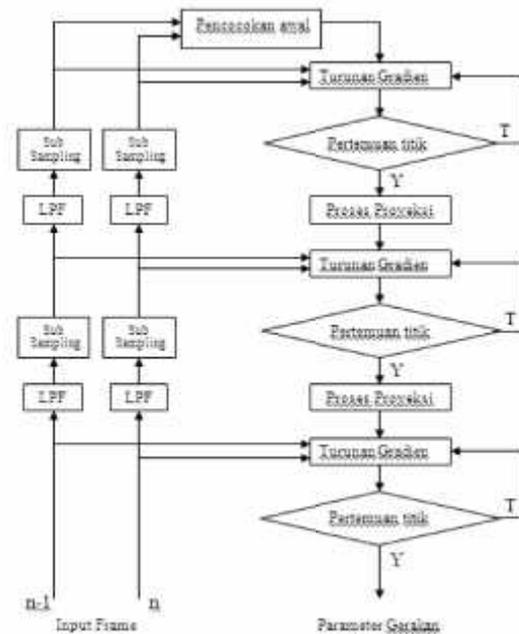
- a adalah parameter gerakan pada saat t dan t+1.
- H adalah matrik yang sama n x n untuk setengah waktu pertama.
- b adalah n-elemen vector yang sama pada minus setengah waktu pertama gradien pada SSD.
- n adalah jumlah parameter pada model.

Parameter  $a_0$  dan  $a_1$  adalah inialisasi pertama dengan translasi estimasi vector dalam inisial tingkatan *matching*. Posisi gradien di mulai dari level atas pada pyramid, kemudian mengikuti sub urutan level dalam pendekatan dari atas ke bawah. Pada tiap-tiap level, posisi gradien mengikuti langkah-langkah:

1. Perhitungan matrik H dan vector b.
2. Sistem didefinisikan menghitung  $H^{(-1)}b$ .
3. Menghasilkan perubahan pada variable a adalah dengan menambahkan ke dalam parameter :

$$a^{(t+1)} = a^{(t)} + \delta a \quad \dots\dots\dots (6)$$

di mana  $\delta a = H^{(-1)} b$



Gambar 4. Blok Diagram Tiga Level Hirarki Implementasi Estimasi Kompensasi Gerak

**HASIL DAN PEMBAHASAN**

Hasil simulasi estimasi dan kompensasi gerak pada video coding dengan menggunakan metode *Gradient Descent* (penurunan gradien), dapat diperlihatkan pada gambar di bawah untuk sequence Foreman, maka didapatkan rekonstruksi target frame, anchor frame dan motion vector serta prediksi frame.



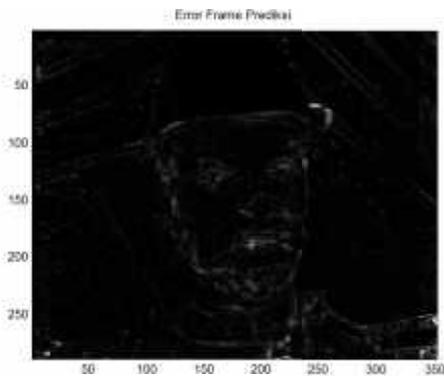
Gambar 5. Rekonstruksi Frame Target



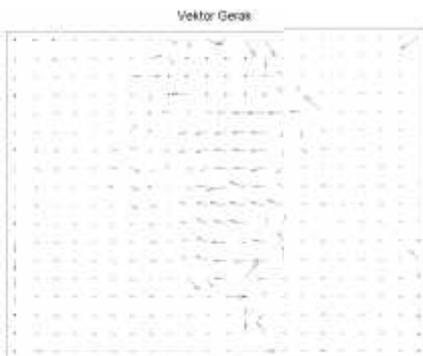
Gambar 6. Rekonstruksi Frame Anchor



Gambar 7. Rekonstruksi Frame Prediksi



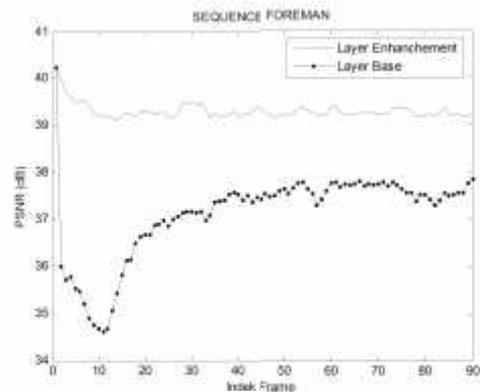
Gambar 8. Rekonstruksi frame Error



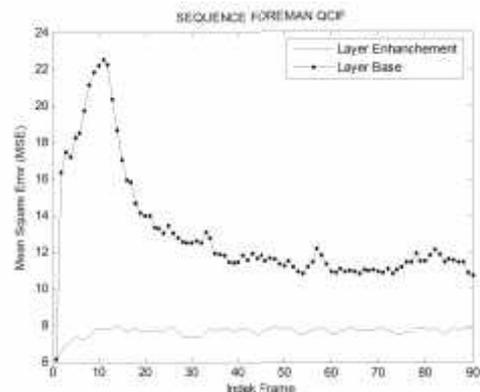
Gambar 9. Arah Vektor Gerak

Pada gambar diatas dapat dilihat banyaknya arah motion vektor untuk *sequence* “Foreman” akan mempengaruhi nilai PSNR, dari hasil simulasi didapat nilai PSNR estimasi dan kompensasi gerak pada frame 66 sampai 69 *sequence* “Foreman” mempunyai nilai PSNR pada frame prediksi sebesar 29,88 dB.

Dari simulasi *sequence* video yang dipergunakan hanya diambil 90 frame dari urutan pertama yang dikodekan dengan frame rate 30 fps untuk menunjukkan analisa unjuk kerja. Unjuk kerja dari pada video coding dapat di lihat pada gambar 10 dan 11 di bawah.



Gambar 10. Grafik PSNR Sequence Foreman



Gambar 11. Grafik MSE Sequence Foreman

### KESIMPULAN

Dari hasil simulasi pada *sequence* “Foreman” didapatkan bentuk grafik dengan nilai PSNR yang cenderung turun antara 39 dB - 40 dB

pada layer *Base* dan 35 dB – 40 dB pada layer *Enhancement*, hal ini dikarenakan terlalu banyak motion vektor terutama pada urutan frame paling akhir pada motion estimation sehingga terlalu banyak pixel dalam urutan frame yang dikodekan hal ini akan mempengaruhi penurunan nilai PSNR, demikian juga sebaliknya pada MSE terlalu banyak urutan bit dalam pixel yang dikodekan akan mempengaruhi penurunan nilai MSE.

### SARAN

Dari simulasi dan analisa hasil pengukuran, dapat disarankan sebagai berikut : proses optimasi pada estimasi dan kompensasi gerak pada pengkodean video dapat dikembangkan pula pada berbagai macam teknik pengkodean, misalnya pada PFGS (Progressive Fine Granularity Scalability) yang mempunyai bentuk skalabilitas yang lebih baik dari FGS yang tentunya mempunyai banyak variasi unjuk kerja pengkodean dan juga pada teknik-teknik pengkodean baik video maupun citra.

### DAFTAR PUSTAKA

- D. Wu, Y. T. Hou, and Y.-Q. Zhang, "Transporting real-time video over the Internet: challenges and approaches," Proc. IEEE, vol. 88, pp. 1855-1877, Dec. 2000.
- W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," IEEE Trans. Circuits Syst. Video Technol., vol. 11, pp. 301-317, March 2001.
- Bern Girod, Niko Farber, "Feedback-Based Error Control for Mobile Video Transmission", IEEE Trans, Vol. 87, No. 10, October 1999.
- Fan Zhai, Randall Berry, Thrasyvoulos N, Pappas, Angelos K. Katsaggelos, "A Rate-Distortion Optimized Error Control Scheme For Scalable Video Streaming Over The Internet", ICME 2003.
- Arnold M. Michelson, Allen H. Levesque, "Error-Control Techniques For Digital Communication", John Wiley & Sons, Inc., United State of America, 1985.
- Mohammed Ghanbari, "Video Coding an Introduction to Standart Codecs", The Institut of Electrical Engineers, United Kingdom, 1999.
- John B. Anderson, Seshadr Mohan, "Source And Channel Coding", Kluwer Academic Publishers, United State of America, 1991.
- Dapeng Wu, Yiwei Thomas Hou, Wenwu Zhu, Ya-Qin Zhang, Jon M. Peha, "Streaming Video Over the Internet: Approaches and Directions", IEEE Trans. On Circuit and System for Video Technology, vol.11, No.3, March 2001.
- Marta Mrak, Nikola Sprljan, Sonja Grgic, "Video Coding Schemes for Transporting Video Over The Internet", University of Zagreb, Faculty of Electrical Engineering and Computing Unska 3 / XII, HR-10000 Zagreb, Croatia.
- Gary J. Sullivan, Thomas Wiegand, "Rate-Distortion Optimization for Video Compression", IEEE Signal Processing Magazine, November 1998.
- Guy Cote, Shahram Shirani, Faouzi Kossentini, "Optimal Mode Selection and Synchronization for Robust Video Communication Over Error-Prone Networks", IEEE Journal on Selected Areas in Communications, Vol.18, No.6, June 2000.
- Guido M. Schuster, Aggelos K. Katsaggelos, "Rate-Distortion Based Video Compression Optimal Video Frame Compression and Object Boundary Encoding", Kluwer Academic Publisher, Netherlands, 1997.