

IKHTISAR KOMPRESI CITRA (IMAGE COMPRESSION OVERVIEW)

Erik Iman Heri Ujianto¹, Sri Hartati²

¹Program Studi Teknik Informatika, Fakultas Sains & Teknologi, Universitas Teknologi Yogyakarta, Email: eih.ujianto@gmail.com

²Program Pascasarjana Ilmu Komputer, Fakultas MIPA, Universitas Gadjah Mada, Email: shartati@ugm.ac.id

INTISARI

Metodologi pengembangan perangkat lunak saat ini menghasilkan data yang besar yang membutuhkan ruang penyimpanan yang lebih besar. Semakin besar ukuran data, semakin lama waktu yang diperlukan untuk menyimpan dan transmisi. Untuk mengatasi hal ini diantaranya dilakukan dengan mekanisme kompresi, tujuannya untuk memperkecil *space* dan mempersingkat waktu simpan dan transmisi. Kompresi citra dapat diklasifikasikan menjadi dua teknik, yaitu *lossless image compression* dan *lossy image compression*.

Kata kunci: citra, kompresi, *lossy*, *lossless*

1. PENDAHULUAN

Perkembangan *software* saat ini menghasilkan ukuran data yang semakin besar, sehingga memerlukan ruang penyimpanan yang semakin besar dan memerlukan waktu yang lebih banyak pada saat penyimpanan dan proses transmisinya. Hal ini dikarenakan data yang disimpan atau yang ditransmisikan tidak hanya berupa teks tetapi juga data multimedia yang memerlukan ruang yang lebih besar. Semakin besar ukuran data, semakin banyak ruang dan waktu yang diperlukan, untuk mengatasi permasalahan tersebut diperlukan suatu mekanisme yang disebut kompresi data (*data compression*). Metode kompresi data terbagi menjadi dua kelompok besar, yaitu *dictionary based* dan *statistical based*. *Dictionary based* bekerja dengan mengganti kelompok simbol dalam data menjadi kode-kode dengan panjang tertentu, dengan asumsi kode-kode tersebut secara umum lebih pendek dari kelompok simbol yang digantikan. Sedangkan *statistical based* melakukan kompresi dengan pendekatan lain, yaitu simbol-simbol yang ada di-*encode* satu per satu dan panjang dari kode output akan bervariasi tergantung dari probabilitas / frekuensi pemunculan simbol. Kompresi citra merupakan aplikasi dari teknik kompresi data untuk citra digital dengan tujuan untuk mengurangi redundansi yang ada di dalam citra sehingga lebih efisien dalam proses penyimpanan maupun proses transmisi.

2. SKEMA KOMPRESI CITRA

Ada sifat-sifat penting pada skema kompresi citra (Ujianto & Hartati, 2009), seperti: a). **Skalabilitas** (*scalability*); pada umumnya mengacu pada penurunan kualitas dicapai dengan manipulasi **bitstream** atau file (tanpa dekompresi / *decompression* dan re-kompresi / *re-compression*). Skalabilitas sangat berguna untuk melihat pratinjau citra (*previewing images*) saat men-downloadnya (misalnya dalam sebuah web browser) atau untuk menyediakan kualitas variabel akses (*variable quality access*) misalnya ke database. Jenis-jenis skalabilitas, antara lain : *Quality progressive*, yaitu citra dikompresi secara perlahan dengan penurunan kualitasnya. *Resolution progressive*, yaitu citra dikompresi dengan meng-encode resolusi citra yang lebih rendah lebih dulu, kemudian ke resolusi yang lebih tinggi. *Component progressive*, yaitu citra dikompresi sesuai komponennya, mulanya meng-encode komponen *gray*, kemudian komponen warnanya. b). **Region of interest**

coding; bagian-bagian tertentu dari citra yang disandikan (*encoded*) dengan kualitas yang lebih tinggi daripada yang lain. c).**Meta information**; data terkompresi dapat berisi informasi mengenai citra yang dapat digunakan untuk mengkategorikan (*categorize*), cari (*search*) atau telusuri citra (*browse image*). d).**Processing power**; algoritma kompresi yang berbeda membutuhkan jumlah daya proses (*processing power*) untuk mengkodekan (*encode*) dan *decode*. Beberapa algoritma kompresi tinggi memerlukan kekuatan pemrosesan tinggi.

3. PARAMETER KOMPRESI

Pengukuran Kualitas Kompresi

Kualitas citra terbaik pada suatu *bit-rate* atau tingkat kompresi (*compression rate*) adalah tujuan utama kompresi citra (Ujiyanto & Hartati, 2009). Untuk melihat kualitas hasil suatu metode kompresi dapat diukur dengan menggunakan *Mean Square Error* (MSE) maupun dengan *Peak Signal-to-Noise Ratio* (PSNR). Keduanya merupakan standar pengukuran *error* dalam kompresi citra. MSE adalah sigma dari jumlah *error* antara citra hasil kompresi dan citra asli, sedangkan PSNR digunakan untuk menghitung *peak error*. Pada pengukuran menggunakan MSE, nilai yang rendah adalah lebih baik, sebaliknya pada pengukuran menggunakan PSNR, nilai yang tinggi adalah lebih baik.

Rumus MSE =

$$\frac{1}{MN} \sum_{y=1}^M \sum_{x=1}^N [I(x,y) - I'(x,y)]^2 \quad (1)$$

dimana : $I'(x,y)$ merupakan nilai pixel citra asli
M, N merupakan dimensi citra

Rumus PSNR =

$$20 * \log_{10} (255 / \text{sqrt}(\text{MSE})) \quad (2)$$

Pengukuran Rasio Kompresi

Rasio kompresi data, atau disebut sebagai kekuatan kompresi adalah istilah yang digunakan untuk menghitung pengurangan ukuran representasi data yang dihasilkan oleh algoritma kompresi data.

$$\text{Compression Ratio} = \frac{\text{Compressed Size}}{\text{Uncompressed Size}}$$

Sebagai contoh misalnya pemampatan file 10MB menjadi 2MB, berarti memiliki rasio kompresi : $2/10=0.2$, atau bisa dibaca sebagai 1:5 (baca : satu banding 5), atau 1/5.

Space Saving

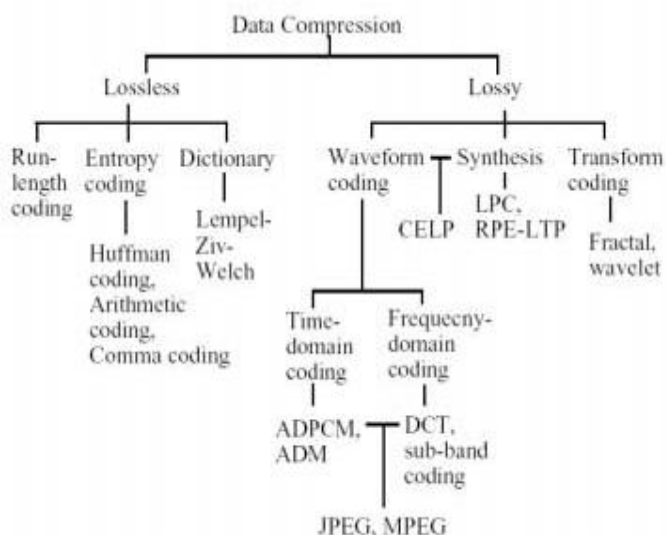
Penghematan ruang (*space saving*) didefinisikan sebagai pengurangan dalam ukuran relatif terhadap ukuran *uncompressed* :

$$\text{Space Savings} = 1 - \frac{\text{Compressed Size}}{\text{Uncompressed Size}}$$

Sehingga, pemampatan file 10MB ke 2MB akan menghasilkan penghematan ruang sebesar $1 - 2/10 = 0.8$, atau dinotasikan dalam prosentase = 80%.

4. KOMPRESI : LOSSLESS .VS. LOSSY

Kompresi citra *lossy* merupakan teknik pemampatan pada citra dengan menghilangkan beberapa informasi dari citra asli sehingga ukuran file citra menjadi relatif lebih kecil. Sedangkan kompresi citra *lossless* merupakan teknik pemampatan pada citra dimana tidak ada informasi di dalam citra yang dihilangkan. Metode kompresi yang termasuk kategori *lossless*, antara lain : Run-Length Coding, Entropy Coding (Huffman coding, Arithmetic coding, Comma coding), dan Adaptive Dictionary Based : Lempel-Ziv-Welch (LZW).



Gambar 1. Metode Kompresi Data

Kompresi Citra Lossless

LZ di LZ1 (LZ77), LZ2 (LZ78), LZS, LZW, dan DCLZ, diciptakan oleh dua ilmuwan Israel: Abraham Lempel dan Jacob Ziv. Mereka menerbitkan paper pada tahun 1977 dan 1978, menggambarkan dua algoritma LZ1 kompresi yang serupa, yang menjadi bentuk dasar GZIP, PKZip, WinZip, ALDC, LZS dan PNG (Summers, 2006). PNG (*Portable Network Graphics*) adalah format file standar untuk gambar direkomendasikan oleh World Wide Web Consortium (Savakis, Roelofs, 1999).

LZW diperkenalkan pada tahun 1984 oleh Terry Welch, yang digunakan untuk meningkatkan LZ2 dan digunakan dalam file TIFF (LZW). Pada awal 1990-an, hardware DCLZ pertama kali diimplementasikan dari algoritma kompresi LZ menggunakan Content Addressable Memory (CAM) DCLZ (LZ2 based), dan hardware ini diimplementasikan oleh Hewlett Packard menggunakan linked list 4k Dictionary (Summers, 2006).

Algoritma modern dan standar untuk kompresi lossless gambar warna kontinu menawarkan hampir *lossless mode* operasi, selain dari modus ketat lossless. Kedua algoritma yang sangat penting untuk dipertimbangkan adalah LOCO-I (*low-complexity context-based lossless image compression*) standar sebagai JPEG-LS (ISO / IEC IS 14495) dan CALIC (*context-based adaptive lossless/nearly lossless image coding*) (Domański & Rakowski, 2001). LOCO-I (*LOW COMPLEXITY LOSSLESS COMPRESSION FOR IMAGES*) memiliki rasio kompresi yang sama atau lebih baik, tapi kurang kompleks untuk aritmatika coding (Weinberger, Seroussi, & Sapiro, 2000). CALIC memiliki kemampuan sebagai prediktor non-linear untuk beradaptasi dengan beberapa sumber statistik (Wu & Memon, 1997). Pendekatan lain dalam kompresi citra lossless adalah dengan menggabungkan metode batas deskripsi dengan Run-Length dan teknik Aritmatika-coding (Grailu, & Lotfizad,

2009). Metode ini lebih baik dalam rasio kompresi dengan kisaran 1,3-2,4 kali dalam kasus lossless pada 300 dpi. Masalah selama implementasi Arithmetic-coding (Santoso, 2001) adalah keterbatasan prosesor *floating point* pada *encoder* dan *decoder*. Solusi yang diusulkan untuk dibahas seperti modifikasi Arithmetic-coding dengan menggunakan bilangan bulat (integer). Secara ringkas, algoritma ini bagus untuk kompresi data untuk alasan-alasan berikut: 1). Coding bit pada Aritmatika-coding lebih sedikit dari Huffman-coding. Modifikasi menggunakan integer juga mampu menghilangkan keterbatasan peralatan *encoder* dan *decoder* saat melakukan pengolahan *floating point* yang panjang. 2). Bit coding adalah lebih sedikit dan dapat diimplementasikan.

Kompresi Citra Lossy

Teknik kompresi data secara khusus tergantung pada jenis data yang harus dikompresi dan unjuk kerja yang diinginkan. Ukuran kinerjanya adalah waktu dan kompleksitas ruang, fidelity, serta rasio kompresi. Jenis datanya dapat berupa teks, citra, suara, video. (Carpentieri, 2010). Kompresi citra *lossy* merupakan teknik pemampatan pada citra dengan menghilangkan beberapa informasi dari citra asli sehingga ukuran file citra menjadi relatif lebih kecil. Metode kompresi yang termasuk kategori *lossy*, antara lain : Waveform coding (Time domain coding, seperti ADPCM, ADM); (Frequency domain coding, seperti DCT, sub-band coding), Synthesis (LPC, RPE-LTP), Transform coding (Fractal, Wavelet).

5. KESIMPULAN

Kompresi citra merupakan bidang kajian yang masih luas untuk diteliti. Paper ini adalah survei tentang metode kompresi citra, penelitian selanjutnya adalah mengkaji lebih dalam salah satu teknik kompresi dengan mengimplementasikannya dalam bentuk aplikasi perangkat lunak sebagai bagian dari *roadmap* penelitian.

DAFTAR PUSTAKA

- Carpentieri, B., (2010). "Interactive Compression of Books", *WSEAS Transactions on Computers*, Issue 3, Volume 9, ISSN: 1109-2750, P. 278-287.
- Domański, M., and K. Rakowski, (2001). "Near-Lossless Color Image Compression for Multimedia Applications", *Digital Signal Processing for Multimedia Communications and Services*, EURASIP ECMCS, 2001, Budapest, September 11-13 2001, pp. 181-184.
- Grailu, Hadi., and Mojtaba Lotfizad. (2009), "An improved pattern matching technique for lossy/lossless compression of binary printed Farsi and Arabic textual images" *International Journal of Intelligent Computing and Cybernetics* Vol. 2 No. 1, 2009 pp. 120-147.
- Roelofs, G., (1999). *PNG: The Definitive Guide*, Publisher O'Reilly Associates.
- Santoso, P.,(2001). *Jurnal Teknik Elektro* Vol. 1, No. 1, p.: 14-18. <http://puslit.petra.ac.id/journals/electrical>.
- Savakis, Andreas E., "Evaluation of lossless Compression Methods for Grya Scale document Images", *Department of Computer Engineering Rochester Institute of Technology Rochester*, New York.
- Summers, Tony., (2006). *Lossless Data Compression in Storage Network*, Storage Networking Industry Association (SNIA), Comtech AHA.
- Ujianto, E.I.H., and Sri Hartati., (2009). "Lossless Image Compression Overview". *Sesindo ITS*. Surabaya, Indonesia.
- Weinberger, M.J., Seroussi, G., and Sapiro, G. (2000). "The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS", *Image Processing IEEE Transactions* on Volume 9, Issue 8, Aug 2000 Page(s):1309 – 1324.
- Wu, X., and N. D. Memon, (1997). "Context-based adaptive lossless image coding," *IEEE Trans. Commun.*, vol. 45, pp.437-444.