

Kompresi Gambar Digital Bitmap dengan Kode Huffman

Kanisius Kenneth Halim – 13515008

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia

13515008@std.stei.itb.ac.id

Abstract—Bitmap atau raster adalah suatu gambar digital yang disimpan dalam bentuk matriks yang berisikan titik, dimana titik tersebut tersebut diisi dengan berbagai warna. Sehingga dengan demikian titik-titik tersebut membentuk suatu gambar yang berbentuk persegi atau persegi panjang. BMP merupakan suatu format file yang biasa digunakan untuk menyimpan sebuah gambar digital yang umum digunakan dan didukung oleh banyak sistem operasi, BMP disimpan tanpa kompresi sehingga membutuhkan cukup banyak ruang dalam memori ketika disimpan.

Keywords— bitmap, digital, gambar, huffman, kompresi, pixel .

I. PENDAHULUAN

Pada zaman perkembangan teknologi ini, hampir seluruh aspek kehidupan manusia telah disokong oleh teknologi, salah satunya adalah keberadaan sebuah gambar digital. Gambar digital adalah sebuah data yang bisa didapatkan dengan cara mengambil gambar dari sebuah kamera digital, menggunakan pemindai (*scanner*) yang dapat mengubah data analog menjadi sebuah data digital berupa gambar, ataupun dibuat dengan program di komputer. Keberadaan gambar digital telah sangat membantu manusia untuk menyebarkan informasi dengan lebih baik dan menarik, serta memungkinkan seseorang mengabadikan sebuah momen dan menyimpannya dalam bentuk digital. Sebuah gambar digital bitmap disimpan di memori dalam bentuk *array* 2 dimensi, sehingga ruang memori yang dibutuhkan untuk menyimpan cukup besar. Hal ini menjadi sebuah kendala dikarenakan semakin banyaknya gambar digital yang digunakan maka semakin banyak ruang memori yang dibutuhkan, sehingga banyak pengguna *gadget* mengalami memori penuh, meskipun pada era digital ini sudah banyak sistem penyimpanan berbasis *cloud*, tetapi pengaksesan data yang besar dari internet membutuhkan biaya yang tinggi sehingga banyak masyarakat yang masih enggan untuk menggunakannya dan masih menyimpan data dalam *local memory*.

Dari kendala diatas, jika dikaji dari penggunaan sebuah gambar digital yang berukuran cukup besar dapat mengurangi efisiensi baik secara waktu jika ditransaksikan

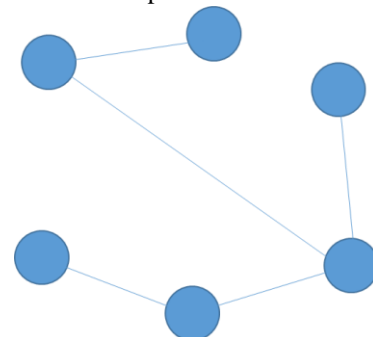
melalui jaringan internet dan secara ruang jika disimpan dalam memori komputer. Karena hal ini maka metode untuk mengurangi kebutuhan ruang gambar digital dibutuhkan.

Sudah banyak metode yang digunakan untuk mengkompresi sebuah gambar digital menjadi lebih tidak boros tempat, salah satunya JPEG, yang merupakan sebuah format gambar digital yang lebih irit tempat tanpa banyak mengorbankan kualitas suatu gambar. Metode kompresi yang digunakan pada gambar bitmap pada umumnya didasarkan dari kode Huffman yang sangat baik dalam mengkompresi sebuah data yang banyak redundansi atau pengulangan data yang sama dan juga keunggulan lainnya dari kompresi kode Huffman yaitu *lossless* yang berarti tidak ada data yang hilang setelah proses kompresi-dekompresi.

II. DASAR TEORI

A. Pohon

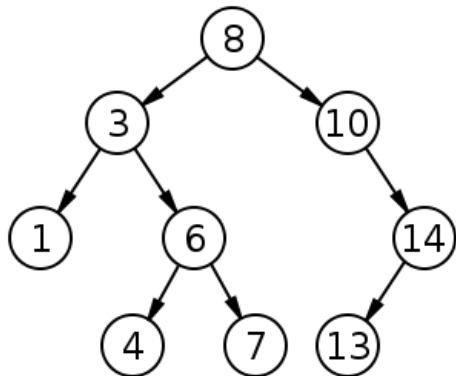
Pohon adalah sebuah struktur data yang pada dasarnya merupakan graf yang tidak berarah dan tidak terdapat sirkuit di dalamnya. Pohon yang memiliki sebuah simpul yang diperlakukan sebagai akar, dan sisi-sisinya diberi arah sehingga terjadi sebuah hubungan yaitu *parent-child* disebut pohon berakar.



Gambar 1. Pohon merentang

Pohon berakar yang pada setiap simpulnya hanya boleh memiliki 2 anak (*2-ary*) disebut sebagai pohon biner. Pohon biner merupakan pohon yang aplikasinya sangat banyak dalam bidang ilmu komputer dan informatika. Salah satu aplikasi dari pohon biner adalah

pohon pencarian biner atau *binary search tree* yang merupakan sebuah struktur data yang memperbolehkan penggunaannya melakukan sebuah pencarian biner yang jika ditinjau dari kompleksitas waktu $O(\log n)$ yang tergolong sangat cepat dibanding dengan algoritma pencarian sekuensial yang membutuhkan waktu $O(n)$.



Gambar 2. Pohon pencarian biner
source: <http://programminggeeks.com/wp-content/uploads/2014/01/nodes-in-binary-search-tree.png>

B. Kode Huffman

Kode Huffman adalah suatu metode pemampatan/kompresi data yang bersifat *lossless* sehingga tidak ada data yang hilang dalam proses mengkompresi data. Kode Huffman merupakan metode yang sangat baik untuk mengkompresi data yang terdapat banyak redundansi atau banyak pengulangan data yang sama. Dengan melakukan kode huffman maka dihasilkan sebuah data yang cenderung memiliki ukuran lebih kecil, dan juga sebuah pohon Huffman.

Berikut merupakan cara untuk membuat kode Huffman dari plain text.

1. Hitung frekuensi kemunculan setiap simbol yang ada dalam teks
2. Anggap terdapat sebuah list atau tabel yang menyatakan simbol dan peluang kemunculan (frekuensi kemunculan dibagi dengan frekuensi total) simbol tersebut .
3. Pilih dua buah simbol dengan peluang terkecil secara bebas lalu hapus kedua simbol tersebut dari list.
4. Gabung kedua simbol tersebut:
 - Jika kedua simbol yang dipilih masih merupakan sebuah karakter, maka gabungkan kedua karakter itu menjadi sebuah simbol baru, contoh terdapat sebuah karakter A dengan peluang $1/7$ dan karakter C dengan peluang $1/7$, maka akan digabungkan menjadi AC dengan peluang $2/7$ dan bentuk sebuah pohon dengan akar AC dan memiliki daun A atau C.
 - Jika simbol-simbol yang dipilih telah membentuk sebuah pohon, maka gabungkanlah kedua pohon tersebut sehingga simbol yang baru yang merupakan gabungan kedua simbol tersebut, misal terdapat simbol AC dengan

peluang $2/8$ dan BD dengan peluang $2/8$, lalu gabungkanlah simbol-simbol tersebut menjadi ABCD dengan peluang $4/8$ yang memiliki subpohon kiri adalah pohon yang dibentuk oleh AC dan subpohon kanan adalah pohon yang dibentuk oleh BD.

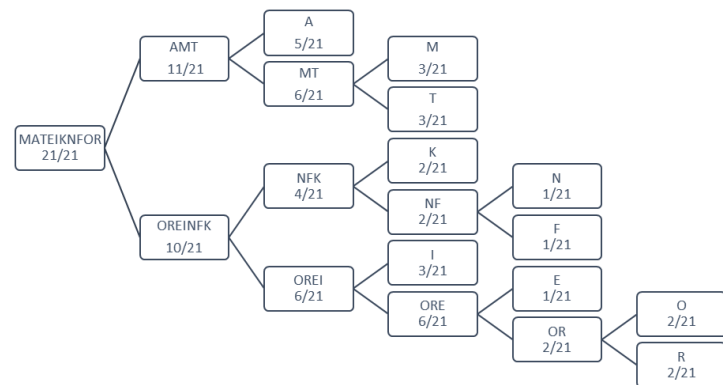
5. Jika seluruh simbol telah terdapat dalam pohon Huffman yang terbentuk maka, lanjut ke langkah 6, jika tidak tambahkan simbol yang baru dibentuk ini ke dalam list tadi, lalu ulangi dari langkah 3.
6. Beri label pada setiap pohon yang dibentuk dengan 0 atau 1.

Contoh kompresi dengan menggunakan kode Huffman, misal terdapat sebuah teks berisikan “MATEMATIKA INFORMATIKA”, spasi diabaikan, maka akan terbentuk sebuah tabel peluang kemunculan seperti ini:

Simbol	Frekuensi
M	3
A	5
T	3
E	1
I	3
K	2
N	1
F	1
O	1
R	1

Tabel 1. Tabel frekuensi kemunculan huruf

Dari tabel diatas dapat dibentuk sebuah pohon huffman seperti ini



Gambar 3. Pohon Huffman yang terbentuk

Pohon huffman diatas dibuat dengan asumsi sisi kanan diberi label 1 dan anak kiri diberi label 0, sehingga dengan demikian, kita dapat menentukan kode huffman dari masing masing simbol.

Simbol	Kode Huffman
M	101
A	11
T	100
E	0011
I	001
K	011
N	0101
F	0100
O	00001
R	00000

Tabel 2. Kode Huffman

Sehingga dari tabel diatas, string biner yang terbentuk dari string “MATEMATIKA INFORMATIKA” adalah

10111100001110111100001011110010101010000001000001011110000101111 (9 byte)

Sebelum dikompresi menggunakan kode Huffman, string tersebut membutuhkan 21 byte memory.

C. Bitmap/Raster graphic

Bitmap atau Raster graphic yang dibahas dalam makalah ini merupakan sebuah gambar digital yang disimpan dengan cara menggunakan matriks sebesar resolusi dari gambar tersebut. Setiap titik pada matriks tersebut merepresentasikan sebuah pixel pada gambar. Pixel tersebut dapat diisi dengan berbagai macam warna yang akan membentuk sebuah gambar. Semakin banyak jumlah pixel maka resolusi gambar akan semakin baik.

Teknologi penyimpanan gambar berupa bitmap ini sudah sangat umum digunakan, seperti hasil gambar dari kamera merupakan sebuah gambar bitmap, hasil pemindaian dari alat pemindai juga akan menghasilkan sebuah gambar bitmap. Pada umumnya penggunaan metode bitmap pada menyimpan dan menampilkan gambar digital lebih banyak digunakan dibanding metode lain yaitu vektor, karena bitmap telah didukung oleh seluruh sistem operasi dan penjelajah internet (browser) yang umum digunakan. Namun disisi lain, bitmap memiliki kekurangan yaitu penggunaan ruang memori yang relatif lebih besar dikarenakan file bitmap menyimpan seluruh informasi warna yang terdapat dalam matriks, setiap pixel dapat berisi 1 bit (hitam dan putih), 2 bit (4 warna), 8 bit, 24 bit (RGB), atau lebih banyak lagi. Bitmap pada umumnya disimpan dalam format BMP, yang telah di dukung oleh banyak sistem operasi. BMP dapat menyimpan data bitmap tanpa kompresi maupun dengan kompresi.

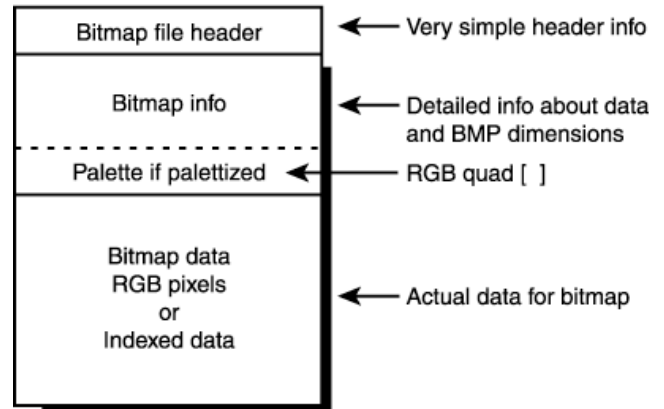
a. Struktur data BMP

Sebuah file BMP terdiri dari 4 bagian utama yaitu

1. Header File
Berisi instruksi dasar agar file dikenali sebagai file bitmap, menyediakan informasi berupa besar file dan posisi
2. Header Bitmap
Berisi seberapa besar gambar yang akan

ditampilkan (baris dan kolom), format pewarnaan apa yang digunakan.

3. Tabel Warna
Berisikan informasi tentang warna yang digunakan
4. Pixel data
Berisikan matriks pixel yang merupakan bagian yang menyimpan matriks gambar.
Bitmap.Bmp file



Gambar 4. Struktur Data pada BMP
source: http://www.yaldex.com/games-programming/0672323699_ch07lev1sec7.html

Dengan mengetahui struktur data BMP kita dapat menghitung perkiraan ukuran dari file yang akan terbentuk, dengan rumus dasar seperti ini

$$FileSize = 54 + (NBrs * NKol) * ColorDepth$$

Seluruh hitungan dalam byte, NBrs merupakan banyaknya baris pixel, NKol merupakan banyaknya kolom pixel, ColorDepth merupakan jumlah bit yang digunakan untuk setiap pixel.

- Color depth yang umum digunakan untuk gambar
- 1 bit (2 warna) untuk gambar monokrom setiap 8 pixel membutuhkan 1 byte
 - 4 bit (16 warna) setiap 2 pixel membutuhkan 1 byte
 - 8 bit (256 warna) setiap 1 pixel membutuhkan 1 byte
 - 16 bit setiap 1 pixel membutuhkan 2 byte
 - 32 bit (4 milyar warna, namun biasa digunakan hanya 1 milyar warna atau 16 juta warna), setiap 1 pixel membutuhkan 4 byte memori.

D. Kompresi

Kompresi atau pemampatan merupakan suatu metode untuk mengurangi jumlah bits data sehingga data yang dihasilkan dapat lebih kecil. Banyak metode yang dapat digunakan untuk kompresi. Kompresi terbagi menjadi 2 jenis

a. Kompresi lossless

Kompresi lossless adalah sebuah teknik kompresi dimana hasil data yang dikompres dapat didekompres kembali (reversible) dan tidak ada data yang hilang dalam melakukan kompresi.

Contoh kompresi lossless : zip, rar.

b. Kompresi lossy

Sistem kompresi lossy adalah sistem kompresi yang irreversible sehingga data yang saat dikompres dianggap tidak berguna akan hilang secara permanen, contoh kompresi lossy : JPEG, MP3.

Kedua jenis kompresi diatas memiliki keunggulan dan kekurangannya masing masing, sistem kompresi ini berlaku untuk segala jenis data yang ada, seperti teks, gambar, audio, video dan lain-lain.

III. PENERAPAN KODE HUFFMAN PADA KOMPRESI GAMBAR BITMAP

A. Penggunaan kode Huffman pada Kompresi

Sebuah gambar digital, dapat direpresentasikan dengan menggunakan palet warna menjadi suatu matriks yang berisikan indeks warna yang ditemukan dalam palet tersebut, misal dengan menggunakan palet IBM's Color Graphic Array 4 bit, warna warna dapat dipetakan menjadi sebuah digit hexadesimal,



Gambar 5. Palet Warna

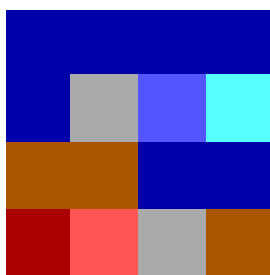
Palet warna diatas kemudian dipetakan menjadi

$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & A & B \\ C & D & E & F \end{pmatrix}$$

Gambar 6. Matriks Palet Warna

Sehingga dengan demikian gambar di bawah ini dapat diubah menjadi sebuah matriks.

Gambar :



Gambar 7. Gambar 4bit

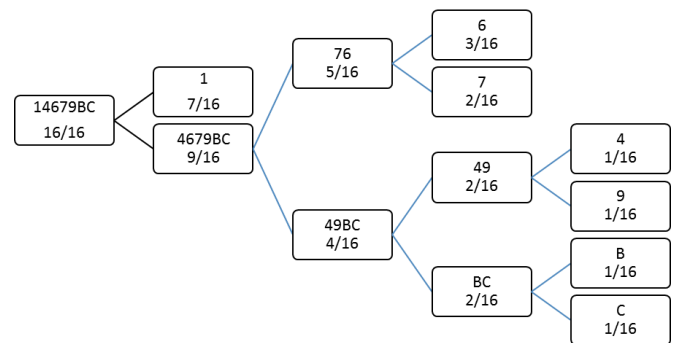
Matriks:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 7 & 9 & B \\ 6 & 6 & 1 & 1 \\ 4 & C & 7 & 6 \end{pmatrix}$$

Gambar 8. Matriks hasil pemetaan warna

Jika matriks tersebut ditulis kembali dengan dalam bentuk array : [1 1 1 1 1 7 9 B 6 6 1 1 4 C 7 6] (64 bit).

Lalu dilakukan kode Huffman pada array tersebut, maka akan didapat kan sebuah pohon Huffman



Gambar 9. Pohon Huffman 2

Setelah didapatkan pohon huffman maka, dapat string biner sebagai berikut

11111010001000010110111100110000010011(37 bit)

Dengan dilakukannya kompresi dengan kode huffman terbukti bahwa jumlah bit yang dibutuhkan dalam menyimpan gambar menjadi lebih pendek.

B. Penggunaan kode Huffman pada Kompresi Gambar digital berformat JPEG

JPEG merupakan sebuah format gambar yang cukup umum digunakan, dikarenakan didukung oleh seluruh sistem operasi, dan juga sizenya yang cukup kecil tanpa mengurangi banyak kualitas. JPEG adalah sebuah gambar bitmap yang telah melewati suatu proses kompresi lossy. Banyak langkah yang dilakukan dalam mengkompresi sebuah gambar agar menjadi JPEG. Salah satu tahap yaitu Huffman Coding, dimana setelah dilakukan berbagai tahap kompresi, dihasilkan suatu matriks 8x8 hasil dari kuantisasi gambar, pada matriks ini terdapat banyak angka 0, hal ini dimanfaatkan dengan menggunakan kode Huffman yang sangat baik dalam mengkompresi data yang redundan. Berikut adalah contoh matriks yang terjadi setelah dibagi dengan matriks kuantisasi.

$$B = \begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -2 & -4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -3 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Gambar 10. Matriks 8x8 setelah kuantisasi
sumber: <http://stackoverflow.com/questions/4931621/jpeg-encoding-technique>

Setelah diperoleh matriks diatas, maka dilakukan enkoding dengan kode Huffman, sehingga pada hasil akhirnya diperoleh sebuah Gambar JPEG yang memiliki kualitas yang sama baiknya namun dapat dikompres hingga 25% dari besar awalnya.

IV. KESIMPULAN

Pohon merupakan suatu teori matematika yang memiliki terapan yang cukup banyak di dunia ilmu komputer, salah satunya adalah pohon Huffman yang memiliki aplikasi sangat banyak, bukan hanya dalam kompresi data berbasis teks saja, namun kode Huffman juga dapat digunakan untuk melakukan kompresi pada sebuah gambar digital.

Pembahasan diatas mengenai pohon Huffman untuk melakukan pemampatan atau kompresi bukanlah suatu pembahasan secara teknis dan lengkap, dikarenakan sistem yang sangat kompleks dalam praktiknya. Pembahasan diatas adalah sebuah analisis bagaimana sebuah kode Huffman dapat diterapkan dalam melakukan kompresi data digital, dan membuktikan bahwa meskipun sudah banyak teknik untuk melakukan kompresi yang lebih baru, namun pohon Huffman masih menjadi suatu teknik untuk menkompresi dan enkoding yang cukup baik dikarenakan waktu untuk kompresi dan dekompresi tidak membutuhkan waktu yang banyak.

REFERENCES

- [1] Munir, Rinaldi. *Diktat Kuliah IF2120 Matematika Diskrit*. Bandung
- [2] <http://www.dragonwins.com/domains/getteched/bmp/bmpfileform.htm>, diakses pada 8 Desember 2016 pukul 21.35.
- [3] Rosen, Kenneth H. 2012. *Discrete Mathematics and Its Applications*. McGraw-Hill : New York.
- [4] <http://milankie.huffmancoding.com/chollowayjpegimagecompression.pdf> diakses pada 9 Desember 2016 pukul 12.37
- [5] <http://techterms.com/definition/bitmap>, diakses pada 8 Desember 2016 pukul 22.32
- [6] <http://www.cs.duke.edu/csed/poop/huff/info/>, diakses pada 8 Desember 2016 pukul 23.21

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 8 Desember 2016



ttd

Kanisius Kenneth Halim, 13515008