

ANALISA BERBAGAI JENIS HURUF KOMPUTER MENGUNAKAN ALGORITMA BERBASIS CHAIN CODE DALAM BENTUK RUN LENGTH ENCODING

Grace Levina Dewi dan Hendrawan Armanto

Teknik Informatika

Sekolah Tinggi Teknik Surabaya

gracelevina@stts.edu dan hendrawan@stts.edu

ABSTRAK

Representasi dari huruf komputer merupakan salah satu topik penting di dalam memproses suatu gambar. Representasi yang digunakan pada penelitian huruf komputer menggunakan algoritma chain code yang bermanfaat untuk memberikan identitas suatu huruf komputer. Hasil dari chain codenya ditampilkan dalam bentuk Run Length Encoding (RLE) yang merupakan suatu kompresi data sehingga dapat menghemat tempat penyimpanan. Huruf yang diteliti menggunakan algoritma chain code ini memakai huruf besar dari huruf A sampai huruf Z. Jenis huruf komputer yang diteliti dalam paper ini meliputi Adobe Caslon Pro Bold, Times New Roman, dan Cooper Black. Totalnya ada 26 karakter pada masing-masing jenis huruf komputer yang diteliti. Bagian yang diambil dalam penelitian paper ini hanya bagian luar dari huruf komputer dari A sampai Z. Langkah-langkah yang dilakukan pertama kali, yaitu: mengumpulkan huruf-huruf yang diperlukan dan diletakkan pada ukuran 50x50 pixel (pengambilan data), mengubah gambar-gambar huruf yang sudah disediakan ke dalam gambar biner, mengambil bagian luar dari huruf, menghitung chain code dengan konektivitas 8 arah, mencari perbedaan sesuai dengan konektivitas yang digunakan, menghitung shape number, dan proses terakhir mengompresi hasil shape number ke dalam bentuk Run Length Encoding. Hal ini membantu mempermudah proses gambar huruf komputer selanjutnya. Hasil yang didapatkan ternyata hampir 85% huruf komputer yang dihitung menggunakan algoritma tersebut hasilnya benar.

Kata kunci: chain code, Run Length Encoding

ABSTRACT

Representation font of the computer case is one of the important topics in the process an image. Representation used in research algorithm chain code letters using the computer that are useful to identify a computer font. Results from the chain code themselves displayed in the form of Run Length Encoding (RLE), which is a data compression which saves storage space. The letters were investigated using chain code algorithm is using uppercase letters A to Z. Letter font type computers that are examined in this paper include Adobe Caslon Pro Bold, Times New Roman, and Cooper Black. In total there are 26 characters in each typeface studied computer. Taken part in the research of this paper is only the outside of the computer case from A to Z. The steps are performed first, namely: to collect the necessary letters and placed on a 50x50 pixel size (data collection), convert images letters which has been

provided to the binary image, taking the outside of the letters, counting chain code with 8-way connectivity, look for differences according to the connectivity is used, calculate the shape number, and the process of compressing the last number to shape outcomes in the form of Run Length Encoding. This helps simplify the process of drawing the letter next computer. Results obtained turned out to almost 85% of the letters are computer calculated using the algorithm results are correct.

Keywords: chain code, Run Length Encoding

I. PENDAHULUAN

Bidang pengenalan pola huruf merupakan salah satu bidang yang sangat menarik untuk dikembangkan lebih lanjut. Hal ini termasuk dalam bidang representasi citra. Tujuan dari representasi citra ini untuk mengubah citra menjadi bentuk yang lebih mudah dimengerti oleh proses selanjutnya. Terdapat dua pilihan pada saat melakukan penelitian, yaitu: representasi berdasarkan karakteristik-karakteristik eksternal (meliputi batas dan bentuk) dan representasi berdasarkan karakteristik-karakteristik internal (meliputi warna atau tekstur).

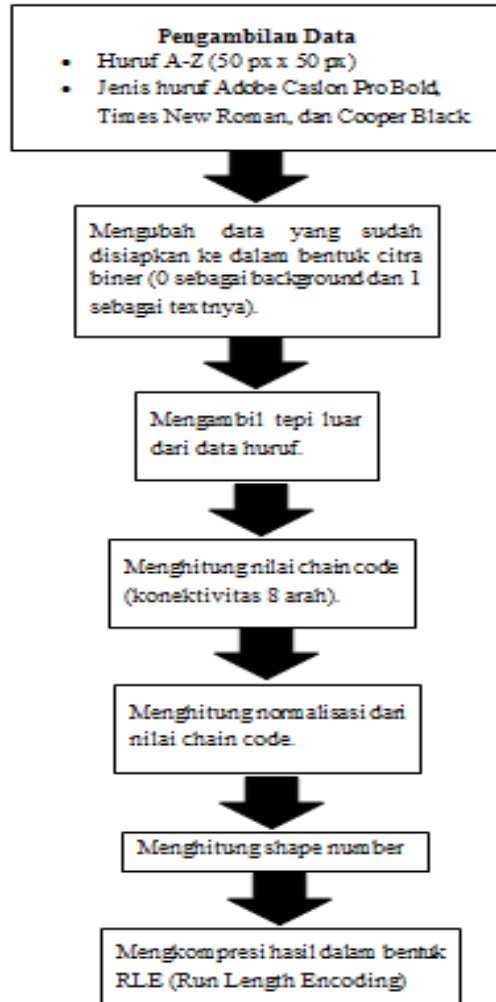
Proses pengenalan huruf yang diteliti dalam paper ini membahas mengenai mekanisme dalam melakukan ekstraksi ciri (representasi eksternal yang membahas tentang batas dan bentuk). Berbagai permasalahan juga ditemukan dalam pembuatan paper ini. Masalah yang dihadapi antara lain teknik apakah yang tepat untuk mengenali berbagai jenis huruf dengan ukuran, tebal suatu huruf, dan bentuk huruf yang berbeda-beda; bagaimana algoritma yang tepat untuk penanganan masalah pada ukuran dan bentuk huruf yang berbeda-beda sehingga huruf tersebut masih dikenali sesuai dengan jenis huruf yang sudah disediakan, dan faktor-faktor apa saja yang menentukan hasil pengenalan huruf komputer.

Solusi yang ditemukan untuk menangani masalah pengenalan huruf tersebut dapat memanfaatkan teknik representasi citra dengan algoritma chain code yang mempunyai salah satu teknik untuk mengekstrak ciri. Tujuan umum dari chain code ini untuk mewakili batas-batas tepi dari object dengan rangkaian segmen yang berupa garis lurus dengan arah dan ukuran yang sudah ditentukan. Skema yang digunakan pada penelitian dengan konektivitas 8 arah dan meneliti dua puluh enam karakter dalam bentuk huruf besar (A-Z). Hasil yang didapatkan dari hasil representasi disediakan dalam bentuk shape number. Supaya data yang disajikan lebih ringkas, ditampilkan pada suatu bentuk yang disebut sebagai Run Length Encoding.

II. METODOLOGI RISET

Proses yang dilakukan dalam penelitian pengerjaan chain code ini dapat dilihat pada gambar 1. Langkah-langkah yang pertama kali dilakukan adalah dengan mengumpulkan huruf komputer dari huruf A-Z (jenis huruf Adobe Caslon Pro Bold, Times New Roman, dan Cooper Black) yang diperlukan dan diletakkan pada ukuran 50x50 pixel (pengambilan data), mengubah gambar-gambar huruf yang sudah disediakan ke dalam gambar biner, mengambil bagian luar dari huruf, menghitung chain code dengan konektivitas 8 arah, mencari perbedaan sesuai dengan konektivitas yang

digunakan, menghitung shape number, dan proses terakhir mengompresi hasil shape number ke dalam bentuk Run Length Encoding. Pembahasan masing-masing metode dapat dilihat pada sub bab 2.1 mengenai metode chain code dan sub bab 2.2. mengenai metode Run Length Encoding.



Gambar 1. Blog Diagram Penelitian

2.1 Metode Freeman Chain Code

Metode yang digunakan pada pembahasan paper ini menggunakan metode Freeman Chain Code. Algoritma dari kode Freeman Chain Code dikenalkan pertama kali oleh Freeman pada tahun 1974. Di Indonesia lebih dikenal dengan istilah kode rantai. Algoritma ini dibentuk dengan tujuan untuk merepresentasikan kontur suatu objek. Representasi yang dimaksud meliputi piksel-piksel dari tepi suatu objek yang saling berhubungan dan memiliki arah tertentu. Terdapat dua jenis Freeman Chain Code, yaitu: menggunakan representasi empat arah dan menggunakan representasi delapan arah (sesuai dengan arah mata angin).

Setiap arah yang ditunjukkan memiliki nilai yang berbeda-beda. Skema penomoran diterapkan pada metode ini, dapat dilihat pada gambar 2. Pada bagian (a) merupakan skema dengan 4 arah mata angin. Pada bagian (b) merupakan skema dengan 8 arah mata angin. Skema 4 arah mata angin memiliki kekurangan, yaitu: tidak ada arah

penghematan yang dilakukan sebesar 84 % pada huruf komputer yang dipakai sebagai contoh tersebut. Angka 255 yang dipakai ditandai sebagai pembatas dan bernilai 1. Penulisan batas ini terjadi jika jumlah dari masing-masing data shape lebih besar dari 4. Jadi, ketika jumlahnya masing-masing data ada 3, ditulis seperti biasa.

III. HASIL PENELITIAN

Pada sub bab ini dibahas mengenai hasil dari penelitian yang dilakukan. Penjelasan lebih lanjut dapat dilihat pada tabel 2. Keterangan yang disajikan di dalam tabel meliputi karakter dari A sampai Z, masing jenis huruf yang digunakan dilengkapi dengan hasil Run Length Encodingnya. Dengan begitu, dapat diketahui berapa besar persentase dari huruf yang benar. Angka 255 yang ada pada bagian RLE merupakan batas dari suatu bilangan yang sudah ditentukan.

Tabel 2. Hasil RLE 3 Jenis Huruf Komputer (Adobe Caslon Pro Bold, Times New Roman, Cooper Black)

Huruf Komputer	Run Length Encoding		
	<i>Adobe Caslon Pro Bold</i>	<i>Times New Roman</i>	<i>Cooper Black</i>
A	255 53 0 255 33 7 255 22 1 4 4 4 3 3 6 2	255 44 0 255 30 7 255 19 1 255 4 4 6 6 2	255 42 0 255 26 7 255 25 1 2 2 2 3
B	255 61 0 255 18 1 255 17 7 4 4 6	255 125 0 255 26 7 255 14 1 255 5 4 6 6 6 2	255 64 0 255 19 1 255 10 7
C	255 54 0 255 32 7 255 31 1 6 6 6 4 4 4 2	255 37 0 255 21 7 255 20 1 4 4	255 54 0 255 31 1 255 26 7 2
D	255 70 0 255 16 7 255 13 1 4 4 2	255 88 0 255 7 7 255 3 4	255 66 0 255 19 1 255 10 7
E	255 115 0 255 27 7 255 13 1 255 5 4 6 6 2	255 129 0 255 18 7 255 5 4 255 4 1 255 3 6 2 2 2	255 99 0 255 25 7 255 24 1 4 4 2
F	255 93 0 255 24 7 255 13 1 255 5 4 6 6 6	255 102 0 255 17 7 255 5 4 255 3 1 6 6 6 2 2	255 78 0 255 19 7 255 17 1 255 4 2
G	255 79 0 255 37 7 255 30 1 4 4 4 2 2 6	255 33 0 255 17 7 255 16 1 4 4 6 2	255 57 0 255 38 1 255 7 2 2
H	255 144 0 255 25 7 255 6 4 255 4 1 6 3	255 156 0 255 23 7 255 7 4	255 95 0 255 23 7 255 16 1 255 7 2
I	255 64 0 255 10 7 4 4 4 1 1 1 3	255 70 0 255 7 7 255 3 4	255 53 0 255 8 1 255 7 7 255 3 2
J	255 73 0 255 13 1 255 10 7 4 4 6	255 6 1 255 5 0 3 7	255 53 0 255 20 1 255 14 7 2 2 6
K	255 95 0 255 31 7 255 16 1 255 3 4 2 2 6 6 3 3	255 119 0 255 26 7 255 8 1 255 7 4 6 6 6	255 72 0 255 26 7 255 25 1 255 5 2 6
L	255 83 0 255 18 7 255 7 1 255 3 4 2	255 89 0 255 9 7 255 3 4 1 1	255 70 0 255 12 1 255 11 7 255 3 2
M	255 134 0 255 55 7 255 41 1 255 5 4 6 6 6 3	255 121 0 255 55 7 255 41 1 255 6 4 6 6 6	255 81 0 255 32 7 255 30 1 255 5 2 6 6 3
N	255 115 0 255 26 7 255 21 1 255 3 6 255 3 4 2 2	255 127 0 255 21 7 255 11 1 255 5 4 255 3 6 2	255 88 0 255 17 7 255 14 1 255 5 2
O	255 36 0 24 1 255 19 7 2 2 6	255 21 0 255 17 1 255 16 7 4 4 6	255 35 0 255 24 1 255 17 7

P	255 66 0 255 13 7 255 8 1 4 4 2 3	255 91 0 255 23 7 255 11 1 255 4 4	255 58 0 255 17 1 255 12 7 2 2
Q	255 49 0 255 40 1 255 39 7 255 3 2 6	255 65 0 255 39 7 255 36 1 4 4 4 6	255 41 0 255 35 1 255 28 7 6 6 6 2 4
R	255 76 0 255 27 7 255 18 1 255 3 4 6 2	255 107 0 255 32 7 255 16 1 255 5 4	255 67 0 255 21 1 255 19 7 255 3 2 6
S	255 62 0 255 29 7 255 3 6 255 3 4 2	255 28 0 255 17 1 255 16 7 4 4 6	255 57 0 255 31 1 255 27 7 2 2 2
T	255 94 0 255 18 7 255 8 1 4 4 4 3 3 6 2	255 87 0 255 11 7 255 4 1 255 3 4 6 6 2 2	255 70 0 255 17 1 255 14 7 2 2
U	255 92 0 255 27 7 255 21 1 255 3 4 6 6 2	255 51 0 255 7 7 255 3 1 4 4 4 6 6 2 2	255 82 0 255 23 1 255 22 7 255 3 2
V	255 45 7 255 36 1 255 31 0 255 4 4 255 3 6 2 2	255 43 0 255 32 7 255 28 1 6 6 6 4 4 3 3	255 35 7 255 33 1 255 28 0 4 4 2 6 3
W	255 69 7 255 63 0 255 57 1 255 5 4 255 4 6 2 2 2	255 71 0 255 62 7 255 48 1 255 7 4 255 5 6 2	255 54 0 255 53 7 255 44 1 2 2 2 4 4 4 6
X	255 61 0 255 45 7 255 32 1 255 5 4 255 3 6 3 2	255 84 0 255 34 7 255 16 1 255 7 4 6 6 6	255 61 0 255 31 7 255 28 1 255 7 2 6 6
Y	255 45 0 255 38 7 255 28 1 255 3 4 6 6 3 2	255 64 0 255 30 7 255 16 1 255 5 4 6	255 58 0 255 18 7 255 17 1 255 5 2 6
Z	255 91 0 255 24 7 255 23 1 255 3 6 3 3 4 4 2	255 78 0 255 21 7 255 19 1 6 6 6 3 3 4 4 2	255 64 0 255 30 1 255 27 7 2 2

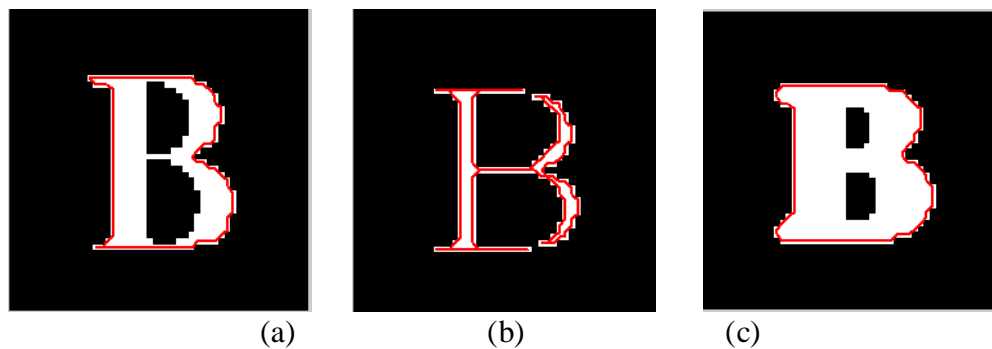
Hasil uji coba dapat dilihat pada tabel 3 mengenai seberapa akurat program algoritma chain code dan RLE yang dilakukan dalam penelitian paper ini. Uji coba yang dilakukan ini berdasarkan 3 jenis huruf komputer (Adobe Caslon Pro Bold, Times New Roman, Cooper Black) dan dihitung berdasarkan persentasenya. Data gambar huruf yang disediakan dari huruf A-Z berjumlah 78 (1 jenis huruf mencangkup 26 karakter). Jumlah total yang sukses dalam uji coba berjumlah 67. Jadi keakuratan yang dicapai sebanyak 85,9%.

Tabel 3. Hasil Uji Coba Keakuratan Hasil Chain Code dan RLE

Huruf Komputer	Adobe Caslon Pro Bold	Times New Roman	Cooper Black	Tingkat Keberhasilan
A	Sukses	Sukses	Sukses	3
B	Sukses	Gagal (terputus)	Sukses	2
C	Sukses	Gagal (terputus)	Sukses	2
D	Sukses	Gagal (terputus)	Sukses	2
E	Sukses	Sukses	Sukses	3
F	Sukses	Sukses	Sukses	3
G	Sukses	Gagal (terputus)	Sukses	2
H	Sukses	Sukses	Sukses	3
I	Sukses	Sukses	Sukses	3
J	Sukses	Gagal (terputus)	Sukses	2
K	Sukses	Sukses	Sukses	3
L	Sukses	Sukses	Sukses	3
M	Sukses	Sukses	Sukses	3
N	Sukses	Sukses	Sukses	3

O	Sukses	Gagal (terputus)	Sukses	2
P	Sukses	Gagal (terputus)	Sukses	2
Q	Sukses	Gagal (terputus)	Sukses	2
R	Sukses	Gagal (terputus)	Sukses	2
S	Sukses	Gagal (terputus)	Sukses	2
T	Sukses	Sukses	Sukses	3
U	Sukses	Gagal (terputus)	Sukses	2
V	Sukses	Sukses	Sukses	3
W	Sukses	Sukses	Sukses	3
X	Sukses	Sukses	Sukses	3
Y	Sukses	Sukses	Sukses	3
Z	Sukses	Sukses	Sukses	3
			Total:	67

Pada gambar 4 menunjukkan hasil dari gambar yang sukses dan gambar yang gagal (terputus). Terjadi kegagalan ini dapat menyebabkan huruf berjenis times new roman tidak dikenali bentuk hurufnya. Petunjuk dari Chain Code dapat dilihat pada gambar garis merah.



Gambar 4. Hasil (a) Adobe Caslon Pro: Sukses (b) Times New Roman: Gagal (terputus) (c) Cooper Black: Sukses

IV. PENUTUP

Setelah didapatkan hasil dari penelitian, ada kesimpulan yang dapat ditarik dan saran yang diberikan. Hal ini bertujuan supaya dari data hasil penelitian ini dapat dikembangkan lebih lanjut. Kesimpulan yang dapat diambil dari penelitian paper ini, yaitu:

- Hasil dari Run Length Encoding dengan ukuran tulisan meskipun berbeda jenis huruf komputernya memiliki tingkat keakuratan yang berbeda. Jenis huruf Adobe Caslon Pro Bold dan Cooper Black memiliki tingkat keakuratan 100% dalam perhitungan chain code dan RLE nya sedangkan Times New Roman memiliki tingkat keakuratan 53,8%. Hal ini disebabkan karena bentuk tulisan dari Times New Roman terputus.
- Chain code memiliki kelemahan yaitu: tidak bisa didapatkan juga batas tepi yang ada kurang dari 30 piksel dan bergantung pada tebal tipisnya jenis tulisan. Tetapi hasil akan sama meskipun posisi chain code berlainan.

V. DAFTAR PUSTAKA

- [1] Prasetyo, Eko. *Pengolahan Citra Digital dan Aplikasinya menggunakan Matlab*. Penerbit Andi, Jakarta. 2011.
- [2] Muntasa, Arif dan Mauridhi Hery Purnomo. *Konsep Pengolahan Citra Digital dan Ekstraksi Fitur*. Graha Ilmu. Agustus. 2010.
- [3] Imran, H Khan, Samitha U.V, dan Suresh Kumar D.S. *Isolated Kannada Character Recognition using Chain Code Features*. International Journal of Science and Research (IJSR). Agustus. 2013.
- [4] Annapurna, Pulipati, Sriraman Kothuri, dan Srikanth Lukka. *Digit Recognition Using Freeman Chain Code*. International Journal of Application or Innovation in Engineering Management. Agustus. 2013.
- [5] Guan Shieng Huang, Jia-Jie Liu, dan Yue-Li Wang. *Sequence Algorithms for Run-Length-Encoded Strings*. 27-29 Juni. 2008. <http://staffweb.ncnu.edu.tw/shieng/slides/COCOON-58.pdf>.