

Jurusan Teknik Informatika

Skripsi Sarjana Komputer

Semester Ganjil tahun 2007/2008

**OTOMATISASI SEGMENTASI DOKUMEN  
DENGAN PENDEKATAN  
GRAY LEVEL CO-OCCURRENCE MATRIX BERBASIS TEXTURE**

**Motilal Chandra            0800759886**

**Kelas/Kelompok : 07PAT/05**

**Abstrak**

Tujuan penelitian adalah untuk melakukan segmentasi dan mengklasifikasi isi dari suatu citra dokumen. Citra dokumen disegmentasi kedalam tiga wilayah: wilayah text, wilayah grafik dan wilayah space. Dalam menyusun skripsi ini, metode penelitian yang digunakan adalah studi pustaka dari artikel – artikel baik dari situs web, tesis, paper dan buku – buku yang berhubungan dalam penulisan ini, pengumpulan data, dan metode perancangan. Hasil dari penulisan ini menunjukkan bahwa texture *Gray Level Co-occurrence Matrix* dapat membedakan suatu wilayah dari suatu citra dokumen. Kesimpulan dari penulisan skripsi ini adalah dapat melakukan segmentasi dan klasifikasi dengan percentase yang cukup tinggi yaitu 91,3% untuk *Gray Level Co-occurrence Matrix* yang menggunakan 8 *gray level*, semakin besar *gray level* yang digunakan semakin tinggi tingkat keberhasilan pengklasifikasian.

Kata Kunci:

Citra Dokumen, Segmentasi Texture, *Gray Level Co-occurrence Matrix* (GLCM), *K-Means Clustering*, Ekstraksi Fitur.

## Prakarta

Puji dan syukur kepada Tuhan yang Maha Esa atas bantuan dan rahmat – Nya hingga penulisan skripsi ini dapat terselesaikan. Setelah memperoleh bekal pendidikan selama masa perkuliahan, dicoba untuk menerapkan ilmu pengetahuan yang telah diperoleh dengan menyusun Laporan Skripsi kami yang berjudul “Otomatisasi Segmentasi Dokumen dengan Pendekatan Gray Level Co-occurrence Matrix Berbasis Texture”. Adapun tujuan penulisan Laporan Skripsi ini adalah untuk memenuhi salah satu syarat menyelesaikan pendidikan di Universitas Bina Nusantara.

Disadari bahwa dalam penyusunan Skripsi ini masih terdapat kekurangan yang tentunya tidak dapat terlepas dari keterbatasan waktu, keterbatasan tingkat pengetahuan serta pengalaman yang dimiliki, maka tidak tertutup adanya saran dan kritik yang bersifat membangun ke arah penyempurnaan dan peningkatan mutu skripsi kami ini.

Pada kesempatan ini, ingin disampaikan terima kasih yang sebesar – besarnya atas bimbingan dan dukungan dari berbagai pihak baik secara langsung maupun secara tidak langsung kepada Yth,

1. Prof. Dr. Gerardus Polla, M.App.Sc, selaku Rektor Universitas Bina Nusantara.
2. Bapak Fredy Purnomo, S.Kom., M.Kom, selaku Ketua Jurusan Teknik Informatika
3. Bapak Diaz D. Santika, Ir, M.Sc, selaku dosen pembimbing Skripsi yang telah bersedia meluangkan waktu, tenaga dan pikiran dalam membimbing dan mengarahkan dalam menyelesaikan skripsi ini.

4. Seluruh Staf pengajar Universitas Bina Nusantara yang telah memberikan ilmu pengetahuan selama masa perkuliahan.
5. Orang tua yang telah memberikan dukungan moral dan materi kepada penulis.
6. Rajbinder Singh S.T yang telah membantu dalam penulisan skripsi ini.
7. Rekan – rekan yang turut membantu dalam memberikan masukan serta informasi dalam masa penyusunan skripsi ini.

Akhir kata diharapkan agar skripsi ini dapat bermanfaat bagi pihak – pihak lain yang berkepentingan maupun sebagai bahan bacaan bagi kepustakaan yang ada.

Jakarta, 28 Januari 2008

Penulis

## DAFTAR ISI

Halaman Judul Luar .....	i
Halaman Judul Dalam .....	ii
Halaman Persetujuan Hard Cover .....	iii
Halaman Pernyataan Dewan Penguji .....	iv
Abstrak .....	v
Prakata .....	vi
Daftar Isi .....	viii
Daftar Table .....	xiv
Daftar Gambar .....	xv
Daftar Lampiran .....	xvii
<b>BAB 1 PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Ruang Lingkup .....	3
1.3 Tujuan dan Manfaat .....	3
1.4 Metodologi.....	4
1.5 Sistematika Penulisan .....	4
<b>BAB 2 LANDASAN TEORI</b>	
2.1 Intelegensia Semu .....	6
2.1.1 Latar Belakang .....	7
2.1.2 Definisi .....	8
2.1.3 Sejarah .....	10

2.1.4	Tujuan .....	12
2.1.5	Beda Intelegensia Semu dan Kecerdasan Alami .....	13
2.1.6	Komputasi Kecerdasan Buatan dan Komputasi Konvensional .....	14
2.1.7	Lingkup Kecerdasan Buatan pada Aplikasi Komersial.....	15
2.2	Computer Vision .....	18
2.3	Soft Computing .....	18
2.3.1	Pengertian .....	19
2.3.2	Sejarah .....	19
2.3.3	Soft Computing sebagai Solusi .....	19
2.3.4	Pencarian Solusi dengan Metoda Soft Computing .....	21
2.3.5	Metode Soft Computing .....	22
2.3.6	Tujuan Soft Computing .....	24
2.3.7	Aplikasi Soft Computing .....	24
2.4	Optical Character Recognition .....	33
2.4.1	Pengenalan Secara On-line.....	33
2.4.1	Pengenalan Secara Off-line.....	34
2.4.1.1	Pengenalan Tulisan yang Dicitak Oleh Mesin .....	34
2.4.1.2	Pengenalan Tulisan Tangan .....	34
2.5	Citra Digital .....	35
2.6	Pixel .....	35
2.7	Image Processing .....	37
2.7.1	Grayscale .....	38
2.7.2	Histogram .....	38
2.7.3	Histogram Normalisasi .....	40

2.7.4	Histogram Equalisation .....	41
2.7.5	Thresholding .....	43
2.7.6	Scaling .....	45
2.7.7	Segmentasi .....	45
2.7.8	Ekstraksi Fitur .....	46
2.8	Gray Level Co-occurrence Matrix .....	46
2.8.1	Definisi “Order” .....	47
2.8.2	Framework Gray Level Co-occurrence Matrix .....	47
2.9	Cluster Analysis .....	50
2.9.1	Aplikasi-aplikasi yang Memungkinkan .....	50
2.9.2	Persyaratan .....	51
2.9.3	Masalah-masalah .....	52
2.9.4	Tipe-tipe Clustering .....	52
2.9.4.1	Hierarchical Clustering .....	52
2.9.4.2	Partitional Clustering .....	58
2.9.4.2.1	K-Means Clustering .....	59
2.9.4.2.2	QT Clustering Algorithm .....	60
2.9.5	Perhitungan Jarak .....	60
2.9.5.1	The Euclidean Distance .....	61
2.9.5.2	City Block Distance .....	61
2.9.5.3	Chebychev Distance .....	62
2.9.5.4	Mahanalobis Distance .....	62
2.9.5.5	Hamming Distance .....	62

## **BAB 3 ANALISIS DAN PERANCANGAN**

3.1	Analisis .....	64
3.1.1	Gambaran Umum .....	64
3.1.2	Identifikasi dan Perumusan Masalah .....	64
3.1.3	Pemecahan Masalah .....	65
3.2	Perancangan .....	66
3.2.1	Pre-processing .....	69
3.2.1.1	Grayscale .....	70
3.2.1.2	Resizing .....	71
3.2.2	Image Subdivision .....	72
3.2.3	Gray Level Co-occurrence Matrix .....	73
3.2.4	Ekstraksi Fitur .....	81
3.2.5	Standarisasi dan Normalisasi Fitur .....	84
3.2.6	K-Mean Clustering .....	86
3.2.7	Mengabungkan Blok-blok Homogen .....	91
3.2.8	Klasifikasi Wilayah .....	93

## **BAB 4 PENGUJIAN DAN EVALUASI**

4.1	Pengujian .....	97
4.1.1	Prosedur Pengujian Segmentasi Dokumen .....	97
4.1.1.1	Ideal Ekspektasi Segmentasi Dokumen .....	98
4.1.1.2	Segmentasi Dokumen dengan Gray Level Co-occurrence Matrix Berbasis Texture .....	101
4.1.2	Hasil Pengujian Segmentasi Citra Dokumen .....	106
4.1.2.1	Pengujian berdasarkan Gray Level .....	106

4.1.2.1.1	Pengujian dengan 8 Gray Level .....	106
4.1.2.1.2	Pengujian dengan 16 Gray Level .....	107
4.1.2.1.3	Pengujian dengan 32 Gray Lecel .....	108
4.1.2.2	Pengujian berdasarkan Penambahan Noise .....	109
4.1.2.2.1	Pengujian tanpa Penambahan Noise .....	110
4.1.2.2.2	Pengujian dengan Penambahan 10% Noise .....	110
4.1.2.2.3	Pengujian dengan Penambahan 30% Noise .....	111
4.1.2.2.4	Pengujian dengan Penambahan 50% Noise .....	112
4.1.2.3	Pengujian berdasarkan Resolusi Scan (dpi) .....	112
4.1.2.3.1	Pengujian dengan 300dpi .....	113
4.1.2.3.2	Pengujian dengan 200dpi .....	113
4.1.2.3.3	Pengujian dengan 96dpi .....	114
4.1.2.3.4	Pengujian dengan 72dpi .....	115
4.2	Evaluasi .....	116
4.2.1	Evaluasi dengan 8 Gray Level .....	116
4.2.2	Evaluasi dengan 16 Gray Level .....	116
4.2.3	Evaluasi dengan 32 Gray Level .....	117
4.2.4	Evaluasi dengan Tanpa Penambahan Noise dan 300dpi .....	117
4.2.5	Evaluasi dengan Penambahan Noise 10% .....	118
4.2.6	Evaluasi dengan Penambahan Noise 30% .....	118
4.2.7	Evaluasi dengan Penambahan Noise 50% .....	119
4.2.8	Evaluasi dengan 200dpi .....	119
4.2.9	Evaluasi dengan 96dpi .....	120
4.2.10	Evaluasi dengan 72dpi .....	120



4.2.11 Perbandingan Keberhasilan dari Gray Level yang digunakan .....	120
4.2.12 Perbandingan Keberhasilan dari Penambahan Noise yang digunakan	121
4.2.13 Perbandingan Keberhasilan dari resolusi scan (dpi) yang digunakan .	121

## **BAB 5 KESIMPULAN DAN SARAN**

5.1 Kesimpulan .....	122
5.2 Saran .....	123
Daftar Pustaka .....	124
Daftar Riwayat Hidup	
Lampiran	

## DAFTAR TABEL

<b>Tabel 2.1</b>	Kecerdasarn Buatan Vs. Pemrograman Konvensional .....	15
<b>Tabel 2.2</b>	Tabel Rating Investasi dari Suatu Bank .....	28
<b>Tabel 3.1</b>	Relasi Fitur .....	94
<b>Tabel 3.2</b>	Contoh Centroid dari Kelas-kelas .....	94
<b>Tabel 3.3</b>	Rangking Kelas-kelas .....	94
<b>Tabel 3.4</b>	Score .....	95
<b>Tabel 4.1</b>	Hasil Pengujian dengan 8 Gray Level .....	107
<b>Tabel 4.2</b>	Hasil Pengujian dengan 16 Gray Level .....	108
<b>Tabel 4.3</b>	Hasil Pengujian dengan 32 Gray Level .....	109
<b>Tabel 4.4</b>	Hasil Pengujian tanpa Penambahan Noise .....	110
<b>Tabel 4.5</b>	Hasil Pengujian dengan Penambahan Noise 10% .....	111
<b>Tabel 4.6</b>	Hasil Pengujian dengan Penambahan Noise 30% .....	112
<b>Tabel 4.7</b>	Hasil Pengujian dengan Penambahan Noise 50% .....	112
<b>Tabel 4.8</b>	Hasil Pengujian dengan 300dpi .....	113
<b>Tabel 4.9</b>	Hasil Pengujian dengan 200dpi .....	114
<b>Tabel 4.10</b>	Hasil Pengujian dengan 96dpi .....	115
<b>Tabel 4.11</b>	Hasil Pengujian dengan 72dpi .....	115

## DAFTAR GAMBAR

<b>Gambar 2.1</b>	Penerapan Konsep Kecerdasaan Buatan di Komputer .....	10
<b>Gambar 2.2</b>	Pemecahan masalah dengan Soft Computing .....	21
<b>Gambar 2.3</b>	Relasi antar 4 Adjacent .....	36
<b>Gambar 2.4</b>	Relasi antar 8 Adjacent .....	37
<b>Gambar 2.5</b>	Contoh gambar berwarna dan greyscale .....	38
<b>Gambar 2.6</b>	Sebuah ban mobil dan histogramnya .....	39
<b>Gambar 2.7</b>	Ban mobil dan histogramnya .....	40
<b>Gambar 2.8</b>	(a) Citra original dan (b) Citra hasil thresholding.....	43
<b>Gambar 2.9</b>	(a)Citra original dan (b)Citra hasil thresholding dengan metode Otsu.....	44
<b>Gambar 2.10</b>	Hasil thresholding dengan metode Otsu .....	45
<b>Gambar 2.11</b>	Citra Contoh .....	46
<b>Gambar 2.12</b>	Grey Level dari Citra Contoh .....	47
<b>Gambar 2.13</b>	Greylevel dari Citra Contoh (atas) dan GLCM dengan relation(1,0) (bawah) .....	48
<b>Gambar 2.14</b>	, Greylevel dari Citra Contoh (atas) dan kombinasi greylevel dari Citra Contoh (bawah) .....	49
<b>Gambar 2.15</b>	Dendogram .....	53
<b>Gambar 2.16</b>	Single Linkage Clustering .....	55
<b>Gambar 2.17</b>	Complete Linkage Clustering .....	56
<b>Gambar 2.18</b>	Average Linkage Clustering .....	57

<b>Gambar 3.1</b>	Flowchart Sistem Dokumen Segmentasi dengan Pendekatan Gray Level Co-occurrence Matrix berbasis Texture .....	68
<b>Gambar 3.2</b>	Blok Diagram Preprocessing .....	69
<b>Gambar 3.3</b>	Flow Chart Modul Grayscale .....	71
<b>Gambar 3.4</b>	Flow Chart Modul Subdivision .....	73
<b>Gambar 3.5</b>	Flow Chart Modul GLCM .....	75
<b>Gambar 3.6</b>	Flow Chart Ekstrak Fitur .....	83
<b>Gambar 3.7</b>	Flow Chart Standarisasi dan Normalisasi Fitur .....	86
<b>Gambar 3.8</b>	Flow Chart K-Means Clustering .....	89
<b>Gambar 3.9</b>	8-Connected Blok .....	92
<b>Gambar 3.10</b>	Flow Chart Modul Joining .....	93
<b>Gambar 3.11</b>	Flow Chart Klasifikasi Wilayah .....	96
<b>Gambar 4.1</b>	Citra Asli Dokumen Ilmiah .....	98
<b>Gambar 4.2</b>	Wilayah Text Menurut Ekspektasi .....	99
<b>Gambar 4.3</b>	Wilayah Grafik Menurut Ekspektasi .....	99
<b>Gambar 4.4</b>	Wilayah Space Menurut Ekspektasi .....	100
<b>Gambar 4.5</b>	(a) Wilayah Text Tunggal Pertama Menurut Ekspektasi (b) Wilayah Text Tunggal Kedua Ekspektasi .....	101
<b>Gambar 4.6</b>	Tampilan Sistem dengan Citra Dokumen Asli .....	102
<b>Gambar 4.7</b>	Wilayah Text Hasil Sistem .....	103
<b>Gambar 4.8</b>	Wilayah Grafik Hasil Sistem .....	103
<b>Gambar 4.9</b>	Wilayah Space Hasil Sistem .....	104
<b>Gambar 4.10</b>	(a) Wilayah Text Tunggal Pertama Hasil Sistem (b) Wilayah Text Tunggal Kedua Hasil Sistem .....	105

## DAFTAR LAMPIRAN

Listing Program .....	L1
Gambar yang Diuji .....	L17
<b>Gambar 1</b> , Dokumen1 .....	L17
<b>Gambar 2</b> , Dokumen2 .....	L18
<b>Gambar 3</b> , Dokumen3 .....	L19
<b>Gambar 4</b> , Dokumen4 .....	L20
<b>Gambar 5</b> , Dokumen5 .....	L21
<b>Gambar 6</b> , Dokumen6 .....	L22
<b>Gambar 7</b> , Dokumen8 .....	L22
<b>Gambar 8</b> , Dokumen9 .....	L23
<b>Gambar 9</b> , Dokumen10 .....	L24