

**KONVERSI CITRA PAPAN PENUNJUK JALAN MENJADI
KARAKTER ASCII DENGAN OPTICAL CHARACTER
RECOGNITION**

Skripsi



oleh
HANS CHRISTIAN HARTONO
22094656

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2013

**KONVERSI CITRA PAPAN PENUNJUK JALAN MENJADI
KARAKTER ASCII DENGAN OPTICAL CHARACTER
RECOGNITION**

Skripsi



©
Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun oleh

HANS CHRISTIAN HARTONO
22094656

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2013

PERNYATAAN KEASLIAN SKRIPSI

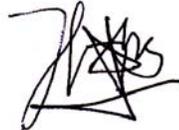
Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

KONVERSI CITRA PAPAN PENUNJUK JALAN MENJADI KARAKTER ASCII DENGAN OPTICAL CHARACTER RECOGNITION

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar kesarjanaan saya.

Yogyakarta, 13 Januari 2013



HANS CHRISTIAN HARTONO
22094656



HALAMAN PERSETUJUAN

Judul Skripsi : KONVERSI CITRA PAPAN PENUNJUK JALAN
MENJADI KARAKTER ASCII DENGAN OPTICAL
CHARACTER RECOGNITION

Nama Mahasiswa : HANS CHRISTIAN HARTONO

N I M : 22094656

Matakuliah : Skripsi (Tugas Akhir)

Kode : TIW276

Semester : Gasal

Tahun Akademik : 2012/2013

Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 13 Januari 2013

Dosen Pembimbing I


Ir. Sri Suwarno, M.Eng.

Dosen Pembimbing II


Lukas Chrisantyo, M.Eng.

HALAMAN PENGESAHAN

**KONVERSI CITRA PAPAN PENUNJUK JALAN MENJADI KARAKTER
ASCII DENGAN OPTICAL CHARACTER RECOGNITION**

Oleh: HANS CHRISTIAN HARTONO / 22094656

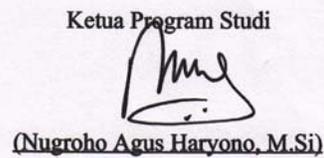
Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 11 Januari 2013

Yogyakarta, 13 Januari 2013
Mengesahkan,

Dewan Penguji:

1. Ir. Sri Suwarno, M.Eng.
2. Lukas Chrisantyo, M.Eng.
3. Willy Sudiarto Raharjo, SKom., M.Cs
4. Aditya Wikan Mahastama, S.Kom


Dekan
(Drs. Wimmie Handiyidjojo, MIT.)


Ketua Program Studi
(Nugroho Agus Haryono, M.Si)

UCAPAN TERIMA KASIH

Puji dan syukur kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat dan karunia kepada penulis dapat menyelesaikan Skripsi dengan judul Konversi Citra Papan Penunjuk Jalan Menjadi Karakter ASCII dengan *Optical Character Recognition* dengan baik dan tepat waktu.

Penulisan dan penyusunan Skripsi ini disusun dalam rangka melengkapi syarat untuk memperoleh gelar Sarjana Komputer. Selain itu bertujuan melatih mahasiswa untuk menghasilkan suatu karya yang dapat dipertanggungjawabkan secara ilmiah, dan dapat bermanfaat bagi penggunanya.

Pada kesempatan ini, penulis mengucapkan terima kasih kepada semua pihak yang telah membantu dalam menyusun skripsi, antara lain :

1. **Bapak Ir. Sri Suwarno, M.Eng.** selaku dosen pembimbing I yang telah memberikan bimbingannya serta memberi masukan yang sangat membantu dari awal hingga akhir selesainya Skripsi ini, juga kepada
2. **Bapak Lukas Chrisantyo, M.Eng.** selaku dosen pembimbing II yang memberikan petunjuk dan masukan dari awal hingga akhir selesainya Skripsi ini.
3. Kedua Orang Tua tercinta yang selalu memberikan semangat, perhatian, dan motivasi serta dukungan agar Skripsi ini selesai.
4. David, Damma, Pitoyo, Budhi dan pihak lain yang tidak dapat penulis sebut satu-persatu yang telah memberikan semangat dan masukan, sehingga Skripsi ini dapat terselesaikan dengan baik dan tepat waktu.

Akhir kata, dengan kerendahan hati, penulis menyadari bahwa Skripsi ini masih jauh dari sempurna, oleh karena itu penulis menerima kritik, saran, semoga Skripsi ini dapat bermanfaat bagi semua pihak.

Yogyakarta, Desember 2012

Hans Christian Hartono

KATA PENGANTAR

Puji syukur penulis ucapkan kepada Tuhan Yesus atas berkat dan penyertaannya sehingga penulis dapat menyelesaikan Skripsi yang berjudul **Konversi Citra Papan Penunjuk Jalan Menjadi Karakter ASCII dengan Optical Chracter Recognition.**

Pengenalan karakter sudah menjadi bahasan umum yang sering diperbincangkan oleh publik. Pengenalan karakter sendiri diawali dengan pengenalan karakter pada dokumen-dokumen, namun seiring perkembangan zaman pengenalan karakter sudah mulai merambah ke berbagai objek.

Melalui penulisan Skripsi ini, penulis berharap agar metode dalam pengenalan karakter dapat semakin dikembangkan. Berbagai metode maupun algoritma yang berbeda dapat digunakan agar pengenalan karakter semakin baik. Tidak menutup kemungkinan juga akan ditemukannya metode baru dalam pengenalan karakter ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, maka saran dan kritik yang konstruktif dari semua pihak sangat diharapkan demi penyempurnaan selanjutnya.

Yogyakarta, Desember 2012

Hans Christian Hartono

INTISARI

Konversi Citra Papan Penunjuk Jalan Menjadi Karakter ASCII dengan Optical Character Recognition

Pembuatan aplikasi untuk pengenalan karakter sudah banyak dibuat, terutama pengenalan karakter di dalam teks dokumen. Namun seiring berjalannya waktu ternyata ada objek-objek lain yang memiliki karakter yang juga dapat dikenali seperti papan penunjuk jalan. Pengenalan karakter pada objek ini akan menjadi dasar yang nantinya dapat digunakan pada kebutuhan pembangunan aplikasi yang lebih lanjut.

Pengenalan karakter pada papan penunjuk jalan ini melewati beberapa tahap, sebelum nantinya karakter pada papan dikenali oleh sistem. Pertama citra papan tersebut akan dirubah menjadi citra *grayscale*, setelah citra *grayscale* didapatkan maka dirubah lagi menjadi citra *binary* menggunakan *thresholding*, lalu dilakukanlah segmentasi citra menggunakan *Connected Components Labeling*. Setelah tersegmentasi maka didapatlah komponen pada citra dimana tiap komponen tersebut akan dikenali dengan menggunakan *Euclidean distance* berdasarkan pada karakter yang sudah ada pada *database*.

Hasil dari penelitian ini, pengenalan karakter dengan langkah diatas tidak terlalu baik dalam mengenali karakter yang ada pada papan petunjuk jalan yang ada. Hal ini terjadi dikarenakan adanya beberapa faktor seperti, kurang akuratnya *Euclidean distance* dalam pengenalan karakter, masih adanya *noise* yang mengganggu pengenalan karakter, pengaruh nilai *threshold*, segmentasi yang kurang sempurna, serta kondisi papan petunjuk jalan yang tidak baik.

Kata kunci : *Optical Character Recognition*, citra *grayscale*, citra biner, *thresholding*, *connected components labeling*, *Euclidean distance*

DAFTAR ISI

HALAMAN JUDUL	
PERNYATAAN KEASLIAN SKRIPSI	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PENGESAHAN	v
UCAPAN TERIMA KASIH	vi
KATA PENGANTAR	vii
INTISARI	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB 1 Pendahuluan.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Perumusan Masalah.....	1
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	2
1.5. Metode Penelitian.....	3
1.6. Sistematika Penulisan.....	4
BAB 2 Tinjauan Pustaka.....	5
2.1. Tinjauan Pustaka	5
2.2. Landasan Teori.....	5
2.2.1. Optical Character Recognition.....	5
2.2.2. Grayscale Image.....	6
2.2.3. Binary Image.....	7
2.2.4. Thresholding	7
2.2.5. Connected Components Labeling	8
2.2.6. <i>Euclidean Distance</i>	8
2.3. Implementasi <i>Euclidean distance</i> dalam <i>Optical Character Recognition</i>	9
BAB 3 Analisis dan Perancangan Sistem	11

3.1.	Alat Penelitian	11
3.1.1.	Perangkat Keras	11
3.1.2.	Perangkat Lunak	11
3.2.	Rancangan Sistem	11
3.2.1.	Usecase.....	11
3.2.2.	Diagram Alir (<i>flowchart</i>)	12
3.2.3.	Algoritma Program	16
3.2.4.	Perancangan Antarmuka	17
Bab 4 Implementasi dan Analisis Sistem.....		19
4.1.	Implementasi Sistem	19
4.1.1.	Antarmuka Program.....	19
4.1.2.	Implementasi Input dan Output.....	21
4.1.3.	Implementasi Algoritma.....	24
4.2.	Analisis Sistem	28
4.3.	Uji Coba Sistem.....	29
BAB 5 Kesimpulan dan Saran.....		39
5.1.	Kesimpulan	39
5.2.	Saran.....	39
DAFTAR PUSTAKA		40
LAMPIRAN		

DAFTAR TABEL

Tabel 4.1 Hasil pengenalan karakter	30
Tabel 4.2 Pengaruh nilai <i>threshold</i>	31
Tabel 4.3 Pengaruh <i>connected components labeling</i>	32
Tabel 4.4 Pengaruh sensitifitas <i>Euclidean distance</i>	33
Tabel 4.5 Kondisi karakter yang sama di papan yang berbeda	34
Tabel 4.6 Jumlah komponen	35
Tabel 4.7 Perhitungan presisi pengenalan karakter untuk nilai <i>threshold</i> 120.	36
Tabel 4.8 Perhitungan presisi pengenalan karakter untuk nilai <i>threshold</i> 150.	37
Tabel 4.9 Perhitungan presisi pengenalan karakter untuk nilai <i>threshold</i> 180.	38



UKDM

DAFTAR GAMBAR

Gambar 2.1 Similaritas dan dissimilaritas citra	9
Gambar 3.1 Usecase sistem	12
Gambar 3.2 Diagram alir utama	13
Gambar 3.3 Diagram alir Connected Components Labeling	14
Gambar 3.4 Diagram alir Euclidean distance	15
Gambar 3.5 Rancangan Halaman Utama 1	17
Gambar 3.6 Rancangan Halaman Utama 2	18
Gambar 4.1 Halaman Utama 1	19
Gambar 4.2 Halaman Utama 2	20
Gambar 4.3 Input citra	21
Gambar 4.4 Resize citra	21
Gambar 4.5 Mengubah citra ke grayscale	22
Gambar 4.6 Mengubah citra grayscale ke citra binary	22
Gambar 4.7 Hasil connected components labeling pada editbox	22
Gambar 4.8 Komponen yang didapat	23
Gambar 4.9 Pengenalan karakter	23
Gambar 4.10 Potongan program dari proses grayscale	24
Gambar 4.11 Potongan program proses thresholding	25
Gambar 4.12 Kondisi pertama	25
Gambar 4.13 Kondisi kedua	26
Gambar 4.14 Kondisi ketiga	27
Gambar 4.15 Connected components labeling tahap 2	27
Gambar 4.16 Proses Euclidean distance	28

INTISARI

Konversi Citra Papan Penunjuk Jalan Menjadi Karakter ASCII dengan Optical Character Recognition

Pembuatan aplikasi untuk pengenalan karakter sudah banyak dibuat, terutama pengenalan karakter di dalam teks dokumen. Namun seiring berjalannya waktu ternyata ada objek-objek lain yang memiliki karakter yang juga dapat dikenali seperti papan penunjuk jalan. Pengenalan karakter pada objek ini akan menjadi dasar yang nantinya dapat digunakan pada kebutuhan pembangunan aplikasi yang lebih lanjut.

Pengenalan karakter pada papan penunjuk jalan ini melewati beberapa tahap, sebelum nantinya karakter pada papan dikenali oleh sistem. Pertama citra papan tersebut akan dirubah menjadi citra *grayscale*, setelah citra *grayscale* didapatkan maka dirubah lagi menjadi citra *binary* menggunakan *thresholding*, lalu dilakukanlah segmentasi citra menggunakan *Connected Components Labeling*. Setelah tersegmentasi maka didapatlah komponen pada citra dimana tiap komponen tersebut akan dikenali dengan menggunakan *Euclidean distance* berdasarkan pada karakter yang sudah ada pada *database*.

Hasil dari penelitian ini, pengenalan karakter dengan langkah diatas tidak terlalu baik dalam mengenali karakter yang ada pada papan petunjuk jalan yang ada. Hal ini terjadi dikarenakan adanya beberapa faktor seperti, kurang akuratnya *Euclidean distance* dalam pengenalan karakter, masih adanya *noise* yang mengganggu pengenalan karakter, pengaruh nilai *threshold*, segmentasi yang kurang sempurna, serta kondisi papan petunjuk jalan yang tidak baik.

Kata kunci : *Optical Character Recognition*, citra *grayscale*, citra biner, *thresholding*, *connected components labeling*, *Euclidean distance*

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Barbagai macam aplikasi untuk pengenalan karakter yang ada saat ini sudah sangat banyak dibuat, terutama aplikasi pengenalan karakter pada sebuah teks dokumen yang ada. Namun pengenalan karakter pada sebuah objek tertentu masih jarang ditemui. Sebagai contoh adalah objek seperti plat nomor, papan nama jalan, baliho, dsb. Jarak dan pembesaran saat pengambilan citra merupakan masalah yang dihadapi saat melakukan pengenalan karakter pada sebuah objek.

Salah satu solusi yang dapat digunakan untuk dapat mengenali karakter adalah dengan *Optical Character Recognition* (OCR). *Optical Character Recognition* pada dasarnya ialah pengenalan karakter alfanumeric dari karakter tulisan tangan atau file maupun citra menjadi teks yang dapat diedit. Karakter yang didapat dari hasil pengenalan karakter tersebut selanjutnya dapat dikembangkan untuk dapat berbagai keperluan.

Dengan selesainya Skripsi ini diharapkan dapat dibangun sebuah sistem yang dapat mengenali karakter pada sebuah citra dengan menggunakan *Optical Character Recognition*. Pada Skripsi ini akan digunakan objek yaitu papan penunjuk jalan yang akan diambil karakter-karakternya.

1.2. Perumusan Masalah

Berdasarkan uraian pada Sub Bab 1.1 maka akan dilakukan sebuah pengenalan karakter pada sebuah papan penunjuk jalan menggunakan metode *Optical Character Recognition*(OCR), sistem yang akan dibangun akan memiliki perumusan masalah sebagai berikut:

1. Bagaimanakah pengambilan karakter dari papan penunjuk jalan yang ada?
2. Bagaimanakah pengenalan karakter pada papan penunjuk jalan tersebut dilakukan?

1.3. Batasan Masalah

Dalam penelitian ini, penulis memberikan batasan masalah untuk sistem yang akan dibuat:

1. Papan penunjuk jalan yang diambil dari sisi depan, dengan jarak kamera dari tanah yaitu 2 meter dan jarak 7-8 meter dari tiang, dengan perbesaran optik 2,5-3 kali pada kamera agar didapat hasil yang optimal, foto papan diambil menggunakan kamera digital *pocket*, yang spesifikasinya akan dijelaskan di Bab 3.
2. Papan penunjuk jalan yang digunakan adalah papan penunjuk jalan yang berwarna hijau, memiliki font bertipe *Sans Serif*, dan papan yang seluruh karakternya berupa karakter kapital.
3. Karakter yang diambil adalah karakter alfabet.
4. Foto yang digunakan memiliki ekstensi .BMP.
5. Foto yang diambil sudah mengalami perubahan secara manual, yaitu hanya berupa papan penunjuk jalan, sedangkan gambar yang berada di sekitar papan sudah dihilangkan.
6. Foto yang digunakan berukuran 640x480 piksel.
7. Foto diambil pada siang hari.
8. *Database* karakter yang digunakan berjenis arial dengan ukuran 20x20 piksel.

1.4. Tujuan Penelitian

Skripsi ini bertujuan untuk melakukan pengenalan karakter alfabet pada sebuah citra papan penunjuk jalan dengan menggunakan *Optical Character Recognition*(OCR).

1.5. Metode Penelitian

Dalam penulisan Skripsi ini, digunakan beberapa metode sebagai acuan dalam perancangan, implementasi dan penelitian terhadap sistem yang dibuat. Metode tersebut sebagai berikut:

1. Studi Pustaka

Studi Pustaka dilakukan dengan mempelajari teori-teori melalui buku, artikel, jurnal dan bahan lain yang mendukung yang berhubungan dengan *character recognition*, *segmentation*, dan metode-metode pendukung lainnya yang dibutuhkan.

2. Perancangan sistem

Terdapat beberapa tahap dalam pembuatan sistem ini, pertama yaitu membuat citra menjadi *grayscale*, lalu diubah ke *binary image*, lalu dilakukan segmentasi dimana akan dilakukan secara baris dan kolom, hal ini digunakan untuk mengambil karakter yang ada. Terakhir dilakukan pembacaan karakter menggunakan karakter yang ada di *database*.

3. Implementasi dan testing

Pada tahap ini akan dilakukan percobaan berupa memberikan *input* sesuai batasan masalah yang ada. Lalu akan dilihat *Output* yang diinginkan adalah berupa karakter alfabet.

4. Konsultasi

Melakukan konsultasi dengan dosen pembimbing berhubungan dengan perancangan dan pelaporan Skripsi.

1.6. Sistematika Penulisan

Skripsi ini disusun dalam sebuah laporan dengan sistematika atau spesifikasi terdiri dari 5 bab:

Bab 1 PENDAHULUAN yang berisi latar belakang masalah, perumusan masalah, batasan masalah, hipotesis, tujuan penelitian, metodologi, dan sistematika penulisan Skripsi.

Bab 2 TINJAUAN PUSTAKA yang berisi gagasan-gagasan yang muncul dengan memberikan landasan teori yang akurat dari berbagai sumber dan konsep-konsep yang dibutuhkan dalam pengenalan karakter menggunakan *Optical Character Recognition*.

Bab 3 ANALISIS DAN PERANCANGAN SISTEM yang berisi perancangan sistem yang akan memberikan gambaran sistem yang akan dibuat serta prosedur-prosedur yang digunakan dalam sistem.

Bab 4 IMPLEMENTASI DAN ANALISIS SISTEM yang berisi implementasi dari hasil perancangan sistem dan pengujian terhadap sistem yang telah dibuat.

Bab 5 KESIMPULAN DAN SARAN YANG berisi kesimpulan atas sistem yang telah dibuat serta saran-saran dalam pengembangan dari Skripsi ini agar dapat dikembangkan kembali.

BAB 5

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan maka dapat disimpulkan bahwa tidak semua papan penunjuk jalan dapat dikenali dengan baik, sedangkan untuk nilai *threshold* yang paling baik digunakan adalah 150, dengan nilai presisi 73,72% berdasarkan percobaan yang dilakukan. Hal ini dikarenakan dalam tahap menuju pengenalan karakter masih terdapat banyak kemungkinan yang terjadi. Sebagai contoh ialah proses dari *binary* dengan *thresholding*, nilai *thresholding* akan mempengaruhi hasil yang ada. Pengambilan karakter dengan menggunakan *connected components labeling* juga masih terdapat kekurangan dalam pengambilan karakter. Selain itu tidak adanya standar dalam pembuatan papan penunjuk jalan juga akan menyebabkan buruknya pengenalan, karena *database* karakter tidak dapat dibuat sebaik mungkin. Hal ini karena *Euclidean distance* dirasa sangat sensitif jika digunakan dalam pengenalan karakter. Selain itu adanya garis baik garis tepi maupun garis pemisah pada papan juga mengganggu pengenalan karena garis tersebut akan dikenali juga sebagai karakter.

5.2. Saran

Sistem yang digunakan merupakan sistem pengenalan karakter pada papan petunjuk jalan, maka dari itu dalam pengembangan sistem kedepannya adalah perlunya algoritma yang lebih baik sebagai proses pengenalan karakter. Selain itu perlunya lagi preprocessing dalam pengambilan papan penunjuk jalan. Selain itu tidak menutup kemungkinan jika diperlukan proses lagi seperti pengurangan *noise* ataupun algoritma lain yang berhubungan dengan jaringan saraf tiruan yang dapat digunakan untuk pengenalan karakter.

DAFTAR PUSTAKA

- Castleman, K. R. (1996). *Digital Image Processing*. New Jersey: Prentice Hall, Inc.
- Hacker, C. (2008, Agustus 5). Retrieved November 3, 2012, from <http://www.borlandtalk.com/resize-a-jpeg-image-re-post-vt113875.html>
- Kumar, T., & Verma, K. (2012). A Theory Based on Conversion of RGB image to Gray image. *International Journal of Computer Application* , 7-10.
- Ozbay, S., & Ercelebi, E. (2005). Automatic Vehicle Identification by Plate Regocnotion. *World Academy of Science, Engineering and Technology* , 222-225.
- Qadri, M. T., & Asif, M. (2009). Automatic Number Plate Recognition System for Vehicle Identification Using Optical Chracter Recognition. *International Conference on Education Technology and Computer* , 335-338.
- Santi, C. N. (2011). Mengubah Citra Berwarna Menjadi Gray0Scale dan Citra biner. *Jurnal Teknologi Informasi DINAMIK Volume 16, No.1* , 14-19.
- Stefano, L. D., & Bulgarelli, A. (1999). A Simple and Efficient Connected Components Labeling Algorithm. *Image Analysis and Processing, 1999. Proceedings. International Conference on* , 322-327.
- Vadivel, A., Majumdar, A. K., & Sural, S. (2003). Performance comparison of distance metrics in content-based Image retrieval applications.
- Wang, L., Zhang, Y., & Feng, J. (2005). On the Euclidean Distance of Images. *Center for Information Sciences* .