

SIMULASI DAN ANALISIS SEGMENTASI CITRA TULISAN TANGAN ANGKA YANG SALING BERSENTUHAN MENGGUNAKAN METODE ZHANG SUEN

Mohammad Adzif Saputra¹, Suryo Adhi Wibowo², Nur Andini³

¹Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

Abstrak

Proses segmentasi merupakan suatu bagian yang sangat penting dalam analisis citra. Tidak hanya pada pengolahan citra objek gambar tetapi juga pada citra tulisan tangan. Akan tetapi seringkali segmentasi pada citra tulisan tangan hanya pada segmentasi kalimat dan segmentasi kata saja. Segmentasi citra tulisan tangan masih perlu dikembangkan hingga pada segmentasi tiap karakter huruf atau angka agar penggunaannya pada pengolahan citra lebih akurat dan lebih baik. Permasalahan untuk mengembangkan segmentasi tulisan tangan ini jika tulisan tersebut saling bersentuhan. Untuk memisahkan 2 karakter tulisan tangan yang saling bersentuhan tersebut diperlukannya suatu proses segmentasi yang bisa memisahkan karakter tersebut.

Dalam Tugas Akhir ini dilakukan simulasi tentang segmentasi citra tulisan tangan yang saling bersentuhan menggunakan Algoritma Zhang Suen. Algoritma ini merupakan salah satu algoritma thinning. Tahap pada Tugas Akhir ini terdiri dari Thinning, ekstrasi fitur points untuk menentukan titik pemotongan, melakukan pemotongan. Pada Tugas Akhir ini akan berfokus pada analisis segmentasi tulisan tangan angka yang saling bersentuhan tidak hanya single touching tapi juga multi touching.

Dari simulasi ini dihasilkan akurasi segmentasi angka yang bersentuhan pada Single Touching 79.599% dengan nilai Sigma = 1, dan Multi Touching 37.943% pada nilai Sigma yang sama. Nilai Sigma sangat mempengaruhi ketebalan tulisan tangan pada citra. Thinning Zhang Suen masih memerlukan peningkatan agar mendapatkan ekstrasi fitur Points yang tepat. Waktu yang diperlukan untuk melakukan segmentasi rata-rata 2.70316 detik per karakter untuk Single Touching dan 2.51648 detik untuk Multi Touching.

Kata Kunci : segmentasi, tulisan tangan, Thinning, algoritma Zhang Suen, single touching, multi touching

Telkom
University

Abstract

the process of segmentation is a very important part in the analysis of the imagery. Not just on image object image processing but also on images of handwriting. But often the segmentation of handwritten image only on the sentence and Word segmentation segmentation. Segmentation of handwritten images still need to be developed to the segmentation of each characters letters or numbers so that its use on image processing more accurate and better. Problems to develop segmentation handwriting glyphs if the glyphs touched. To separate 2 characters handwriting touched the segmentation as a process that can separate character.

In this final task done a simulation about segmentation image penmanship touched use algorithms zhang suen. These algorithms is one algorithms thinning. Its stages on duty this final consisting of thinning, ekstrasi features points to figure out the points cutting, slaughter. On duty this final will focus on analysis segmentation handwriting a figure touched not only single touching but also multi touching.

From this simulation produced the best accuracy of touching numerical segmentation in the Single Touching 79.599%with the Sigma value = 1, and the Multiple Touching 37.943% on the value of Sigma. Sigma values greatly affects the thickness of the handwriting on the image. Thinning Zhang suen still require improvement in order to get the right features extraction Points. The time required to perform segmentation 2.70316 average seconds per character for single touching and 2.51648 seconds per character for multi touching.

Keywords : segmentation, handwriting, Thinning, algorithm Zhang Suen, single touching, multi touching

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pengolahan citra pada masa sekarang mempunyai suatu aplikasi yang sangat luas dalam berbagai bidang kehidupan antara lain bidang arkeologi, astronomi, biomedis, bidang industri dan penginderaan jauh yang menggunakan teknologi citra satelit. Salah satu contoh aplikasi yang sedang berkembang adalah OCR (*Optical Character Recognition*). OCR (*Optical Character Recognition*) adalah sebuah sistem komputer yang dapat membaca huruf, baik yang berasal dari sebuah pencetak (printer atau mesin ketik) maupun yang berasal dari tulisan tangan. Adanya OCR juga akan memudahkan penanganan pekerjaan yang memakai input tulisan seperti penyortiran surat di kantor pos, pemasukan data buku di perpustakaan, dll.

Adanya sistem pengenalan huruf yang cerdas akan sangat membantu usaha besar-besaran yang saat ini dilakukan banyak pihak yakni usaha digitalisasi informasi dan pengetahuan, misalnya dalam pembuatan koleksi pustaka digital, koleksi sastra kuno digital, dll. Segmentasi ini akan mengubah suatu citra masukan yang kompleks menjadi citra yang lebih sederhana, berdasarkan peninjauan terhadap komponen citra. Dengan demikian akan memudahkan pengamat citra untuk melakukan analisis. Segmentasi citra merupakan proses penting yang dapat menentukan kualitas pengenalan pola. Semakin baik hasil segmentasi maka akan sangat membantu dan mempermudah pengenalan pola. Tetapi masih ada kekurangan pada sistem OCR ini, salah satunya tidak bisa membaca karakter yang saling bersentuhan.

Berdasarkan masalah tersebut pada tugas akhir ini dibuat simulasi tentang analisis dan segmentasi tulisan tangan angka yang saling bersentuhan dengan algoritma *Thinning* Zhang Suen untuk membantu meningkatkan kualitas segmentasi khususnya pada citra tulisan tangan angka. Berdasarkan Tugas Akhir sebelumnya^[6] Salah satu tahapnya adalah tahap segmentasi memisahkan tulisan

BAB 1 Pendahuluan

tangan angka yang saling bersentuhan (hanya pada *Single Touching*), tidak berfokus pada analisis segmentasinya tapi fokus pada keakuratan pengenalan kode pos tulisan tangan. Pada Tugas Akhir ini dilakukan Analisis parameter untuk meningkatkan segmentasi tulisan tangan angka yang saling bersentuhan, tidak hanya *Single Touching* tapi juga *Multi Touching*.

1.2 Rumusan Masalah

Dari latar belakang yang telah disampaikan sebelumnya, ada beberapa rumusan masalah yang dibahas pada laporan Tugas Akhir ini sebagai berikut:

- Bagaimana ekstraksi fitur *Points* yang dihasilkan dari citra *Thinning Zhang* suen.
- Bagaimana menentukan jalur segmentasi terbaik untuk memisahkan angka yang bersentuhan dari ekstraksi fitur *Points* yang didapatkan
- Bagaimana tingkat akurasi letak pemisahan dan waktu komputasi yang dibutuhkan

1.3 Tujuan

Hasil yang ingin dicapai melalui penelitian yang dilakukan pada tugas akhir ini adalah:

- Menganalisis hasil segmentasi tulisan tangan angka yang saling bersentuhan menggunakan metode Zhang Suen
- Menganalisis ekstraksi fitur *Points* hasil dari citra *Thinning Zhang* suen
- Menganalisis performansi sistem dengan parameter akurasi titik pemotongan berdasarkan MOS dan waktu komputasinya

1.4 Batasan Masalah

Batasan masalah yang ada pada laporan Tugas Akhir ini meliputi:

- Tulisan tangan yang digunakan dalam Tugas Akhir ini adalah tulisan tangan angka dengan 2 angka yang saling bersentuhan *single touching*

- maupun *multi touching* dengan spidol hitam dan toleransi ketebalan tulisan yang di perbolehkan 1mm - 2,7mm(millimeter).
- b. Pada *Multi Touching* dikususkan pada angka yang bersentuhan di 2 titik
 - c. Masukan sistem ini merupakan hasil *scan* tulisan tangan angka yang saling bersentuhan pada lembar A4 dengan format JPG 400dpi.
 - d. Data diambil dari sampel 30 orang tulisan tangan
 - e. Setiap data terdiri dari 10 baris tulisan tangan dan tiap baris terdiri dari 5 s/d 11 karakter
 - f. Tiap baris terdapat 1 atau 2 angka yang bersentuhan
 - g. Tulisan tangan ditulis tegak.
 - h. Membedakan tulisan tangan yang bersentuhan dengan yang tidak dengan menetapkan jumlah kolom tertentu dan *bwlabel*
 - i. Tulisan tangan yang digunakan dalam Tugas akhir ini penulis batasi hanya pada tulisan tangan angka yang saling bersentuhan bukan tulisan bersambung.
 - j. Hanya membahas segmentasi/pemisahan tulisan tangan angka yang bersentuhan bukan pengenalan karakter.
 - k. Simulasi menggunakan program Matlab.

1.5 Metode Penelitian

Penelitian ini dilakukan dengan metode-metode sebagai berikut.

1 Studi Literatur

Bertujuan untuk mempelajari dasar teori dari berbagai sumber literatur mengenai pengolahan citra digital, pengenalan karakter angka pada tulisan tangan, metode ekstraksi ciri dan algoritmanya pada program MATLAB.

2 Pencarian dan Pengumpulan Data

Pencarian data tulisan tangan yang saling bersentuhan dari tulisan tangan dari beberapa orang

3 Pembuatan Program

Penelitian dilakukan dalam bentuk perancangan, simulasi dan pengujian parameter segmentasi tulisan tangan yang saling bersentuhan dalam *platform* Matlab.

1.6 Sistematika Penulisan

Sistematika Penulisan Laporan tugas akhir ini terdiri atas lima bab yang disusun sebagai berikut

BAB I PENDAHULUAN

Bab ini membahas latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, metode penelitian dan sistematika penulisan.

BAB II DASAR TEORI

Bab ini membahas teori mengenai masalah pengenalan tulisan, citra digital, pengolahan citra digital, dasar teori segmentasi, algoritma Zhang Suen

BAB III PERANCANGAN DAN SIMULASI SISTEM

Bab ini menguraikan tentang tahap proses perancangan dalam mengimplementasikan ke dalam sebuah aplikasi yang dapat memisahkan tulisan tangan angka yang saling bersentuhan dan menganalisisnya.

BAB IV PENGUJIAN DAN ANALISIS HASIL SIMULASI

Bab ini berisi hasil dari penelitian dan menguraikan analisis dari keluaran sistem berdasarkan pada berbagai macam parameter yang mempengaruhi .

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil Tugas Akhir dan saran untuk pengembangan pengembangan lebih lanjut.

BAB V

PENUTUP

5.1. Kesimpulan

Dari hasil analisis terhadap pengujian yang dilakukan, maka dapat ditarik kesimpulan sebagai berikut :

1. Tingkat akurasi paling tinggi untuk segmentasi tulisan tangan angka yang bersentuhan adalah:
 - a. *Single Touching*, didapatkan tingkat akurasi 79.599% dengan nilai $\Sigma = 1$.
 - b. *Multi Touching*, didapatkan tingkat akurasi 37.943% dengan nilai $\Sigma = 1$.
2. Tingkat Akurasi segmentasi tiap karakter mencapai 97.14% untuk *Single touching* dan 97.00% untuk *Multi touching*
3. Tingkat akurasi segmentasi baris 100% untuk *Single* dan *Multi Touching*
4. Citra tulisan tangan hasil Preprocessing tanpa filter median dan Gaussian akan kurang halus dan hasil Thinning menjadi kurang baik sehingga mengakibatkan timbulnya beberapa ekstrasi fitur Points yang tidak dibutuhkan.
5. Semakin besar nilai sigma maka semakin tebal karakter tulisan tangan pada citra sehingga koordinat fitur points akan berubah.
6. Untuk mendapatkan Ekstrasi Fitur *Points* yang tepat algoritma *Thinning* Zhang suen masih memerlukan algoritma peningkatan.
7. Akurasi segmentasi pada *Single Touching* jauh lebih baik daripada *Multi Touching*. Hal ini disebabkan karena fitur *Points* yang didapatkan tidak memenuhi syarat segmentasi *Multi Touching*. *Hole segment* yang didapat terkadang hilang karena terlalu kecil sehingga mengganggu segmentasi.
8. Waktu yang diperlukan untuk melakukan segmenmtasi rata-rata **2.70316** detik per karakter untuk *Single Touching* dan **2.51648** detik untuk *Multi Touching*.

5.2. Saran

Pengembangan yang dapat dilakukan pada tugas akhir ini antara lain :

1. Pada proses pengambilan gambar sebaiknya menggunakan webcam agar sistem bisa digunakan secara *realtime*
2. Penggunaan algoritma segmentasi yang bisa memperbaiki kelemahan pada tugas akhir ini.
3. Kembangkan pada segmentasi huruf/angka yang bersentuhan tidak hanya 2 karakter tapi lebih dari 2.
4. Tambahkan ekstrasi Fitur *points* lainnya agar segmentasi lebih akurat.



DAFTAR PUSTAKA

- [1]. Chen, Yi-Kai, dan Fa Wang Jhing. (2000). *Segmentation of Single or Multiple Touching Handwritten Numeral String Using Background and Foreground analysis*. IEEE Transactions on Pattern Analysis and Machine Intelligence. Taiwan.
- [2]. Hariwibowo, Aditya. 2012. *Pengenalan Pengukuran Huruf pada Tulisan Tangan sebagai Bahan Analisis Grafologi Berbasis Citra Digital dengan Metode Rule Based Sistem*. Bandung :IT Telkom
- [3]. Hery Mauridhi, P dan Arif Muntasa. 2010. *Konsep pengolahan citra digital dan ekstrasi Fitur*. Yogyakarta: Graha Ilmu.
- [4]. J. Sadri, C. Y. Suen, dan T.D. Bui. (2001). *Automatic Segmentation of Unconstrained Handwritten Numeral Strings*. Center of pattern Recognition and Machine Intelligence Concordia University. Canada.
- [5]. Martha Gayatri, Ciendy. 2008. *Analisis Segmentasi Image menggunakan Metode Iteratively Mean Shift Filtering*. Bandung :IT Telkom
- [6]. Megasari, Beby Nur. 2012. *Pengenalan Kode Pos Berbasis Citra Tulisan Tangan yang Saling Bersentuhan dengan Algoritma Segmentasi Background dan Foreground*. Bandung: IT Telkom.
- [7]. Putra Darma. 2010. *Pengolahan Citra Digital*. Yogyakarta : Penerbit Andi.
- [8]. Rupard, Jason. 2005. *An Approach to Chinese Number Recognition*. Florida : University of Nort Florida
- [9]. Septiarini, Anindita. 2012. *Segmentasi Karakter Menggunakan Profil Proyeksi*. Samarinda :Universitas Mulawarman
- [10]. SUEN, ZHANG. 1984. *A Fast Parallel Algorithm for Thinning Digital Patterns*. Canada: Communications of the ACM
- [11]. Wahyu Ismoyo, Ginanjar. 2012. *Realisasi Pencocokan Aantara Label Dan Bentuk Barang Inventaris IT Telkom Menggunakan Metoda Transformasi Wavelet*. Bandung: IT Telkom
- [12]. Zurnawita dan Zulharbi Suar. 2009. *ALGORITMA IMAGE THINNING*. Padang :politekni negeri padang