

EKSTRAKSI KARAKTER BERDASARKAN MULTILEVEL THRESHOLDING DAN REGION GROWING

Sikna Famalia¹, Retno Novi Dayawati², Tjokorda Agung Budi Wirayuda³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Ekstraksi karakter sebagai preprocessing dari proses character pattern recognition telah banyak dimanfaatkan dalam kehidupan sehari-hari. Sebagai contoh, untuk pengenalan karakter pada kertas cek bank, plat nomor kendaraan, barcode, dll. Namun, belum banyak system yang mampu menangani citra dengan background yang kompleks atau bervariasi warna sebagai citra inputan. Sehingga dibutuhkan suatu metode untuk menangani citra jenis tersebut. Tugas akhir ini menerapkan metode Multilevel Thresholding dan Region Growing dalam mengekstraksi citra jenis tersebut. Metode Multilevel Thresholding digunakan untuk menangani citra dengan background yang kompleks, sedangkan Region Growing digunakan untuk mensegmentasi karakter dari backgroundnya. Untuk analisis performansi dilakukan penilaian secara objektif dan subjektif. Penilaian secara objektif dilakukan dengan menghitung nilai akurasi segmentasi, dengan cara membandingkan tingkat kemiripan hasil segmentasi citra uji dengan citra banding. Sedangkan penilaian secara subjektif dilakukan dengan menggunakan Mean Opinion Score (MOS) melalui sepuluh orang pengamat.

Kata Kunci : ekstraksi karakter, segmentasi, Multilevel Thresholding, Region Growing, Mean Opinion Score

Abstract

Extraction character as a preprocessing from the character pattern recognition process have been usefulable in daily live. For the example, for character recognition that use in bank for check paper, transportation plate number, barcode, and so on. But there is not much system that can handle image with complex background or with variation of colour as inputs image, so that needs a method for handle the image. This final project aplying Multilevel Thresholding dan Region Growing method for extracts that image. Metode Multilevel Thresholding method is used for handle complex background image, while Region Growing method is used for segment a character from the background. For performance analyze used Objective and subjective scoring. Objective scoring used with calculating segmentation accuration score, that compare the similarity rate of segmentation test image result with the compare image. Subjective scoring used Mean Opinion Score (MOS) with ten observers.

Keywords : character extraction, segmentation, Multilevel Thresholding, Region Growing, Mean Opinion Score

1. Pendahuluan

1.1 Latar belakang masalah

Pemanfaatan ilmu komputer telah banyak digunakan untuk dunia teknologi. Salah satunya ialah pengenalan pola karakter dari suatu dokumen, pengenalan plat nomor kendaraan, pengenalan tulisan tangan pada kertas cek bank, dll. Salah satu proses yang mendukung kesuksesan pengenalan pola karakter atau *character pattern recognition* tersebut ialah adanya proses ekstraksi karakter.

Maksud ekstraksi karakter di sini ialah suatu proses pemisahan karakter dari backgroundnya yang ada pada suatu citra. Atau bisa disebut juga dengan segmentasi citra. Ekstraksi karakter ini merupakan preprocessing dari *character pattern recognition*. Banyak metode yang dapat digunakan untuk proses ekstraksi karakter, diantaranya ialah metode *Multilevel Thresholding* dan *Region Growing* yang akan penulis implementasikan dalam Tugas Akhir ini.

Multilevel thresholding merupakan proses yang memotong gray level pada image menjadi beberapa region yang jelas. Teknik thresholding jenis ini membutuhkan lebih dari satu threshold untuk input image dan memotong image menjadi region-region yang pasti. Metode ini bekerja sangat baik pada objek dengan background yang kompleks atau bervariasi warna.

Region growing merupakan suatu pendekatan pada segmentasi image dimana pixel-pixel yang bertetangga pada image diuji dan ditambahkan ke dalam suatu kelas region jika tidak terdeteksi adanya edge. Tepi atau edge yang dimaksud adalah perubahan nilai derajat keabuan atau brightness value (BV) pada citra yang besar dalam jarak yang kecil. Proses Region Growing ini berfungsi untuk mengubah citra hasil Multilevel Thresholding menjadi citra biner dan mensegmentasi karakter dari backgroundnya.

Penggunaan kedua metode tersebut sangat dibutuhkan untuk mengoptimalkan hasil segmentasi karakter. Penggunaan hanya salah satu dari kedua metode tersebut tidak akan mensukseskan proses ekstraksi karakter. Penggunaan Multilevel Thresholding saja tanpa Region Growing tidak akan mampu mensegmentasi karakter dari backgroundnya, karena proses Multilevel Thresholding hanya melakukan pengelompokan nilai-nilai intensitas pixel menjadi nilai intensitas pixel yang lebih sesuai. Sedangkan penggunaan proses Region Growing tanpa Multilevel Thresholding tidak mampu mengoptimalkan hasil segmentasi karakter.

1.2 Perumusan masalah

Dalam tugas akhir ini dirumuskan beberapa masalah sebagai berikut :

1. Bagaimana mengimplementasikan metode Multilevel Thresholding dan Region Growing dalam mengekstraksi karakter pada suatu image.
2. Bagaimana menguji kemampuan metode Multilevel Thresholding dan Region Growing dalam mengekstraksi karakter pada image dengan background yang kompleks.
3. Bagaimana cara mengukur keakuratan sistem dalam mengekstraksi suatu karakter pada image.

Adapun batasan masalah agar tidak meluasnya materi pembahasan dalam tugas akhir ini adalah sebagai berikut :

1. Citra masukan berupa citra digital dengan format citra dalam bentuk *.bmp (bitmap) 8 bit berukuran 200x200 pixel.
2. Inputan citra yang akan digunakan tidak memiliki noise.
3. Terdapat tiga tipe citra yang digunakan untuk proses segmentasi dengan Multilevel Thresholding dan Region Growing, yaitu citra dengan warna karakter yang berbeda tajam dari background, citra dengan warna karakter yang menyerupai warna background, citra dengan karakter yang memiliki gradasi warna.
4. Inputan citra untuk proses pengenalan karakter menggunakan font Arial.

1.3 Tujuan

Dalam tugas akhir ini, hal-hal yang diharapkan untuk dicapai adalah sebagai berikut :

1. Mengimplementasikan metode Multilevel Thresholding dan Region Growing dalam mengekstraksi karakter pada suatu image
2. Menguji kemampuan metode Multilevel Thresholding dan Region Growing dalam mengekstraksi karakter pada image dengan background yang kompleks (bervariasi warna)
3. Menganalisa keakuratan sistem dalam mengekstraksi karakter pada suatu image

1.4 Metodologi penyelesaian masalah

Secara umum metode penelitian yang dilakukan adalah sebagai berikut :

1. Studi literatur
Mengumpulkan data berupa referensi-referensi untuk mempelajari konsep dari penggunaan metode Multilevel Thresholding dan Region Growing dalam mengekstraksi karakter pada suatu image yang memiliki background yang kompleks.
2. Analisa dan desain
Menganalisa kebutuhan, mengumpulkan data yang diperlukan, dan melakukan perancangan untuk tahap implementasi
3. Implementasi
Tahap ini meliputi implementasi metode Multilevel Thresholding dan Region growing dengan bahasa pemrograman Matlab.
4. Pengujian dan analisa hasil
Melakukan pengujian secara subjektif dan objektif serta melakukan analisa terhadap keluaran dari aplikasi.
5. Penyusunan laporan
Tahap terakhir ini adalah penyusunan laporan hasil penelitian dan kesimpulan serta saran terhadap penelitian ini.

1.5 Sistematika Penulisan

Dalam tugas akhir kali ini, penulisan disusun berdasarkan sistematika sebagai berikut:

1. Pendahuluan

Pendahuluan berisikan latar belakang, perumusan masalah, batasan masalah, tujuan pembahasan, metodologi dan sistematika penulisan dari Tugas Akhir.

2. Dasar Teori

Dasar teori berisikan seluruh landasan teori yang digunakan untuk mendukung pembuatan Tugas Akhir ini. Teori yang dibahas antara lain tentang pengertian citra digital, pengolahan citra, segmentasi citra, *grayscale*, *multilevel thresholding*, *region growing*, serta perhitungan performansi system secara subjektif maupun objektif.

3. Implementasi Sistem

Analisis dan perancangan sistem merupakan rangkaian seluruh kebutuhan sistem mulai dari gambaran umum sistem, analisis input-output sistem, pemodelan system, serta perhitungan manual.

4. Implementasi dan Analisis Hasil Pengujian

Implementasi dan analisis hasil pengujian berupa implementasi dari sistem, pengujian sistem dan analisis hasil pengujian.

5. Kesimpulan dan Saran

Kesimpulan dan saran berupa suatu rangkuman dari hasil keseluruhan proses implementasi dan pengujian yang dilengkapi dengan saran untuk proses pengembangan lebih lanjut.

setelah penentuan seed point, karakter dengan gray level yang berbeda jauh dari gray level seed dan tidak memenuhi criteria, tidak akan dimasukkan ke dalam region. Hal ini lah yang menyebabkan karakter tak terlihat.

- Hasil pengujian secara objektif menunjukkan nilai yang baik, yaitu di atas 80%. Sedangkan pengujian secara subjektif menunjukkan nilai yang kurang baik. Hal ini dipengaruhi oleh penilaian serta persepsi yang berbeda-beda dari masing-masing penilai. Ada yang menyebutkan bahwa hasilnya kurang bagus karena warnanya hitam dan putih, ada yang menyebutkan karena jenis font yang dipilih kurang jelas, dll
- Ketepatan hasil dari proses pengenalan karakter pada citra berkarakter dengan jenis font Arial mencapai 100%. Ketepatan ini sangat dipengaruhi oleh bagus tidaknya hasil segmentasi. Hasil segmentasi yang *bernoise* bisa merubah nilai x_{min} , x_{max} , y_{min} , dan y_{max} untuk proses pemotongan masing-masing karakter pada citra.

5. Kesimpulan

5.1 Kesimpulan

Setelah melakukan pengujian dan analisis terhadap ekstraksi karakter berdasarkan Multilevel Thresholding dan Region Growing, dapat disimpulkan beberapa hal yaitu:

1. Metode *Multilevel Thresholding* dan *Region Growing* terbukti mampu mengekstraksi karakter dengan baik dari background yang kompleks, dengan memilih parameter $level=9$ dan $t=0.01$. Dibuktikan dengan rata2 hasil akurasi terhadap tujuh citra uji mencapai 91.573 %
2. Untuk inputan citra yang memiliki karakteristik warna karakter berbeda cukup jauh dengan background, pemilihan parameter $level=3$ pada proses *Multilevel Thresholding* akan menghasilkan hasil segmentasi yang lebih baik. Namun parameter $level=3$ tersebut kurang menguntungkan ketika digunakan untuk inputan citra dengan karakteristik warna karakter yang menyerupai warna background. Sehingga, setelah melakukan testing terhadap citra uji disimpulkan bahwa $level$ yang digunakan bernilai 9 dengan harapan mampu menangani kedua kasus di atas.
3. Pengembangan region pada proses *Region Growing* bergantung pada nilai parameter $threshold (t)$ yang dipilih. Semakin besar nilai $threshold$ maka semakin luas wilayah region untuk berkembang dan semakin sedikit jumlah region yang dihasilkan. Namun untuk mendapatkan hasil segmentasi karakter yang optimal dipilih nilai $threshold$ yang kecil (dalam tugas akhir ini dipilih nilai 0.01) agar karakter yang tersegmentasi tampak jelas.
4. Penilaian subjektif dengan MOS menghasilkan rata-rata penilaian yang baik untuk setiap hasil segmentasi karakter pada citra uji.
5. Pengenalan karakter terhadap dua citra uji sangat baik dilakukan, dibuktikan dengan tercapainya ketepatan dalam mengenali karakter mencapai 100%.

5.2 Saran

Tujuan penggunaan metode Multilevel Thresholding dan Region Growing adalah untuk menghasilkan segmentasi karakter pada citra. Namun hasil segmentasi yang didapat masih kurang sesuai dengan yang diharapkan. Sehingga diperlukan metode tambahan untuk memperbaiki hasil segmentasi agar menghasilkan citra hasil yang lebih optimal, dengan kata lain tidak terdapat suatu *noise* yang tak diharapkan yang membuat hasil segmentasi karakter pada citra kurang maksimal.



Referensi

[1]	Gonzales, Rafael C and Woods, Richard. "Digital Image Processing", Addison-Wesley Publishing Company, USA, 1993.
[2]	Schalkoff, Robert J. (1989). Digital Image Processing and Computer Vision. John Wiley & Sons, New York.
[3]	Munir, Rinaldi. Pengolahan Citra Digital dengan Pendekatan Algoritmik. Penerbit Informatika. Bandung. 2004.
[4]	http://en.wikipedia.org/wiki/Region_growing diakses tanggal 22 Desember 2008
[5]	srini.staff.gunadarma.ac.id/Downloads/files/4881/8+Olah+Citra-Konsep+Dasar.pdf
[6]	Arora, S., J. Acharya, A. Verma, Prasanta K. Panigrahi. "Multilevel Thresholding for Image segmentation through a Fast Statistical Recursive Algorithm"
[7]	Victor Hariadi. "PEMANFATAAN GRAPH DALAM PENYELESAIAN PERMASALAHAN SEGMENTASI CITRA" (teori-def citra digital.pdf)
[8]	Acharya, Tinku and Ray, Ajoy K. "Image Processing Principles and Application", Wiley Interscience : September 2005.
[9]	Goto, Hideaki and Hiroto Aso. 2000. <i>Character Pattern Extraction Based on Local Multilevel Thresholding and Region Growing</i> . Tohoku University.
[10]	Goto, Hideaki and Hiroto Aso. 2000. <i>A Character Pattern Extraction Method Applicable to Various Types of Images</i> . Tohoku University.
[11]	Goto, Hideaki and Hiroto Aso. 2002. <i>Character Pattern Extraction from Colorful Document with Complex Backgrounds</i> . Tohoku University.
[12]	Mat-Isa, N. A., M. Y. Mashor, and N. H. Othman. 2005. <i>Seeded Region Growing Features Extraction Algorithm</i> . Universiti Sains Malaysia
[13]	Liao, Ping-sung, Tse-sheng chen, and Pau-choo chung. 1999. <i>A Fast Algorithm for Multilevel Thresholding</i> . Taiwan
[14]	http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/LIB/bandb5.pdf diakses bulan November 2008

5.2 Saran

Tujuan penggunaan metode Multilevel Thresholding dan Region Growing adalah untuk menghasilkan segmentasi karakter pada citra. Namun hasil segmentasi yang didapat masih kurang sesuai dengan yang diharapkan. Sehingga diperlukan metode tambahan untuk memperbaiki hasil segmentasi agar menghasilkan citra hasil yang lebih optimal, dengan kata lain tidak terdapat suatu *noise* yang tak diharapkan yang membuat hasil segmentasi karakter pada citra kurang maksimal.



Referensi

[1]	Gonzales, Rafael C and Woods, Richard. "Digital Image Processing", Addison-Wesley Publishing Company, USA, 1993.
[2]	Schalkoff, Robert J. (1989). Digital Image Processing and Computer Vision. John Wiley & Sons, New York.
[3]	Munir, Rinaldi. Pengolahan Citra Digital dengan Pendekatan Algoritmik. Penerbit Informatika. Bandung. 2004.
[4]	http://en.wikipedia.org/wiki/Region_growing diakses tanggal 22 Desember 2008
[5]	srini.staff.gunadarma.ac.id/Downloads/files/4881/8+Olah+Citra-Konsep+Dasar.pdf
[6]	Arora, S., J. Acharya, A. Verma, Prasanta K. Panigrahi. "Multilevel Thresholding for Image segmentation through a Fast Statistical Recursive Algorithm"
[7]	Victor Hariadi. "PEMANFATAAN GRAPH DALAM PENYELESAIAN PERMASALAHAN SEGMENTASI CITRA" (teori-def citra digital.pdf)
[8]	Acharya, Tinku and Ray, Ajoy K. "Image Processing Principles and Application", Wiley Interscience : September 2005.
[9]	Goto, Hideaki and Hiroto Aso. 2000. <i>Character Pattern Extraction Based on Local Multilevel Thresholding and Region Growing</i> . Tohoku University.
[10]	Goto, Hideaki and Hiroto Aso. 2000. <i>A Character Pattern Extraction Method Applicable to Various Types of Images</i> . Tohoku University.
[11]	Goto, Hideaki and Hiroto Aso. 2002. <i>Character Pattern Extraction from Colorful Document with Complex Backgrounds</i> . Tohoku University.
[12]	Mat-Isa, N. A., M. Y. Mashor, and N. H. Othman. 2005. <i>Seeded Region Growing Features Extraction Algorithm</i> . Universiti Sains Malaysia
[13]	Liao, Ping-sung, Tse-sheng chen, and Pau-choo chung. 1999. <i>A Fast Algorithm for Multilevel Thresholding</i> . Taiwan
[14]	http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/LIB/bandb5.pdf diakses bulan November 2008