

Aplikasi String Matching dalam Analisis Cap Bibir

Khoirunnisa Afifah (13512077)
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13512077@std.stei.itb.ac.id

Abstrak—Cap bibir merupakan salah satu karakteristik yang dapat digunakan dalam identifikasi biometrik. Cap bibir ini sama seperti sidik jari yaitu unik pada setiap manusia. Cap bibir memiliki pola-pola tertentu yang dapat digunakan untuk menentukan jenis kelamin dan pengidentifikasian individu. Makalah ini akan mengalisa cap bibir yang dianalogikan sebagai string untuk mengidentifikasi jenis kelamin menggunakan hasil analisa tersebut.

Kata Kunci—cap bibir, pencocokan string.

I. PENDAHULUAN

Dalam dunia teknologi informasi ada yang disebut teknologi biometrik. Studi ini bertujuan untuk mengidentifikasi karakteristik fisik dan kelakuan manusia untuk keperluan autentikasi. Karakteristik fisik berhubungan dengan bentuk tubuh seperti pendeteksian sidik jari, wajah, iris mata, gigi, cap bibir, dan DNA. Karakteristik yang kelakuan bisa dikenali lewat suara, ritme mengetik dan tanda tangan.

Pada makalah ini akan dijelaskan pencocokan pola cap bibir jika cap bibir di asumsikan sebagai string. Pencocokan dilakukan menggunakan tiga algoritma pencocokan string yaitu Brute Force, KMP, dan Boyer-Moore. Dari pencocokan tersebut akan dihasilkan pola cap bibir yang sesuai dan selanjutnya dapat digunakan untuk menentukan jenis kelamin seseorang.

II. DASAR TEORI

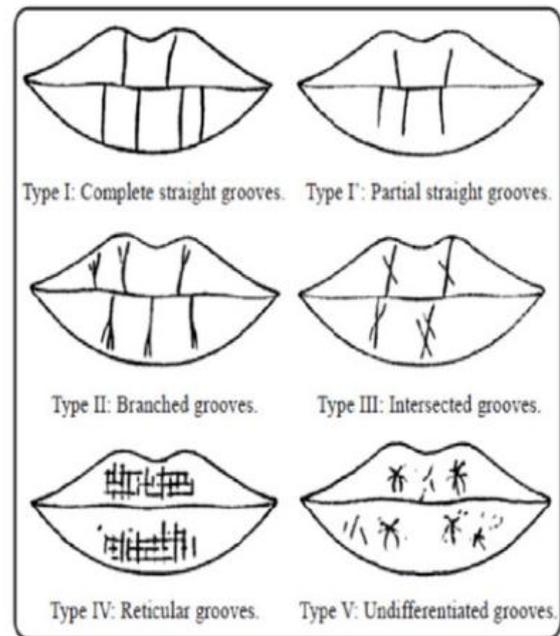
A. Cap Bibir

Seperti sidik jari, alur pada cap bibir merupakan karakter fisik yang unik antara satu orang dengan orang lain. Hal ini disadari pertama kali oleh seorang anthropologist bernama R Fischer pada 1902. Ilmu yang khusus mempelajari tentang cap bibir dinamakan Cheiloscopy. Meski dapat berfungsi sebagai alat identifikasi manusia sama seperti sidik jari, penggunaan cap bibir dalam kasus kriminal masih jarang digunakan.

Selain untuk identifikasi personal, cap bibir juga dapat digunakan untuk memperkirakan jenis kelamin seseorang. Suzuki dan Tsuchihashi pada tahun 1970 telah mengklasifikasikan pola garis cap bibir menjadi enam pola. Pola-pola inilah yang menjadi acuan dalam

penelitian terhadap cap bibir selanjutnya.

Keenam pola menurut Suzuki dan Tsuchihashi tersebut adalah :

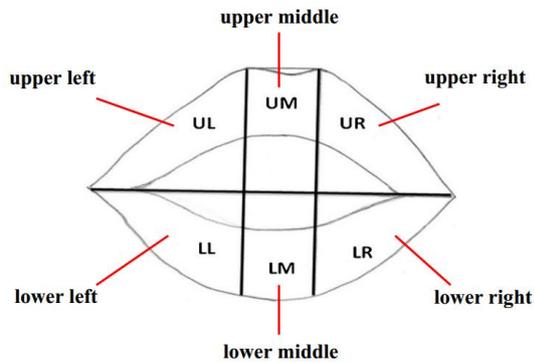


Gambar tipe-tipe pola cap bibir ^[2]

Pola-pola cap bibir tersebut dapat digunakan untuk meneliti jenis kelamin seseorang. M.A.E. Domiaty dan Vahanmala meneliti kecenderungan pola cap bibir terhadap jenis kelamin dan memberikna hasil yang relatif sama yaitu

- Type I : Perempuan
- Type I' : Perempuan
- Type II : Laki-laki
- Type III : Laki-laki
- Type IV : Laki-laki
- Type V : Laki-laki

Selain untuk menentukan jenis kelamin cap bibir dapat dimanfaatkan untuk mengidentifikasi seseorang. Cap bibir dibagi menjadi enam bagian Upper Left, Upper Right, Upper Middle, Lower Left, Lower Right, dan Lower Middle. Setelah itu dihitung jumlah masing-masing jenis alur seperti garis horizontal, vertikal, bercabang, dan berpotongan pada setiap bagian. Data ini kemudian diolah untuk mengidentifikasi seseorang.



Gambar pembagian area identifikasi cap bibir [3]

B. Konsep String

Misal S adalah string dengan ukuran m

$$S = x_1x_2x_3\dots x_m$$

Yang dimaksud prefix dari S adalah substring $S[1..k-1]$. Sedangkan suffix dari S adalah substring $S[k-1..m]$. Dengan k adalah sembarang index antara 1 dan m . $S[0]$ adalah null karakter.

Pada pencocokan string ada yang dinamakan Teks yaitu string dengan panjang m dan Pattern adalah string berukuran n yang akan dicocokkan dengan Teks ($m \geq n$). Algoritma pencocokan string ada tiga macam, yaitu Brute Force, Boyer-Moore, dan Knutt-Morris-Pratt. Selain itu dapat juga menggunakan teknik regex untuk mendapatkan informasi dari text.

C. Algoritma pencocokan string : Brute Force

Inti algoritmanya adalah mencocokkan setiap karakter dari string di teks dengan setiap karakter dari string pattern. Kasus terburuknya ada $m(n-m+1)$ kali pencocokan. Dalam notasi big O kompleksitasnya $O(mn)$. Kasus terbaik terjadi jika karakter pertama pattern tidak pernah sama dengan karakter teks T yang dicocokkan. Jumlah perbandingan maksimal n kali. Kompleksitasnya $O(n)$. Secara umum kompleksitas algoritma pencocokan string dengan Brute Force adalah $O(m+n)$. Algoritma Brute Force baik digunakan jika alfabet penyusun pattern dan teks banyak, tapi menjadi buruk jika alfabet penyusunnya kecil.

D. Algoritma pencocokan string : Knutt-Morris Pratt

Algoritma ini menelusuri pattern di teks dari kiri ke kanan seperti Brute Force, tapi pergeseran pattern dalam teksnya lebih baik dari pada Brute Force. Jika terjadi ketidakcocokan antara teks dan pattern pada $P[j]$ kita menggeser P sejumlah prefix dari p yang juga merupakan suffix dari P . Jadi jumlah pergeseran pattern adalah sebesar panjang pattern dikurangi jumlah prefix yang juga merupakan suffix terbesar. Untuk menghitung jumlah prefix yang juga merupakan suffix kita menggunakan fungsi bernama border function.

Kompleksitas waktu algoritma ini dibagi menjadi dua. Untuk menghitung border function kompleksitasnya $O(m)$, sedangkan untuk pencarian string kompleksitasnya $O(n)$. Kompleksitas waktu totalnya adalah $O(m+n)$.

Algoritma Knutt-Morris-Pratt tidak bekerja baik jika ukuran alfabet dari teks besar.

E. Algoritma pencocokan string : Boyer-Moore

Algoritma Boyer-Moore didasarkan pada dua teknik, yaitu teknik *Looking-Glass* dan teknik *Character-Jump*. Teknik *Looking-Glass* mencocokkan P di T mulai dari belakang. Teknik *Character-Jump* dipakai untuk menggeser pattern saat ada ketidakcocokan antara $T[i]$ dan $P[j]$. Kasus penggeseran string ada 3 kemungkinan. Yang pertama jika karakter $T[i]$ muncul di sebelah kiri $P[j]$, pergeseran yang dilakukan adalah menggeser pattern ke kanan sampai karakter $T[i]$ sama dengan $P[j]$. Kasus kedua jika karakter $T[i]$ muncul di sebelah kanan $P[j]$, pergeseran yang dilakukan adalah menggeser pattern 1 karakter. Kasus ketiga jika karakter $T[i]$ tidak ada di pattern, pergeseran yang dilakukan adalah menggeser $P[j]$ sampai apada posisi $T[i+1]$.

Kompleksitas waktu algoritma untuk kasus terburuk adalah $O(nm+A)$. Boyer Moore baik digunakan jika alfabet (A) nya kecil.

III. HASIL DAN PEMBAHASAN

A. Metode yang digunakan

Dalam makalah ini setiap pola cap bibir akan dianalogikan dengan string yang terdiri dari alfabet 0-9, a-z, dan A-Z yang disusun secara acak.

Cap bibir yang akan dideteksi merupakan teks dengan panjang 30. Teks ini kemudian dibagi menjadi 6 bagian yang merepresentasikan bagian-bagian dari bibir (UL, UM, UR, LL, LM, LR).

Setiap pola cap bibir dianggap sebagai pattern dengan panjang 5 karakter. Pattern ini akan dicocokkan pada teks untuk mengetahui jenis pola dari cap bibir. Pattern dicocokkan menggunakan 3 algoritma pencocokan string yang ada pada dasar teori, kemudian dibandingkan mana yang lebih baik menurut jumlah perbandingan yang dilakukan.

Pembagian pattern sepanjang 5 karakter untuk masing-masing pola diasumsikan sebagai berikut :

- Type I : E8Y5x
- Type I' : E8Y9x
- Type II : yuGz1
- Type III : 1NuLL
- Type IV : utYzk
- Type V : E9y5x

Pada setiap analisis, semua pattern akan dicocokkan dengan teks sampai akhir kemudian dilihat pattern mana sajakah yang cocok. Dari pattern yang di dapat baru ditentukan jenis kelamin dari seseorang.

B. Contoh pengaplikasian

Diberikan contoh dua gambar cap bibir laki-laki dan perempuan sebagai berikut :

IV. KESIMPULAN

Algoritma brute force, KMP, dan Boyer-Moore dapat digunakan untuk menentukan pola cap bibir seseorang jika ditemukan suatu cara mengonversi cap bibir menjadi string yang standar. Dari hasil analisa contoh gambar cap bibir diatas menggunakan string acak diketahui bahwa cap bibir diatas adalah cap bibir perempuan.

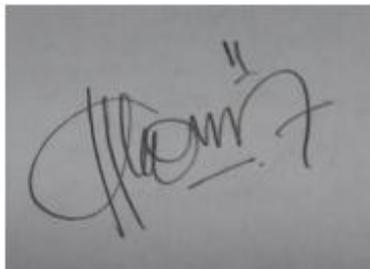
REFERENCES

- [1] Rohit M, Sumit G (2011) Cheiloscopy : A Deterministic Aid for Forensic Sex Determination. Journal of Indian academy of Oral Medicine and Radiology.
- [2] Dongawar et al (2013) Cheiloscopy-Method of Person Identification and Sex Determination. Open Access Scientific Reports
- [3] M.A.E. Domiaty et al. (2010) Morphological Patterns of lip prints in Saudi Arabiaat Almadiyah Almonawarah province. Forensic Science International.
- [4] M. Rinaldi, "Diktat Kuliah IF 2091 Struktur Diskrit", Program Studi Teknik Informatika, 2008, Bandung, Indonesia.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 17 Mei 2014

A photograph of a handwritten signature in black ink on a light-colored surface. The signature is stylized and appears to read 'Khoirunnisa Afifah'.

Khoirunnisa Afifah/13512077