

Fingerprint Recognition

Deadline 01.06.2008

Descriere

Recunoasterea pe baza de amprente este folosita la scara larga pentru identificarea infractorilor. Procedeu este simplu: de la locul savarsirii infractiunii se culeg amprente, se transforma in format electronic si se cauta intr-o baza de date continand amprentele unor persoane cu dosar penal sau ale unui grup de suspecti.

Ampreentele umane au un set de caracteristici care fac posibila identificarea cu o grad ridicat de probabilitate a unei persoane pe baza unei amprente.

Pentru verificarea potrivirii dintre doua amprente pot fi aplicate doua metode: prima bazata pe pattern-uri iar cea de-a doua pe minutia features. Vom considera in continuare un model simplificat pentru verificarea potrivirii bazat pe cea de-a doua metoda. In cadrul acestei metode o amprenta este reprezentata de o multime trasaturi de trei tipuri: linie terminata (line ending), bifurcatie de linii (line bifurcation) si linie scurta (short line). De asemenea fiecare trasatura are asociate un set de coordonate (x, y) , o orientare reprezentata prin valoarea unui unghi si o masura a calitatii amprente. Pentru simplificare, se considera ca o amprenta este reprezentata de o multime de trasaturi de acelasi tip ocupand diferite pozitii in plan.

Recunoasterea pe baza de amprente presupune potrivire iar potrivirea depinde de pozitie si rotatie. Daca pozitia unui punct este reprezentata relativ la un set de alte trei puncte, se obtine un invariant la rotatie si pozitionare. Dar pentru a obtine o pozitionare unica a unui punct in functie de alte trei puncte, cele trei puncte considerate nu trebuie sa fie coliniare. Toate amprente, cele stocate in baza de date si cele de cautat, sunt reprezentate atat ca multimi de puncte cu coordonate in plan dar si ca puncte cu coordonate relative la alte puncte. Problema consta in faptul ca amprente de cautat au pozitii, rotatii si numerotari diferite fata de cele stocate in baza de date. Compararea a doua amprente se realizeaza intr-o prima faza pe baza coordonatelor relative. Apoi, daca prima etapa s-a incheiat cu succes, rezultatul va fi confirmat sau infirmat transformand coordonatele relative ale punctelor din amprenta de cautat in coordonate in plan si verificand potrivirea exacta a punctelor in plan din cele doua amprente.

Sistemul propus de recunoastere a amprentelor are si functia "Partners in Crime". Exista posibilitatea ca la locul infractiunii sa se gaseasca mai multe amprente, apartinand diferitor persoane. In baza de date a politiei exista o asocierea 1 la 1 intre o amprenta si o persoana si de asemenea o persoana are asociat un grup de alte persoane (fiecare detinand o amprenta) impreuna cu care a mai savarsit fapte. Sistemul primeste la intrare o multime de amprente si le cauta pe rand in baza de date. Dupa ce a fost identificata o persoana se cunoaste grupul "Partners in Crime" asociat iar restul de amprente de cautat se compara mai intai cu amprentele persoanelor din acest grup. Daca s-a facut o noua identificare se procedeaza la fel ca mai sus. Daca nu, se cauta in continuare in cadrul bazei de date.

Cerinte

- Se dau o multime de persoane si amprente asociate precum si o multime de grupuri de infractiuni, "Partners in Crime". Fisierul in care sunt stocate aceste date (ex.: database1.in) respecta urmatorul format: pe prima linie sunt doi intregi reprezentand numarul total de amprente si numarul de persoane, in continuare pentru fiecare persoana (P_i) sunt date identificatorul grupului si numarul de amprente, pentru fiecare amprenta A_{ij} sunt date numarul de minutiaes, coordonatele carteziene (x, y) si coordonatele relative (r_1, r_2, r_3) pentru fiecare minutia. La sfarsitul fisierului se afla numarul de grupuri "Partners in Crime", pentru fiecare grup se dau numarul de persoane si multimea persoanelor.

- Se da de asemenea o multime de amprente de cautat (ex.: suspects1.in). Fisierul suspects1.in are urmatorul format: pe prima linie se afla numarul de amprente de cautat, in continuare pentru fiecare amprenta se dau numarul de minutiae-s impreuna cu coordonatele carteziane (x, y) pentru fiecare minutiae.

- Se cere gasirea persoanelor asociate cu amprentele de cautat. Fisierul de iesire (ex.: results1.out) arata in felul urmator: pentru fiecare amprenta i de cautat se va afisa "Fingerprint i is similar to fingerprint j from the database", unde j este indicele amprentei din baza de date. Daca amprenta i nu seamana cu nicio amprenta din baza de date atunci j va fi -1.

Testare

- Pentru a va testa tema puteti descarca un set de teste de [aici](#).

- Un generator de amprente si de grupuri "Partners in Crime" precum si o implementare seriala a solutiei se gasesc la acest [link](#) (ex.: executand "make test1" vor fi generate doua fisiere de intrare, database1.in si suspects1.in si un fisier de iesire, results1.out).

Indicatii

- O descriere simplificata a algoritmul de verificare a potrivirii a doua amprente arata astfel:

1) Pentru fiecare minutiae din amprenta de cautat se determina (r1, r2, r3), unde r1, r2, r3 sunt distantele catre cele mai apropiate minutiae-s din cadrul amprentei de cautat (vezi functia compute_r1r2r3 din fingerprint.c).

2) Utilizand (r1, r2, r3) se cauta amprente din baza de date similare cu amprenta de cautat (vezi compare_fingerprints din fingerprint.c).

3) Daca s-a gasit o amprenta similara se cauta trei minutiae-s necoliniare pentru care (r1, r2, r3) este unic (vezi find_fingerprint si find_fingerprint_using_mask). Altfel spus, se cauta trei asocieri 1 la 1 intre minutiae-s din amprenta de cautat si cea din baza de date. Daca nu exista, sari la pasul 2.

4) Cele trei minutiae-s au fiecare doua seturi de coordonate carteziane: cele din amprenta de cautat si cele din amprenta din baza de date. Astfel, coordonatele carteziane ale celorlalte minutiae-s din amprenta de cautat pot fi reprezentate in sistemul xoy al amprentei din baza de date (vezi compute_rrr_coordinates si compute_xy_coordinates). Pentru fiecare minutiae se determina distantele pana la cele trei minutiae-s de referinta folosind coordonatele carteziane din amprenta de cautat (compute_rrr_coordinates). Apoi, din coordonatele carteziane ale minutiae-s de referinta in cadrul amprentei din baza de date si cele trei distante rezulta (x, y) ale minutiae curente in planul amprentei din baza de date (compute_xy_coordinates).

5) Comparand coordonatele carteziane ale celor doua amprente, se poate decide cu un grad ridicat de probabilitate daca amprentele sunt identice sau nu.

- Ampretele din baza de date sunt impartite pe SPE-uri.

- Ampretele de cautat se cauta pe rand.

- Atunci cand un SPU a gasit potrivire, toate SPU-urile inceteaza sa mai caute.

- Folositi "Partners in Crime" pentru a optimiza cautarea.

Observatii

- Vor fi depunctate solutiile care folosesc busy-waiting, care nu folosesc instructiuni Single Instruction Multiple Data (SIMD) atunci cand este cazul si care utilizeaza ineficient resursele oferite de arhitectura Cell.
- Se acorda un bonus acelor solutii care vor considera ca o amprenta de cautat poate fi o submultime de puncte ale unei amprente din baza de date.
- Se acorda bonus solutiilor originale optimizate.