

Cvičení z předmětu Biometrie

cvičení 3: rozpoznání dynamického podpisu

Pavel Vostatek, vostapav@fel.cvut.cz

15. října 2012

Úvod

Úloha se zabývá praktickým seznámením se s dynamickým podpisem v biometrii a možností jeho rozpoznávání. K dispozici je připravený systém, který modeluje dynamický podpis pomocí gaussian mixture model (GMM). Úkolem bude nejdříve seznámit se s daným systémem a upravit systém tak, aby pro porovnání ohodnocení podobnosti podpisů využíval dynamic time warping (DTW). Pro ohodnocení výkonu algoritmů je k dispozici databáze podpisů SVC2004 (pro popis dat viz [3]).

Rozpoznání podpisu pomocí GMM.

Princip systému založeném na GMM se dá shrnout v následujících bodech:

- Z podpisu je extrahováno n příznaků $x_{1,t} \dots x_{n,t}$, každý snímaný v časech $t = 1 \dots T$.
- Při učení modelu podpisu je pomocí GMM odhadnuto pravděpodobnostní rozdělení zvlášť pro každý příznak $x_1 \dots x_n$ ve všech jeho instancích (časových okamžicích). Pokud je v trénovací množině více podpisů, jsou do odhadu p. rozdělení postupně řazeny vzorky daného příznaku od všech podpisů. Tímto získáme n pravděpodobnostních rozdělení $P_1 \dots P_n$.
- Pokud poté rozpoznáváme neznámý podpis, ohodnotíme jej hodnotící funkcí pomocí log-likelihood: $score = \sum_{i=1}^n \sum_{j=1}^T \ln[P_i(X = x_{i,j})]$
- *Score* je poté v trénovací části využito k nalezení vhodné hodnoty prahu pro klasifikaci podpisu na pravý a padělek.

Úloha je závislá na množství faktorů. Těmi jsou předzpracování signálu, volba konkrétních příznaků, počet gaussových křivek v GMM, počet podpisů použitý na tvorbu modelu. První část úlohy se bude týkat vyzkoušení nastavení různých parametrů a jejich vlivu na výslednou kvalitu rozpoznání podpisu.

Rozpoznání podpisu pomocí DTW.

Při použití míry podobnosti podpisů DTW nedochází k učení modelu, ale pouze uložení zvoleného množství podpisů (trénovací podpisy) do registru. Při rozpoznávání podpisu je poté neznámý podpis porovnáván s trénovacími podpisy pomocí míry DTW (viz přednášky nebo [2]). Algoritmus výpočtu podobnosti dvou (obecně n-rozměrných) vektorů pomocí DTW je následující:

- Máme dva podpisy s_1, s_2 , délky n , resp. m , u kterých jsou změřeny 3 (obecně libovolný počet) příznaky — 2 souřadnice polohy a přítlak.

$$s_1 = \begin{bmatrix} x_1, y_1, p_1 \\ x_2, y_2, p_2 \\ \dots \\ x_n, y_n, p_n \end{bmatrix}, s_2 = \begin{bmatrix} x_1, y_1, p_1 \\ x_2, y_2, p_2 \\ \dots \\ x_m, y_m, p_m \end{bmatrix}.$$

- Porovnání pomocí DTW je možné spočítat následujícím algoritmem: *i*) Je vytvořena matice D o rozměrech $(n + 1, m + 1)$, která je iniciována $D(1, 1) = 0, D(i, 1) = \inf, D(1, j) = \inf, i = 2 \dots n, j = 2 \dots m$. *ii*) Další pole v matici D jsou počítána následovně:

$$D(i, j) = \|s_1(i - 1, :) - s_2(j - 1, :)\| + \min \begin{cases} D(i, j - 1) \\ D(i - 1, j) \\ D(i - 1, j - 1) \end{cases},$$

$\|s_1(i - 1, :) - s_2(j - 1, :)\|$ označuje eukleidovskou vzdálenost mezi (i-1)-tým řádkem (vzorkem) podpisu s_1 a (j-1)-tým řádkem v podpisu s_2 , tedy $\sqrt{\sum_{d=1}^3 [s_1(i - 1, d) - s_2(j - 1, d)]^2}$.

- Vzájemná vzdálenost obou vektorů je na konci výpočtu naakumulována v $dist = D(n + 1, m + 1)$.

Po porovnání testovaného podpisu se všemi trénovacími podpisy spočítáme průměrnou vzdálenost a tuto vzdálenost použijeme jako míru hodnocení kvality podpisu.

Struktura rozpoznávacího systému.

Ke cvičení je přiložen částečně vytvořený systém s již implementovaným porovnáváním podpisu v Matlabu pomocí GMM a sadou testovacích podpisů. Celý systém je možné používat pomocí následujících funkcí:

- *load_data* — funkce k načítání podpisů tak, jak byly uloženy od počítače. Výstupem je cell-array s načtenými podpisy, např. $S = \{s_1, s_2, \dots\}$, kdy s_i odpovídá podpisu tak, jak je definován v předchozí kapitole o DTW. K jednotlivým podpisům se v Matlabu přistupuje následovně: $s_i = S\{i\}$.
- *preprocess* — funkce k předzpracování vstupních dat.
- *extract_features* — funkce k získání příznaků ze vstupních dat.
- *make_model* — funkce k naučení modelu.

- *score* — funkce pro výpočet ohodnocení rozpoznávaného podpisu.
- *ukazka* — funkce s ukázkovým kódem použití.

Všechny funkce i jejich použití jsou dokumentovány v nápovědě funkcí.

Úkoly

V systému rozpoznávání podpisu pomocí GMM vyzkoušejte vždy na několika náhodně vybraných subjektech. Model podpisu natrénujte na několika příkladech a poté zbytek pomocí modelu klasifikujte:

1. Jaký má vliv na výsledek klasifikace (equal error rate — EER, pomocí funkce *get_eer* resp. *get_eer_dtw*) normalizace vstupních příznaků (normalizaci implementujte — nulová střední hodnota příznaků a jednotkový rozptyl).
2. Zkuste přidat mezi příznaky (ve funkci *extract_features*) první derivaci původních příznaků (požijte přímo rozdíl $(x_{n,t+1} - x_{n,t})/dt$) a zhodnoťte vliv přidání těchto příznaků na EER.
3. Zkoumejte vliv počtu trénovacích podpisů na EER. [1. — 3. úkol celkem 8 b]
4. Dále ve funkcích *score* a *make_model* (případně pomocí vlastních externích funkcí) změňte systém tak, aby pro ohodnocení kvality podpisu využíval DTW. Algoritmus pro DTW implementujte a poté znovu otestujte úkoly 2 — 3. [12 b]

Bonusy:

1. Implementujte jinou metodu rozpoznávání podpisu např. z literatury (veliký přehled metod je v [1]). [3 b]
2. Zkuste vlastním falsifikátem vybraného podpisu z databáze prolomit rozpoznávací systém. [1 b]

Odevzdání

Z práce vypracujte stručný protokol a spolu se zdrojovými kódy implementovaných funkcí zazipujte do balíku a nahrajte do odevzdávacího systému courseware. Datum odevzdání je středa 17. 10. 2012, 23:59.

Reference

- [1] Donato Impedovo. Automatic signature verification: The state of the art. Ke stažení: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4603099 (dostupne pres dialog.cvut.cz).

- [2] J M Pascual-Gaspar. Practical on-line signature verification. Ke stazeni: <http://www.springerlink.com/content/n0x73333061702u4/> (dostupne pres dialog.cvut.cz).
- [3] SVC2004. Svc 2004: First international signature verification competition, detailed instructions for participants. Ke stazeni: <http://www.cse.ust.hk/svc2004/instructions.pdf>.