Materiał z książki *Paweł Zając, Stanisław Kwaśniowski (Red.), "Automatyczna identyfikacja w systemach logistycznych", Oficyna Wydawnicza Politechniki Wrocławskiej, 2004.* Kopiowanie i udostępnianie całości lub części materiału bez zgody autorów zabronione.

Rozdział 7.7. Biometria podpisu odręcznego

Adam Czajka i Andrzej Pacut

Instytut Automatyki i Informatyki Stosowanej
Politechnika Warszawska
oraz
Pracownia Biometrii
Naukowa i Akademicka Sieć Komputerowa NASK

{A.Czajka,A.Pacut}@ia.pw.edu.pl

1 Wstęp do biometrii podpisu

Podpis odręczny należy do biometrii zachowania (biometrii behawioralnych), a jego zastosowanie ma szereg zalet. Przede wszystkim, podpis jako znak identyfikujący ma bogatą tradycję, zatem jego stosowanie w systemach automatycznej weryfikacji budzi niewiele zastrzeżeń, inaczej niż np. dla odcisków palca. Stosunkowo łatwo podpis może być włączony w istniejące procedury weryfikacyjne, gdyż obecnie szereg procedur, jak np. procedury bankowe, wykorzystuje podpis. Kolejną ważną zaletą jest możliwość zmiany podpisu w przypadku jego fałszowania w systemach weryfikacji, podczas gdy zmiana odcisków palca czy wzoru tęczówki jest krańcowo trudna. Wykorzystanie podpisu ma jednak również wady. Należy do nich trudność tworzenia baz danych właśnie ze względu na powszechne traktowanie podpisu jako danych osobowych łatwo identyfikujących podpisującego. Do istotnych wad należy również mała dokładność systemów wykorzystujących podpis, w porównaniu z systemami wykorzystującymi np. tęczówkę czy odcisk palca.

Według Słownika Języka Polskiego [1] lub Najnowszego Słownika Języka Polskiego [2] *podpis* oznacza (1) *nazwisko (i imię) napisane własnoręcznie*, (2) *potwierdzenie pisma, nadanie mu ważności przez napisanie własnego nazwiska* lub (3) *podpisanie lub wyrażone pismem (rzadziej znakiem lub znakami) określenie tożsamości piszącego lub stawiającego znaki.* Z punktu widzenia biometrycznego rozpoznawania tożsamości, istotne jest znaczenie (3), które wprowadza akt podpisywania zachodzący w czasie, jak i znaczenie (2), które podkreśla możliwość określenia tożsamości piszącego. Dla jasności sformułowań tej pracy konieczne jest rozróżnienie znaczeń słowa podpis. *Podpis* w niniejszej pracy rozumiany jest jako zbiór *wszystkich możliwych realizacji podpisu*, otrzymywanych w wyniku wykonania aktu podpisywania się. Realizacje podpisu charakteryzują podpis w podobny sposób, w jaki próba statystyczna charakteryzuje populację generalną. Dodajmy, że jedna osoba może mieć kilka podpisów, np. "czytelny", "nieczytelny", "długi" (imię i nazwisko), "krótki" (tylko nazwisko), "inicjały", itd.

W niniejszej pracy założymy, że proces podpisywania może być obserwowany, w trakcie czego dokonywana jest rejestracja *składowych* realizacji podpisu. Należą do nich oczywiście współrzędne końca pióra, ale mogą to być dodatkowo – w zależności od urządzenia rejestrującego – również nacisk pióra na papier, fakt podniesienia pióra, kąty położenia pióra, siła i miejsce uchwycenia pióra etc. W rezultacie, obserwowana realizacja podpisu jest wielowymiarową funkcją

$$\mathbf{S} : \mathcal{T} \mapsto \mathbb{R}^M \tag{1}$$

gdzie M oznacza liczbę obserwowanych wielkości, zaś $\mathcal{T}$ jest zbiorem chwil obserwacji. Wykres tej funkcji jest krzywą w przestrzeni $M+1$-wymiarowej, gdzie dodatkową współrzędną jest czas podpisywania $\mathcal{T}$. Wykres ten wygodnie jest przedstawić w postaci M jednowymiarowych wykresów składowych $S_i, i = 1, \ldots, M$. Dla wygody będziemy tu utożsamiać samą funkcję $\mathbf{S}$ z jej wykresem.

Ta ogólna reprezentacja pozwala łatwo scharakteryzować inne możliwe sposoby zobrazowania realizacji podpisu. W szczególności, rysunek wykonywany przez koniec pióra podczas podpisywania (podpis na papierze) jest identyczny z rzutem wykresu funkcji $\mathbf{S}$ na podprzestrzeń wyznaczoną przez czas i płaszczyznę podpisywania $X \times Y$. Rzut ten jest oczywiście krzywą w przestrzeni trójwymiarowej. Zwróćmy jeszcze uwagę, że rzutowanie wykresu na płaszczyznę $X \times Y$, w wyniku którego "pozbywamy się" czasu daje w wyniku obraz podpisu na płaszczyźnie podpisywania, tzn. podpis "na kartce papieru" zwany również *podpisem skanowanym*.

Ze względu na własności urządzeń pomiarowych, chwile dokonywania pomiarów tworzą zbiór dyskretny $\mathcal{T}$. W rezultacie, wykres realizacji podpisu sprowadza się do ciągu $M+1$-elementowych wektorów obserwacji $s(i)$ o elementach

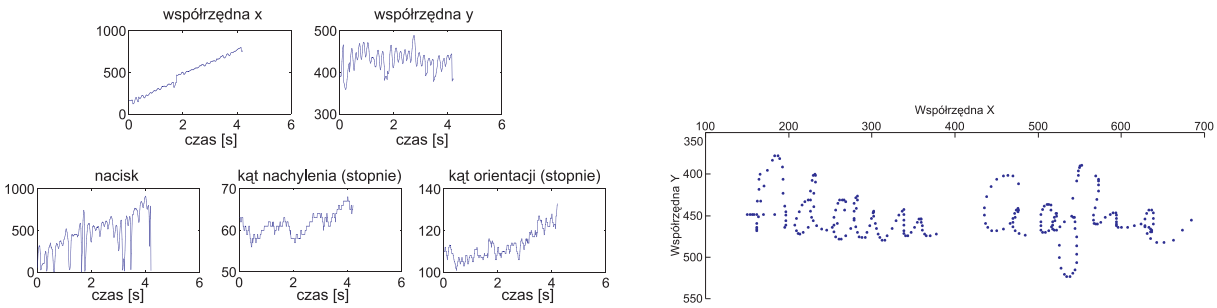
$$t(i), s_1(i), \ldots, s_M(i) \tag{2}$$

dla $i = 1, \ldots, N$, gdzie N oznacza liczność $\mathcal{T}$ dla danej realizacji. Ciąg obserwacji wygodnie zapisać w postaci macierzy o rozmiarze $M + 1 \times N$

$$\mathbf{S} = \begin{bmatrix} t(1) & s_1(1) & \cdots & s_M(1) \\ t(2) & s_1(2) & \cdots & s_M(2) \\ \vdots & \vdots & \vdots & \vdots \\ t(N) & s_1(N) & \cdots & s_M(N) \end{bmatrix} \tag{3}$$

którą dla wygody oznaczać będziemy również przez $\mathbf{S}$. Zauważmy, że rzutowanie na odpowiednie płaszczyzny odpowiada "usuwaniu" kolumn macierzy obserwacji. W szczególności, eliminacja czasu odpowiada eliminacji pierwszej kolumny macierzy, tzn. obraz realizacji podpisu "na papierze" jest reprezentowany przez drugą i trzecią kolumnę (rys. 1). Dodajmy, że "eliminacja" czasu pozbawia nas możliwości przyporządkowywania punktom wykresu wartości czasu.

Przypomnijmy, że przez *podpis* rozumiemy zbiór wszystkich możliwych realizacji podpisu. Podzbiorem tego zbioru będzie *podpis referencyjny*, rozumiany jako zbiór (jednej lub kilku) *realizacji referencyjnych* tego samego podpisu, względem których należy oceniać inne realizacje tego podpisu. Przez *podpis testujący* rozumiemy zbiór realizacji podpisu nie zawartych w podpisie referencyjnym, a wykorzystywany do badania metody rozpoznawania tożsamości. Przez *podpis sfałszowany* będziemy rozumieć zbiór realizacji podpisu, wykonanych przez inna osobę w celu imitacji danego podpisu.



Rysunek 1: Wykresy składowych realizacji podpisu (z lewej) i obraz podpisu na papierze (podpis skanowany) (z prawej)

W pracy rozważamy zagadnienia identyfikacji i zagadnienia weryfikacji tożsamości. W zadaniach *weryfikacji tożsamości* oceniamy przynależność danej realizacji do zadanego podpisu, natomiast w zadaniach *identyfikacji tożsamości* znajdujemy podpis, do którego należy dana realizacja. Dla porównania wyników weryfikacji stosujemy zwykle wskaźnik fałszywych odrzuceń FRR (ang. *False Rejection Rate*), który określa stopień błędnego odrzucania przez system podpisów właściwych, oraz wskaźnik fałszywej akceptacji FAR (ang. *False Acceptance Rate*) określający stopień błędnej akceptacji podpisów sfałszowanych. Błąd, odpowiadający punktowi przecięcia wykresów FAR oraz FRR w funkcji parametru (zwanego zwykle progiem) oznaczany jest jako wskaźnik równego błędu ERR (ang. *Equal Error Rate*). Dla systemów klasyfikacji stosujemy wskaźnik poprawnej klasyfikacji CCR (ang. *Correct Classification Rate*) określający stopień poprawnej klasyfikacji podpisów.

2 Istniejące rozwiązania

Metodyka biometrii podpisu ma już bogatą literaturę. Przegląd metod biometrii, a w szczególności technik opartych na podpisie odręcznym, znaleźć można np. w monografii [3]. Tutaj przedstawimy jedynie niektóre rozwiązania, ważne dla rozwoju technik bazujących na podpisie. Metody podzielimy na dwie grupy, związane z doborem wielkości obserwowanych. Do pierwszej należą *metody statyczne*, tzn. metody oparte na analizie obrazu podpisu na papierze (podpisu skanowanego), zaś do drugiej – *metody dynamiczne*, wykorzystujące informację czasową i wymagające rejestracji procesu podpisywania.

W pracy G. Dimauro [4] zaprezentowano system weryfikacji podpisów składanych na czekach bankowych. Pierwszym krokiem jest lokalizacja skanowanego podpisu w obrębie obrazu, następnie uniezależnienie obrazu od orientacji i skali. Każdy obraz dzielony jest następnie na spójne segmenty. Weryfikacja przeprowadzana jest na bazie 6 globalnych cech (wyliczanych dla całego obrazu) i 16 cech wyliczanych w obrębie każdego z segmentów. System rozpoznający składa się z trzech "ekspertów" "głosujących" za lub przeciw przyjęciu hipotezy o oryginalności złożonego podpisu. Rezultatem jest FRR = 22% oraz FAR = 3.9%. W pracy S. Lee i J. C. Pan [5] zaprezentowano metodę śledzenia procesu podpisywania na podstawie skanowanego podpisu. Pierwszym krokiem jest zamiana obrazu realizacji podpisu na obraz szkieletowy o szerokości jednego punktu, przekształcony

następnie na sekwencje krzywych, w obrębie których znajdowane sa punkty charakterystyczne, takie jak maksima, minima, punkty początkowe i końcowe. Empiryczna macierz kowariancji powstała na bazie punktów charakterystycznych jest wykorzystywana podczas normalizacji każdego z podpisów (uniezależnienie od obrotu, przesunięcia oraz skali).

Podejścia dynamiczne klasyfikować można na bazie wymiarowości obserwacji. Brault i Plamondon [6] zaproponowali model wykorzystujący tylko pomiary położenia końcówki pióra, oparty na koncepcji systemu nerwowo-mięśniowego. Fałszowanie podpisu jest tutaj modelowane jako sekwencja zadań, w obrębie których imitowany jest element fundamentalny podpisu, definiowany jako sekwencja trzech pociągnięć pióra: krzywoliniowe - ostrokątne - krzywoliniowe. Na każde zadanie składa się przestrzenna percepcja celu zadania (układ nerwowo-wzrokowy), jego przygotowanie (mózg) oraz jego wykonanie (układ nerwowo-mięśniowy). W pracy Lee [7] zaprezentowano system bazujący na indywidualnych cechach każdego położenia pióra (10-15 cech), dających ich najlepsze zróżnicowanie w przestrzeni cech. Cechy indywidualne wybierane sa ze zbioru 40 cech wspólnych dla wszystkich podpisów, opisujących zarówno ich właściwości statyczne jak i dynamiczne. Weryfikacja wykonywana jest przez "głosujących ekspertów".

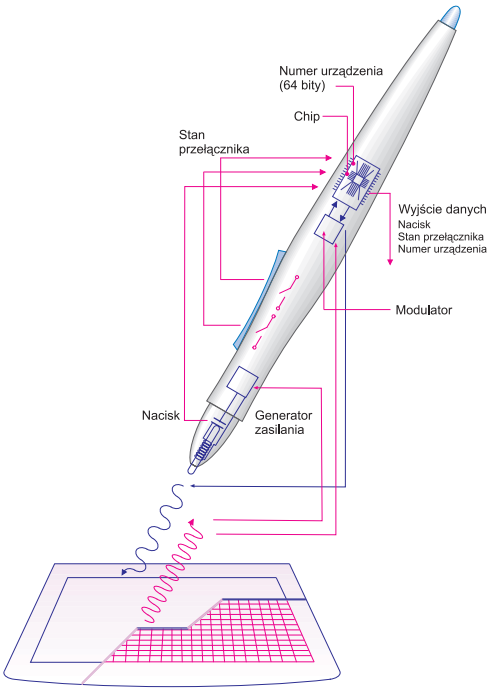
Crane i Ostrem [8] wykorzystali pomiary położenia pióra i jego nacisku na powierzchnię do konstrukcji systemu bazującego na wyborze indywidualnych cech spośród 44 cech wspólnych. Optymalny wybór cech indywidualnych dokonuje się przez minimalizację ERR przeprowadzaną dla każdego podpisu. Rozwiązanie, choć prowadzące do dobrych rezultatów (ERR=0.5%) jest dość kłopotliwe, gdyż wymaga dużej liczby realizacji (około 100 dla jednego podpisu) do prawidłowego rozpoznania najlepszego (w sensie minimalizacji ERR) zestawu cech. Komercyjne systemy weryfikacji podpisów, produkcji firm SoftPro i Cybersign, również bazują na obserwacjach trzech składników podpisu w czasie.

Największą wiedzę o procesie podpisywania otrzymujemy obserwując dodatkowo również orientację pióra w przestrzeni. Tego typu pięciowymiarowe obserwacje zastosowano w pracy Wessels i Omlin [9], w której zaproponowano system hybrydowy składający się z ukrytych modeli Markowa oraz samoorganizujących się sieci Kohonena. Dla FRR=0%, system zaakceptował 13% fałszywych podpisów (FAR=13%).

3 Stanowisko badawcze i rejestracja podpisów

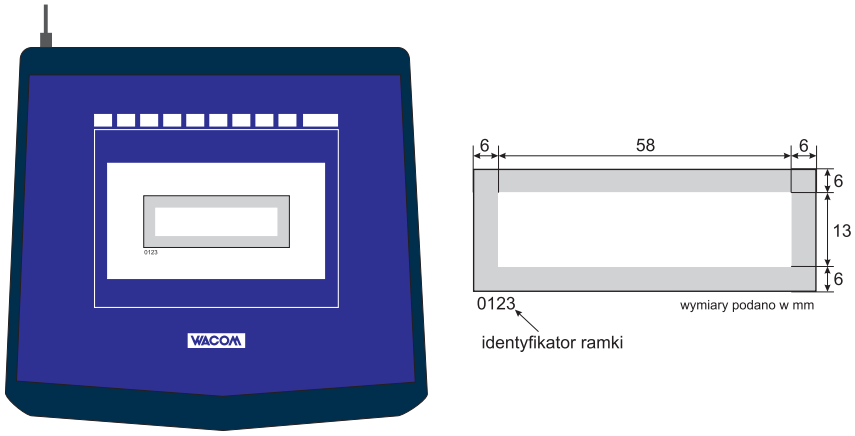
Składowymi podpisu, obserwowanymi w badaniach autorów tej pracy, były współrzędne x, y końca pióra, nacisk p pióra na papier (fakt podniesienia pióra sygnalizowany był przez $p = 0$), oraz dwa kąty (orientacja i odchylenie od pionu) określające położenie pióra. Realizacje podpisu rejestrowane są za pomocą tabletu graficznego WACOM oraz piszącego, bezprzewodowego pióra (rys. 2). Urządzenie to umożliwia pomiary położenia, nacisku i orientacji przestrzennej pióra z częstotliwością 100Hz. Położenie i orientacja pióra mierzone są również dla pióra uniesionego do wysokości ok. 1 cm nad powierzchnia tabletu, co pozwala na rejestrację trajektorii podpisu nawet w momencie odrywania końcówki pióra od kartki.

Na tablecie umieszczono specjalną ramkę, tworząc w ten sposób *panel podpisu* (rys. 3). Wymiary ramki zbliżone są do wymiarów pól podpisu używanych przy potwierdzaniu transakcji dokonywanych karta płatniczą. Biały prostokąt wewnątrz większego szarego



Rysunek 2: Schemat działania tabletu graficznego WACOM wyposażonego w bezprzewodowe pióro

oraz czarna ramka obejmująca oba prostokąty sugerują tylko miejsce złożenia podpisu. Każdy mógł jednak złożyć podpis w dowolnym miejscu obszaru aktywnego tabletu. Kilku uczestników złożyło podpisy, które wyraźnie wykraczały poza czarna ramkę. Przy każdej rejestracji osoba podpisująca się na kilka sposobów deklarowała rodzaj podpisu (np. tylko nazwisko, imię + nazwisko lub inicjały) i każdy wariant był traktowany jako odrębny podpis tej samej osoby.



Rysunek 3: Panel podpisu: tablet (z lewej) oraz ramka podpisu (z prawej)

Chociaż proces podpisywania zależny jest od osobowości, przyzwyczajeń, czy tez podświadomego wyćwiczenia pewnych ruchów realizowanych przez układ nerwowo-mięśniowy,

kolejne realizacje podpisu różnią się od siebie ze względu na warunki podpisywania. Aby charakteryzacja podpisu poprzez jego zarejestrowane obserwacje była jak najpełniejsza, realizacje te powinny być gromadzone w ciągu możliwie długiego okresu i być pobierane w różnych sytuacjach. Kolejne realizacje podpisu zbierano więc w różnych warunkach, uwzględniających fizyczne zmęczenie, stres, pośpiech, radość, odprężenie, itp. Dzięki temu, proces podpisywania był lepiej eksplorowany, prowadząc do lepszej tolerancji systemu na naturalne modyfikacje realizacji podpisu. Zbytnie zwiększenie wariacji realizacji, uwzględniające rzadko występujące warunki, może jednak niewłaściwie reprezentować podpis i w efekcie prowadzić do zwiększenia poziomu fałszywych akceptacji.

System weryfikacji podpisu powinien być przygotowany na ograniczoną dostępność realizacji, wykorzystując np. tylko 3 do 6 realizacji referencyjnych jednego podpisu. Może się zdarzyć, że osoba podpisuje się na kilka sposobów (czyli odpowiadają jej kilka różnych podpisów, nie realizacje podpisu). W takim przypadku definiujemy kilka różnych podpisów dla tej samej osoby. Obecnie baza podpisów zwiera podpisy ok. 120 osób (ponad 800 realizacji), część eksperymentów dyskutowana w tej pracy wykonana była jednak dla mniejszej (tworzonej jeszcze) bazy 48 podpisów. Na potrzeby badań systemów rozpoznawania, baza danych podzielona została na zbiór uczący, składający się z podpisów referencyjnych (ok. 6 realizacji dla każdego podpisu) i zbiór testowy (1 realizacja dla każdego podpisu). Dodatkowo, utworzono 9 podpisów sfałszowanych. Do fałszowania podpisów poproszono kilka osób, które mogły przygotować się do ostatecznej próby fałszowania przez dowolny okres czasu (nawet kilka dni) dysponując jednak tylko bazą podpisów skanowanych (podpisów "na papierze") i wybierając jeden lub dwa podpisy, które z ich punktu widzenia były najłatwiejsze do imitacji.

4 Dobór cech podpisu

Konstrukcja cech podpisu leży u podstaw sukcesu lub porażki metody. Oryginalne dane pomiarowe zwykle są wstępnie przetwarzane w celu lepszego wydobycia cech umożliwiających weryfikację piszącego. Do takich przekształceń należy filtracja pomiarów, określanie dodatkowych cech takich jak lokalna krzywizna, usuwanie trendu etc. Poniżej omówimy prostą procedurę doboru cech, zastosowaną z sukcesem w jednym z systemów weryfikacji podpisu opracowanych przez autorów tej pracy.

4.1 Zestaw cech globalnych

Konstrukcję cech rozpoczęto od wyodrębnienia podstawowych cech globalnych, do których należały

- Długość podpisu, rozumiana jako liczba punktów N realizacji podpisu
- Wartości średnie i empiryczne odchylenia standardowe składowych s_m, $m = 1, \ldots, 5$,

realizacji podpisu, tzn. odpowiednich kolumn macierzy obserwacji, czyli

$$\overline{s}_m = \frac{1}{N}\sum_{k=1}^{N} s_m(k) \tag{4}$$

$$\sigma_m = \sqrt{\overline{s_m^2} - \overline{s_m}^2} \tag{5}$$

- Współczynniki trendu składowych realizacji, tzn. współczynniki nachylenia α_m prostej regresji dla każdego m-tego składnika realizacji podpisu, traktowanego jako funkcja czasu

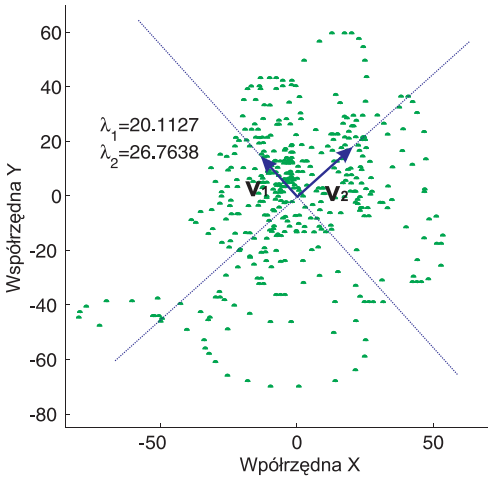
$$s'_m(k) = \overline{s}_m - k\,\alpha_m, \quad m = 1, \ldots, M \tag{6}$$

gdzie

$$\alpha_m = 12\,\frac{\overline{\overline{s}}_m - \frac{1}{2}(N+1)\overline{s}_m}{N^2-1}, \quad \overline{\overline{s}}_m = \frac{1}{N}\sum_{k=1}^{N} k\, s_m(k) \tag{7}$$

Powyższe wielkości posłużyły do transformacji oryginalnych składowych do postaci z usuniętym trendem (rys. 4), tzn.

$$s^*_m(k) = s_m(k) - s'_m(k), \quad k = 1, \ldots, N, \quad m = 1, \ldots, M \tag{8}$$



Rysunek 4: Realizacja z rys. 1 zrzutowana na płaszczyznę współrzędnych końca pióra po usunięciu trendu liniowego oraz wektory własne empirycznej macierzy kowariancji tych składowych

Dla składowych z usuniętym trendem wyodrębniono kolejne cechy globalne:

- Wartości i wektory własne empirycznej macierzy kowariancji Γ składowych realizacji podpisu, tzn. *macierzy inercji* o elementach

$$\gamma_{i,j} = \overline{s^*_i s^*_j} \tag{9}$$

Jako cechy podpisów używane są pierwiastki kwadratowe wartości własnych $\lambda_1 \geqslant \lambda_2$ macierzy Γ (rys. 4).

- Trzecie momenty centralne położenia (s_1^*, s_2^*) i nacisku (s_3^*), mianowicie

$$\beta^3_{p,q,r} = \frac{1}{N}\sum_{k=1}^{N}\left[s_1^*(k)\right]^p\left[s_2^*(k)\right]^q\left[s_3^*(k)\right]^r \tag{10}$$

  gdzie $p+q+r = 3$, $p, q, r = 0, 1, 2, 3$, które związane są z orientacją krzywej w trójwymiarowej przestrzeni obserwacji. Dla właściwego skalowania stosowano pierwiastki sześcienne trzecich momentów $(\beta_{p,q,r})$.

Zauważmy, że tylko cztery cechy (wartości średnie oraz odchylenia standardowe współrzędnych końca pióra) mogą być oszacowane na podstawie obserwacji podpisów skanowanych. Te cechy nazywać będziemy *cechami widocznymi*. Cechy widoczne będą wykorzystywane do grupowania podpisów w skupienia. Obliczenie pozostałych cech, zwanych dalej *cechami niewidocznymi*, wymaga obserwacji procesu podpisywania, co znacznie utrudnia ich oszacowanie przez osobę fałszującą podpis.

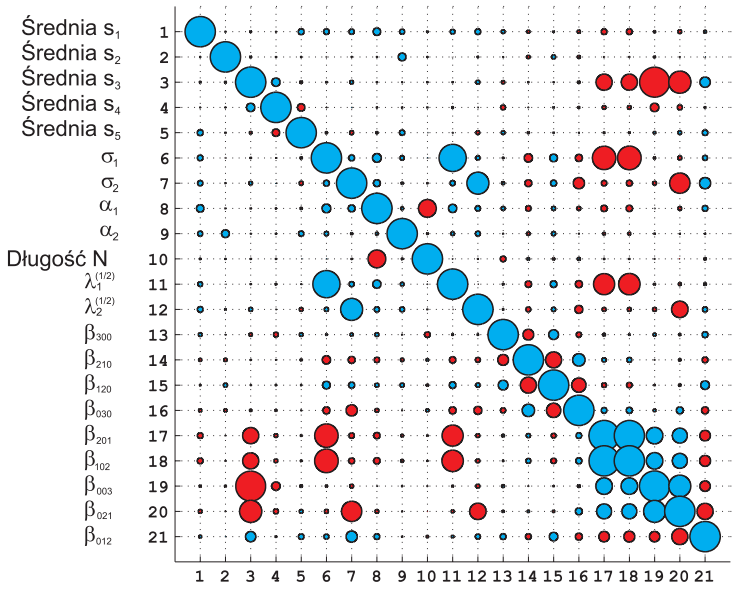
4.2 Usuwanie cech skorelowanych

Pierwotny zestaw cech może oczywiście zostać rozszerzony o cechy opisujące podpis bardziej szczegółowo, np. o cechy reprezentujące lokalne właściwości dynamiczne podpisu. Dla zadanego rozmiaru bazy danych, zbytnie zwiększanie wymiarowości przestrzeni cech powoduje obniżenie jakości działania systemu ze względu na trudności w estymacji funkcji klasyfikującej. Należy więc uwzględniać głównie takie cechy, które wnoszą dodatkową informację. Zadanie to może być częściowo zrealizowane poprzez odrzucenie cech silnie skorelowanych.

Pierwotny zestaw cech w istocie zawiera cechy liniowo skorelowane (rys. 5). Ostateczny zbiór cech powstał przez odrzucenie cech, dla których wartość bezwzględna współczynnika korelacji przekraczała arbitralnie przyjęty próg, wynoszący 0.4. Ze względu na niejednoznaczność wyboru cech "silnie skorelowanych", możemy odrzucać te cechy, których obliczanie jest bardziej złożone. Dla przykładu, wartości własne macierzy inercji oraz odchylenia standardowe odpowiednich składników wykazują korelacje na poziomie 0.9. W tym przypadku zrezygnowaliśmy z wartości własnych macierzy. Procedura ta doprowadziła do finalnego zbioru 12 cech. Powstały zbiór cech używany był zarówno w procesach klasyfikacji jak i weryfikacji podpisów.

4.3 Analiza skupień cech

Rozkład realizacji cech podpisów może okazać się niejednorodny. Część realizacji podpisów może leżeć blisko siebie, wykazując podobieństwa dla pewnego podzbioru cech, inne zaś będą różnić się we wszystkich wymiarach. Funkcja klasyfikująca znajdowana jest na podstawie analizy przestrzeni cech, zatem niejednorodność rozmieszczenia obiektów może prowadzić do złej aproksymacji funkcji klasyfikującej. Z tego powodu istotnym elementem systemu weryfikacji podpisów jest analiza skupień podpisów. Dzieląc przestrzeń cech na podobszary o zbliżonych wartościach cech otrzymujemy klasy podpisów o podobnych właściwościach. W rezultacie tworzona jest rodzina kilku funkcji klasyfikujących, które są aproksymowane w obrębie podobszarów, co zmniejsza błąd oszacowania tych funkcji.



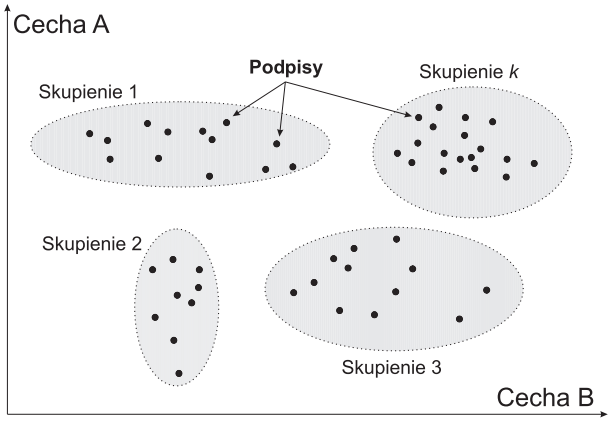
Rysunek 5: Empiryczna macierz korelacji pierwotnego zestawu cech

Ważnym elementem analizy skupień jest wybór podzbiorów cech, na bazie których dokonywana jest klasyfikacja bądź weryfikacja, nie muszą to być jednak te same podzbiory. Najlepsze wyniki weryfikacji otrzymaliśmy przy podziale na grupy bazującym wyłącznie na cechach widocznych, a następnie analizy dyskryminacyjnej w obrębie każdej grupy na bazie cech niewidocznych. Takie postępowanie jest intuicyjnie zrozumiałe: tylko cechy widoczne są możliwe do fałszowania poprzez imitację podpisu widniejącego na papierze. W rezultacie otrzymuje się dobrą aproksymację funkcji klasyfikującej określonej na podprzestrzeni cech niewidocznych w obrębie każdego podpisu.

Algorytm oparty na *drzewach rozpinających* oraz algorytm wykorzystujący *hierarchiczne drzewa skupień* były przez nas wykorzystywane z dobrymi wynikami. W obu przypadkach do wyznaczenia skupień podpisów liczony jest centroid podpisu na podstawie wszystkich składających się na niego realizacji.

Algorytm oparty na drzewach rozpinających konstruuje drzewo rozpinające nieskierowany graf pełny, którego węzłami sa centroidy podpisów a długości łuków odpowiadają odległości (euklidesowej) pomiędzy punktami. Przecięcie $k-1$ najdłuższych łuków w tak powstałym drzewie prowadzi do k podzbiorów podpisów podobnych w wybranej podprzestrzeni cech. Wybranie arbitralnego progu oznaczającego największą dopuszczalną odległość pomiędzy centroidami podpisów oraz przecięcie wszystkich łuków o długości większej od wybranego progu również prowadzi do poddrzew rozpinających, jednakże ich liczba nie może być z góry określona i strategia postępowania pozostaje w gestii projektanta systemu. W naszej pracy stosowaliśmy pierwszą strategię (rys. 7), a przestrzeń cech grupowano w skupienia jedynie na podstawie podprzestrzeni cech widocznych (wartości średnie położenia oraz empiryczne odchylenia standardowe, tzn. $\overline{s}_1, \overline{s}_2, \sigma_1, \sigma_2$). W rezultacie, w procesie weryfikacji bazę podpisów podzielono na cztery skupienia, a w procesie klasyfikacji na sześć skupień.

Drugi algorytm, dający w rozpatrywanym zagadnieniu lepsze rezultaty aproksymacji funkcji klasyfikującej, polega na budowie binarnego hierarchicznego drzewa na bazie

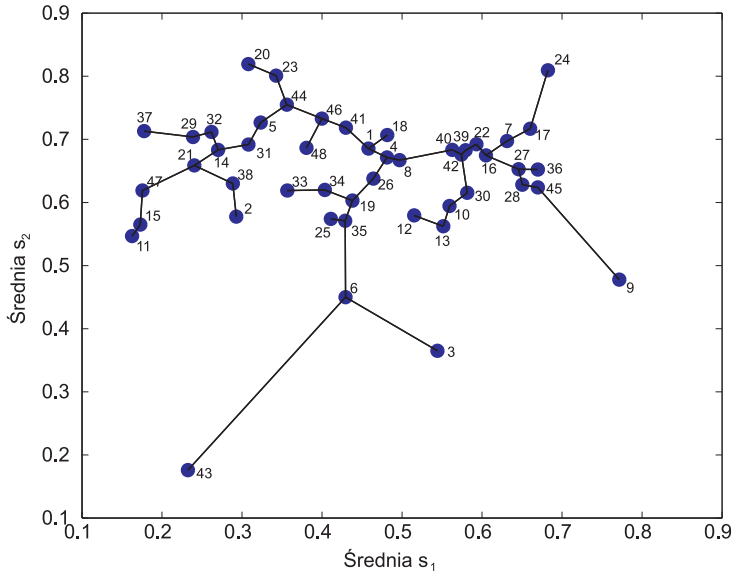


Rysunek 6: Skupienia w przestrzeni cech grupują "podobne" podpisy – w sensie odległości punktów w przestrzeni. Każdy punkt reprezentujący podpis jest średnią cech realizacji referencyjnych.

macierzy odległości. Liście drzewa reprezentują podpisy, stanowiące punkt wyjścia do konstrukcji skupień. Elementy łączone sa parami przy użyciu algorytmu najbliższego sąsiada. Para zgrupowanych podpisów tworzy obiekt, który może zostać dołączony do innego obiektu w dalszych iteracjach algorytmu, tworząc następne skupienie podpisów (rys. 8). Macierz podobieństwa (odległości) podpisów rzutowanych na przestrzeń cech wykorzystuje odległość Mahalanobisa $d(x,y) = \sqrt{(x-y)^T C^{-1}(x-y)}$ gdzie C jest empiryczną macierzą kowariancji cech podpisów. Stosowaliśmy tutaj dwie metody podziału na poddrzewa. W pierwszej z nich, połączenia przekraczające pewną najdłuższą dopuszczalna odległość pomiędzy obiektami zostają przecięte. Powstałe w ten sposób poddrzewa reprezentują kolejne skupienia obiektów. Inny sposób wykorzystuje *współczynnik niespójności* κ (ang. *inconsistent coefficient*), charakteryzujący każde połączenie w dendrogramie poprzez porównanie jego długości ze średnią długością połączeń w poddrzewie znajdującym się bezpośrednio pod danym łukiem, a mianowicie

$$\kappa = \left| \frac{l_k - \overline{l}_k}{\sigma_k} \right| \tag{11}$$

gdzie l_k oznacza długość rozważanego łuku o indeksie k, $\overline{l}_k$ oznacza średnią długość wybranych połączeń w poddrzewie o korzeniu w węźle o indeksie k, natomiast σ_k oznacza odchylenie standardowe wybranych połączeń w poddrzewie pod łukiem k, przy czym połączeniom zakończonym liśćmi drzewa przypisana jest wartość $\kappa = 0$. Algorytm przecina wszystkie łuki wykraczające poza zadany maksymalny współczynnik niespójności, co prowadzi do tworzenia poddrzew binarnych reprezentujących skupienia podpisów.



Rysunek 7: Drzewo rozpinające w przestrzeni dwóch cech (wartości średnie położenia pióra $\overline{s}_1, \overline{s}_2$)

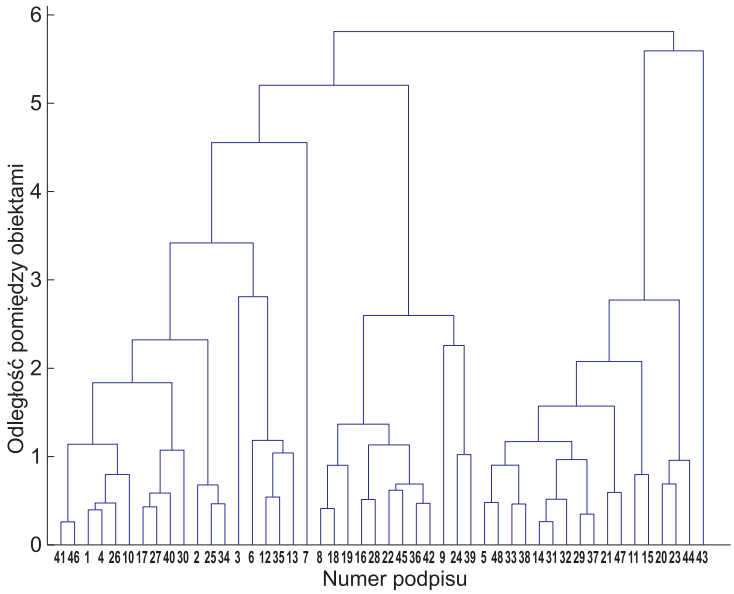
5 Klasyfikacja podpisów

5.1 Zastosowanie wielowarstwowego perceptronu

Jak wiadomo, perceptron dwuwarstwowy o odpowiednio wielkiej liczbie neuronów w nieliniowej warstwie ukrytej może dowolnie dokładnie aproksymować dowolną funkcję, w tym przypadku funkcję decyzyjną. W obu warstwach stosowaliśmy neurony o sigmoidalnych funkcjach aktywacji. Dla każdego skupienia, tak w zadaniu weryfikacji jak i identyfikacji, tworzymy odrębny klasyfikator o C wejściach, gdzie C jest liczbą cech wybranych do danego zadania (dla weryfikacji jest to 8 cech niewidocznych, dla zadania identyfikacji wykorzystano wszystkie 12 cech ze zbioru finalnego). Liczba neuronów ukrytych zależy od C; zwykle było to $2C$ neuronów. Każda sieć posiada L neuronów wyjściowych, gdzie L jest liczbą podpisów w skupieniu przypisanym do danego klasyfikatora. Do treningu sieci używaliśmy podpisy referencyjne odpowiednie dla każdego skupienia, natomiast do testowania wykorzystaliśmy realizacje ze zbioru testowego. W przypadku idealnym wyjście tylko jednego neuronu powinno być bliskie jedności. Neuron ten reprezentuje podpis, którego realizacją pobudzono sieć, natomiast wyjścia pozostałych neuronów są bliskie zeru.

Dla potrzeb identyfikacji, podpisy podzielone zostały na 6 grup (skupień) z wykorzystaniem zestawu 12 cech. W obrębie każdej grupy utworzony został odrębny klasyfikator neuronowy, stąd wymagana była wstępna wiedza *a priori* o przynależności danej realizacji do skupienia. Brak takiej wiedzy można zastąpić identyfikacją w obrębie każdej grupy, a następnie wyborem najlepszego wyniku. Każda realizacja klasyfikowana była zgodnie z zasadą *winner takes all*, czyli wynikiem identyfikacji był indeks neuronu wyjściowego osiągający największą wartość wyjścia dla danej realizacji wejściowej.

Weryfikacja tożsamości bazująca jedynie na indeksie neuronu wyjściowego o najwięk-



Rysunek 8: Drzewo hierarchiczne (dendrogram) podpisów dla 12 cech podpisu

szej wartości wyjścia – czyli z wykorzystaniem techniki *winner takes all* — może nie być wystarczająco skuteczna. W zadaniu tym pomocna jest również analiza wartości wyjścia zwycięskiego neuronu w odniesieniu do wyjść innych neuronów, jak również porównanie tej wartości z wykorzystaniem statystycznych własności realizacji referencyjnych. Oznaczmy przez $[m]$ indeks neuronu m-tego co do wartości wyjścia, licząc od największego. Niech c będzie odpowiedzią sieci na badany podpis i niech $\overline{c}_i^{[m]}$ i $\overline{\overline{c}}_i^{[m]}$ oznaczają, odpowiednio, średnią i odchylenie standardowe $[m]$-tego neuronu obliczoną dla i-tego podpisu referencyjnego. Wielkość

$$h_i^{[m]} = \frac{c^{[m]}(f) - \overline{c}_i^{[m]}}{\overline{\overline{c}}_i^{[m]}} \tag{12}$$

może być wykorzystana do utworzenia dodatkowych statystyk testujących. Dodatkowym warunkiem akceptacji realizacji podpisu może być

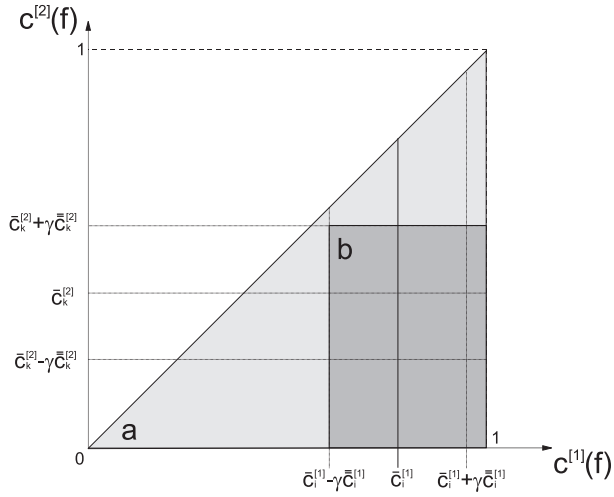
$$h_i^{[1]} > -\gamma \tag{13}$$

gdzie γ jest parametrem określanym np. na bazie nierówności Czebyszewa lub wyznaczonym doświadczalnie. Uwzględnienie drugiego co do wartości wyjścia neuronu prowadzi do dodatkowego warunkiem akceptacji realizacji podpisu (rys. 9)

$$h_i^{[2]} < \gamma \tag{14}$$

5.2 Zastosowanie sieci radialnej RCE

Sieć RCE [10] jest siecią samoorganizującą o radialnych funkcjach aktywacji, która może rozwiązywać zadania nieliniowej klasyfikacji. Sieć tę stosowaliśmy zarówno do zadań weryfikacji jak i zadań identyfikacji podpisów.



Rysunek 9: Obszary akceptacji realizacji podpisu dla dwóch scenariuszy: obszar a odpowiada zastosowaniu metody *winner takes all*, obszar b prezentuje technikę akceptacji podpisów przy uwzględnieniu dodatkowych warunków na wartości wyjść neuronów warstwy wyjściowej(14)

Sieć RCE składa się z jednej warstwy neuronów ukrytych o radialnej funkcji aktywacji, oraz z warstwy wyjściowej. Wyjście każdego neuronu warstwy ukrytej połączone jest wyłącznie z wejściem jednego neuronu wyjściowego. Wyjściowy neuron o indeksie k odpowiada k-tej klasie, zaś wartość jego wyjścia estymuje prawdopodobieństwo przynależności wzorca wejściowego do klasy k. Zbiór neuronów ukrytych, pusty przed rozpoczęciem procesem uczenia, jest uzupełniany na skutek błędnych klasyfikacji. Każdy neuron ukryty charakteryzowany jest przez trójkę (ω, η, λ), gdzie ω oznacza wektor wag, η odpowiada liczbie wzorców uczących poprawnie klasyfikowanych przez neuron, natomiast λ definiuje pole przyciągania neuronu, rozumiane jako kula o środku ω i promieniu λ.

Nowo dodany neuron ukryty połączony zostaje z neuronem wyjściowym reprezentującym klasę wzorca wejściowego, jego wektorowi wag zostaje przypisana wartość wzorca wejściowego, licznikowi wzorców η zostaje przypisana wartość 1, natomiast promień λ otrzymuje wartość maksymalną $\lambda_{\max}$. W odpowiedzi na wzorzec wejściowy, każdy neuron ukryty oblicza odległość wzorca wejściowego od wektorem wag. Jeśli wzorzec znajduje się w zasięgu neuronu ukrytego, zwiększana jest odpowiadająca mu wartość licznika η. Jeśli żaden z istniejących neuronów nie potrafi poprawnie sklasyfikować wzorca, dodawany jest neuronu ukryty, połączony z odpowiednim neuronem wyjściowym reprezentującym klasę wzorca wejściowego. Jeśli wzorzec wejściowy jest w zasięgu kilku neuronów ukrytych połączonych z różnymi neuronami wyjściowymi, co prowadzi do niejednoznaczności klasyfikacji, to promienie pól przyciągania λ neuronów źle klasyfikujących są modyfikowane, a mianowicie

$$\lambda = \begin{cases} \lambda_{\min} & d \leqslant \lambda_{\min} \\ d & \lambda_{\min} < d < \lambda_{\max} \\ \lambda_{\max} & d \geqslant \lambda_{\max} \end{cases} \tag{15}$$

gdzie d oznacza odległość do najbliższego neuronu ukrytego skojarzonego z inną klasą niż

klasa wzorca wejściowego, natomiast $\lambda_{\min}$ jest najmniejszym dopuszczalnym promieniem przyciągania. Szczegóły realizacyjne można znaleźć w [10]

6 Dynamiczne marszczenie czasu

Dynamiczne marszczenie czasu jest techniką porównywania funkcji. U jej podłoża leży obserwacja, że bezpośrednie porównywanie wartości funkcji x i y dla identycznych chwil czasowych nie jest właściwe, gdyż czas może "biec nierównomiernie", choć z zachowaniem następstwa zdarzeń. Model tego zjawiska polega na założeniu, że "indywidualny czas" jest funkcją g "rzeczywistego czasu" t. Rodzinę $\mathcal{G}^o$ możliwych "funkcji marszczących" tworzą funkcje ciągłe niemalejące. W rezultacie, porównywane są funkcje

$$x(g^x(t)), \quad y(g^y(t)) \tag{16}$$

dla argumentów $\mathcal{T} = t \in [0, t_{\max}]$, a miarą ich odległości może być np.

$$\inf_{g^x, g^y \in \mathcal{G}^o} \int_{\mathcal{T}} \|x(g^x(t) - y(g^y(t))\|^2 \, dt \tag{17}$$

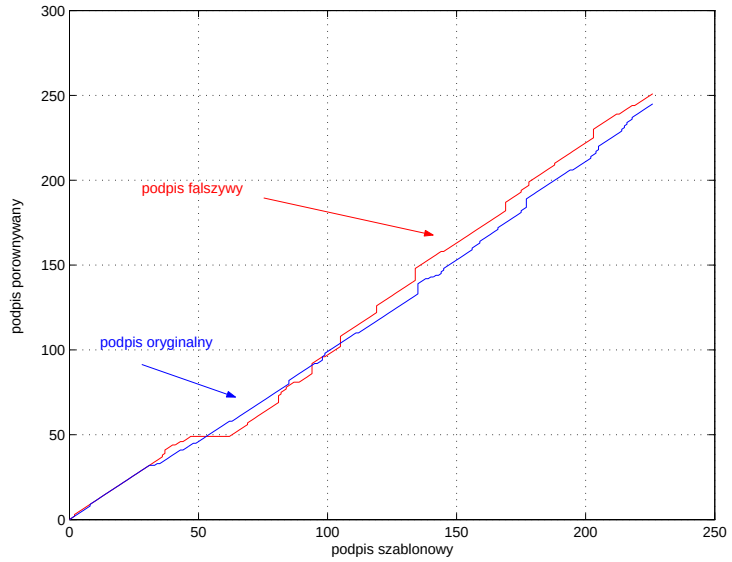
Można dla wygody przyjąć, że czas indywidualny wzorca x jest identyczny z czasem rzeczywistym. Wówczas miarą bliskości badanej funkcji x i wzorca jest

$$\inf_{g \in \mathcal{G}} \int_{\mathcal{T}} \|x(t) - y(g(t))\|^2 \, dt \tag{18}$$

gdzie (ze względu na to, że funkcja marszcząca wzorca została przekształcona na czas rzeczywisty) rodzina $\mathcal{G}$ funkcji marszczących musi być uzupełniona o funkcje przedziałami ciągłe. Dodatkowo można wymagać, by oba czasy "niezbyt" różniły się, tzn. by $g(t) \approx t$, $\dot{g}(t) \approx 1$. Odpowiednio dobrana miara odległości g od funkcji tożsamościowej $f(t) = t$ może zatem służyć jako dodatkowa miara odstępstwa podpisu badanego od podpisu rzeczywistego.

Dla problemów z czasem dyskretnym funkcje marszczące mają nie tylko dyskretną dziedzinę, ale dyskretny zbiór wartości. Mogą być one reprezentowane przez "ścieżki czasu" łączące punkt (0,0) z punktem $(t_{\max}, t^y\max)$ poprzez punkty sąsiednie, przy czym sąsiadami punktu (i, j) są punkty $(i+1, j)$, $(i+1, j+1)$, oraz $(i, j+1)$ (rys. 10). Zauważmy, że przejście diagonalne (z (i, j) do $(i+1, j+1)$ odpowiada "zgodnemu" przebiegowi czasu dla obu ciągów, zaś przejścia "poziome" lub 'pionowe" odpowiadają jednostkowemu wydłużeniu czasu spędzonego na zapisie fragmentu podpisu referencyjnego lub podpisu badanego. Dla każdego podpisu przebieg ścieżki może być inny, tzn. nie jest on bezpośrednio zależny od rodzaju zapisywanego fragmentu, co różni tę metodę od wariantów ukrytych modeli Markowa. Warunek bliskości g i funkcji tożsamościowej może sprowadzić się np. do żądania, by ścieżki czasu nie wiodły "zbyt długo" w kierunku poziomym lub pionowym, oraz by $t_{\max}$ i $t^y_{\max}$ nie różniły się "zbyt bardzo".

Dowodzi się, że poszukiwanie najlepszej ścieżki czasu można wykonać w czasie wielomianowym, jednak uogólnienia do problemów marszczenia w więcej niż jednym wymiarze jest NP-zupełne.



Rysunek 10: Ścieżka czasu dla realizacji testowej i realizacji fałszywej tego samego podpisu

7 Przykładowe wyniki weryfikacji i identyfikacji podpisów

7.1 Jakość identyfikacji

Uczenie perceptronu z zastosowaniem 1,2,3,4 lub wszystkich dostępnych ($\geqslant 4$) realizacji referencyjnych dla każdego podpisu (tab. 1) prowadziło do znacznych różnic w poziomie prawidłowych klasyfikacji. Przy odpowiednim doborze współczynników, sieć RCE wykazała jednak większe zdolności uogólniające od perceptronu, dając poprawną klasyfikację wszystkich realizacji testowych, tzn. CCR=100%.

Tabela 1: Wyniki klasyfikacji za pomocą perceptronu uczonego dla różnej liczby realizacji referencyjnych reprezentujących podpis

liczba realizacji referencyjnych	procent poprawnych klasyfikacji (CCR)
1	63.5 %
2	75.0 %
3	83.3 %
4	87.5 %
$\geqslant 4$	95.8 %

7.2 Jakość weryfikacji

Wyniki weryfikacji przy użyciu sieci RCE są gorsze niż przy zastosowaniu sieci perceptronowej. Sieć RCE wykazała lepsze właściwości uogólniające niż dwuwarstwowy perceptron

o sigmoidalnych funkcjach aktywacji, nie wykrywała jednak wszystkich fałszerstw. Dla sieci RCE otrzymano FAR = 8.33% i FRR = 11.11%. Dla perceptronu przy zastosowaniu zbiorów $\mathcal{S}^r$ i $\mathcal{S}^t$ otrzymano natomiast FAR=0% przy FRR=11%.

Znacznie lepsze wyniki uzyskano przez zastosowanie techniki marszczenia czasu. Bezpośrednie wykorzystanie technik marszczenia prowadzi do FAR=0% przy FRR=1.7%. Zauważmy jeszcze, że metoda ta może być rozbudowana poprzez wykorzystanie dodatkowych cech podpisu przekształconego przez marszczenie, gdyż przekształcenie to zbliża cechy globalne do cech podpisu referencyjnego.

Powyższe wyniki otrzymane zostały dla jednokrotnej próby weryfikacji każdego podpisu. W praktyce można dopuszczać ponowienie weryfikacji i ponowne złożenie podpisu po negatywnej weryfikacji. Takie postępowanie znacznie zmniejsza współczynnik FRR ale równocześnie zwiększa FAR.

8 Podsumowanie

Chociaż dokładność systemów weryfikacji tożsamości wykorzystujących podpis odręczny jest niższa niż systemów wykorzystujących obraz tęczówki czy odciski palca, to jednak na pewno mogą być one stosowane jako systemy doradcze. Rozpatrzmy np. rolę podpisu przy zakupie z użyciem karty kredytowej. Weryfikacja podpisu wykonywana jest obecnie przez osobę dokonującą sprzedaży i jest jednym z elementów systemu bezpieczeństwa. Jakość takiej weryfikacji jest jednak stosunkowo niska i z łatwością może być ulepszona przez zautomatyzowane systemy weryfikacji podpisu. Podobne, doradcze znaczenie może mieć system weryfikacji podpisu przy transakcjach przeprowadzanych w bankach. Rozwój prac nad zwiększeniem niezawodności systemów weryfikacji bazujących na podpisie mają więc zasadnicze znaczenie.

Podziękowanie

Autorzy dziękują Pani Joannie Putz-Leszczyńskiej za udostępnienie wykresów związanych z marszczeniem czasu, otrzymanych w ramach wykonywanej przez nią pracy magisterskiej.

Literatura

[1] *Słownik Języka Polskiego*, Witold Doroszewski (Red.), PWN, Warszawa 1964

[2] *Najnowszy Słownik Języka Polskiego*, Boguslaw Dunaj (Red.), WILGA 1999

[3] S. Nanavati, M. Thieme, R. Nanavati *Biometrics. Identity Verification in a Networked World*, Wiley Computer Publishing, New York 2002

[4] G. Dimauro, S. Impedevo, G. Pirlo, A. Salzo, "A Multi-Expert Signature Verification System for Bankcheck Processing", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 11, nr 5 (1992) str. 827-844, 1992

[5] S. Lee, J. C. Pan, "Offline Tracing and Representation of Signatures", *IEEE Trans. On Systems, Man and Cybernetics*, vol. 18, nr 6 (June 1996), 1996

[6] J.-J. Brault, R. Plamondon, "A Complexity Measure of Handwritten Curves: Modelling of Dynamic Signature Forgery", *IEEE Trans. On Systems, Man and Cybernetics*, vol. 23, nr 2 (March/April 1993), 1993

[7] L. L. Lee, T. Berger, E. Aviczer, "Reliable On-Line Human Signature Verification System", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 18, nr 6, (June 1996), 1996

[8] H. Crane, J. S. Ostrem, "Automatic Signature Verification Using a Three-Axis Force-Sensitive Pen", *IEEE Trans. On Systems, Man and Cybernetics*, vol. SMC-13, nr 3 (May/June 1983), 1983

[9] T. Wessels, C. W. Omlin, "A Hybrid System for Signature Verification", *Transactions of IJCNN'2000*, Como, Włochy 2000

[10] D.L. Reilly, "The RCE Neural Network", *Nestor, Inc.*, 1989

[11] Andrzej Pacut, Adam Czajka, "Recognition of Human Signatures", *Proc. of the IEEE-INNS-ENNS International Joint Conference on Neural Networks IJCNN'2001, Washington, D.C., USA*, vol. 2, str. 1560-1564, IEEE, 2001

[12] Adam Czajka, Andrzej Pacut, "Rozpoznawanie podpisów odręcznych jako przykład systemu weryfikacji tożsamości", *Techniki Komputerowe, Biuletyn Informacyjny Instytutu Maszyn Matematycznych*, nr 1/2001, str. 7-24, Warszawa, 2001

[13] Adam Czajka, Andrzej Pacut, "Neural networks for signature classification and identity verification", IT.FORUM SECURE 2002, 6-7 listopoda 2002, Warszawa, vol. 1, str. 1-7, 2002