

Dynamický podpis II

Jim Wild

Databáze podpisů

Dataset	Database	Users	Signatures		Total
			Genuines	Forgeries	
DD	MCYT-A	50	25	25	2500
TD	MCYT-B	50	25	25	2500
	SVC2004	40	20	20	1600
	BIOMET	84	15	17	2688
	MYIDEA	69	18	36	3726
Total		293	5802	7212	13014

Metody rozpoznávání

▶ **Deterministické metody**

- **Dynamic Time Warping (DTW)**
- **Vector Quantization (VQ)**

▶ **Statistické metody**

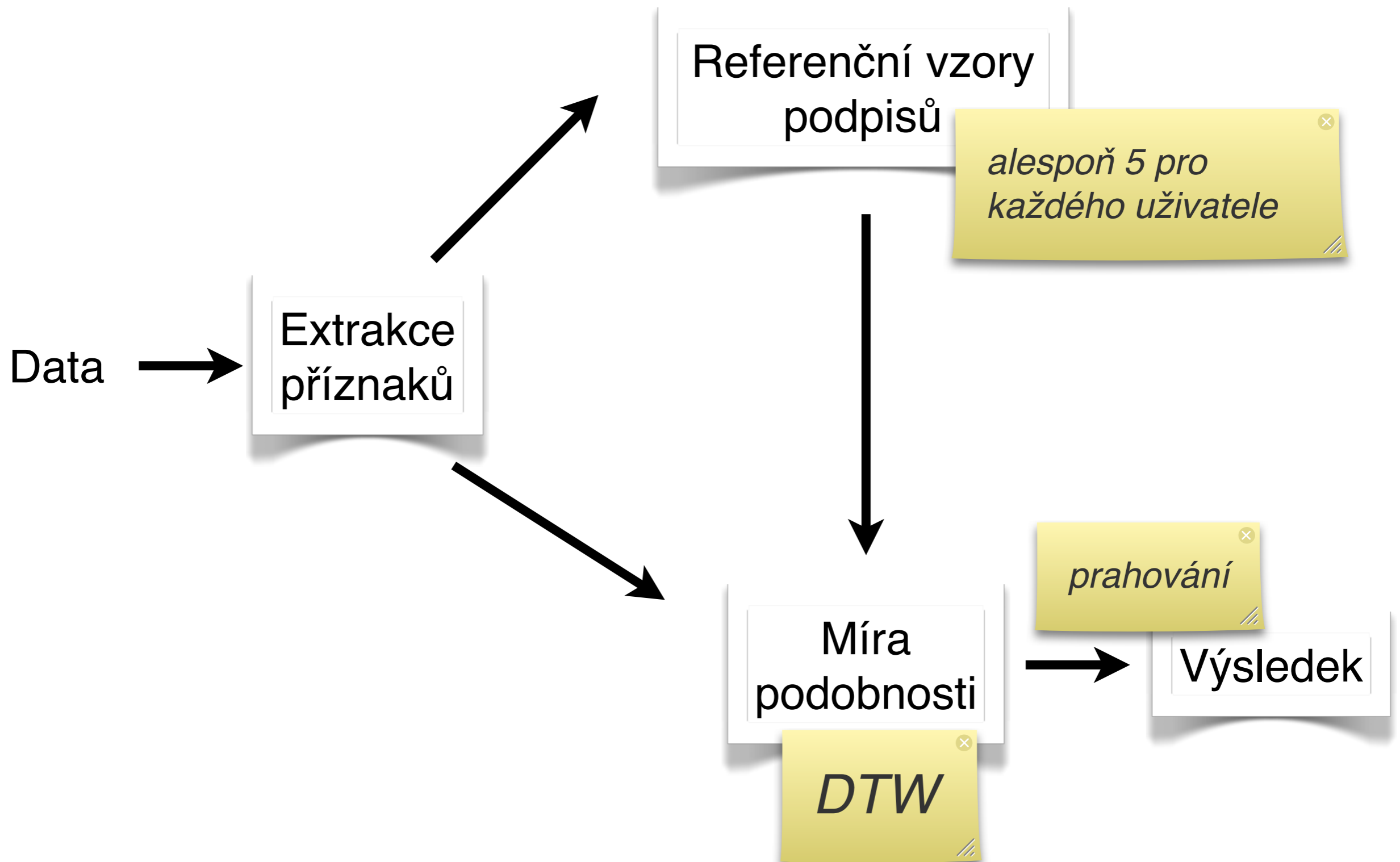
- **Gaussian Mixture Model (GMM)**
- **Hidden Markov Model (HMM)**

DTW zpracování DP

▶ **Practical On-Line Signature Verification**

- J.M. Pascual-Gaspar, V. Cardenoso-Payo, and C.E. Vivaracho-Pascual
- Advances in Biometrics, Lecture Notes in Computer Science 2009

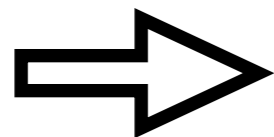
Schéma rozpoznání podpisu



Extrakce příznaků

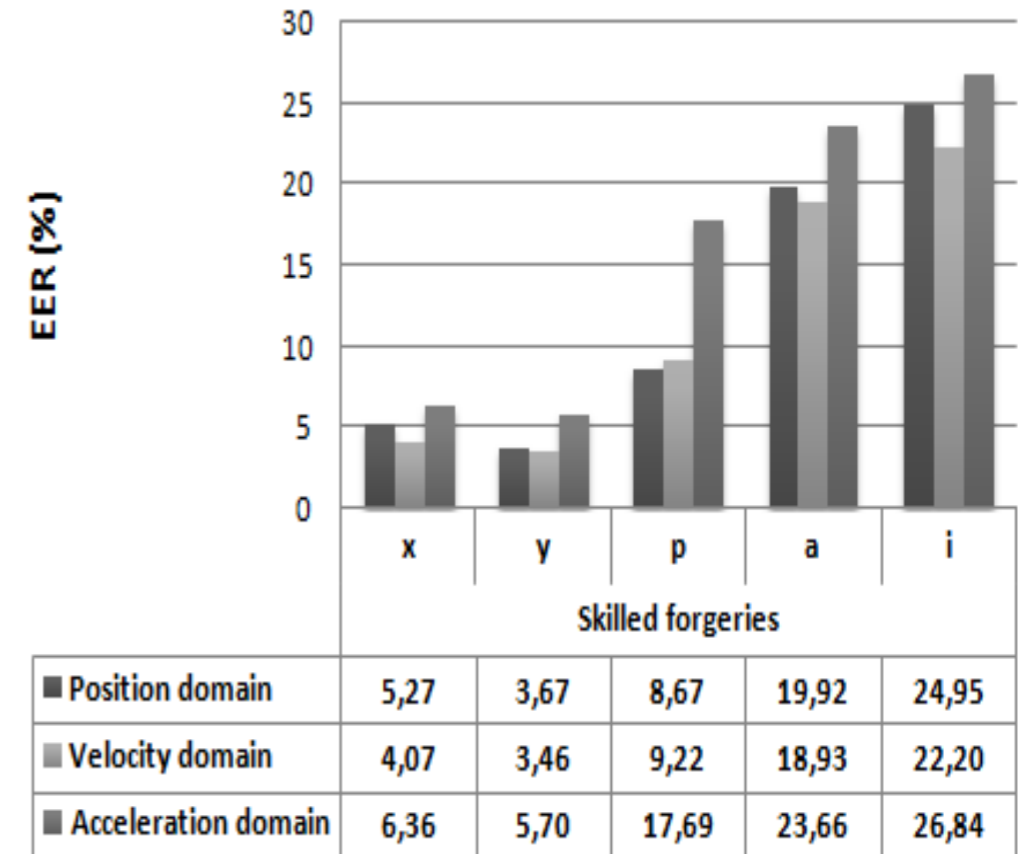
- ▶ **x, y**
- ▶ **p - přítlak**
- ▶ **a - natočení (azimuth)**
- ▶ **i - náklon (inclination)**

+ 1. a 2. derivace



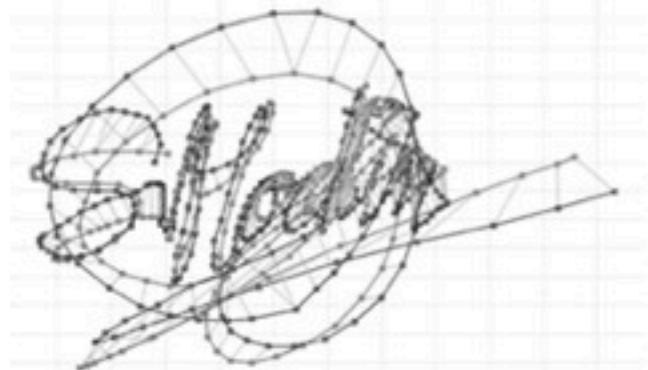
15 příznakových vektorů

(i když následně došlo k redukci)

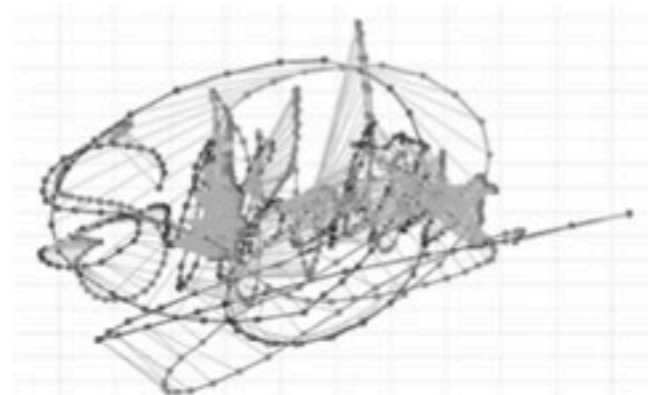


Rozpoznání podpisu

► Dynamic Time Warping



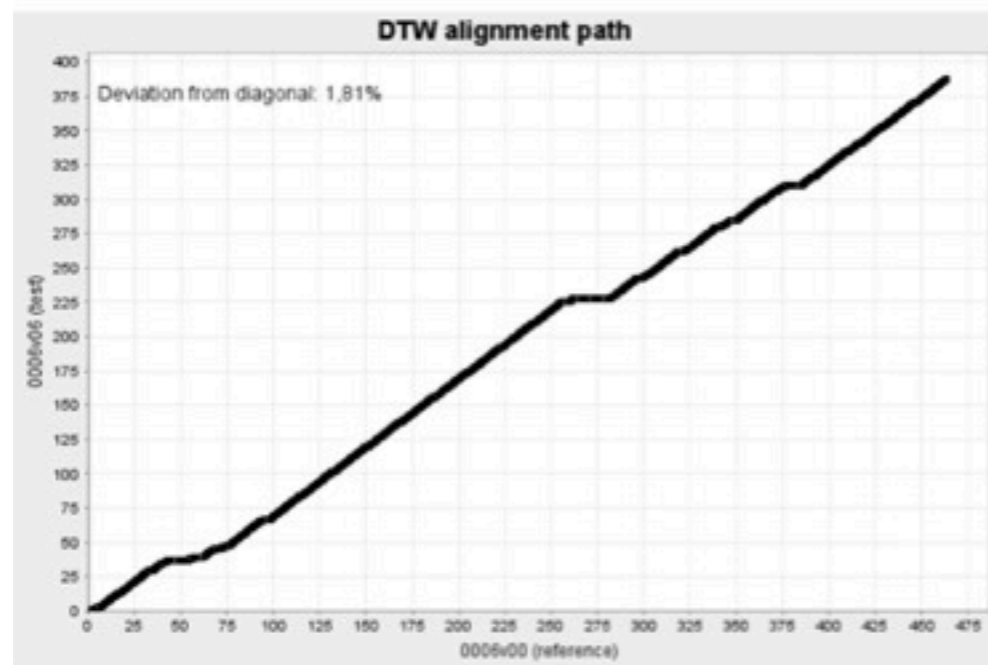
(a) Genuine-genuine



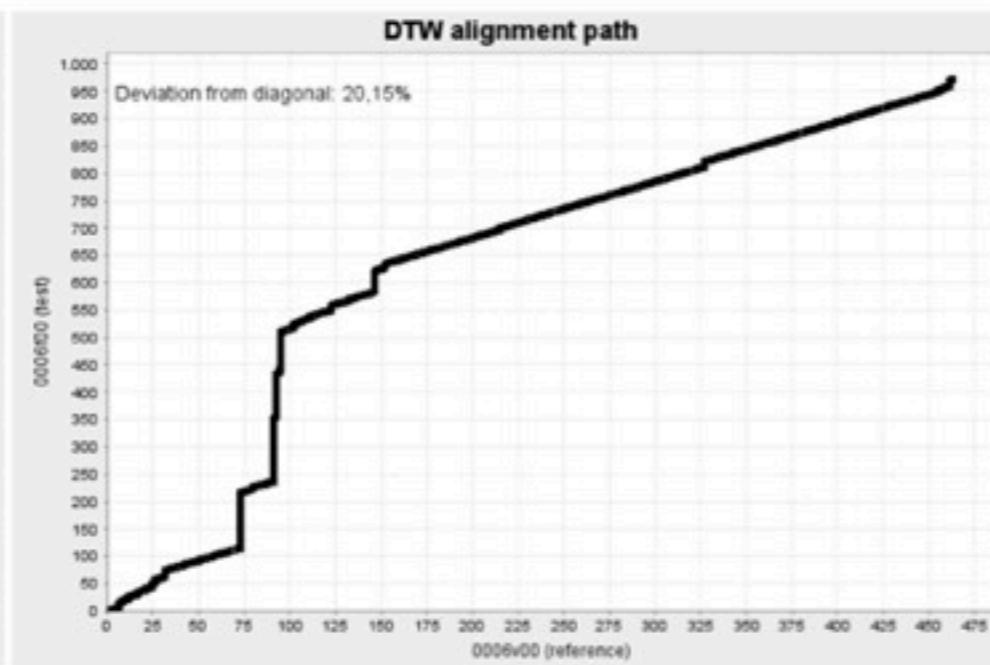
(b) Genuine-forgery



(c) Intra-class variability

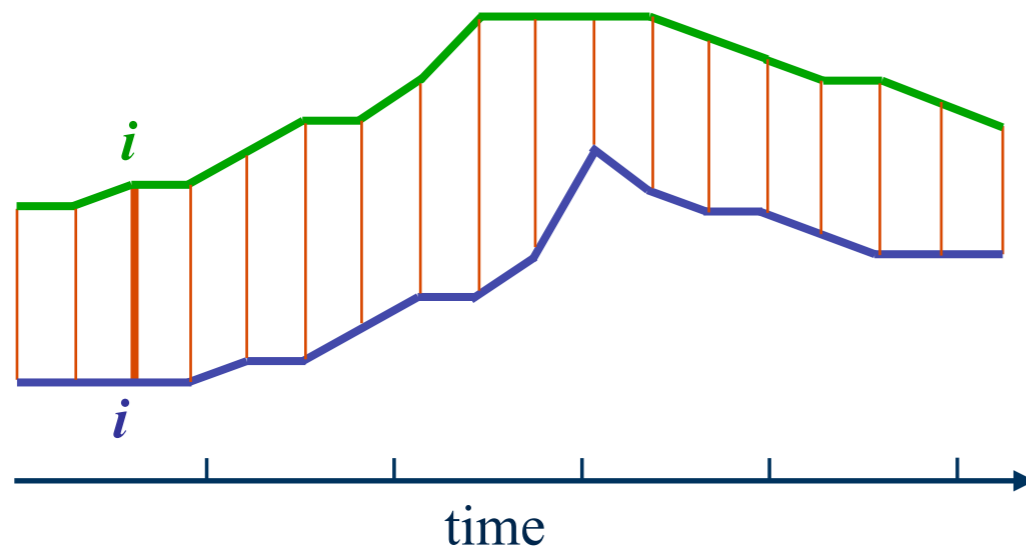


(d) Gen-Gen DTW path

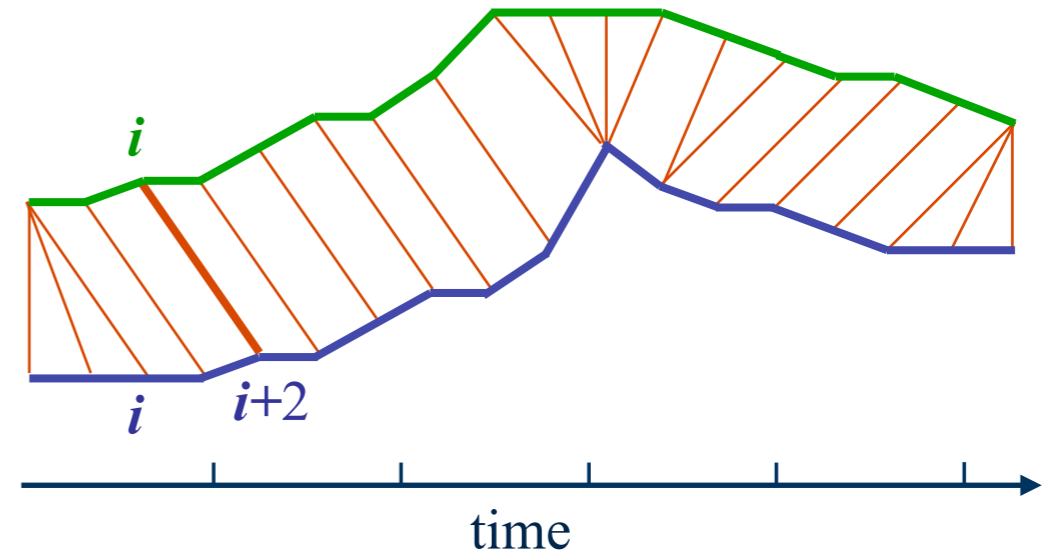


(e) Gen-Forg DTW path

Proč DTW?

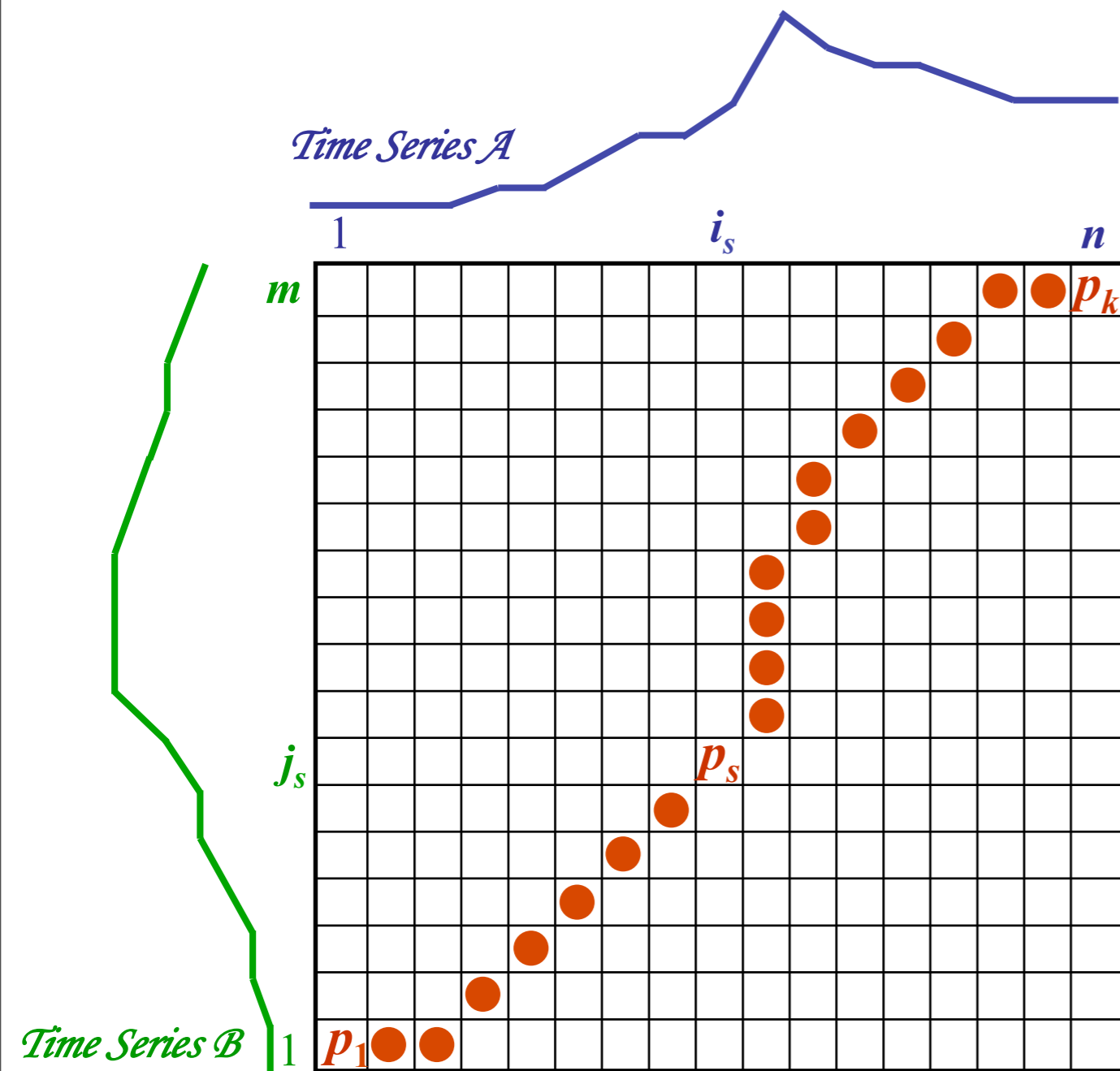


Porovnání *křivek*
(*standardně*)



Porovnání *křivek*
(*DTW*)

Warpovací funkce

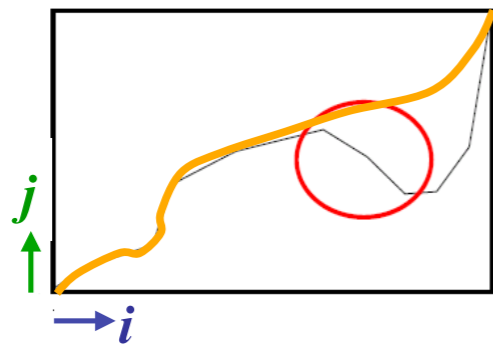


- ▶ mřížka ukazuje vzdálenost (podobnost) jednotlivých bodů (n -rozměrných) křivek
- ▶ snaha o nalezení minimální cesty z $[0,0]$ do $[n,m]$
- ▶ řeší dynamické programování

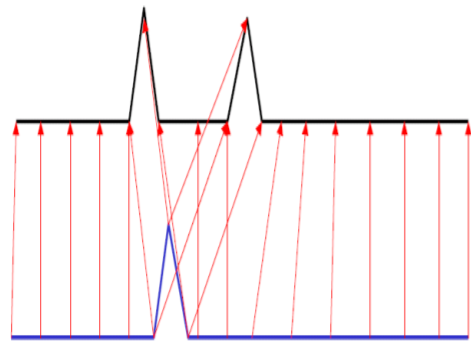
Omezení warpovací funkce

Monotonicity: $i_{s-1} \leq i_s$ and $j_{s-1} \leq j_s$.

The alignment path does not go back in “time” index.

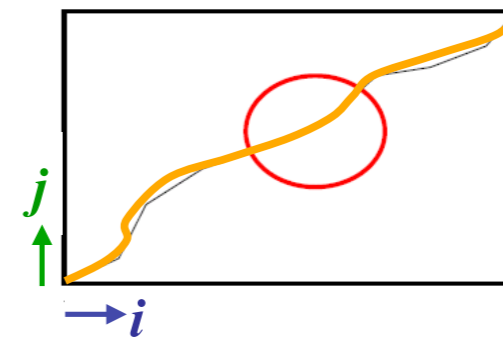


Guarantees that features are not repeated in the alignment.

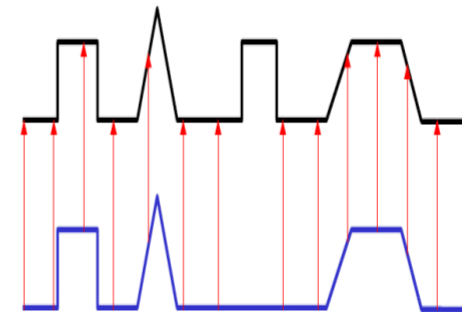


Continuity: $i_s - i_{s-1} \leq 1$ and $j_s - j_{s-1} \leq 1$.

The alignment path does not jump in “time” index.



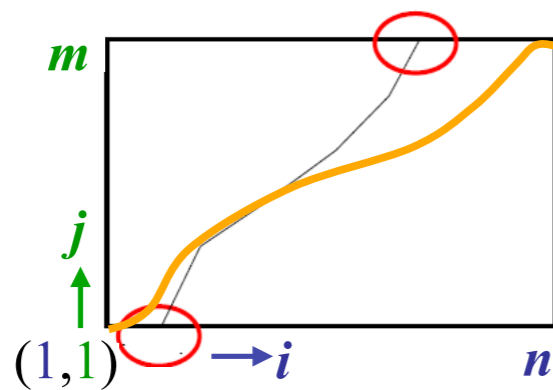
Guarantees that the alignment does not omit important features.



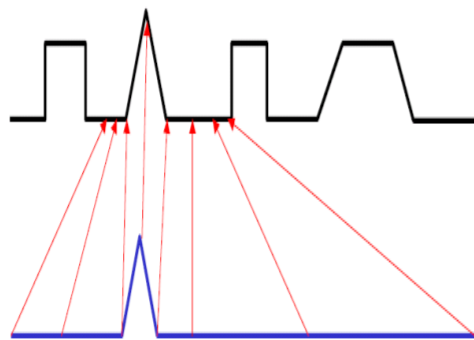
Omezení warpovací funkce

Boundary Conditions: $i_1 = 1, i_k = n$ and $j_1 = 1, j_k = m$.

The alignment path starts at the bottom left and ends at the top right.

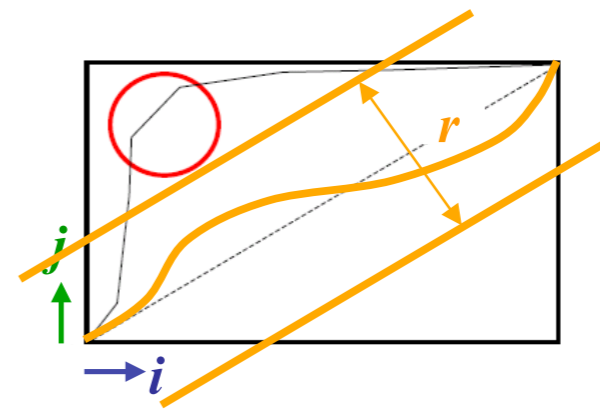


Guarantees that the alignment does not consider partially one of the sequences.

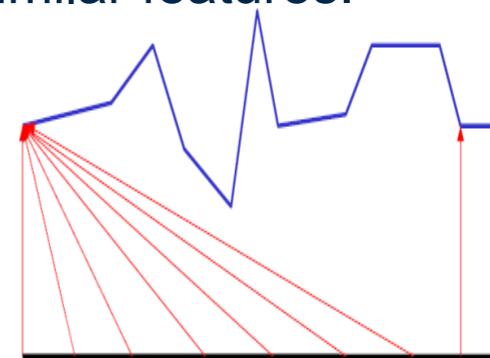


Warping Window: $|i_s - j_s| \leq r$, where $r > 0$ is the window length.

A good alignment path is unlikely to wander too far from the diagonal.



Guarantees that the alignment does not try to skip different features and gets stuck at similar features.



DTW - příklad

Time Series A →

-0.87	-0.84	-0.85	-0.82	-0.23	1.95	1.36	0.60	0.0	-0.29
-0.88	-0.91	-0.84	-0.82	-0.24	1.92	1.41	0.51	0.03	-0.18

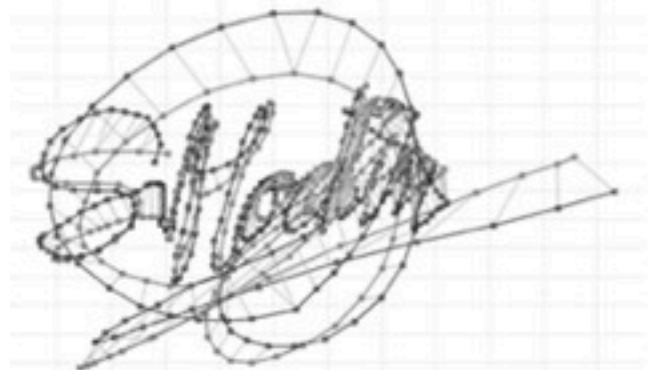
1.94
0.77
-0.17
-0.58
-0.71
-0.65
-0.60
-0.46
1.97
0.74
-0.32
-0.63
-0.68
-0.62

0.51	0.51	0.49	0.49	0.35	0.17	0.21	0.33	0.41	0.49
0.27	0.27	0.26	0.25	0.16	0.18	0.23	0.25	0.31	0.68
0.13	0.13	0.13	0.12	0.08	0.26	0.40	0.47	0.49	0.49
0.08	0.08	0.08	0.08	0.10	0.31	0.47	0.57	0.62	0.65
0.06	0.06	0.06	0.07	0.11	0.32	0.50	0.60	0.65	0.68
0.04	0.04	0.06	0.08	0.11	0.32	0.49	0.59	0.64	0.66
0.02	0.05	0.08	0.11	0.13	0.34	0.49	0.58	0.63	0.66

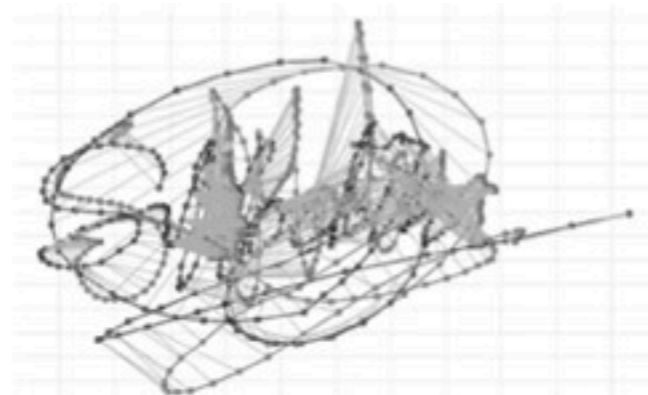
↑ *Time Series B*

Rozpoznání podpisu

► Dynamic Time Warping



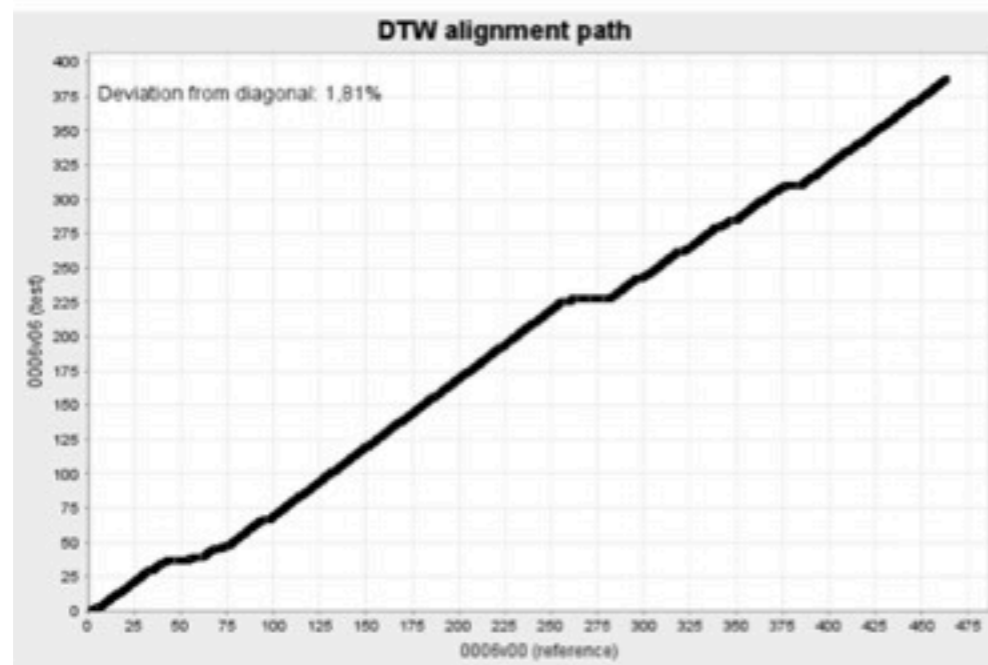
(a) Genuine-genuine



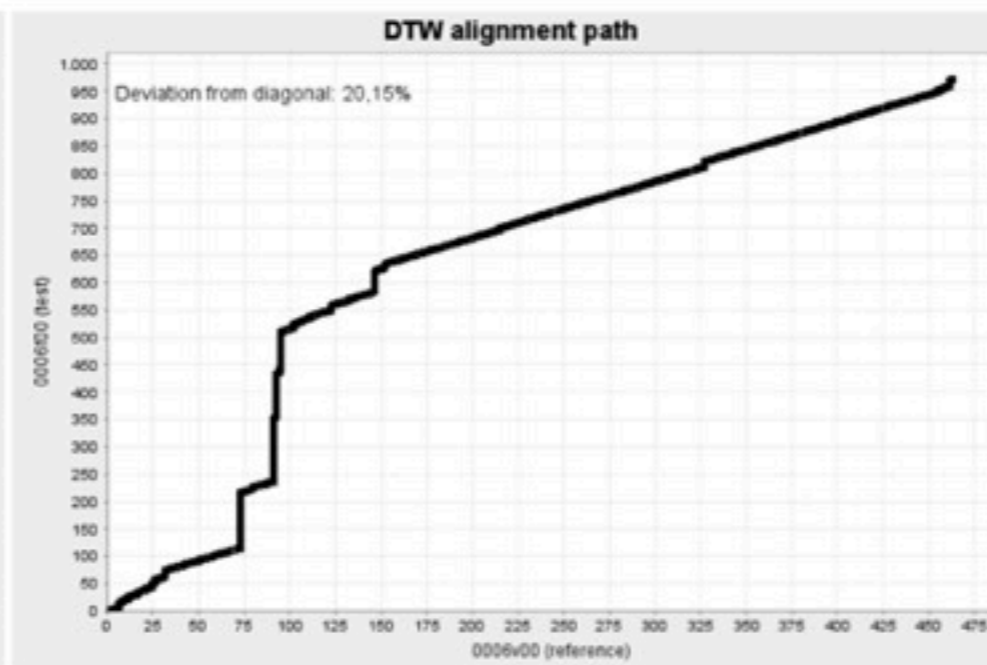
(b) Genuine-forgery



(c) Intra-class variability

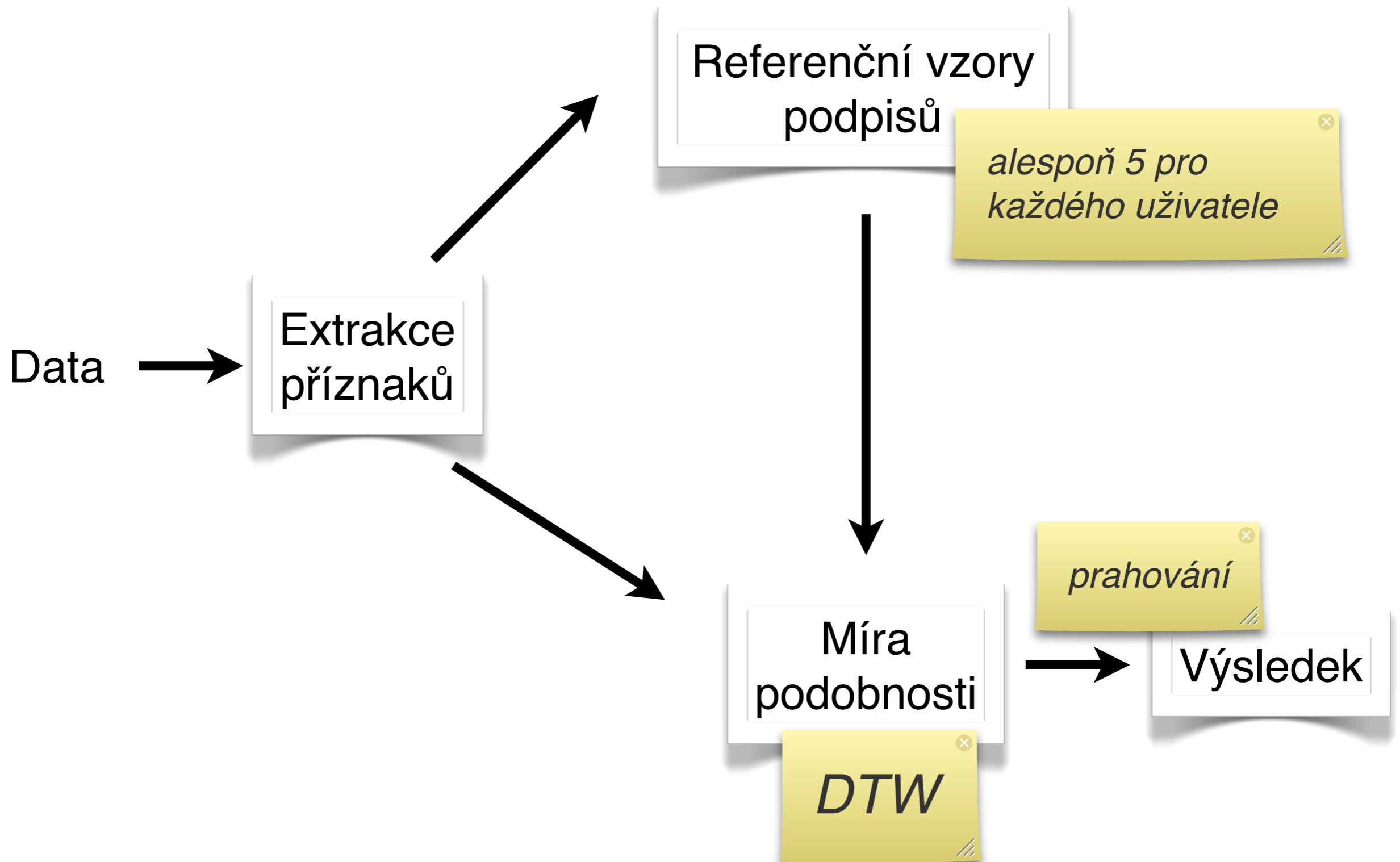


(d) Gen-Gen DTW path



(e) Gen-Forg DTW path

Schéma rozpoznání podpisu



Výsledky

	Random forgery	Skilled forgery
EER	0.41%	2.26%

EER = Equal error rate

udává chybovost (jak chybného přijetí podpisu za pravý, tak i chybného zamítnutí pravého podpisu)

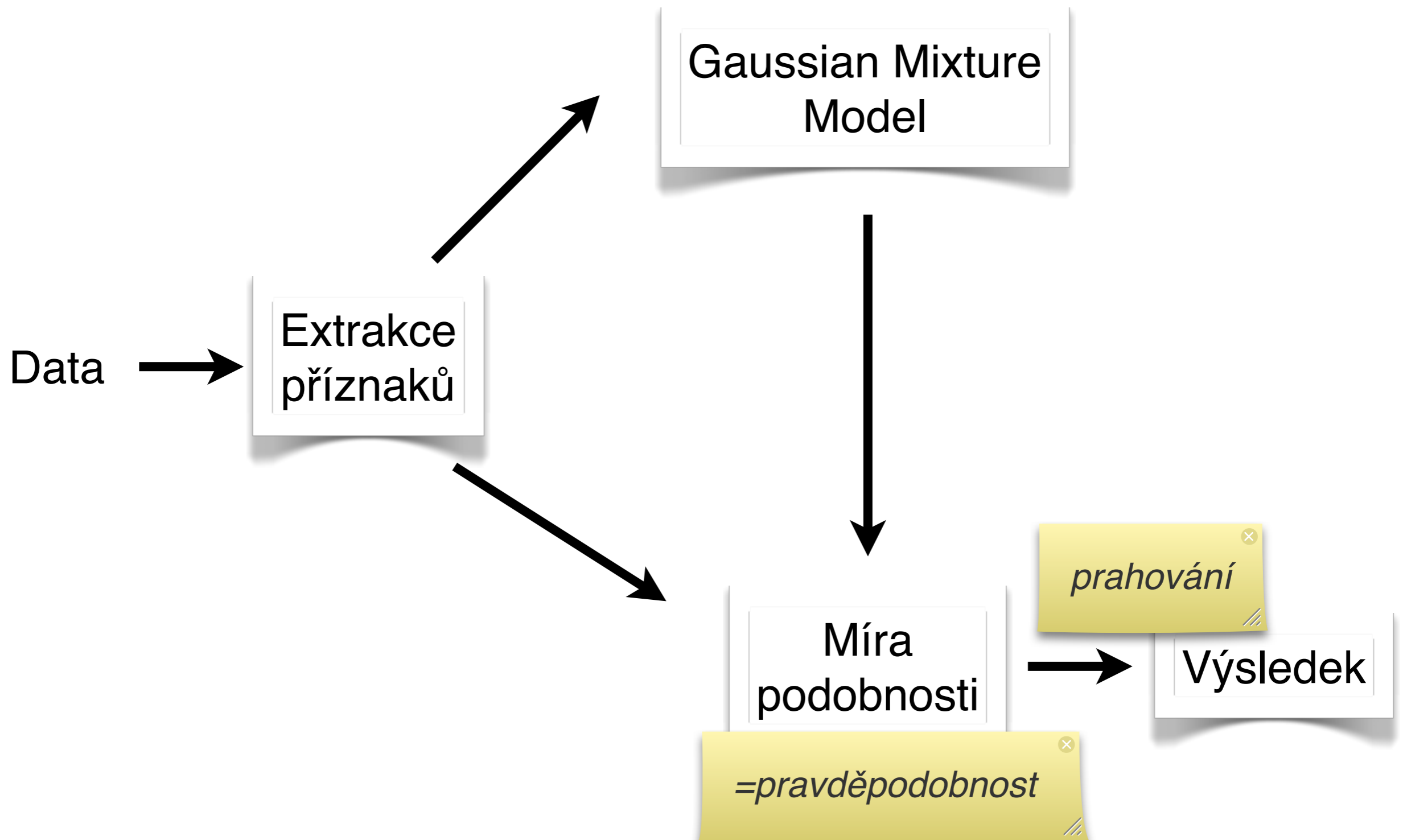
Cvičení

► Úkol:

- Napište funkci v Matlabu, která na základě předloženého podpisu rozhodne, zda se jedná o originál či padělek.
 - K dispozici je několik originálních podpisů (od jednoho člověka) a jejich padělků.
 - Testování bude probíhat na příkladech, které **nemáte** k dispozici.
 - definice funkce:

```
function [pravy]=rozpoznej_podpis(podpis)
% podpis - NxT matice, kde N je počet vzorků a T počet
% naměřených veličin (x, y, p, ...)
% pravy - 0 nebo 1 v závislosti na pravosti podpisu
end
```


Návrh řešení - schéma



Návrh řešení

- 1. Předzpracování dat (+ extrakce příznaků)**
- 2. Natrénování a uložení GMM modelu**
- 3. Míra podobnosti**
- 4. Stanovení rozhodovacího prahu**
- 5. Rozhodovací funkce**
= poskládání jednotlivých částí dohromady

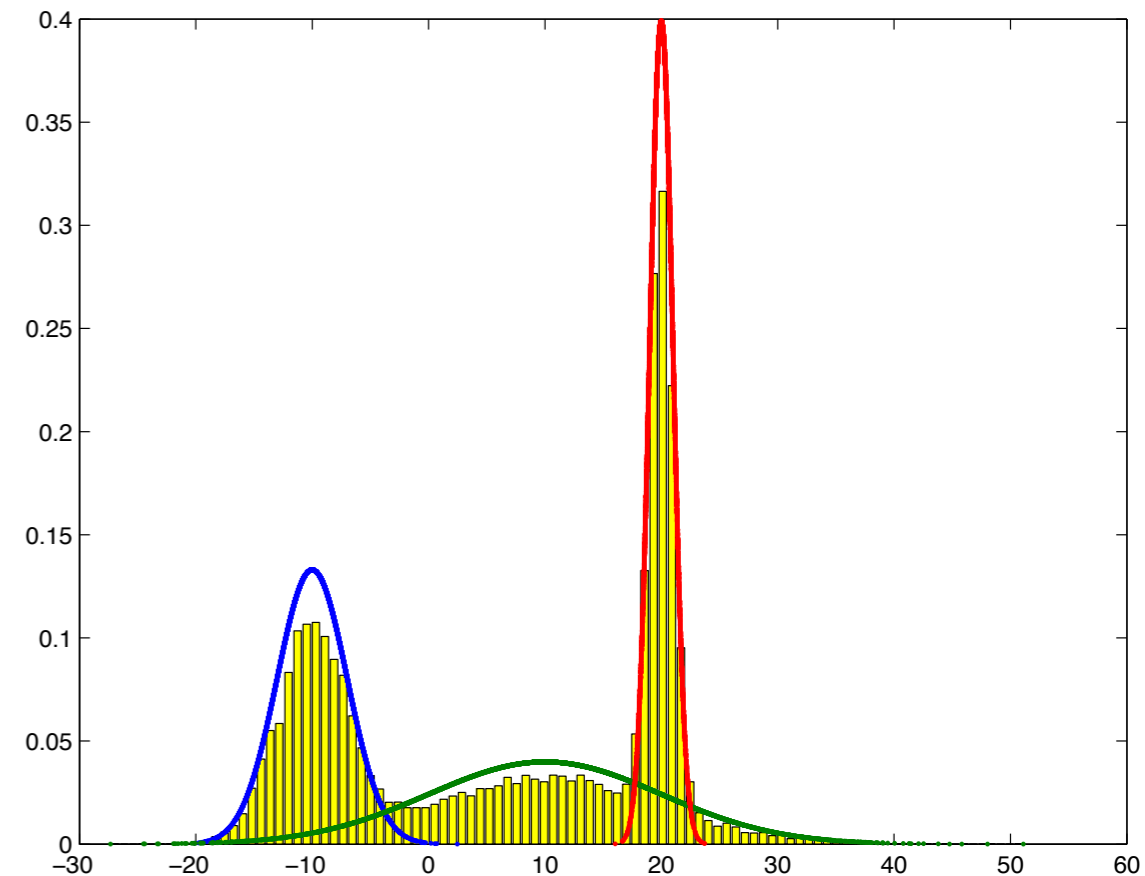
1. Předzpracování dat

- ▶ **Oříznutí**
- ▶ **Normalizace**
- ▶ **Výpočet příznakových vektorů (1., 2. derivace)**
- ▶ **definice m-funkce:**

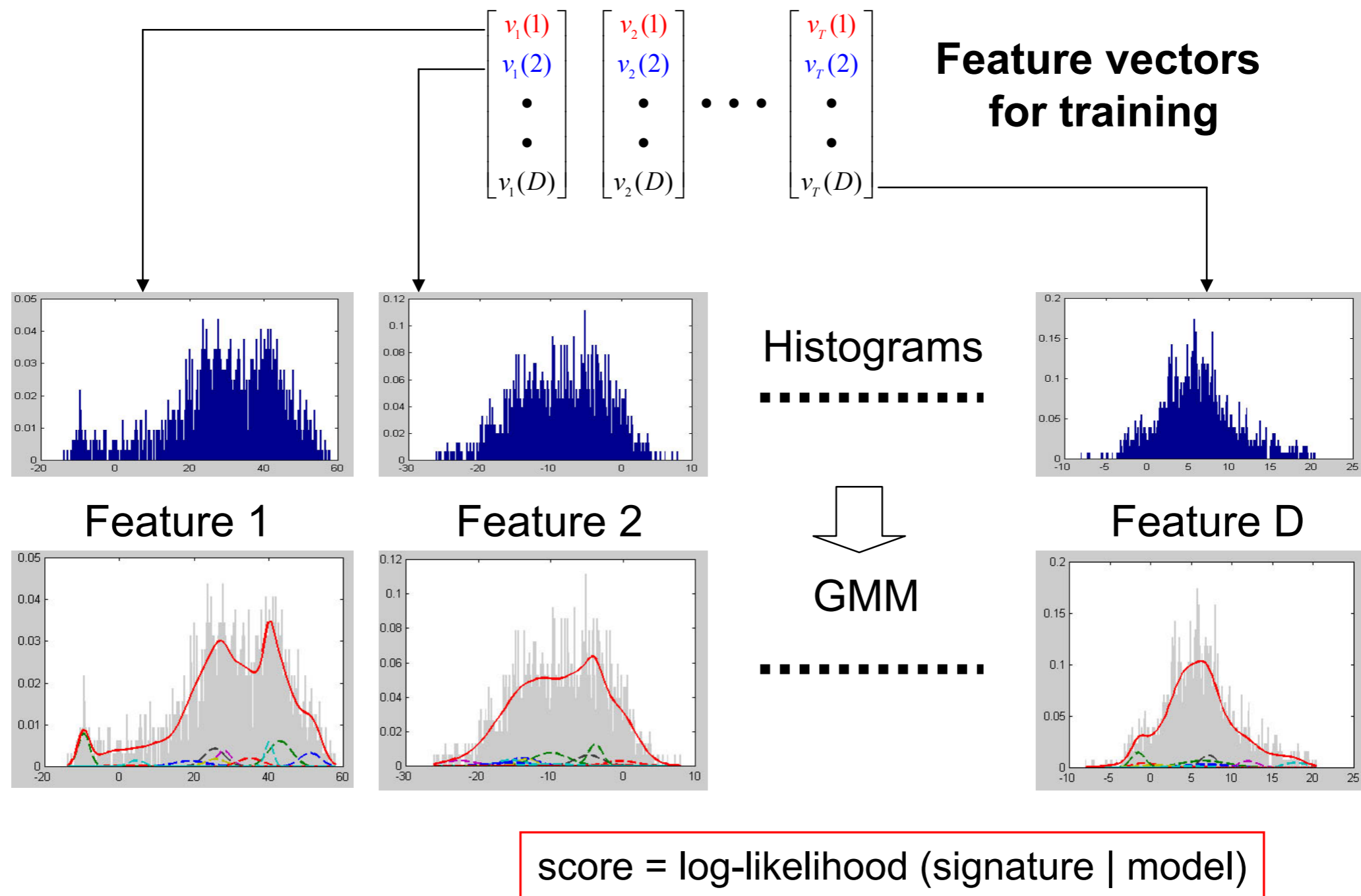
```
function [priznaky]=predzpracovani(podpis)
% podpis - N x T matice, kde N je počet vzorků a T počet
% naměřených veličin (x, y, p, ...)
% priznaky - M x L matice, kde L je počet příznakových
% vektorů a M je jejich délka (M<=N)
end
```

2. Trénování GMM

- ▶ Pro každý příznakový vektor pomocí E-M algoritmu odhadneme GMM
 - použijeme třídu `gmdistribution` ze Statistics toolboxu (funkce `fit`)
- ▶ Uložíme si výsledný `gmdistribution` objekt pro pozdější použití
- ▶ E-M algoritmus



Gaussian Mixture Model



3. Míra podobnosti = skóre

- ▶ **Míra podobnosti = pravděpodobnost, že náš příznakový vektor $x_{1,2,\dots,M}$ pochází z daného GMM**

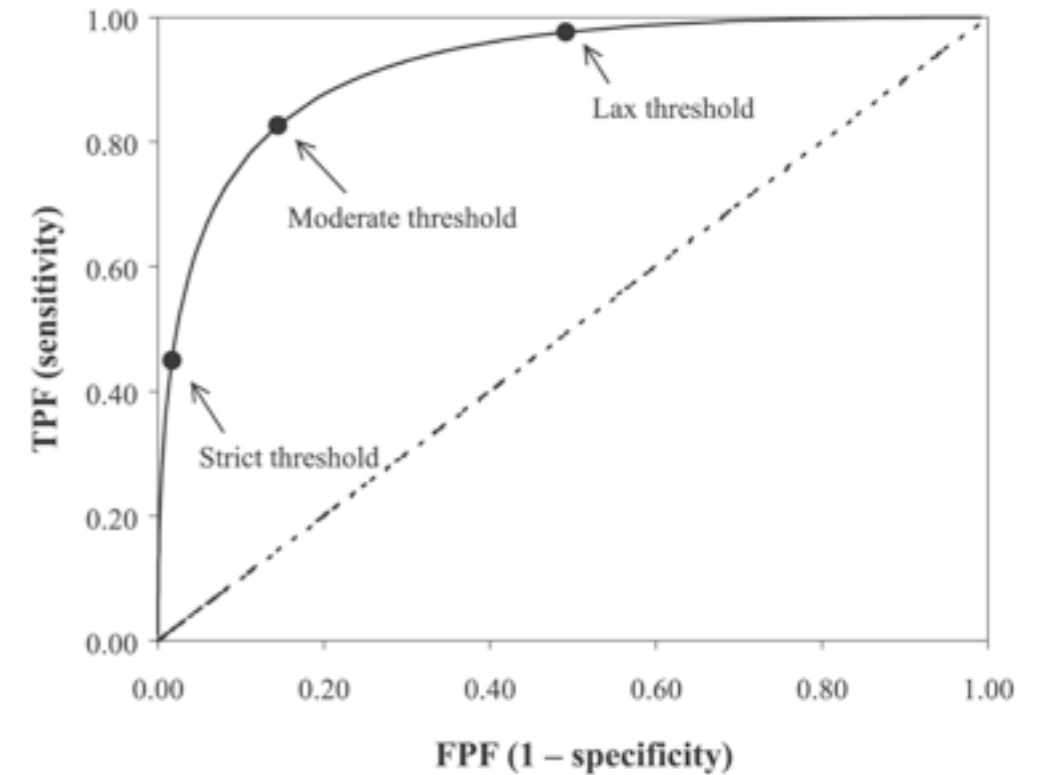
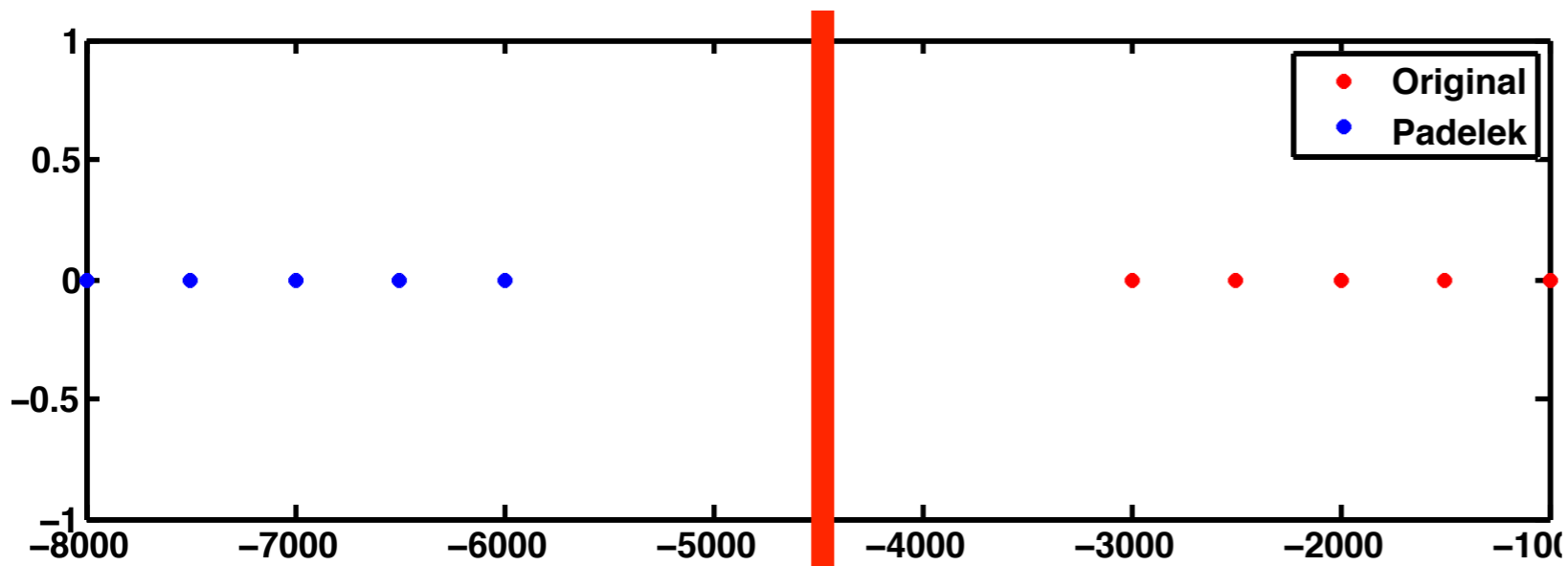
$$p(x_{1,2,\dots,M} | GMM) = \prod_{i=1}^M p(x_i | GMM)$$

- pro zjištění $p(x_i | GMM)$ použijeme funkci `gmdistribution.pdf(x)`
- vypočítáme celkovou $\log(p(x_{1,2,\dots,M} | GMM))$ příznakového vektoru:

$$\log(p(x_{1,2,\dots,M} | GMM)) = \sum_{i=1}^M \log(p(x_i | GMM))$$

- ▶ **Skóre jednotlivých vektorů sečteme (proč?)**

4. Rozhodovací práh



► Experimentálně:

- spočítáme skóre pro několik originálů a padělků, které vyneseme na osu a určíme práh

► ROC křivkou:

- pro každou hodnotu prahu spočítáme sensitivitu a specificitu - vybíráme "optimum"

► Bayes

► ...

Bonus

- ▶ **Identifikace - přiřazení podpisu ke správnému originálu**
- ▶ **Využití globálních příznaků + lepší klasifikátor**
- ▶ **Otestování na databázi podpisů**
 - (viz slide 2)

Děkuji za pozornost