

Fourierovský popis ploch – klasifikátor velkých písmen

Semestrální práce z předmětu KIV/AOS

Jan Bařtipán / A03043 / bartipan@students.cz

18. 12. 2006

Úvod

Popis ploch dvourozměrných objektů hraje důležitou roli například při klasifikaci takovýchto objektů do jednotlivých tříd. Aby klasifikace byla co nejpoužitelnější, je třeba nalézt takový popis objektů, který je invariantní vůči posunutí, natočení a změně měřítka.

To znamená, že když máme dva různé objekty O_1, O_2 a jejich popisy

$P(O_1)=P_1, P(O_2)=P_2$, jsou tyto popisy shodné ($P_1=P_2$) pro každé dva objekty,

které vzniknou jako: $O_1=T_{x,y}(O_2)$ (invariantní vůči posunutí), $O_1=R_\phi(O_2)$

(invariantní vůči natočení) nebo $O_1=S_{u,v}(O_2)$ (invariantní vůči změně měřítka), tedy

$P_1=P_2$ právě když $O_1=T_{x,y}(R_\phi(S_{u,v}(O_2))) \forall x, y, \phi, u, v$.

X

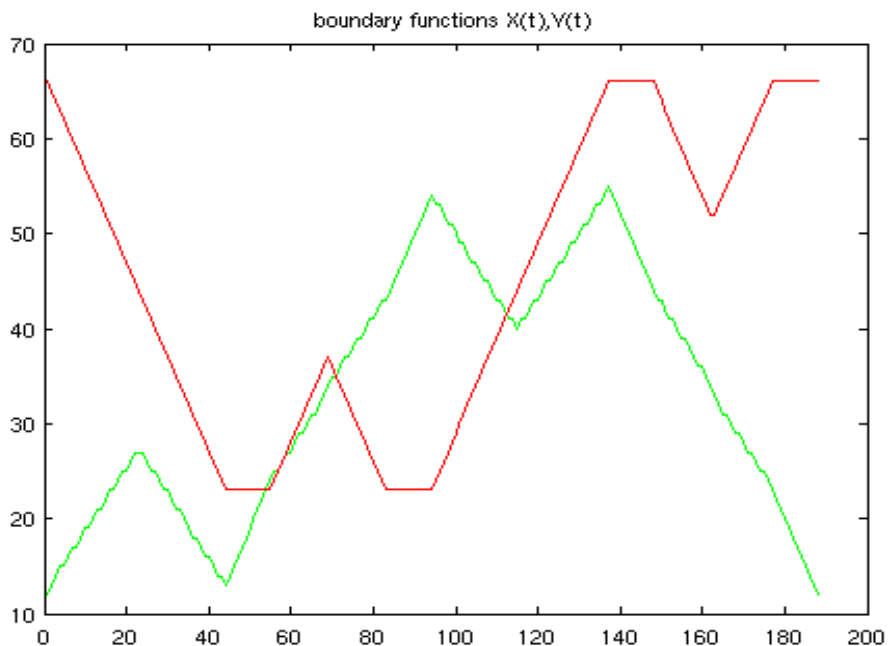
Obrázek 1:
Náš objekt,
písmeno X.

Fourierovské popisovače

Pro popis oblasti plochy pro účely klasifikace lze použít Fourierovské popisovače. To jsou koeficienty odvozené od Fourierových koeficientů bodů vnější hranice oblasti určitého objektu. Takovýto popis je invariantní vůči posunutí, natočení a změně měřítka.

Mějme vnější hranice oblasti $B=\{B_0, B_1, \dots, B_n\}$, tedy seznam souřadnic jednotlivých hraničních bodů $B_i=(x_i, y_i)$. Je zřejmé, že hranici lze popsat pomocí dvou funkcí pro jednotlivé souřadnice, tedy

$B=(X(i), Y(i))$, kde $i=0, 1, 2, \dots, n-1$.



Obrázek 2: Souřadnice hraničních bodů jako funkce $X(t), Y(t)$ pro písmeno X.

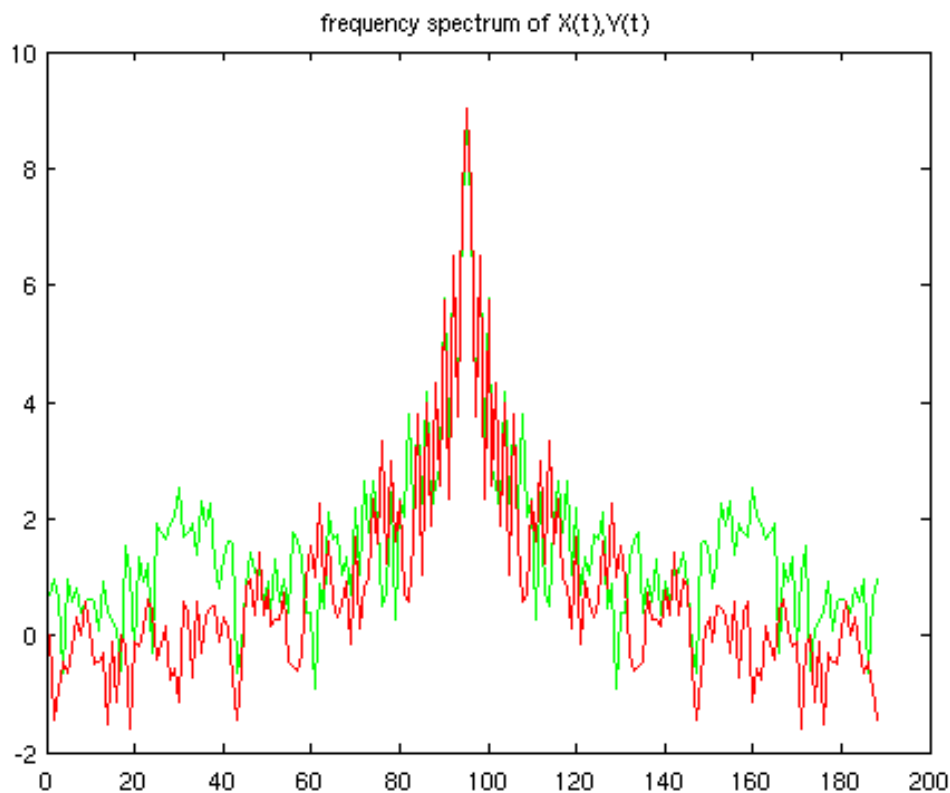
Spočítáme Fourierovu transformaci jednotlivých souřadnicových funkcí. Získáme tedy tyto koeficienty:

$$a_{xi} = \sum_{k=1}^n x_k \cdot \cos\left(\frac{2\pi k i}{n}\right)$$

$$b_{xi} = \sum_{k=1}^n x_k \cdot \sin\left(\frac{2\pi k i}{n}\right)$$

$$a_{yi} = \sum_{k=1}^n y_k \cdot \cos\left(\frac{2\pi k i}{n}\right)$$

$$b_{yi} = \sum_{k=1}^n y_k \cdot \sin\left(\frac{2\pi k i}{n}\right)$$



Obrázek 3: Spektrum funkcí $X(t)$, $Y(t)$

A naopak při znalosti jednotlivých koeficientů $a_{xi}, b_{xi}, a_{yi}, b_{yi}$ jsme schopni rekonstruovat původní souřadnicové křivky, pomocí zpětné Fourierovy transformace, tedy:

$$x_i = \frac{1}{2} \cdot a_{x0} + \sum_{k=1}^n \left[a_{xk} \cdot \cos\left(\frac{2\pi k i}{n}\right) + b_{xk} \cdot \sin\left(\frac{2\pi k i}{n}\right) \right]$$

$$y_i = \frac{1}{2} \cdot a_{y0} + \sum_{k=1}^n \left[a_{yk} \cdot \cos\left(\frac{2\pi k i}{n}\right) + b_{yk} \cdot \sin\left(\frac{2\pi k i}{n}\right) \right]$$

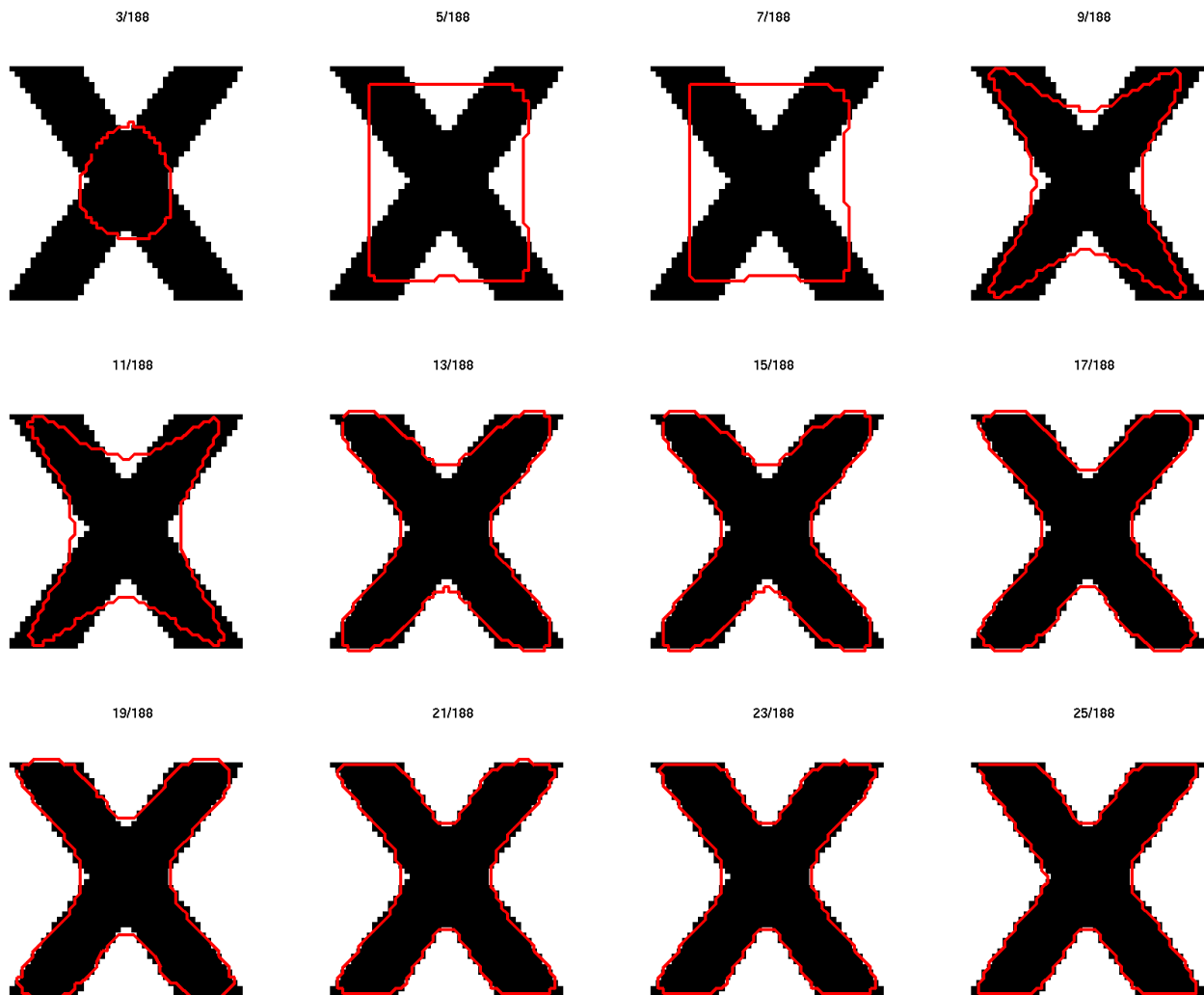
Podívejme se na význam jednotlivých koeficientů podrobněji. Pokud použijeme pouze koeficienty pro $i=0,1$ získáme definici elipsy předpisem:

$$x(t) = \frac{1}{2} \cdot a_{x0} + a_{x1} \cdot \cos(2\pi \cdot t) + b_{x1} \cdot \sin(2\pi \cdot t)$$

$$y(t) = \frac{1}{2} \cdot a_{y0} + a_{y1} \cdot \cos(2\pi \cdot t) + b_{y1} \cdot \sin(2\pi \cdot t)$$

$$t \in \langle 0,1 \rangle$$

První dva koeficienty a_{x0}, a_{y0} mají význam těžiště objektu, určují střed elipsy a tedy jsou závislé na poloze. Druhá čtveřice koeficientů určuje tvar elipsy (hlavní a vedlejší poloměr). Každá další čtveřice koeficientů $a_{xi}, b_{xi}, a_{yi}, b_{yi}$ „zpřesňuje“ obrys objektu. Všechny koeficienty $a_{xi}, b_{xi}, a_{yi}, b_{yi}, i=1,2,\dots,n-1$ jsou nezávislé na poloze.



Obrázek 4: Rekonstrukce vnější hranice za použití 3 až 25 čtveřic koeficientů.

Protože dvojice koeficientů a_{xi}, b_{xi} resp. a_{yi}, b_{yi} ve skutečnosti reprezentuje jedno komplexní číslo $a_{xi} + b_{xi}j$ resp. $a_{yi} + b_{yi}j$, kde j je imaginární konstanta, můžeme tyto koeficienty vyjádřit v polární reprezentaci jako amplitudu a fázové posunutí:

$$A_{xi} = \sqrt{a_{xi}^2 + b_{xi}^2}, \phi_{xi} = \arccos\left(\frac{a_{xi}}{A_{xi}}\right)$$

$$A_{yi} = \sqrt{a_{yi}^2 + b_{yi}^2}, \phi_{yi} = \arcsin\left(\frac{a_{yi}}{A_{yi}}\right)$$

Koeficienty A_{x0}, A_{y0} jsou závislé na posunutí ze zřejmých důvodů, avšak koeficienty $A_{xi}, A_{yi}, i=1 \dots n-1$ jsou invariantní vůči posunutí, jsou však závislé na změně měřítka i rotaci. Pokud však tyto koeficienty upravím takto:

$$AA_{xi} = \frac{A_{xi}}{A_{x1}}$$

$$AA_{yi} = \frac{A_{yi}}{A_{y1}}$$

$$i=2 \dots n-1$$

získáme popis vnější hranice objektu, který je invariantní i vůči změně měřítka. Poslední úpravou získáme popis, který je invariantní vůči rotaci:

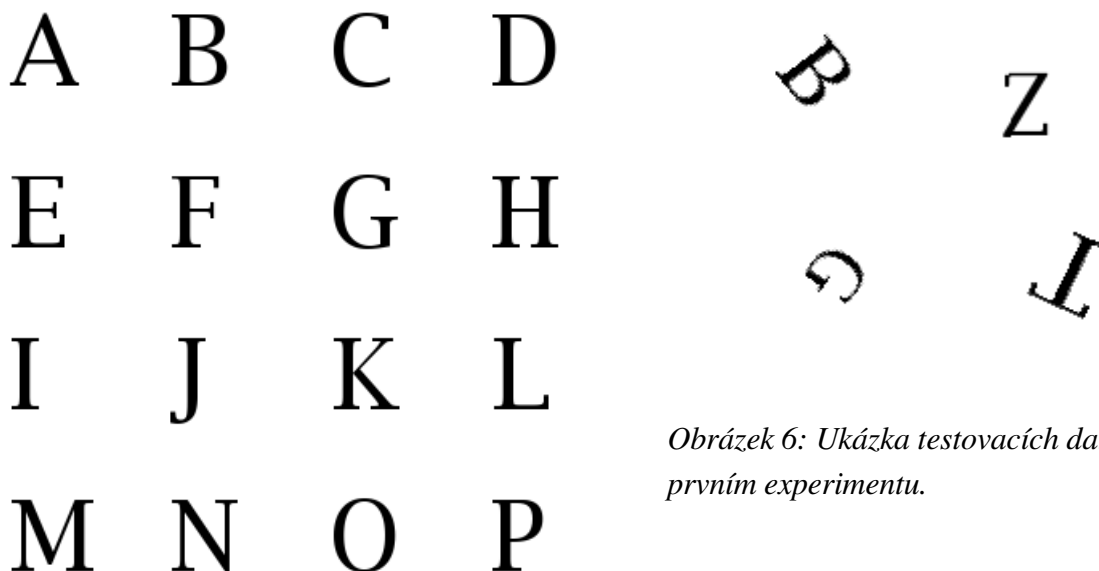
$$D_i = \sqrt{AA_{xi}^2 + AA_{yi}^2}$$

$$i=2 \dots n-1$$

Výše popsaným způsobem lze tedy pro n bodů vnější hranice získat $n-2$ koeficientů, které tuto hraniční křivku popisují takovým způsobem, který je invariantní vůči posunutí, změně měřítka i rotaci. Na obrázku 4 je vidět, že pokud použijeme jen prvních několik koeficientů $a_{xi}, b_{xi}, a_{yi}, b_{yi}, i=0 \dots k$ pro zpětnou rekonstrukci, výsledek může být velmi dobrý již pro velmi malá k . Pro klasifikaci proto použijeme nikoliv všechny možné koeficienty ale jen prvních několik ($D_i, i=2 \dots k: k < n-1$) a můžeme očekávat, že zvětšení konstanty k nebude mít vliv na zlepšení výsledků klasifikace.

Experimenty

Moje implementace klasifikátoru je napsaná v MatLabu a využívá jeho funkcionální pro provedení segmentace konstantním prahem ($h=0.5$), a nalezením hranice v černobílém obrázku (funkce `bwboundaries`). Úkolem klasifikace byla klasifikace velkých písmen anglické abecedy (26 znaků). Vzorová data představovala sada 26 obrázků (každá skupina měla právě jeden vzor) – viz obrázek 5.



Obrázek 5: Několik vzorů prvního experimentu.

Obrázek 6: Ukázka testovacích dat v prvním experimentu.

Jako font pro generování vzorových a testovacích dat posloužil patkový font Luxi regular. Vzorová data byla generována v konstantní velikosti, umístěna na stejném místě v obrázku, nenatočená. Testovací data byla generována s různou velikostí (v rozmezí 80% - 200% velikosti vzorového fontu), byla umísťována na různé pozice v obrázku, různě natočená. Písmeno bylo vždy vykresleno tak, aby bylo na obrázku celé. Soubor testovacích dat čítal 100 různých obrázků.

Klasifikace probíhala dle následujícího algoritmu:

1. pro jednotlivé vzory i si uchováme jejich Fourierovské popisovače

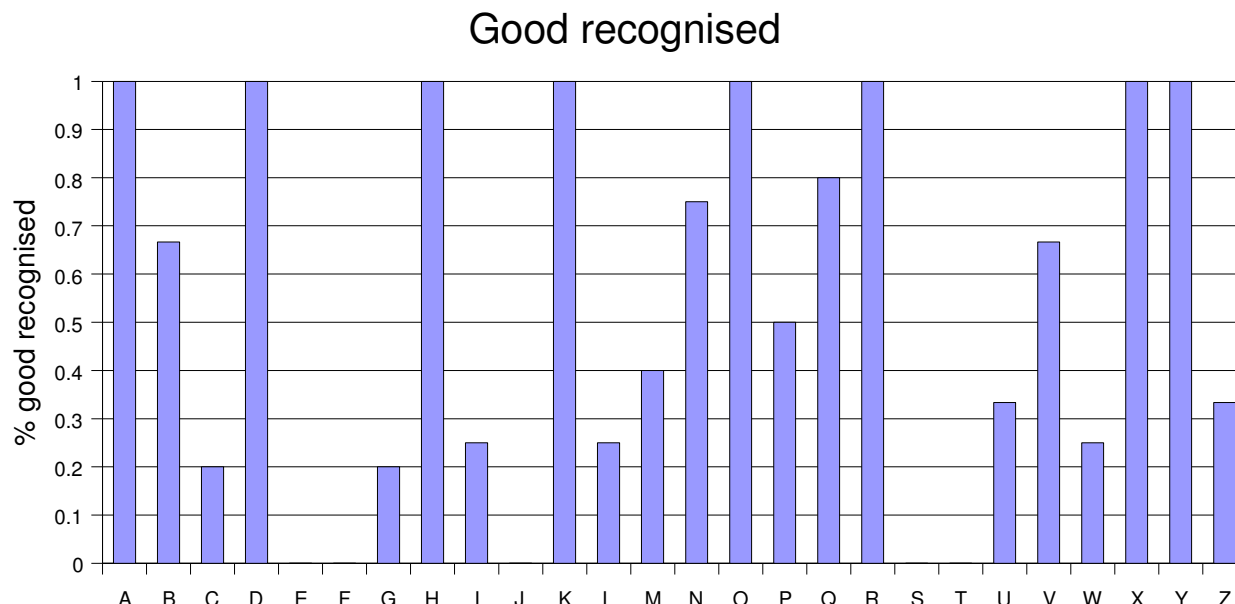
$$d_i = (D_2^i, D_3^i, \dots, D_k^i), k < n-1, \text{ k jsem zvolil jako } k=9, (d_i \text{ je příznakový vektor}).$$

2. nalezneme objekt na testovaném obrázku, zjistíme jeho souřadnice a pro ně spočteme Fourierovské popisovače $\bar{d} = (\bar{D}_2, \bar{D}_3, \dots, \bar{D}_k)$

3. Klasifikujeme objekt do třídy $x : x = \text{minarg}(\|d_i - \bar{d}\|), i=1, \dots, 26$ (tj. podle kritéria minimální vzdálenosti)

Výsledky tohoto experimentu pro jednotlivé třídy jsou na obrázku 7. Celková úspěšnost klasifikace byla 54%. Jak je vidět, pro některé třídy je klasifikace stoprocentně úspěšná (třídy A, D, H, K, O, R, X, Y). Selhání klasifikátoru u tříd N a Z resp. M a W, lze vysvětlit jejich vzájemnou podobností při otočení. Selhání v ostatních případech lze přičíst na vrub problému rasterizace, kterou umocňuje fakt, že jsem pro

generování vzorových a testovacích dat použil patkový řez písma.

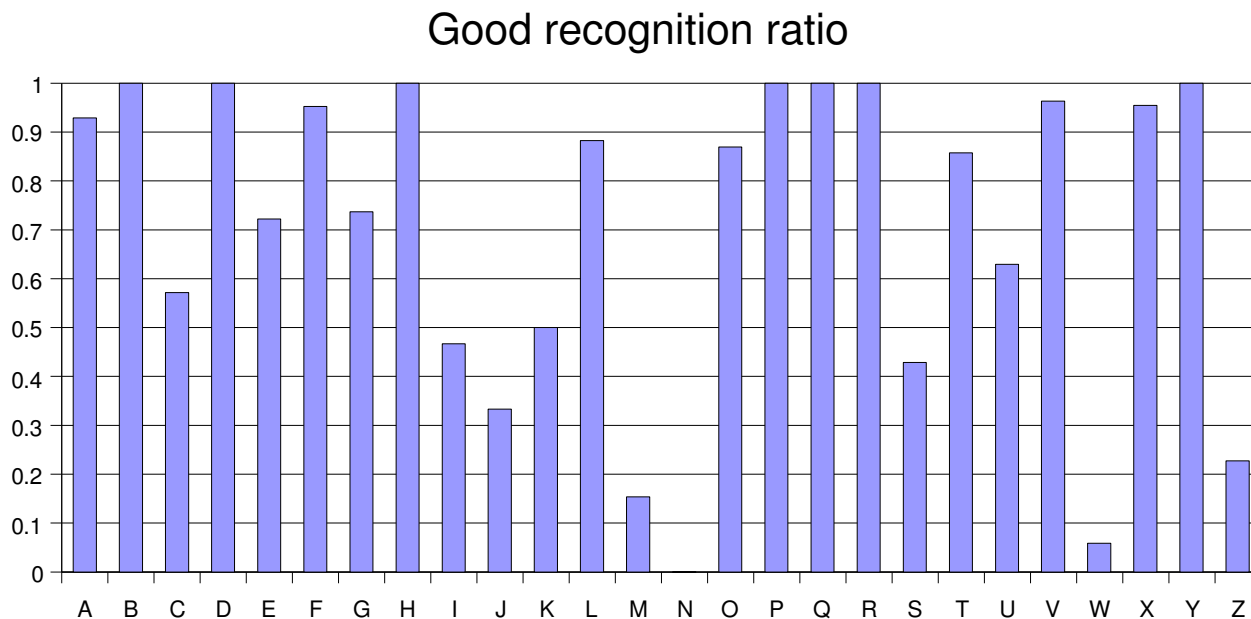


Obrázek 7: Procentuální úspěšnost klasifikace v prvním experimentu pro jednotlivé písmena.

Pro druhý experiment jsem upravil podmínky generování vzorových a testovacích dat takto:

- byl použit bezpatkový řez písma – Deja Vu Sans Bold,
- velikost testovacích písmen odpovídala 100% - 250% velikosti vzorového písma,
- byla zvětšen počet testovacích dat (na 500), aby výsledky byly přesnější.

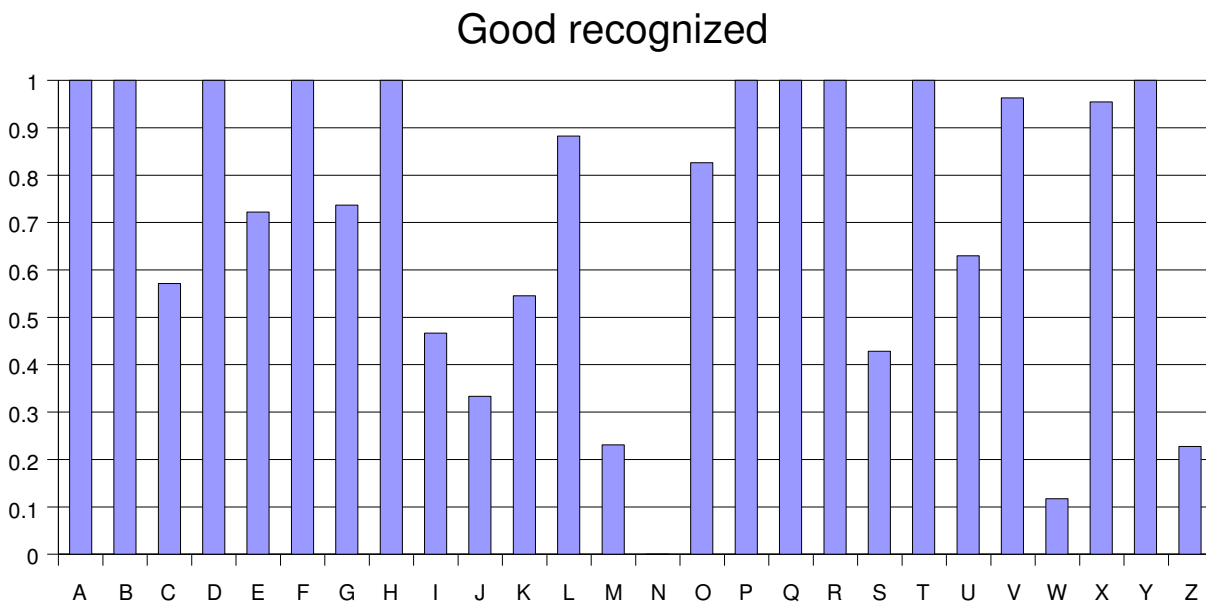
Algoritmus klasifikace i konstanta k zůstaly nezměněny. Výsledky tohoto experimentu pro jednotlivé třídy je vidět na obrázku 8, celkové úspěšnost klasifikace se zvýšila na 71%.



Obrázek 8: Procentuální úspěšnost klasifikace v druhém experimentu pro jednotlivé písmena.

Jak je vidět z obrázku, výsledky se znatelně zlepšily, celková úspěšnost klasifikace byla 71%. Množina písmen, které byly klasifikovány se 100% úspěšností obsahuje stále 8 prvků (tedy B, D, H, P, Q, R, Y), pokud nám bude stačit 90% úspěšnost, tato skupina se rozroste o další tři písmena (A, X a V).

Ve třetím experimentu jsem zkoumal, zda zvýšení konstanty k bude mít za následek zlepšení klasifikace. Doposud bylo $k=9$ (tedy příznakový vektor měl osm prvků), ve třetím experimentu jsem počet prvků příznakového vektoru zdvojnásobil, tedy nastavil konstantu $k=17$. Výsledky experimentu pro jednotlivé třídy jsou zobrazeny na obrázku 9.



Obrázek 9: Výsledky třetího experimentu pro jednotlivé třídy.

Celková úspěšnost klasifikace byla 72%, což vypadá jako nepatrné zlepšení vzhledem k tomu, že jsem zdvojnásobil příznakový vektor. Když se však podíváme na výsledky tohoto experimentu podle kritéria počtu tříd, jež jsou stoprocentně rozpoznány, dojdeme k číslu 10 (A, B, D, F, H, P, Q, R, T, Y), pokud naše kritérium, změkčíme na 90% úspěšnost rozpoznání, skupina se nám rozroste o další dvě třídy (V a X).

Závěr

Fourierovské popisovače poskytují nástroj pro popis hranice objektu invariantně vůči posunutí, změně měřítka a natočení. Tento popis lze použít pro klasifikaci. Pro experimenty jsem sestrojil jednoduchý klasifikátor. I přesto se mi podařilo akceptovatelně rozpoznat téměř polovinu skupin.

Pro zlepšení výsledků klasifikace by bylo třeba vylepšit kritérium klasifikace. Pro účely experimentů každá třída měla pouze jednoho zástupce. Lepších výsledků bychom dosáhli při použití více vzorů pro jednu třídu.