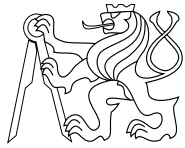




CENTER FOR
MACHINE PERCEPTION



CZECH TECHNICAL
UNIVERSITY

RESEARCH REPORT

Analýza textury sonografických obrazů difúzních procesů parenchymu štítné žlázy

Martin Švec, Radim Šára

xsvecm@cmp.felk.cvut.cz

sara@cmp.felk.cvut.cz

N^o CTU-CMP-1999-12

12.listopadu 1999

Lze získat na

<ftp://cmp.felk.cvut.cz/pub/cmp/articles/sara/Svec-TR-1999-12.pdf>

Projekt podpořen grantem MZ ČR – NB 5472-3: Analýza textury sonografických obrazů endokrinopatií a metabolických onemocnění, zejména difúzních procesů parenchymu štítné žlázy a grantem MŠ ČR – VS 96049: Centrum strojového vnímání.

Centrum strojového vnímání, Katedra kybernetiky
Fakulta elektrotechnická ČVUT

Karlovo náměstí 13, 121 35 Praha 2

fax: (02) 2435 7385, tel: (02) 2435 7637, www: <http://cmp.felk.cvut.cz>

Obsah

Cíl projektu	1
Použitá terminologie	1
1 Data použitá pro první experimenty	3
2 Texturní příznaky	4
3 Výběr příznaků	6
3.1 Fisherův lineární diskriminant	6
3.2 Stabilita klasifikátoru podle minimální vzdálenosti	9
4 Stav projektu ke dni 12.listopadu 1999	12
4.1 Shrnutí	12
4.2 Dílčí úkoly během projektu	12
4.3 Nejbližší úkoly	13
A Příloha	14
A.1 Fisherův lineární diskriminant	14
A.2 Stabilita	21
A.3 Pořadí k-tic příznaků podle jednotlivých kritérií	23
A.4 Seznam používaných funkcí	29
A.5 Účastníci projektu	30

Cíl projektu

Projekt „Analýza textury sonografických obrazů endokrinopatií a metabolických onemocnění, zejména difúzních procesů parenchymu štítné žlázy“ se zabývá podporou diagnózy štítné žlázy pomocí metod automatického rozpoznávání. Odpovědným řešitelem projektu je MuDr. Petr Sucharda z 3. interní kliniky 1. LF UK v Praze. Jména ostatních odborných spolupracovníků včetně jejich e-mailových adres jsou uvedena v příloze A.5. Tříletý projekt je podporován IGA MZ ČR. Tato zpráva shrnuje stav úvodních experimentů provedených v prvním roce řešení tříletého grantového úkolu na katedře kybernetiky FEL ČVUT. Cílem našeho projektu je nalézt klasifikátor vhodný pro klasifikaci sonografických obrazů. Na začátku projektu uvažujeme tři lehce dosažitelné diagnózy: zdravá tkáň, atrofická thyroitida (AT) a lymfocytární thyroitida (LT). Předpokládáme, že během projektu se počet diagnóz zvýší.

Data jsou získána digitalizací televizního signálu, což je kopie obrazovky sonografu, kterou sleduje expert při určování diagnózy. V našem případě jde o ultrasonografický přístroj Toshiba ECCOCEE (model SSA-340A, rok výroby 1996).

Obrazy jsou snímány za různých podmínek. Jsou jimi zvětšení (4cm nebo 6cm) a tři standardní polohy sondy vzhledem ke štítné žláze (příčně, podélně vlevo nebo podélně vpravo). Podmínky, které jsou pro automatickou klasifikaci nejvhodnější, nemusí být totožné s podmínkami nejlepšími pro vizuální diagnózu.

Použitá terminologie

Obrazem rozumíme digitalizovaná vstupní data, tak jak jsou znázorněna na obr. 1.

Oblast zájmu (zkráceně oblast) je pouze část obrazu, ve které se vyskytuje štítná žláza. Pokud budeme dále hovořit o určité klasifikaci obrázku, bude to zároveň znamenat stejnou klasifikaci oblasti a naopak.

Vzorek je část oblasti. Je to čtverec, jehož strana má délku n . Názornou ukázkou oblasti zaplněné vzorky můžete vidět na obr. 2.

Diagnóza je v našem případě to, co o oblasti prohlásí expert, tedy zda jde o zdravou tkáň, LT či AT tkáň.

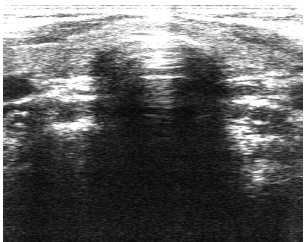
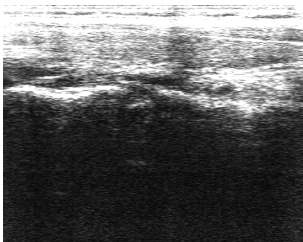
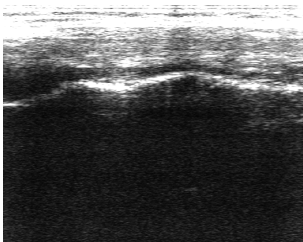
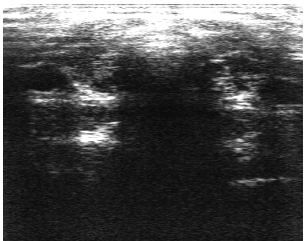
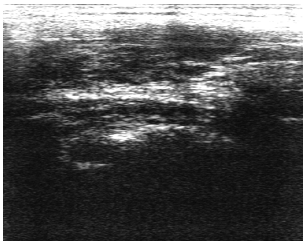
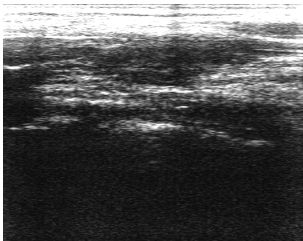
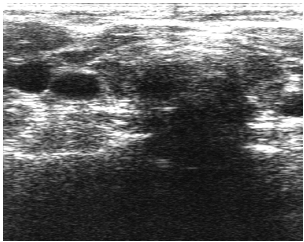
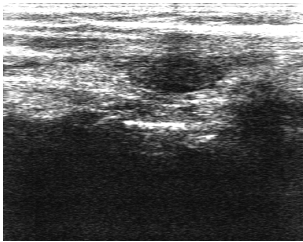
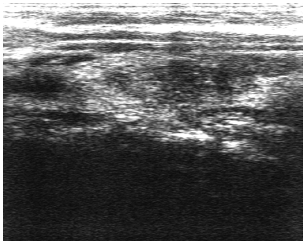
Třída je označení obrazů provedené klasifikátorem. Oblasti zařazené do jedné třídy by měly mít určitou společnou vlastnost, odlišnou od ob-

lastí jiných tříd. V některé literatuře se používá pojem *indikátor třídy*. V našem případě tedy máme zatím 3 třídy: „zdravá“, LT a AT.

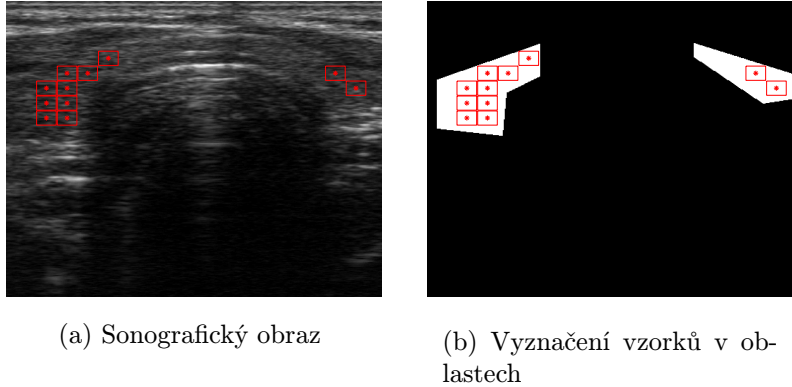
Stabilita značí procentuální četnost, se kterou se při náhodném výběru trénovací a testovací množiny projeví kombinace příznaků jako nejúspěšnější (bližší vysvětlení v kapitole 3.2).

V dalším textu již budou tyto pojmy uváděny normálním písmem.

Cílem této zprávy je vytvořit si názor na vhodnost použitých příznaků a na klasifikátor, který byl zatím v projektu realizován.

	Příčně	Podélně vpravo	Podélně vlevo
Zdravá	 martin.0	 martin.10	 martin.20
LT	 7859214726.0	 7859214726.10	 7859214726.20
AT	 366113013.34	 366113013.44	 366113013.54

Obrázek 1: Vstupní data – sonografické obrázky štítné žlázy (zvětšení 4cm) pro různé diagnózy a polohy sondy.



Obrázek 2: Oblast obrazu pokrytá sítí vzorků 21×21 .

1 Data použitá pro první experimenty

Všechny datové množiny sejmuté digitizérem DT3155 jsou rozbalené v podadresářích `/data/Sono` na linuxové síti v CMP FEL. Adresáře jsou uspořádány podle data snímání ve formátu `dd-mm-rr`. V adresáři jsou pak soubory se jmény typu `id.nn`, kde `id` je identifikační označení pacienta (číslo nebo jméno) a `nn` je číslo obrázku. Jméno každého obrázku je pak dáno zápisem `dd-mm-rr/id.nn`. Pro experimenty popsané v této zprávě byla použita data těchto pacientů:

- Zdravá tkáň – `06-04-99/martin`, `31-03-99/dana`, od obou pacientů obrázky číslo 30–39.
- LT tkáň – `31-03-99/435428093`, `30-03-99/7859214726`, od obou pacientů obrázky číslo 30–39.

Data pro třídu AT zatím nejsou k dispozici.

Množství získaných a zpracovaných dat udává tabulka 1.

Tabulka 1: Množství dat.

Diagnóza	Počet obrazů.	Počet obrazů označených expertem
zdravá tkáň	210	60
LT	153	60
AT	23	0

Podmínky sejmutí pro jednotlivé obrázky jsou:

- *06-04-99/martin.nn*
nn 0..9 poloha sondy pricne, 6 cm
nn 10..19 poloha sondy podelne vpravo, 6 cm
nn 20..29 poloha sondy podelne vlevo, 6 cm
nn 30..39 poloha sondy pricne, 4 cm
nn 40..49 poloha sondy podelne vpravo, 4 cm
nn 50..59 poloha sondy podelne vlevo, 4 cm
- *31-03-99/dana, 31-03-99/435428093, 30-03-99/7859214726* – stejně jako *06-04-99/martin*

Zatím se tedy soustředíme na podmínky poloha sondy příčně a zvětšení 4 cm, protože je důležitější zjistit nejprve vhodnou metodiku postupu a tu pak použít i pro ostatní podmínky.

2 Texturní příznaky

Z obr. 1 je patrné, že obraz má silně texturní charakter. Určení diagnózy expertem není otázkou rozlišení obrysů, nýbrž rozpoznání odlišné textury. Z textury lze extrahovat velké množství vlastností, ne všechny jsou však významné z hlediska naší klasifikační úlohy.

Rozhodli jsme se pro statistické vlastnosti. Jde o 9 Haralickových příznaků¹ počítaných na kookurenční matici (C) vzorku [HS92]. Tuto matici jsme navíc ještě normalizovali (vydělili) součtem všech jejích prvků.

Význam kookurenční matice lze zhruba vysvětlit takto: Obraz se skládá z bodů (pixelů), z nichž každý má svou jasovou hodnotu. Každý prvek matice C pak udává kolikrát spolu v celém obraze sousedí dva pixely o jasových hodnotách daných právě indexem tohoto prvku v C. Sousedství dvou pixelů je zde definováno směrem. Například pro pixel o souřadnicích $[x, y]$ a směr $[a, b]$ je určen jeho soused, který má souřadnice $[x + a, y + b]$.

My jsme se rozhodli pracovat s kookurenční maticí vypočtenou² pro směr $[1\ 0]$ (v obraze svisle, v polárním souřadnicovém systému jde o úhel -90°), který odpovídá směru šíření zvukové vlny dovnitř tkáně. Neznamená to však, že Haralickovy příznaky vypočtené pro tento směr budou těmi nejlepšími.

¹Výčet funkcí provádějících výpočet jednotlivých příznaků lze obdržet v Matlabu příkazem `help Texture`.

²Příkaz v Matlabu `C=cooccur(vzorek, [1 0])`.

V jiných směrech mohou být příznaky, které samy nebo v kombinaci s příznaky jiného směru odliší třídy lépe. Jedním z našich dalších úkolů bude tedy vyzkoušet i jiné vhodné směry, například [2 0], [1 1] nebo [0 1].

Příznaky jsou vypočteny ze vzorců uvedených v tabulce č. 2, kde

i je řádkový index prvku v matici C ,
 j je sloupcový index prvku v matici C .

$$\mu_i = \sum_{ij} iC_{ij} \quad \mu_j = \sum_{ij} jC_{ij} \quad C_i = \sum_j C_{ij}$$

$$\text{var}(i) = \sum_{ij} (i - \mu_i)^2 C_{ij} \quad \text{var}(j) = \sum_{ij} (j - \mu_j)^2 C_{ij}$$

Na čísla příznaků uvedená v této tabulce se budeme odkazovat v dalších částech textu.

Tabulka 2: Příznaky.

Číslo	Název	Výpočet
H1	texture cluster tendency	$\sum_{ij} (i - \mu_i + j - \mu_j)^2 C_{ij}$
H2	texture entropy	$-\sum_{ij} C_{ij} \log C_{ij}$
H3	texture contrast	$\sum_{ij} i - j C_{ij}$
H4	texture correlation	$\frac{\sum_{ij} (i - \mu_i)(j - \mu_j) C_{ij}}{\sqrt{\text{var}(i)\text{var}(j)}}$
H5	texture homogeneity	$\sum_{ij} \frac{C_{ij}}{1 + i - j }$
H6	texture inverse difference moment	$\sum_{ij, i \neq j} \frac{C_{ij}}{ i - j }$
H7	maximum texture probability	$\max_{ij} C_{ij}$
H8	texture probability of run length of 2	$\sum_i \frac{(C_i - C_{ii})^2 C_{ii}}{C_i^2}$
H9	uniformity of texture energy	$\sum_{ij} C_{ij}^2$

Pomocí několika funkcí, které byly v Matlabu vytvořeny³, jsme získali datovou strukturu vhodnou ke konečným výpočtům či jako vstup do klasifikátoru. Délku strany čtverce jsme vzhledem k velikostem oblastí zvolili 21 (obr. 2) a 31 pixelů a podle dalších výsledků se mezi těmito hodnotami rozhodneme.

3 Výběr příznaků

3.1 Fisherův lineární diskriminant

Nejprve se pokusíme seřadit příznaky podle toho, jakou nám poskytují míru informace o vzorcích z hlediska rozdělení vzorků do tříd. K tomu slouží jednoduchá metoda – počítání Fisherova lineárního diskriminantu (F). Můžeme si ji snadno představit v příznakovém prostoru. Jsou zde určité shluky bodů (body odpovídají vzorkům), které se mohou různě překrývat. Každý shluk bodů znázorňuje vzorky příslušící k jedné třídě. Ideální stav je ten, kdy příznaky (souřadnice bodů v příznakovém prostoru) jsou vybrány tak, aby shluky odpovídající všem třídám byly navzájem oddělené. My předpokládáme, že shluky mají normální rozdělení hustoty. Potom je můžeme jednoznačně reprezentovat střední hodnotou a rozptylem (pro 1D, ve vícedimenzionálním prostoru počítáme místo rozptylu kovarianční matici). Je zřejmé, že užitečnost příznaků je tím větší, čím menší budou rozptyly uvnitř shluků a čím více budou jednotlivé shluky od sebe vzdáleny (čím větší bude rozptyl mezi shluky). Zajímají nás tedy ty příznaky, pro něž je hodnota poměru

$$\frac{\text{rozptyl mezi třídami}}{\text{rozptyl uvnitř tříd}}$$

větší než pro ostatní. Rozptyl uvnitř jednotlivých tříd může být různý, proto pro zajištění mezního případu použijeme poměr

$$\frac{\text{rozptyl mezi třídami}}{\text{větší z rozptylů uvnitř tříd}}$$

Rozptyl mezi třídami jsme nepočítali jako kovarianci středních hodnot tříd [KBCJ80], ale jako kovarianci všech vzorků, což je v našem případě přesnější. Obecně předpokládejme, že máme data roztríděná do n tříd. Každá třída je reprezentována právě maticí \mathbf{X}_i , tvořenou po sloupcích příznakovými vektory.

³Podrobnější popis naleznete v kapitole 4 – *Seznam používaných funkcí*.

Zaveďme označení:

$$\mathbf{A} = \text{cov}([\mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3 \ \dots \ \mathbf{X}_n]) , \quad (1)$$

$$\mathbf{B}_i = \text{cov}(\mathbf{X}_i) , \quad (2)$$

$$\lambda_i = \max(\text{eig}(\mathbf{B}_i^{-1}\mathbf{A})) . \quad (3)$$

Pak Fisherův lineární diskriminant

$$F = \sqrt{\min_{i=1}^n \lambda_i} . \quad (4)$$

Odvození vztahu (4) je uvedeno v příloze A.1.

Pro názornější objasnění hodnot, kterých může Fisherův lineární diskriminant nabývat, provedeme následující úvahu. Definujme

$$p = \frac{k_i}{\sum_{j=1}^N k_j} ,$$

kde k_i je počet vektorů v matici \mathbf{X}_i ,
 $\sum_{j=1}^N k_j$ je počet vektorů v $[\mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3 \ \dots \ \mathbf{X}_n]$.

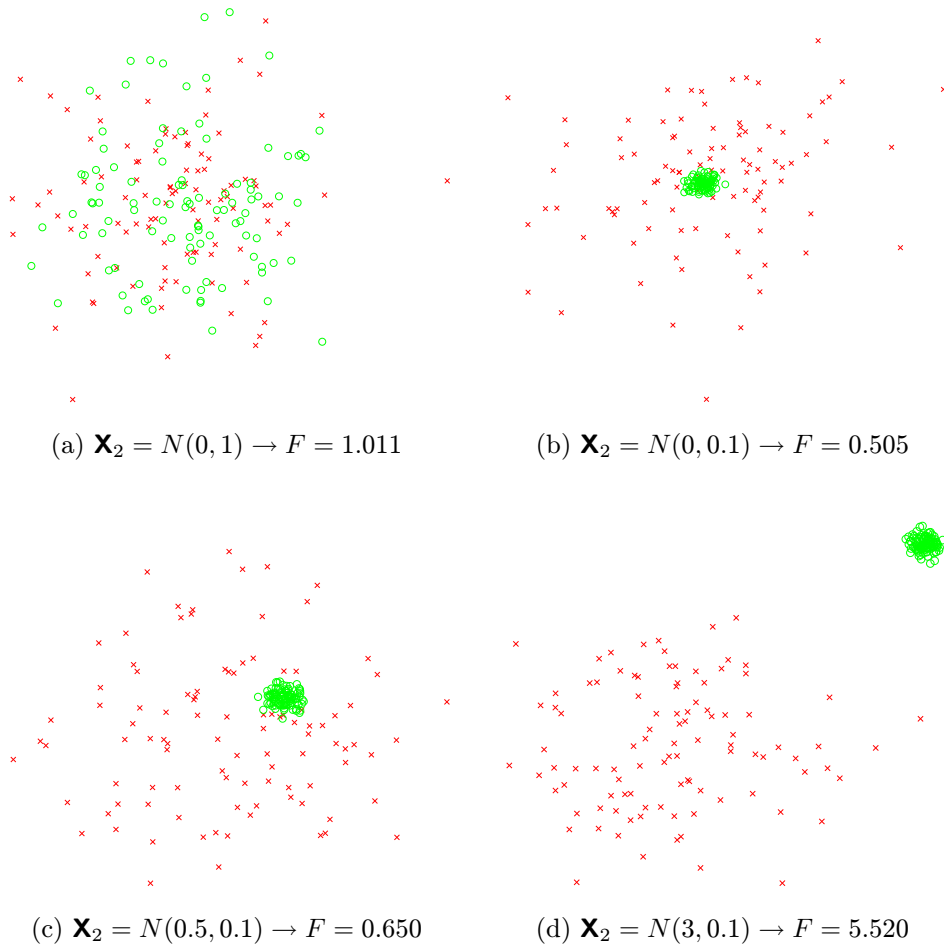
Parametr p nabývá hodnot z intervalu $< 0, 1 >$.

Experimentálně bylo zjištěno:

1. Jestliže třídy jsou dokonale vnořené (mají stejné střední hodnoty a rozptyl jedné ze tříd je nulový), pak $F = p$.
2. Jestliže třídy jsou identické (mají stejné střední hodnoty a stejný rozptyl (kovarianční matice)), pak $F = 1$.
3. Jestliže třídy jsou dokonale separovatelné, pak $F = \infty$.
4. Jestliže třídy jsou pouze vnořené, pak $p < F < 1$.
5. Jestliže třídy jsou pouze separovatelné, pak $1 < F < \infty$. Čím větších hodnot F nabývá, tím lépe jsou separovatelné.

Názorný příklad vzájemných poloh tříd ve dvou-dimenzionálním prostoru je uveden na obr. 3.

Podle hodnoty Fisherova lineárního diskriminantu pro $n = 2$ byly seřazeny jednotlivé kombinace příznaků. V tabulce č. 3 je uveden interval, ve



Obrázek 3: Základní vzájemné polohy dvou tříd ve 2D prostoru, parametr $p = 0.5$, $k_1 = k_2 = 100$, $\mathbf{X}_1 = N(0, 1)$ (má normální rozdělení, střední hodnota je nulová, rozptyl je roven jedné).

kterém se pohybuje F pro všechny kombinace příznaků vzhledem k parametru N , kde N značí počet příznaků ($N = 1$ – samotné příznaky, $N = 2$ – dvojice příznaků, atd.).

Protože se snažíme nalézt klasifikátor, který by pro správnou funkci potřeboval minimální počet příznaků, a protože můžeme snadno zobrazit a vizuálně kontrolovat maximálně 3D prostor, budeme nadále uvažovat pouze kombinace příznaků pro $N \leq 3$. Jejich podrobný výčet je spolu s hodnotou F uveden v příloze A.3 v tabulkách číslo 1, 2, 3. Pro názornost je uveden příznakový prostor pro nejlepší a nejhorší příznaky ve 2D a ve 3D a pro vzorky 21×21 a 31×31 (obr. 6 až obr. 9).

Tabulka 3: Zjištěný interval pro Fisherův lineární diskriminant podle počtu příznaků (N).

N	Vzorky 21×21 Interval F	Vzorky 31×31 Interval F
1	$\langle 0.7852, 0.9659 \rangle$	$\langle 0.7892, 1.0950 \rangle$
2	$\langle 0.8850, 1.1936 \rangle$	$\langle 0.8748, 1.4101 \rangle$
3	$\langle 0.9033, 1.2474 \rangle$	$\langle 0.9130, 1.6036 \rangle$
4	$\langle 0.9522, 1.2719 \rangle$	$\langle 1.0333, 1.6703 \rangle$
5	$\langle 0.9956, 1.2759 \rangle$	$\langle 1.1828, 1.7026 \rangle$
6	$\langle 1.0254, 1.2781 \rangle$	$\langle 1.2884, 1.7158 \rangle$
7	$\langle 1.1635, 1.2794 \rangle$	$\langle 1.3826, 1.7224 \rangle$
8	$\langle 1.2453, 1.2798 \rangle$	$\langle 1.6013, 1.7277 \rangle$
9	1.2801	1.7278

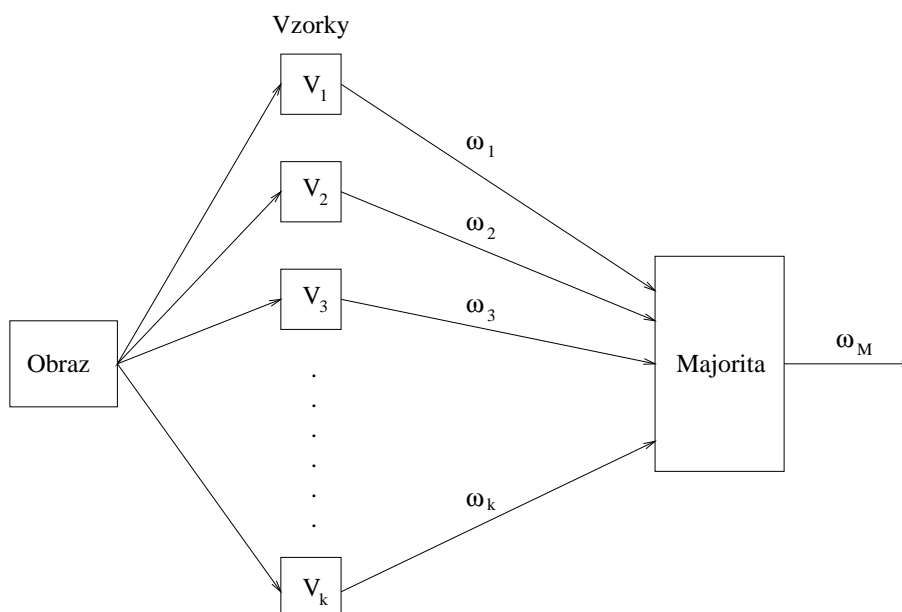
Metoda Fisherova lineárního diskriminantu je vhodnější spíše pro rozhodování mezi dvěma třídami, což nám ale nebrání ji použít i pro více tříd. Pokud se totiž tímto způsobem podaří nalézt takovou množinu příznaků, že všechny dvojice tříd budou vzájemně dobře rozlišitelné, dospějeme k dobrým výsledkům i v klasifikaci.

3.2 Stabilita klasifikátoru podle minimální vzdálenosti

Nyní již z předchozího víme, jaké příznaky (kombinace příznaků) by měly být dobré pro naši klasifikaci. Tyto výsledky se pokusíme ověřit na klasifikaci do zatím dostupných tříd, tedy „zdravé“ a LT.

Všechny vzorky byly rozděleny náhodně na dvě poloviny – na testovací a trénovací množinu, přičemž vzorky z jednoho obrazu se mohly vyskytovat pouze v jedné z množin. Z trénovací množiny jsme určili střední hodnoty obou tříd. Byl naprogramován klasifikátor podle minimální vzdálenosti od těchto středních hodnot. Vzdálenost je euklidovská a není vážena kovarianční maticí. Klasifikátor pracuje v podstatě na dvou úrovních. V první úrovni provádí třídění vzorků a v druhé třídění obrazů. Obrazu je přiřazena taková třída, do které patří většina jeho vzorků (majorita, viz obr. 4). Výstupem je pak úspěšnost daná porovnáním testovací množiny a výsledku druhé úrovně klasifikátoru. Při naprosté shodě klasifikace s testovací množinou by tedy byla úspěšnost 100%.

Zjistili jsme, že pro stejně velké trénovací a testovací množiny, do kterých mnohokrát za sebou vybíráme náhodně přibližně stejný počet „zdravých“



Obrázek 4: Rozhodování o třídě obrazu podle třídy, do které je přiřazena většina vzorků (ω je označení třídy).

i LT vzorků, se úspěšnost klasifikátoru pohybuje mezi 70 % – 100 %.

Nabízejí se zde tři možné příčiny:

1. Je k dispozici zatím jen malé množství dat. To by znamenalo, že ve třídách je střední hodnota zatím získaných vzorků poměrně hodně vzdálena od střední hodnoty, kterou bychom získali z nekonečného počtu vzorků.
2. V oblasti s určitou diagnózou předpokládáme pouze výskyt vzorků patřících do té samé třídy. Je však možné, že oblast třídy LT obsahuje také vzorky zdravé, a dokonce jich tu může být i většina. Potom při výběru špatných dat do trénovací množiny je možné, že klasifikátor zařadí „zdravé“ vzorky z LT oblasti také do třídy LT. Tím by se v podstatě kompenzovala chyba chybou a je tak vysvětlen i případ, kdy náhodně vyšla úspěšnost klasifikace 100 %.
3. Gausovský model popsaný střední hodnotou a rozptylem (kovarianční maticí) není dostatečný.

Vzhledem k tomu, že ani jednu z těchto příčin nelze ihned odstranit, zaměřili jsme se zatím na zjištění nejlepší kombinace tří příznaků (příznakový prostor dimenze 3D). Postupně jsme pro vzorky 21×21 a 31×31 provedli

1000 klasifikací na různých trénovacích a testovacích množinách a to pro všechny kombinace příznaků ve 3D. Zjišťovali jsme, kolikrát se v daném počtu pokusů projeví každá kombinace jako nejuspěšnější. Tuto relativní četnost pak nazýváme **stabilitou klasifikátoru vzhledem k testovací množině** (zkráceně **stabilita**) a udáváme ji v procentech (viz terminologie na začátku této zprávy). Její maximální hodnoty jsou uvedeny v tabulce 4. Příznakový prostor pro trojici příznaků s největší stabilitou znázorňuje obr. 10.

Podrobnější informace o všech kombinacích lze opět najít v příloze – tabulka 4. Zde si můžeme všimnout rozporu mezi pořadím podle Fisherova lineárního diskriminantu a podle stability. Je to dáno tím, že rozložení vzorků uvnitř shluků v příznakovém prostoru nemá gausovský charakter. Kdyby tomu tak bylo, tak by rozptyl velice dobře charakterizoval tento shluk a pořadí příznaků podle Fisherova lineárního diskriminantu by se blížilo pořadí podle stability klasifikátoru.

Tabulka 4: Kombinace příznaků s největší stabilitou.

Vzorek	Kombinace	Stabilita [%]
21 × 21	H2 H4 H9	42.7
31 × 31	H5 H7 H9	27.6

Pro kombinace z tabulky 4 jsme provedli opět 1000 klasifikací na různých trénovacích a testovacích množinách a zjistili průměrnou úspěšnost. Zaměřili jsme se na jednotlivé obrázky a zjistili, kolikrát byl v daném počtu klasifikací zařazen obrázek do špatné třídy. Tuto relativní četnost, pak nazýváme **chybovost klasifikátoru vzhledem k obrázku** (zkráceně **CH**, udáváme ji opět v procentech) a je uvedena v tabulce 5 (symbol 30-03-99 je zde označen jako A, 06-04-99 jako B).

Vidíme, že nejčastěji jsou chybně klasifikovány některé obrázky patřící ke třídě LT, což podporuje hypotézu, že i v těchto obrázcích existují vzorky s diagnózou „zdravý“. Dokládá to i obr. 11. Ten jsme získali zvětšením příznakového prostoru pro kombinace příznaků z tabulky 5 (srovnejte s obr. 10). Navíc jsou v něm ještě zvýrazněny vzorky obrazu s největší chybností (30-03-99/7859214726.31 pro obr. 11(a) a 30-03-99/7859214726.39 pro obr. 11(b)). V příznakovém prostoru jsou tyto vzorky poměrně náhodně rozmístěny ve shluku jak LT, tak i „zdravé“ třídy.

Tabulka 5: Chybně klasifikované obrázky.

Vzorky 21 × 21		Vzorky 31 × 31	
Kombinace příznaků: H2 H4 H9		Kombinace příznaků: H5 H7 H9	
Průměrná úspěšnost = 85.82 %		Průměrná úspěšnost = 89.115 %	
Jméno - diagnóza	CH [%]	Jméno - diagnóza	CH [%]
A/7859214726.31 - LT	52.4	A/7859214726.39 - LT	56.4
A/7859214726.32 - LT	48.2	A/7859214726.33 - LT	49.2
A/7859214726.34 - LT	48.1	A/7859214726.32 - LT	48.7
A/7859214726.30 - LT	46.1	A/7859214726.38 - LT	47.8
A/7859214726.33 - LT	45.7	A/7859214726.36 - LT	14.3
A/7859214726.39 - LT	24.5	A/7859214726.30 - LT	0.7
A/7859214726.38 - LT	13.4	B/martin.37 - zdravý	0.3
A/7859214726.37 - LT	2.8	A/7859214726.31 - LT	0.2
A/7859214726.36 - LT	1.5	B/martin.30 - zdravý	0.1
A/7859214726.35 - LT	0.9		

4 Stav projektu ke dni 12.listopadu 1999

- Kombinace příznaků pro příznakový prostor dimenze 1, 2 a 3 byly seřazeny podle velikosti Fisherova lineárního diskriminantu.
- Byl vyzkoušen klasifikátor podle minimální vzdálenosti od střední hodnoty.

4.1 Shrnutí

Z experimentů, které byly dosud provedeny, vyplývají následující závěry:

- Variabilita třídy LT je větší než variabilita třídy „zdravá“.
- Uspořádání příznaků podle Fisherova lineárního diskriminantu a podle stability se liší. Pro větší množiny dat a pro spolehlivější klasifikátory předpokládáme postupné sblížení tohoto pořadí.
- Separovatelnost tříd v příznakovém prostoru je malá, třídy jsou do sebe vnořené.

4.2 Dílčí úkoly během projektu

- Je třeba shromáždit co možná největší množství dat, jak zdravých, tak i nemocných tkání.

- Ve všech obrázcích vyznačit oblast zájmu (štítnou žlázu) a tu pro unifikaci dat pokrýt sítí čtverců s jednotnou plochou. Každý čtverec pak bude považován za samostatný vzorek dat.
- Určit za pomoci experta, zda pro oblast s určitým onemocněním nesou také všechny její čtverce charakteristiku tohoto onemocnění. V případě, že tomu tak není, bude třeba aby expert označil v každém obrázku čtverce s nemocnou tkání. Určitá třída pak bude přiřazena jednomu obrázku podle toho, kolik v něm bude zastoupeno čtverců „zdravých“, LT či AT.
- Je třeba určit optimální velikost strany čtverce.
- Pro každý čtverec vypočítat vhodné příznaky a z nich vybrat takové, pro které vzniknou v příznakovém prostoru pokud možno snadno oddělitelné shluky.
- Postupně zkoušet klasifikátory od nejjednodušších po složitější.

4.3 Nejbližší úkoly

1. Pokud by se zjistilo, že v jednom obrázku například s diagnózou LT by bylo srovnatelné množství vzorků s vlastnostmi třídy „zdravý“ a třídy LT, pak bude třeba napsat funkci, pomocí které expert tyto čtverce označí. Tím by se s velkou pravděpodobností zvýšila úspěšnost jakéhokoli klasifikátoru.
2. Zpracovat příznaky pro jiný směr v kookurenční matici C .
3. Zkoušet další klasifikátory pro klasifikaci jednotlivých vzorků.
- 3a. Napsat takovou funkci, která z jednotlivých oklasifikovaných vzorků nejlépe určí příslušnost celého obrázku k jedné z tříd.

Reference

- [HS92] Robert M. Haralick and Linda G. Shapiro. *Computer and Robot Vision*. Addison-Wesley Publishing Company, 1992.
- [KBCJ80] Zdeněk Kotek, Ivan Brůha, Václav Chalupa, and Jan Jelínek. *Adaptivní a učící se systémy*. SNTL, 1980.

A Příloha

A.1 Fisherův lineární diskriminant

Jak již bylo uvedeno v kapitole 3.1, vycházíme ze vztahů (1) a (2).

Rozptylu mezi třídami ve směru určeném vektorem \mathbf{x} odpovídá zápis $\mathbf{x}^T \mathbf{A} \mathbf{x}$, kde \mathbf{x} je směrový vektor, pro který platí $\|\mathbf{x}\| > 0$. Podobně rozptyl uvnitř třídy můžeme nahradit zápisem $\mathbf{x}^T \mathbf{B}_i \mathbf{x}$. Naším cílem je nalézt takový směr daný vektorem \mathbf{x} , pro který bude maximální poměr

$$\frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{B}_i \mathbf{x}} . \quad (5)$$

1. Matice \mathbf{B}_i je pozitivně definitní a symetrická.
Lze ji rozložit na

$$\mathbf{B}_i = \mathbf{B}_i^{\frac{1}{2}} \mathbf{B}_i^{\frac{1}{2}} . \quad (6)$$

2. Definujme substituci

$$\mathbf{y} = \mathbf{B}_i^{\frac{1}{2}} \mathbf{x} \quad \longrightarrow \quad \mathbf{x} = \mathbf{B}_i^{-\frac{1}{2}} \mathbf{y} . \quad (7)$$

3. Potom

$$\begin{aligned} \max_{\mathbf{x}} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{B}_i \mathbf{x}} &= \max_{\mathbf{x}} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{B}_i^{\frac{1}{2}} \mathbf{B}_i^{\frac{1}{2}} \mathbf{x}} = \max_{\mathbf{y}} \frac{(\mathbf{B}_i^{-\frac{1}{2}} \mathbf{y})^T \mathbf{A} \mathbf{B}_i^{-\frac{1}{2}} \mathbf{y}}{(\mathbf{B}_i^{-\frac{1}{2}} \mathbf{y})^T \mathbf{B}_i^{\frac{1}{2}} \mathbf{B}_i^{\frac{1}{2}} \mathbf{B}_i^{-\frac{1}{2}} \mathbf{y}} = \\ &= \max_{\mathbf{y}} \frac{\mathbf{y}^T \mathbf{B}_i^{-\frac{1}{2}T} \mathbf{A} \mathbf{B}_i^{-\frac{1}{2}} \mathbf{y}}{\mathbf{y}^T \mathbf{B}_i^{-\frac{1}{2}T} \mathbf{B}_i^{\frac{1}{2}} \mathbf{B}_i^{\frac{1}{2}} \mathbf{B}_i^{-\frac{1}{2}} \mathbf{y}} = \max_{\mathbf{y}} \frac{\mathbf{y}^T \mathbf{B}_i^{-\frac{1}{2}} \mathbf{A} \mathbf{B}_i^{-\frac{1}{2}} \mathbf{y}}{\mathbf{y}^T \mathbf{B}_i^{-\frac{1}{2}} \mathbf{B}_i^{\frac{1}{2}} \mathbf{B}_i^{\frac{1}{2}} \mathbf{B}_i^{-\frac{1}{2}} \mathbf{y}} = \\ &= \max_{\mathbf{y}} \frac{\mathbf{y}^T \mathbf{B}_i^{-\frac{1}{2}} \mathbf{A} \mathbf{B}_i^{-\frac{1}{2}} \mathbf{y}}{\mathbf{y}^T \mathbf{y}} . \end{aligned} \quad (8)$$

4. Výraz (8) je definicí maximálního vlastního čísla matice $\mathbf{B}_i^{-\frac{1}{2}} \mathbf{A} \mathbf{B}_i^{-\frac{1}{2}}$ (označme ho λ_i), takže

$$\max_{\mathbf{x}} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{B}_i \mathbf{x}} = \max(\text{eig}(\mathbf{B}_i^{-\frac{1}{2}} \mathbf{A} \mathbf{B}_i^{-\frac{1}{2}})) = \lambda_i . \quad (9)$$

5. Vzhledem k tomu, že \mathbf{A} i $\mathbf{B}_i^{-\frac{1}{2}}$ jsou symetrické a součin symetrických matic je opět symetrická matice, platí

$$\mathbf{A}\mathbf{B}_i^{-\frac{1}{2}} = (\mathbf{A}\mathbf{B}_i^{-\frac{1}{2}})^T = \mathbf{B}_i^{-\frac{1}{2}T}\mathbf{A}^T = \mathbf{B}_i^{-\frac{1}{2}}\mathbf{A} \quad (10)$$

a vzhledem ke vztahu (6) můžeme zapsat výraz (9) jako

$$\lambda_i = \max(\text{eig}(\mathbf{B}_i^{-1}\mathbf{A})) . \quad (11)$$

Pro Fisherův lineární diskriminant platí

$$F = \sqrt{\min_{i=1}^n \lambda_i} . \quad (12)$$

Výběr minima ze všech vlastních čísel λ_i znamená výše uvedené zajištění mezního případu, kdy rozptyly uvnitř shluků nejsou stejné. Fisherův lineární diskriminant je odmocněn, protože kovariance dává druhé mocniny rozptylů.

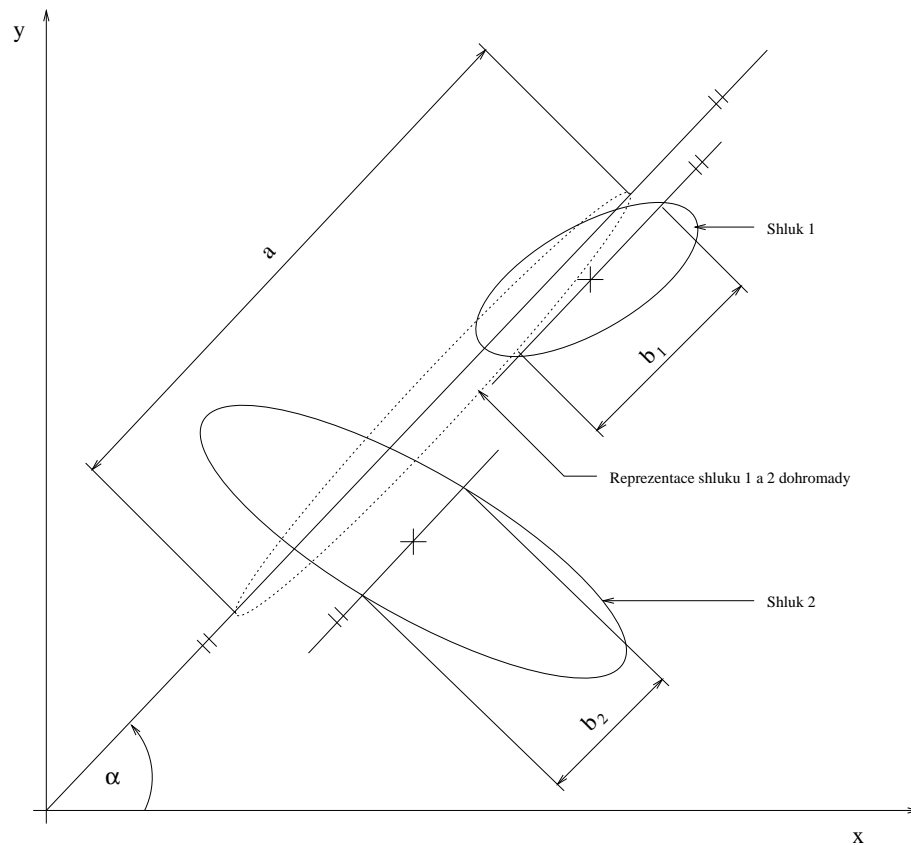
Pro lepší představu součinu $\mathbf{B}_i^{-1}\mathbf{A}$ ze vztahu (11) uvažujme dva shluky, jejichž vektory středních hodnot označíme $\bar{\mathbf{x}}_1$ a $\bar{\mathbf{x}}_2$. Pak z výrazů (1) a (2) lze odvodit, že

$$\mathbf{B}_1^{-1}\mathbf{A} = \frac{n_1}{n_1 + n_2} + \frac{n_2}{n_1 + n_2} \mathbf{B}_1^{-1}\mathbf{B}_2 + \frac{1}{4} \mathbf{B}_1^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^T, \quad (13)$$

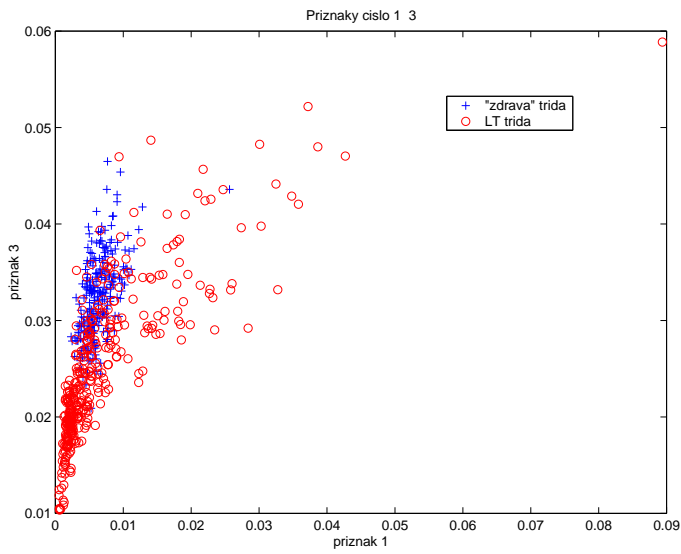
kde n_1, n_2 je počet měření v jednotlivých shlucích (měření je zde reprezentováno vektorem, jehož délku udává počet příznaků). Snadno můžeme odvodit jednoduché případy:

1. Jestliže jsou shluky identické ($\bar{\mathbf{x}}_1 = \bar{\mathbf{x}}_2$, $\mathbf{B}_1 = \mathbf{B}_2$), pak $\mathbf{B}_1^{-1}\mathbf{A} = 1$ (neplatí naopak).
2. Jestliže shluky nejsou identické, $\bar{\mathbf{x}}_1 = \bar{\mathbf{x}}_2$ a $\mathbf{B}_2 = 0$, pak $\mathbf{B}_1^{-1}\mathbf{A} = \frac{n_1}{n_1+n_2}$.
3. Jestliže $\mathbf{B}_1 \rightarrow \infty$, pak $\mathbf{B}_1^{-1}\mathbf{A} = \frac{n_1}{n_1+n_2}$.
4. Jestliže $\mathbf{B}_2 \rightarrow 0$ a rozdíl $\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2$ je jedním z bázových vektorů \mathbf{B}_1^{-1} , pak $\mathbf{B}_1^{-1}\mathbf{A} = \frac{n_1}{n_1+n_2}$.

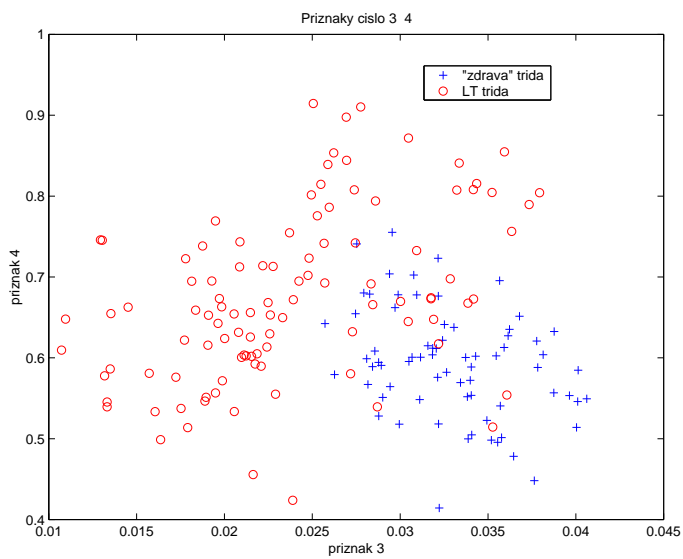
Význam směru daného maximálním vlastním číslem je patrný na obr. 5. Úhel α zde určuje takový směr, v němž bude minimum z poměrů $\{\frac{a}{b_1}, \frac{a}{b_2}\}$ maximální. Rozptyl mezi třídami je zde označen a , rozptyly uvnitř tříd b_1 , b_2 .



Obrázek 5: Znázornění rozptylů daných úhlem α ve dvou-dimenzionálním prostoru.

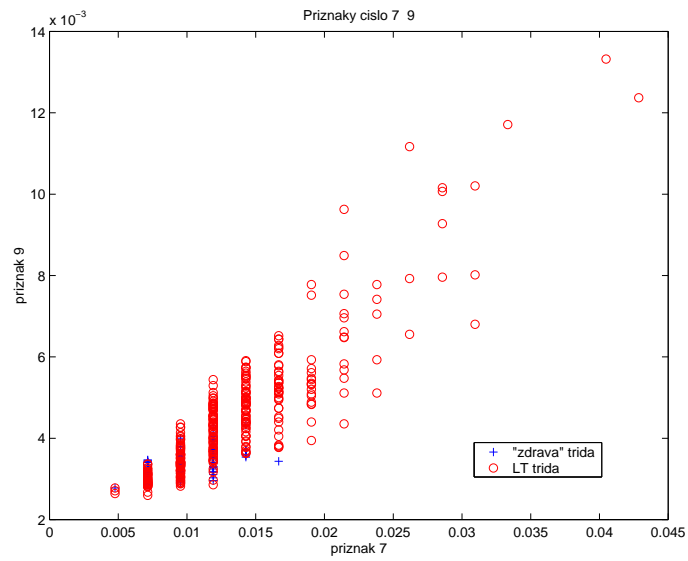


(a) Vzorky 21×21 , kombinace H1 H3

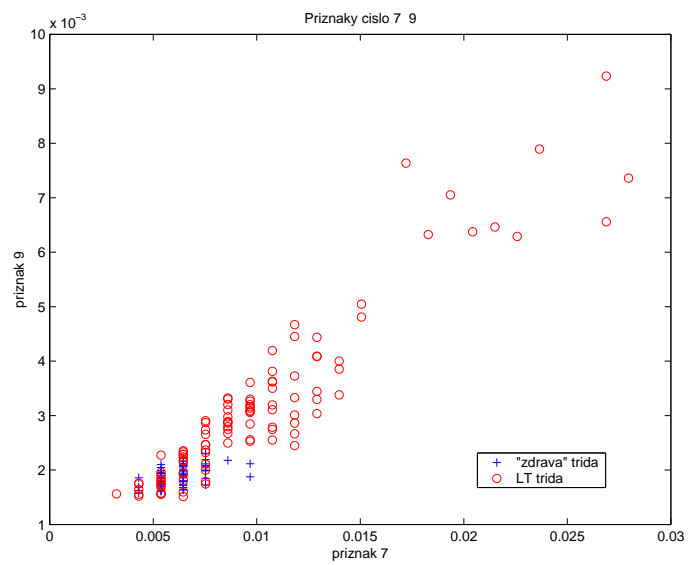


(b) Vzorky 31×31 , kombinace H3 H4

Obrázek 6: Příznakový prostor pro nejlepší dvojici příznaků podle Fisherova lineárního diskriminantu.

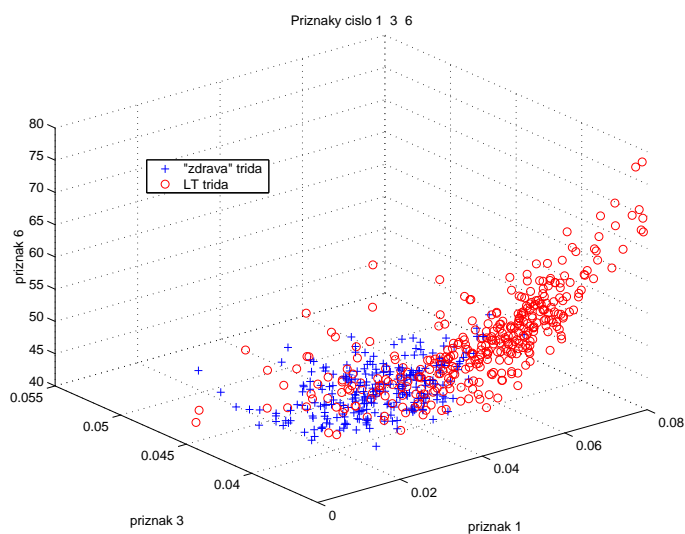


(a) Vzorky 21×21 , kombinace H7 H9

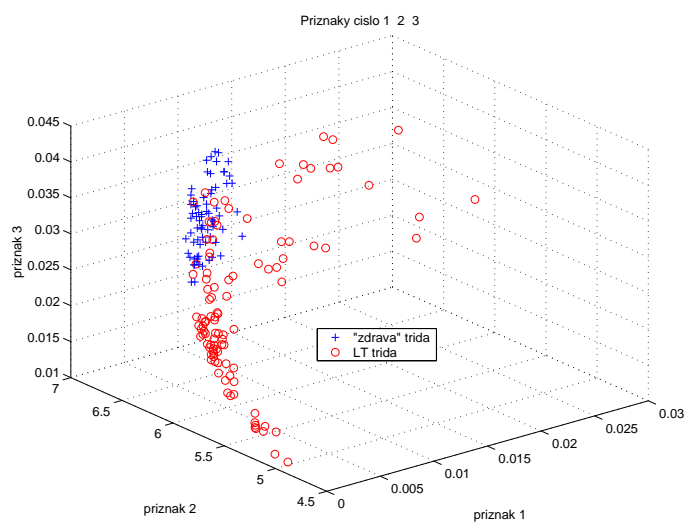


(b) Vzorky 31×31 , kombinace H7 H9

Obrázek 7: Příznakový prostor pro nejhorší dvojici příznaků podle Fisherova lineárního diskriminantu.

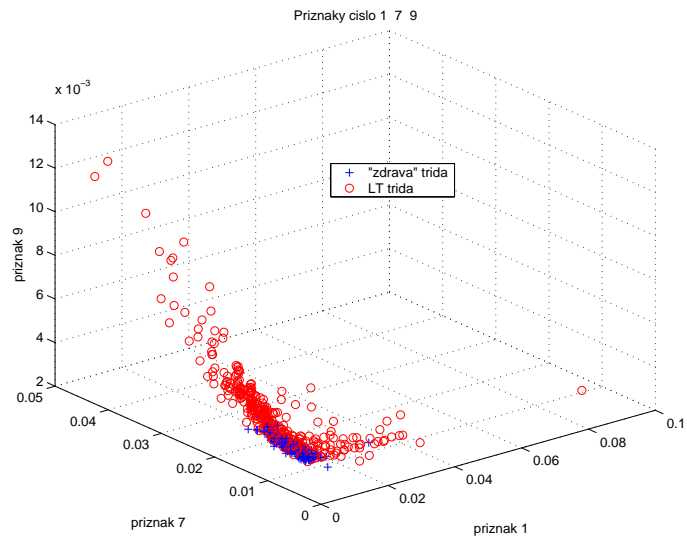


(a) Vzorok 21 × 21, kombinace H1 H3 H6

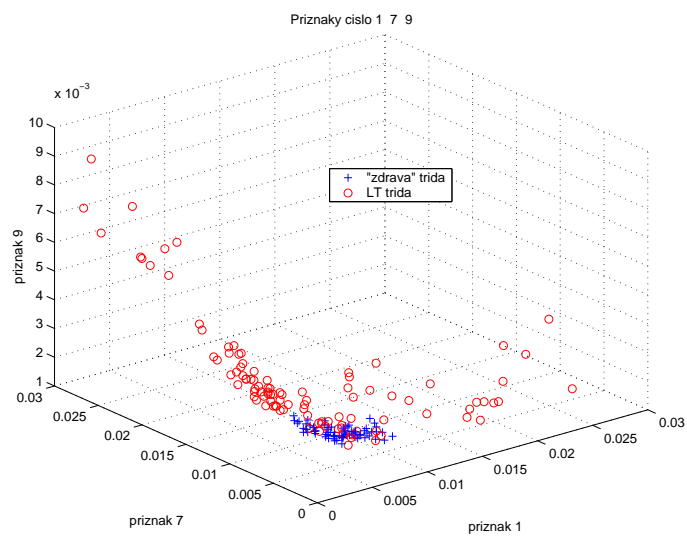


(b) Vzorok 31 × 31, kombinace H1 H2 H3

Obrázek 8: Příznakový prostor pro nejlepší trojici příznaků podle Fisherova lineárního diskriminantu.



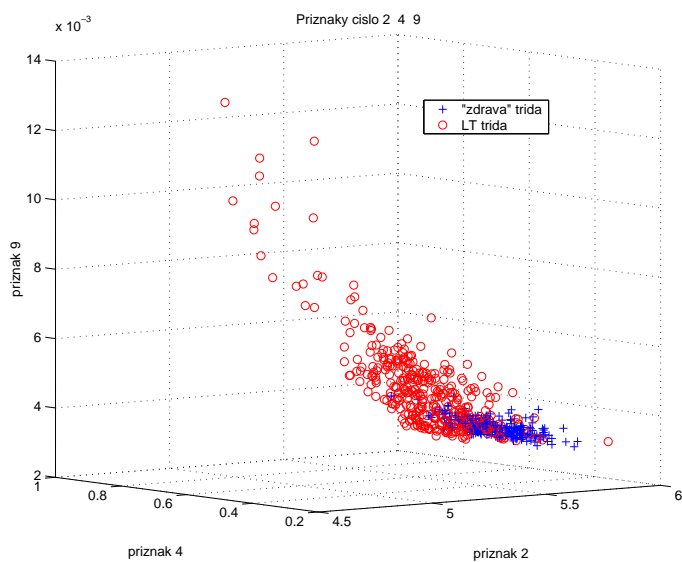
(a) Vzorok 21 × 21, kombinace H1 H7 H9



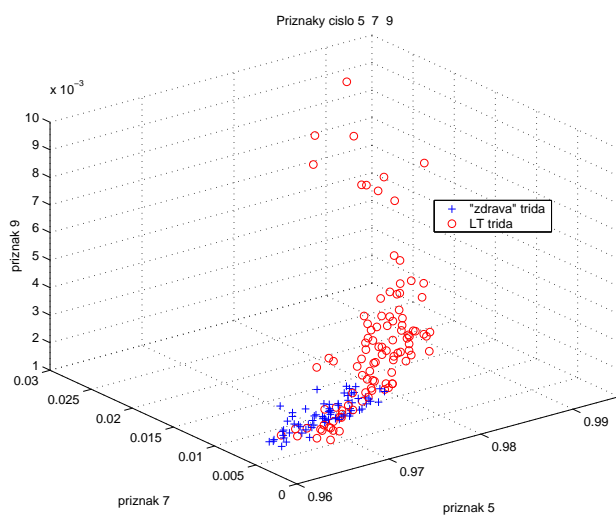
(b) Vzorok 31 × 31, kombinace H1 H7 H9

Obrázek 9: Příznakový prostor pro nejhorší trojici příznaků podle Fisherova lineárního diskriminantu.

A.2 Stabilita

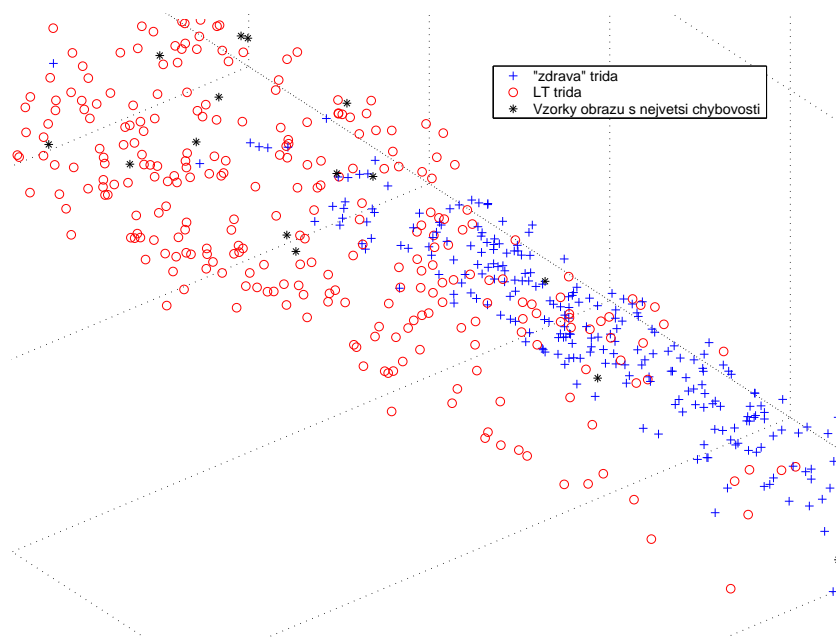


(a) Vzorky 21×21 , kombinace H2 H4 H9

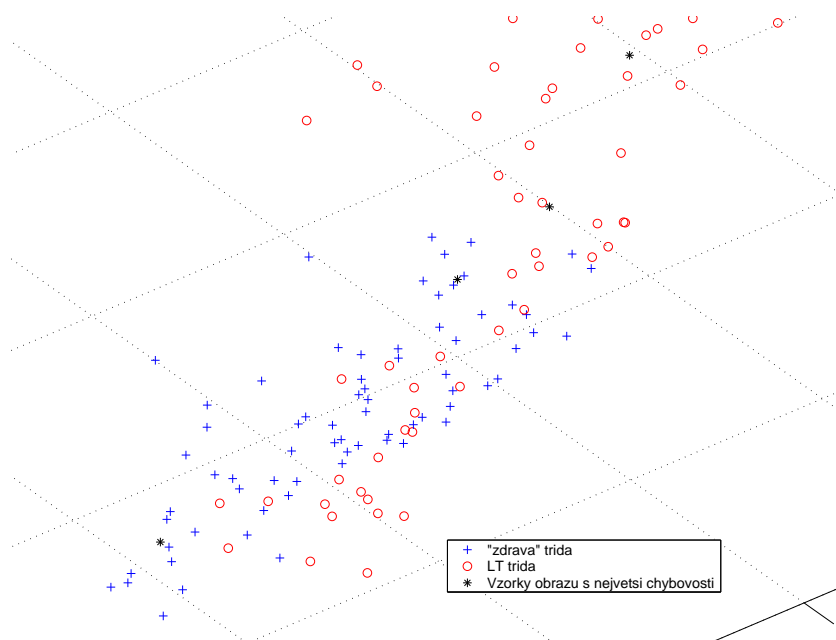


(b) Vzorky 31×31 , kombinace H5 H7 H9

Obrázek 10: Příkladový prostor pro trojici příznaků s největší stabilitou (klasifikace podle minimální vzdálenosti).



(a) Kombinace příznaků H2 H4 H9, vzorky 21×21 .



(b) Kombinace příznaků H5 H7 H9, vzorky 31×31 .

Obrázek 11: Zvětšený příznakový prostor obrázku 10 se zvýrazněnými vzorky obrazu s největší chybovostí.

A.3 Pořadí k-tic příznaků podle jednotlivých kritérií

Tabulka 1: Hodnoty F pro samotné příznaky.

	Vzorky 21×21		Vzorky 31×31	
Pořadí	Příznak	F	Příznak	F
1	H5	0.9659	H3	1.0950
2	H6	0.9636	H5	1.0926
3	H3	0.9634	H6	1.0268
4	H4	0.9305	H4	0.9550
5	H8	0.9220	H8	0.9517
6	H2	0.8881	H2	0.9249
7	H7	0.8770	H9	0.8746
8	H9	0.8569	H7	0.8724
9	H1	0.7852	H1	0.7892

Tabulka 2: Hodnoty F pro kombinace 2 příznaků.

Pořadí	Vzorky 21 × 21		Vzorky 31 × 31	
	Příznaky-H	F	Příznaky-H	F
1	1 3	1.1936	3 4	1.4101
2	1 5	1.1883	4 5	1.4043
3	4 5	1.1768	2 5	1.3915
4	3 4	1.1749	2 3	1.3845
5	4 6	1.1371	1 3	1.3259
6	2 4	1.1075	1 5	1.3168
7	4 8	1.0935	4 6	1.2766
8	1 6	1.0748	5 9	1.2340
9	4 9	1.0480	3 9	1.2325
10	2 6	1.0221	4 8	1.2107
11	4 7	1.0205	2 4	1.2103
12	6 9	1.0086	3 8	1.2015
13	3 5	1.0041	5 8	1.2012
14	2 5	0.9948	2 6	1.2006
15	1 8	0.9909	3 7	1.1864
16	2 3	0.9881	5 7	1.1859
17	5 9	0.9858	1 4	1.1764
18	1 4	0.9838	6 9	1.1633
19	3 9	0.9807	3 6	1.1576
20	6 7	0.9782	5 6	1.1563
21	6 8	0.9734	3 5	1.1557
22	5 8	0.9722	1 6	1.1479
23	5 7	0.9715	4 9	1.1077
24	3 8	0.9682	6 7	1.0939
25	3 7	0.9679	4 7	1.0720
26	5 6	0.9663	6 8	1.0695
27	3 6	0.9650	1 8	1.0408
28	1 2	0.9516	8 9	1.0308
29	8 9	0.9440	1 2	1.0123
30	2 8	0.9352	7 8	1.0085
31	7 8	0.9301	2 8	1.0000
32	2 9	0.9236	2 9	0.9839
33	1 7	0.9033	2 7	0.9449
34	1 9	0.8907	1 9	0.9124
35	2 7	0.8881	1 7	0.8978
36	7 9	0.8850	7 9	0.8748

Tabulka 3: Hodnoty F pro kombinace 3 příznaků.

První tři kombinace s největší stabilitou jsou v tabulce zvýrazněny.

Pořadí	Vzorky 21×21		Vzorky 31×31	
	Příznaky–H	F	Příznaky–H	F
1	1 3 6	1.2474	1 2 3	1.6036
2	1 2 3	1.2465	1 2 5	1.5996
3	1 2 5	1.2450	2 3 4	1.5659
4	1 3 9	1.2443	2 4 5	1.5636
5	1 5 6	1.2441	2 5 9	1.5471
6	1 5 9 ⁴	1.2422	2 3 9	1.5458
7	1 3 5	1.2382	1 3 5	1.5442
8	1 3 4	1.2335	3 4 5	1.5437
9	1 4 5	1.2309	3 4 6	1.5197
10	1 3 7	1.2223	3 4 9	1.5177
11	1 5 7	1.2183	4 5 6	1.5147
12	1 3 8	1.2112	4 5 9	1.5145
13	1 5 8	1.2071	3 4 7	1.5077
14	3 4 5	1.1853	4 5 7	1.5034
15	2 4 9 ⁵	1.1824	1 3 9	1.4983
16	4 5 6	1.1812	1 5 9	1.4932
17	4 5 7	1.1802	3 4 8	1.4760
18	4 5 9	1.1788	4 5 8	1.4705
19	3 4 7	1.1776	1 3 7	1.4667
20	3 4 6	1.1775	1 3 6	1.4629
21	4 5 8	1.1774	1 5 7	1.4592
22	2 4 5	1.1769	1 5 6	1.4533
23	3 4 9	1.1762	2 3 5	1.4286
24	3 4 8	1.1752	1 3 8	1.4278
25	2 3 4	1.1749	1 5 8	1.4195
27	4 6 9	1.1425	1 3 4	1.4148
28	4 6 7	1.1417	2 3 7	1.4115
29	4 6 8	1.1380	1 4 5	1.4088
30	2 4 6	1.1371	2 5 6	1.4051
31	2 4 7	1.1264	2 5 8	1.3958
32	2 3 5	1.1260	2 3 6	1.3950
33	1 6 9	1.1202	2 3 8	1.3890

⁴Druhá kombinace podle stability, vzorek 21×21 .

⁵První kombinace podle stability, vzorek 21×21 .

Pořadí	Vzorky 21 × 21		Vzorky 31 × 31	
	Příznaky–H	F	Příznaky–H	F
34	2 4 8	1.1133	2 4 9 ⁶	1.3531
35	1 2 4	1.1080	4 6 9	1.3519
36	1 2 6	1.1072	4 6 7	1.3286
37	3 5 9	1.0968	2 4 6	1.3048
38	4 7 8	1.0939	1 4 6	1.2907
39	4 8 9	1.0937	2 4 7	1.2842
40	1 4 8	1.0935	4 6 8	1.2783
41	1 6 7	1.0922	1 6 9	1.2775
42	1 6 8	1.0751	4 7 8	1.2554
43	1 4 9	1.0580	1 2 4	1.2516
44	4 7 9	1.0489	1 4 8	1.2503
45	3 5 8	1.0445	4 8 9	1.2467
46	1 2 9	1.0436	1 2 6	1.2452
47	3 5 7	1.0406	3 5 9	1.2448
48	1 4 7	1.0338	5 8 9	1.2406
49	3 5 6	1.0292	3 8 9	1.2386
50	2 6 7	1.0245	5 6 9	1.2343
51	2 6 9	1.0234	5 7 9 ⁷	1.2341
52	2 6 8	1.0231	3 7 9	1.2326
53	2 5 6	1.0224	3 6 9	1.2326
54	2 3 6	1.0221	1 6 7	1.2218
55	6 7 9	1.0131	5 6 8	1.2213
56	3 6 9	1.0101	2 6 8	1.2194
57	5 6 9	1.0092	3 6 8	1.2189
58	6 8 9	1.0088	2 4 8	1.2163
59	2 5 9	1.0023	5 7 8	1.2149
60	1 8 9	1.0012	3 7 8	1.2142
61	2 5 7	1.0004	1 4 9	1.2055
62	2 5 8	0.9956	5 6 7	1.2033
63	1 7 8	0.9952	2 6 9	1.2022
64	2 3 9	0.9943	3 6 7	1.2020
65	2 3 7	0.9936	3 5 8	1.2016
66	1 2 8	0.9910	2 6 7	1.2011
67	5 7 9 ⁸	0.9908	1 4 7	1.1928

⁶Druhá kombinace podle stability, vzorek 31 × 31.

⁷První kombinace podle stability, vzorek 31 × 31.

⁸Třetí kombinace podle stability, vzorek 21 × 21.

Pořadí	Vzorky 21 × 21		Vzorky 31 × 31	
	Příznaky–H	F	Příznaky–H	F
68	2 3 8	0.9895	3 5 7⁹	1.1875
69	5 8 9	0.9860	6 7 9	1.1671
70	3 7 9	0.9859	6 8 9	1.1665
71	3 8 9	0.9813	3 5 6	1.1624
72	6 7 8	0.9802	1 6 8	1.1596
73	5 6 7	0.9782	1 2 9	1.1520
74	3 6 7	0.9782	4 7 9	1.1139
75	5 6 8	0.9759	1 8 9	1.1108
76	3 6 8	0.9746	1 7 8	1.1040
77	5 7 8	0.9733	6 7 8	1.0959
78	3 7 8	0.9691	1 2 7	1.0667
79	1 2 7	0.9559	1 2 8	1.0438
80	2 8 9	0.9474	7 8 9	1.0333
81	7 8 9	0.9453	2 8 9	1.0310
82	2 7 9	0.9402	2 7 8	1.0234
83	2 7 8	0.9360	2 7 9	0.9851
84	1 7 9	0.9033	1 7 9	0.9130

⁹Třetí kombinace podle stability, vzorek 31 × 31.

Tabulka 4: Stabilita nejlepších kombinací příznaků ve 3D.

Vzorky 21×21			Vzorky 31×31		
H	Stabilita [%]	F	H	Stabilita [%]	F
2 4 9	42.7	1.1824	5 7 9	27.6	1.2341
1 5 9	15.3	1.2422	2 4 9	22.8	1.3531
5 7 9	11.0	0.9908	3 5 7	12.6	1.1875
1 5 7	7.0	1.2183	6 8 9	10.7	1.1665
6 8 9	6.0	1.0088	3 5 9	8.5	1.2448
3 5 7	2.7	1.0406	3 7 9	6.5	1.2326
3 7 9	2.6	0.9859	2 8 9	2.8	1.0310
2 4 8	2.4	1.1133	1 5 9	2.2	1.4932
1 3 9	2.2	1.2443	4 8 9	1.5	1.2467
3 5 8	2.1	1.0445	4 5 8	0.9	1.4705
1 3 7	2.0	1.2223	5 8 9	0.7	1.2406
3 5 9	0.6	1.0968	1 3 7	0.5	1.4667
2 4 5	0.5	1.1769	1 5 7	0.5	1.4592
2 4 7	0.5	1.1264	7 8 9	0.4	1.0333
4 5 8	0.4	1.1774	3 5 8	0.3	1.2016
5 7 8	0.4	0.9733	5 7 8	0.3	1.2149
1 3 5	0.2	1.2382	1 3 5	0.2	1.5442
2 3 4	0.2	1.1749	1 3 9	0.2	1.4983
2 7 8	0.2	0.9360	4 7 9	0.2	1.1139
1 2 4	0.1	1.1080	1 7 9	0.1	0.9130
1 3 8	0.1	1.2112	2 4 8	0.1	1.2163
2 6 9	0.1	1.0234	2 6 9	0.1	1.2022
2 8 9	0.1	0.9474	3 4 8	0.1	1.4760
3 4 8	0.1	1.1752	3 7 8	0.1	1.2142
3 7 8	0.1	0.9691	3 8 9	0.1	1.2386
4 6 9	0.1	1.1425			
4 8 9	0.1	1.0937			
5 8 9	0.1	0.9860			
7 8 9	0.1	0.9453			

A.4 Seznam používaných funkcí

Byla napsána funkce (`gpolyline`), která umožňuje expertovi označit v obrázku hranice štítné žlázy (oblast zájmu). Výstupem funkce je textový soubor, ve kterém je popsána poloha vyznačené oblasti. Tento soubor je potom použit k pokrytí štítné žlázy sítí nepřekrývajících se čtverců (funkce `sonocover1`). Funkce `sonocover` pak pro všechny obrázky se stejnými vlastnostmi (diagnóza, zvětšení, orientace sondy) vytvoří datovou strukturu obsahující odděleně jednotlivé čtverce z těchto oblastí. Do této struktury je pak pomocí funkce `calcfeat` přidána položka, ve které jsou pro každý čtverec vypočtené hodnoty 9 příznaků. Dále byly implementovány funkce pro výpočet Fisherova lineárního diskriminantu a pro klasifikaci podle minimální vzdálenosti. Funkce `fspace` umožňuje prohlížení příznakového prostoru.

`gpolyline` – Prostředí pro kreslení kontur oblastí.

`coverroi` – Pokrytí uzavřených oblastí obrázku sítí nepřekrývajících se čtverců.

`sonocover1` – Aplikování funkce `coverroi` na sonografický obrázek.

`sonocover` – Aplikování funkce `sonocover1` na všechny specifikované sonografické obrázky (získání vzorků).

`calcfeat` – Výpočet 9 Haralickových příznaků ze vzorků získaných funkcí `sonocover` nebo `sonocover1`.

`fisher` – Výpočet Fisherova lineárního diskriminantu.

`combfisher` – Aplikace funkce `fisher` na specifikované kombinace 9 Haralickových příznaků.

`setsim` – Náhodně vybírá trénovací a testovací množinu z výstupu funkce `calcfeat`.

`mindist` – Klasifikátor podle minimální vzdálenosti (pro 2 třídy vzorků).

`combmindist` – Aplikování funkce `mindist` na specifikované kombinace příznaků.

`fspace` – Zobrazuje příznakový prostor pro 2 třídy vzorků.

Předchozí seznam můžete vyvolat v Matlabu příkazem `help Sono`

A.5 Účastníci projektu

Petr Sucharda (petr.sucharda@lf1.cuni.cz)

Daniel Smutek (smutek@cesnet.cz)

Radim Šára (sara@cmp.felk.cvut.cz)

Ladislava Janků (jankul@lab.felk.cvut.cz)

Martin Švec (xsvecm@cmp.felk.cvut.cz)