

Algoritmy a struktury neuropočítačů

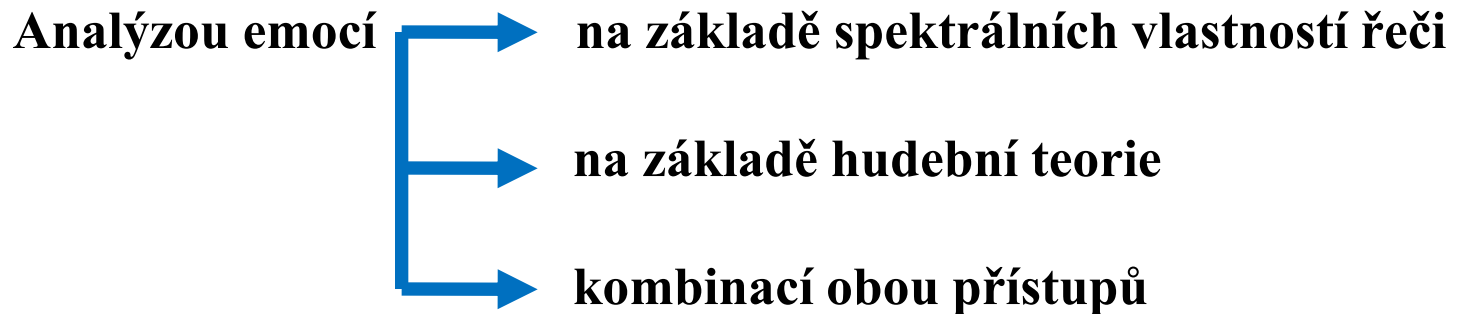
ASN – P09

Analýza emocionální řeči neuronovými sítěmi

Proč?

Pomocí emocí zlepšit kvalitu syntetické řeči a zvýšit přirozenost

Jak?

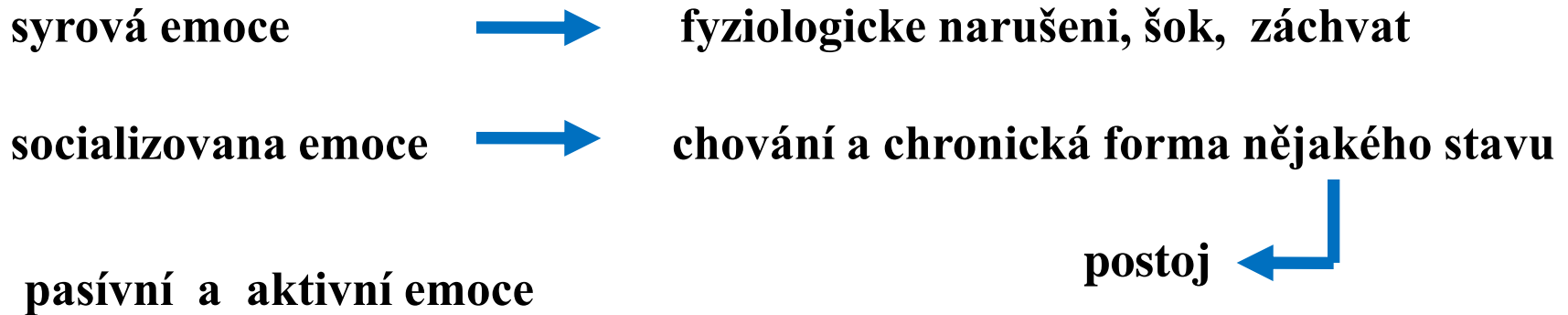


Pomocí MLNN a SOM

Algoritmy a struktury neuropočítačů

ASN – P09

Dělení emocí:



Akustické rysy spojené s emocemi

- Velikost, rozsah a tvar obálky základní hlasivkové frekvence f_0 .
- Úroveň hlasové energie a distribuce energie na jednotlivých frekvencích.
- Frekvence, velikost a tvar obálky řečových formantů, které mají vliv při artikulaci.
- Kadence, která udává informaci o rychlosti promluvy.

Radost

- **Reakce na zisk či úspěch, má silný sociální kontext.**
- **Je vyvolána různými druhy podnětů a dosahuje různých stupňů intenzity od spokojenosti po štěstí.**
- **Fyziologicky je provázena zrychlováním srdeční činnosti i dýchání, zvyšováním krevního tlaku a teploty, rozšiřováním krevních vlásečnic (objevuje se červenání).**

Algoritmy a struktury neuropočítačů

ASN – P09

Nuda

- **Reakce nervového systému na nedostatek fyzické nebo mentální aktivity.**
- **Intenzita této emoce záleží na vnějším prostředí a na době trvání této emoce.**
- **Vyznačuje se ztrátou zájmu.**
- **Fyziologicky se projevuje zpomalením srdečního tepu, poklesem krevního tlaku a zvýšeným sliněním.**

Smutek

- **Je to reakce na neúspěch či ztrátu.**
- **Intenzita pocíťovaného smutku závisí na významu neúspěchu či ztráty.**
- **Je to utlumující cit, provázený inaktivitou a ztrátou zájmů.**
- **Fyziologické projevy jsou podobné jako u fyzické bolesti, dochází ke zpomalení srdeční činnosti a ke změnám v rytmu dechu.**

Algoritmy a struktury neuropočítačů

ASN – P09

Vztek

- **Stejně jako zlost, hněv, zuřivost, je to emocionální reakci na nějakou překážku.**
- **Aktivizuje se a zvyšuje fyzická síla objektu. Původně byl vztek spojen s agresí, tj. s obrannou reakcí a sociálním postojem nepřátelství.**
- **Fyziologicky je vztek provázen zvýšeným svalovým napětím a zvýšenou motorikou, zrychluje se frekvence dechu a tep srdce, zvyšuje se krevní tlak, snižuje se pohyb střev a sekrece žaludečních šťáv.**

Algoritmy a struktury neuropočítačů

ASN – P09

Citové a volní postoje mluvčího spoluvytváří tzv. **barva hlasu**

řekl smutně, vesele, nadšeně, rozčileně, naléhavě

Emoce se vyjadřují také **tempem řeči** - mimořádně rychlé nebo pomalé

Tvorba databáze promluv

Nahrávání promluv – v profesionálním studiu, namluvili divadelní herci

*Jé.
Má?
Nevím.
Vidíš?
Povídej!
Poezie.*

*To mi nevadí.
Neumím to vysvětlit.
To bude světový rekord.
Jak se Ti to líbí?
Podívej se na nebe!
Až přijdeš uvidíš.*

Algoritmy a struktury neuropočítačů

ASN – P09

Formát „wav“, vzorkovací kmitočet 44,1 kHz, 24 bittů

Jednotlivé věty byly namluveny 4 emocemi v 10ti opakováních

Převzorkováno na 16 bittů

Subjektivní poslechové testy pro ověření věrohodnosti vyjádřené emoce.

Celkem 720 vzorků emocí - 360 jednoslovních + 360 víceslovných

Algoritmy a struktury neuropočítačů

ASN – P09

Experimenty

Software: Matlab a NN Toolbox
SOMToolbox

Emoce: 1- hněv, 2- nuda, 3- radost 4- smutek

Časova oblast

Aritmetický průměr absolutní hodnoty

Směrodatná odchylka

Maximum

Minimum

Frekvenční oblast

Základní hlasivková frekvence

Frekvence 1. - 4. formantu

vstupy
do UNS

Princip hudební teorie

tónovou příbuznost

f_0



Harmonická řada tónů je percepčně je vnímána jako jediný tón o výšce „x“

Amplituda vyšších harmonických obvykle klesá s jejich rostoucím indexem, ale i když je amplituda f_0 nižší, případně nulová, je stále jednoznačně percepčně vnímána jako tentýž tón.

Původní výška budícího tónu harmonické řady je ve zbytku frekvenčního pásma kódována frekvenčním rozdílem po sobě jdoucích harmonických, který je vždy roven hodnotě „x“. Tento jev vysvětluje **teorie rozdílových a kombinačních tónů**.

Rozdílové a kombinační tóny - vysvětlení

Při výrobě varhan pro realizaci nejhlubších tónů. Píšťaly by byly příliš veliké (i drahé). Proto se vyrobí dvě píšťaly menší produkující složené tóny o základních frekvencích $2x$ a $3x$, kde „ x “ je frekvenční hodnota požadovaného tónu.

Je dodržen poměr harmonických f_3 a f_2 ($=1,5$) fiktivního tónu „ x “, jejich rozdíl je roven jeho frekvenci.

Pokud zní dva tóny o frekvencích „ a “ a „ b “, vnímá člověk především tóny rozdílové o kmitočtu $f = a - b$, a tóny kombinační (také součtové) o $f = a + b$

Je vnímána příbuznost a příslušnost libovolných dvou (i více) tónů k jejich základu, ať tóny znějí současně nebo postupně.

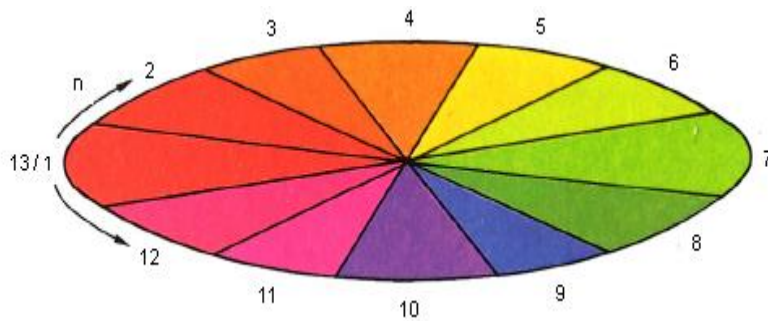
Algoritmy a struktury neuropočítačů

ASN – P09

Z řeči jsou extrahovány f_0 a formanty

Je definován **hudební interval** - poměr frekvence n -tého tonu ku frekvenci tónu, od kterého počítáme interval.

Kvinta je poměr frekvence pátého tónu ku frekvenci prvního tónu : **1.498**.



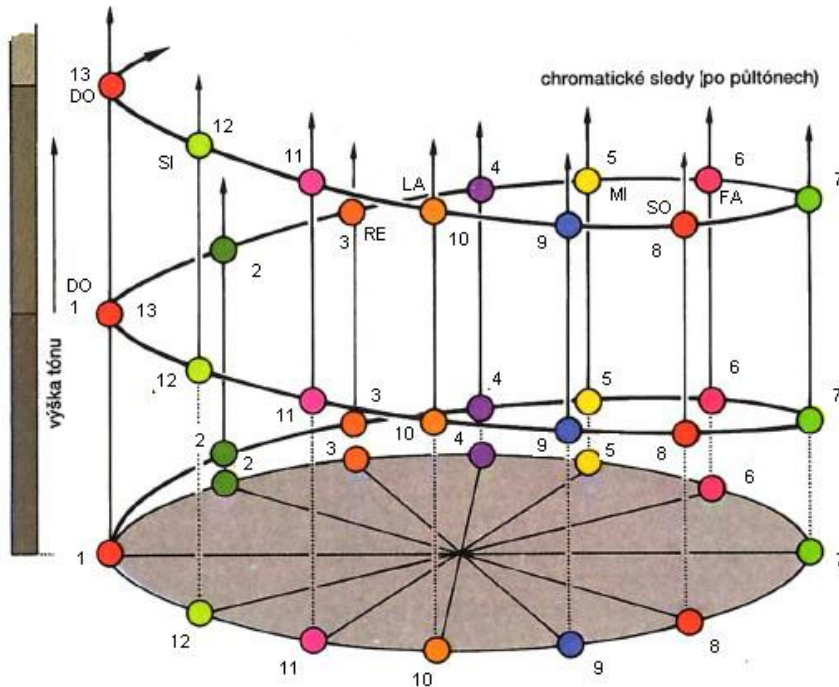
Kvintový kruh

Kvintový kruh vyjadřuje vlastnost, že **příbuznost** s rostoucím n nejdříve klesá, nejmenší hodnoty nabývá pro $k7$, pak znovu roste, kvintovými kroky se blíží k členu $k13$, který je s výchozím tónem $o1 = k1$ v oktávovém poměru (příbuznosti) jako člen řady $o8$.

Algoritmy a struktury neuropočítačů

ASN – P09

Charakter kvintové i oktávové příbuznosti



Oktávová příbuznost zdůrazněna stejnou barvou. Kvintová je respektována blízkostí odstínů. Čísla u tónů označují pořadí v chromatické stupnici.

Vyjdeme z bodu 1 do bodu 2 ----- dostaví se pocit negativní emoce (barvy nejsou příbuzné = vzdálená kvintová příbuznost). Pokud budeme pokračovat do bodu 3, vystřídá záporný vjem pocit pozitivní emoce dané blízkou kvintovou příbuzností (blízký barevný odstín) s výchozím bodem.

Algoritmy a struktury neuropočítačů

ASN – P09

Úspěšná klasifikace emocí:

74% (MLNN)

QE / TE

0.274 / 0.014

1 word sentence

QE / TE

0.275 / 0.017 (SSOM)

multiword sentence

QE ... kvantizační chyba

TE ... topografická chyba

Porovnání se zahraničními výsledky:

Úspěšná klasifikace

54-64% standardní klasifikátory

81 % ANN korejština – tónový jazyk

Algoritmy a struktury neuropočítačů

ASN – P09

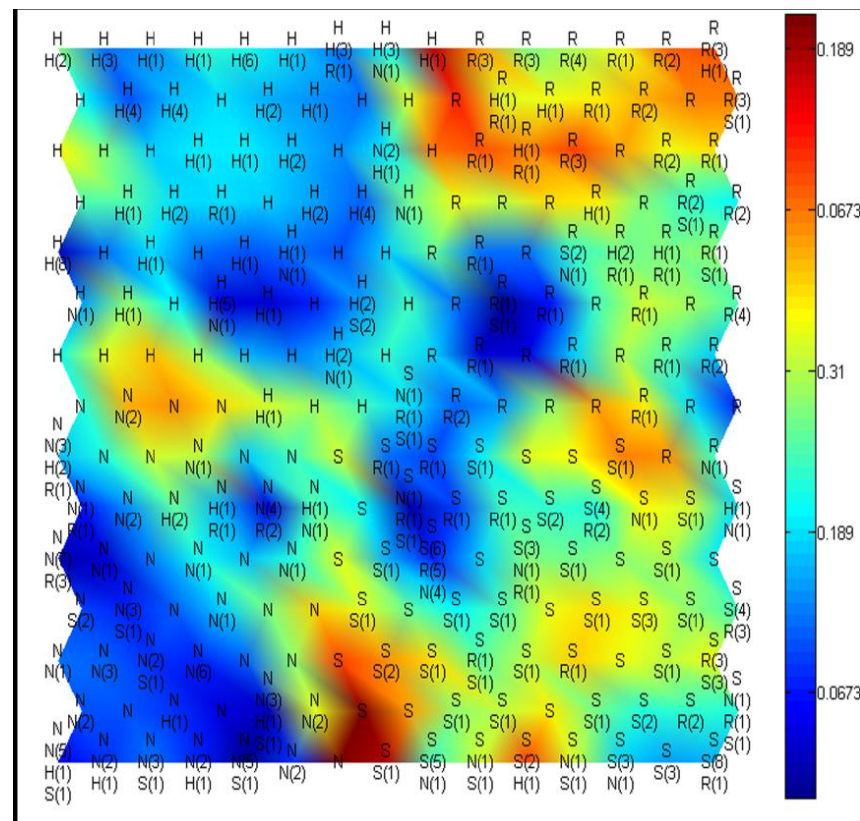
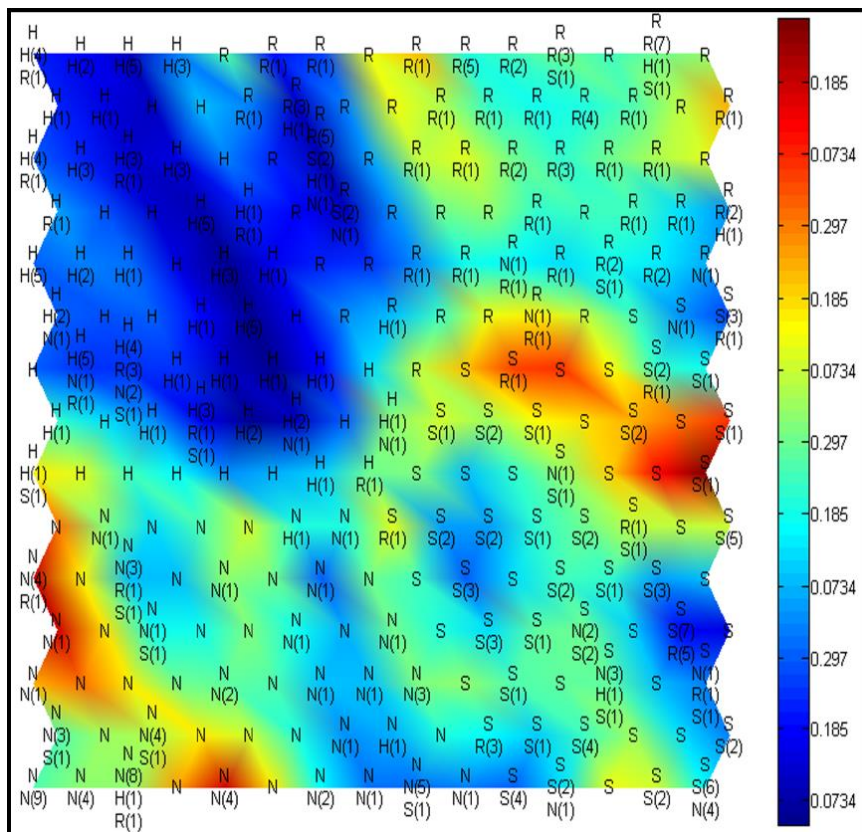
U – matice, KSOM

Hněv

Smutek

Radost

Nuda



Algoritmy a struktury neuropočítačů

ASN – P09

Training Confusion Matrix

Output Class	1	48 22.4%	6 2.8%	8 3.7%	1 0.5%	76.2% 23.8%
	2	1 0.5%	39 18.2%	3 1.4%	6 2.8%	79.6% 20.4%
	3	2 0.9%	6 2.8%	37 17.3%	3 1.4%	77.1% 22.9%
	4	0 0.0%	11 5.1%	3 1.4%	40 18.7%	74.1% 25.9%
		94.1% 5.9%	62.9% 37.1%	72.5% 27.5%	80.0% 20.0%	76.6% 23.4%
	1	2	3	4		Target Class

Validation Confusion Matrix

Output Class	1	15 21.1%	1 1.4%	7 9.9%	0 0.0%	65.2% 34.8%
	2	0 0.0%	10 14.1%	1 1.4%	2 2.8%	76.9% 23.1%
	3	0 0.0%	0 0.0%	18 25.4%	7 9.9%	72.0% 28.0%
	4	0 0.0%	1 1.4%	1 1.4%	8 11.3%	80.0% 20.0%
		100% 0.0%	83.3% 16.7%	66.7% 33.3%	47.1% 52.9%	71.8% 28.2%
	1	2	3	4		Target Class

Test Confusion Matrix

Output Class	1	21 29.6%	3 4.2%	3 4.2%	1 1.4%	75.0% 25.0%
	2	2 2.8%	10 14.1%	0 0.0%	0 0.0%	83.3% 16.7%
	3	0 0.0%	3 4.2%	9 12.7%	9 12.7%	42.9% 57.1%
	4	0 0.0%	0 0.0%	0 0.0%	10 14.1%	100% 0.0%
		91.3% 8.7%	62.5% 37.5%	75.0% 25.0%	50.0% 50.0%	70.4% 29.6%
	1	2	3	4		Target Class

All Confusion Matrix

Output Class	1	84 23.6%	10 2.8%	18 5.1%	2 0.6%	73.7% 26.3%
	2	3 0.8%	59 16.6%	4 1.1%	8 2.2%	79.7% 20.3%
	3	2 0.6%	9 2.5%	64 18.0%	19 5.3%	68.1% 31.9%
	4	0 0.0%	12 3.4%	4 1.1%	58 16.3%	78.4% 21.6%
		94.4% 5.6%	65.6% 34.4%	71.1% 28.9%	66.7% 33.3%	74.4% 25.6%
	1	2	3	4		Target Class

**Matice záměn
pro jednoslovné věty**

**parametry hudební
Teorie**

**MLNN 29-55-4
Trainscg, sigmoida**

**Trénování: 216 vzorků
Test: 72 vzorků
Validace: 72 vzorků**

Algoritmy a struktury neuropočítačů

ASN – P09

Training Confusion Matrix

Output Class	1	42 19.4%	11 5.1%	5 2.3%	1 0.5%	71.2% 28.8%
	2	12 5.6%	34 15.7%	12 5.6%	10 4.6%	50.0% 50.0%
	3	2 0.9%	1 0.5%	26 12.0%	4 1.9%	78.8% 21.2%
	4	2 0.9%	3 1.4%	10 4.6%	41 19.0%	73.2% 26.8%
			72.4% 27.6%	69.4% 30.6%	49.1% 50.9%	73.2% 26.8%
	Target Class	1	2	3	4	

Validation Confusion Matrix

Output Class	1	15 20.8%	2 2.8%	2 2.8%	1 1.4%	75.0% 25.0%
	2	0 0.0%	14 19.4%	4 5.6%	5 6.9%	60.9% 39.1%
	3	1 1.4%	1 1.4%	6 8.3%	3 4.2%	54.5% 45.5%
	4	0 0.0%	3 4.2%	6 8.3%	9 12.5%	50.0% 50.0%
			93.8% 6.3%	70.0% 30.0%	33.3% 66.7%	50.0% 50.0%
	Target Class	1	2	3	4	

**Matice záměn
pro víceslovné věty**

Hudební teorie

MLNN

Test Confusion Matrix

Output Class	1	13 18.1%	2 2.8%	1 1.4%	0 0.0%	81.3% 18.8%
	2	2 2.8%	14 19.4%	4 5.6%	3 4.2%	60.9% 39.1%
	3	1 1.4%	0 0.0%	11 15.3%	2 2.8%	78.6% 21.4%
	4	0 0.0%	6 8.3%	4 5.6%	9 12.5%	47.4% 52.6%
			81.3% 18.8%	63.6% 36.4%	55.0% 45.0%	64.3% 35.7%
	Target Class	1	2	3	4	

All Confusion Matrix

Output Class	1	70 19.4%	15 4.2%	8 2.2%	2 0.6%	73.7% 26.3%
	2	14 3.9%	62 17.2%	20 5.6%	18 5.0%	54.4% 45.6%
	3	4 1.1%	2 0.6%	43 11.9%	9 2.5%	74.1% 25.9%
	4	2 0.6%	12 3.3%	20 5.6%	59 16.4%	63.4% 36.6%
			77.8% 22.2%	68.1% 31.9%	47.3% 52.7%	67.0% 33.0%
	Target Class	1	2	3	4	

Algoritmy a struktury neuropočítačů

ASN – P09

Výsledky pro MLNN:

Metoda	úspěšnost	
	1-sláb.	vice-sláb
I	88.7 %	77.8 %
II	70.4 %	65.3 %
III	85.9 %	84.7 %