

# Rozpoznávanie emócií z reči s použitím príznakov fázového spektra

Meno: Igor Guoth

Vedúci: Ing. Milan Rusko, PhD.

Oddelenie analýzy a syntézy reči

# Obsah

---

- Emócie
- Úvod do rozpoznávania emócií (základné prístupy pri rozpoznávaní emócií z reči, databázy, využitie a aktuálny stav v danej problematike).
- Popis vybraného experimentu a jeho výsledky.

# Emócia

---

- Emócia – Čo to v skutočnosti je?
  - Mentálny stav, krátkeho trvania (sec/min)
  - Nálada (hodiny, dni)
  - Modelovanie osobnosti (mesiace, roky,...)

# Rozpoznávanie emócií

---

- Umelá inteligencia
- Multidisciplinárna oblasť(psychológia, fyziológia, počítačové vedy, atď.)
- Rozpoznávanie, interpretácia a spracovanie emócií
- Aplikácie v praxi(telefónne centrá, psychiatria, dohľadové a bezpečnostné systémy, zlepšenie systémov na identifikáciu rečníka a rozpoznávanie reči, atď.)

# Čo berieme pri vývoji do úvahy

---

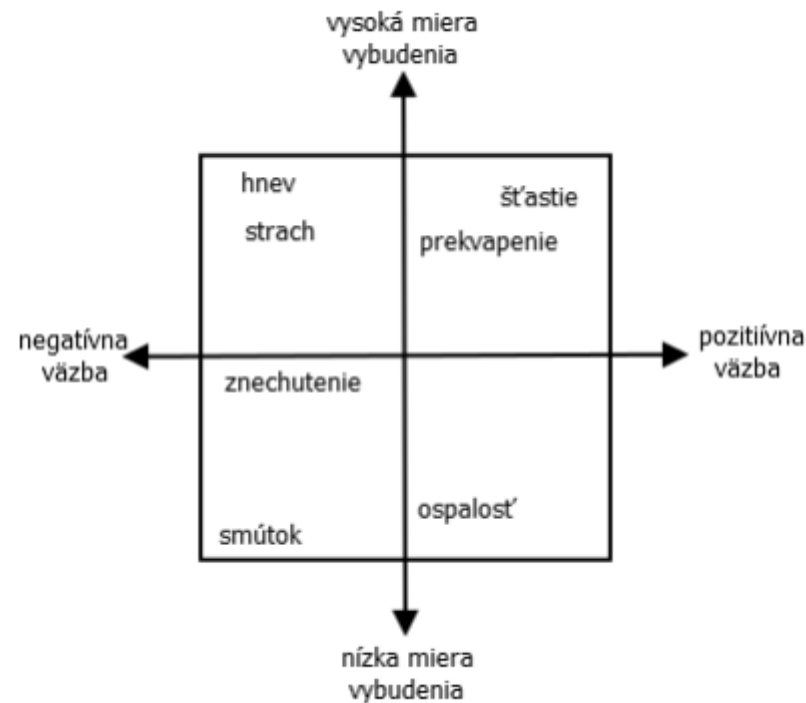
- psychologické a fyziologické hľadisko – úzko späté
  - Psychologický proces spôsobuje fyziologické zmeny v organizme!
  - Značný vplyv na nervový systém – zmena krvného tlaku, tepu, zmeny v dýchacom systéme!
- rečové hľadisko – zmena rečového prejavu jedinca

# Rozpoznávanie diskrétnych emócií - klasifikácia

Author	Emotions number	Basic Emotions
Ekman	6	anger, disgust, fear, joy, sadness, surprise
Frijda	6	desire, happiness, interest, surprise, wonder, sorrow
Plutchik	8	acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise
Tomkins	9	anger, interest, contempt, disgust, distress, fear, joy, shame, surprise
Matsumoto	22	joy, anticipation, anger, disgust, sadness, surprise, fear, acceptance, shy, pride, appreciate, calmness, admire, contempt, love, happiness, exciting, regret, ease, discomfort, respect, like

# Rozpoznávanie emócií v spojitej škále - regresia

---



# Databázy používané na rozpoznávanie emócií - delenie

---

- Simulované databázy:
  - 60 % percent z celkového množstva databáz
  - najrozšírenejšia a najjednoduchšia metóda
  - EmoDB – Diskrétne rozpoznávanie úspešnosť systémov > 80%.
- Indukované databázy:
  - reálnejšie emócie v porovnaní so simulovanými databázami
- Prirodzené databázy:
  - problém s ich získavaním – otázka súkromia



# Trendy pri rozpoznávaní emócií

---

- Vyhľadávanie nových príznakov popisujúcich rečový signál, ktoré môžu byť vhodné pre budovanie systémov na rozpoznávanie emócií z reči.
- Testovanie nových metód učenia strojov pri rozpoznávaní emócií. (hlboké neurónové siete - DNN).
- Budovanie šumovo odolných systémov.
- Jazykovo nezávislý systém.

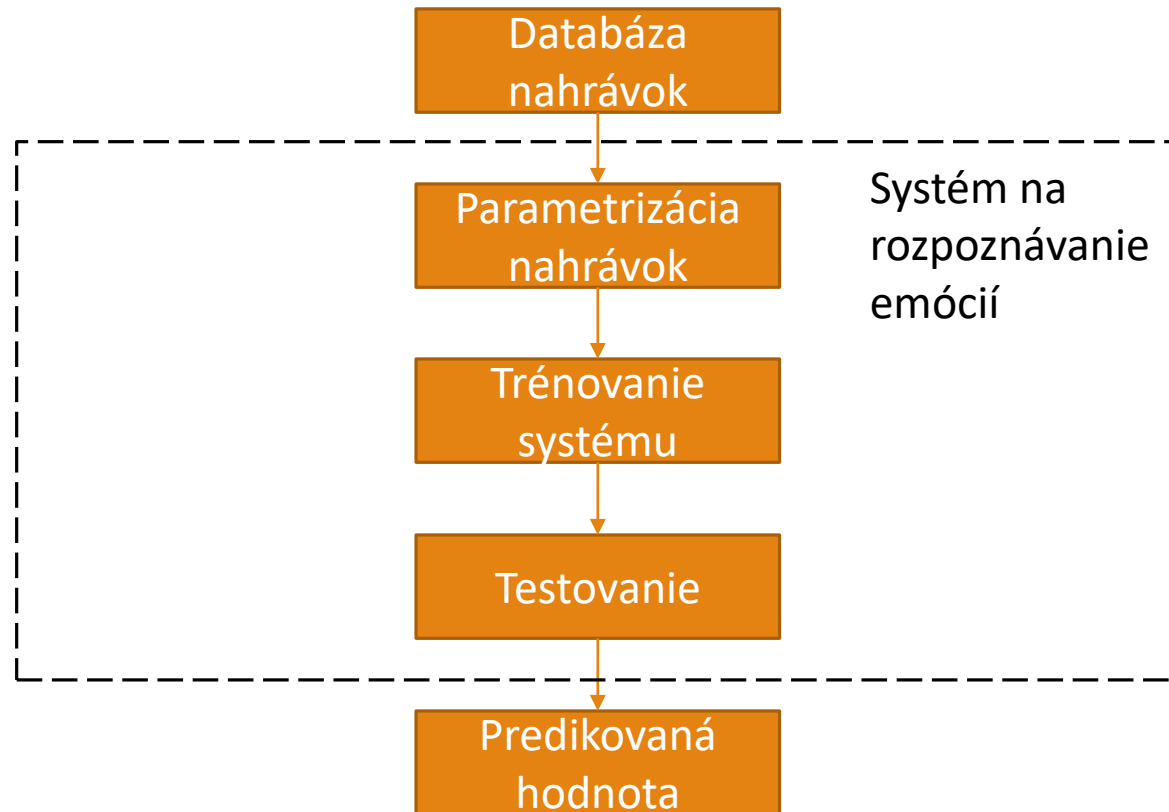
# Stručný popis experimentu

---

- Rozpoznávanie emócií v spojitom priestore.
- SVR (Regresia s použitím metódy podporných vektorov), Neurónová sieť.
- Testovanie možnosti použitia informácie o fázovom spektre a zmenách v ňom pri rozpoznávaní emócií v spojitom priestore.

# Koncept systému

---



# Databáza použitá v experimente

---

- IEMOCAP – multimodálna databáza
- Audio vizuálny emočný korpus
- Anglický jazyk
- 12 hodín dát
- 9 emočných stavov – trojrozmerný emočný priestor
- Dimenzie:
  - Aktivácia - DA
  - Valencia - DV
  - Dominancia -DD

# Experiment

---

- 2 základné prístupy pri rozpoznávaní emócií – extrakcia veľkého množstva rôznych typov príznakov, použitie kepstrálno-spektrálnych príznakov.
- Nový prístup – použitie fázového spektra vo forme APGD koeficientov pri rozpoznávaní v spojitom emočnom priestore a ich možná kombinácia s predchádzajúcimi prístupmi.

# Veľká množina príznakov

---

- Použité príznaky:
  - Prozódia(tempo reči, F0, hlasitosť..)
  - Príznaky z VQ skupiny príznakov (jitter, shimmer, HNR, formanty...) – príznaky, ktoré zvyšujú úspešnosť v dimenzii väzba
  - Spektrálne a kepstrálne príznaky(MFCC, LSP, BBE)
  - Ich deriváty – zmeny v čase

# Kepstrálno spektrálne príznaky

---

- Alternatíva k 1-D kepstrálnej analýze – použitie 2-D DCT
- Globálne, lokálne zmeny spektrálnej obálky, informácie o zmene výšky hlasu alebo zdroja vybudenia atď.

# Fázové spektrum vo forme APGD príznakov

---

- Informácie o fáze a o zmenách v nej v minulosti veľmi zanedbávané z viacerých dôvodov.
- Príznaky počítané s použitím LP analýzy a funkciou skupinového oneskorenia.



# Výsledky experimentu pri kombinácii APGD so základnými typmi príznakov pri rozpoznávaní emócií

---

- Hodnotenie s použitím Paersonovho korelačného koeficientu.

Kombinácia príznakov	DA	DD	DV
openSMILE	0.6797	0.4811	0.4786
openSMILE + APGD	0.6910	0.4874	0.5031
TDC	0.6241	0.4402	0.3841
TDC+APGD	0.6703	0.4684	0.4602

# Experiment s detekciou stresu

---

- Databáza CRISIS - slovenská databáza určená na detekciu stresu.
- Delenie stresu do 3 rôznych úrovní (Neutrálna reč, a dve rôzne úrovne stresu).

# Výsledky experimentu pri detekcii stresu

---

	Sachia - UBM		Igor - SVM	
	MFCC	APGD	MFCC	APGD
Error rate [%]	22,72	31,57	22,37	23,18

