

Rozpoznávání tónů v akordech

Martin Kirschner
[AIL002] – Neuronové sítě
2006

Úloha:

- Rozpoznání tónů, ze kterých je složený vstupní akord.
- Využití
 - Na převod hudby z formátu wav do formátu midi, ...
 - V kombinaci s dalšími nástroji (rozdělení skladby podle nástrojů...)

Trénovací množina dat

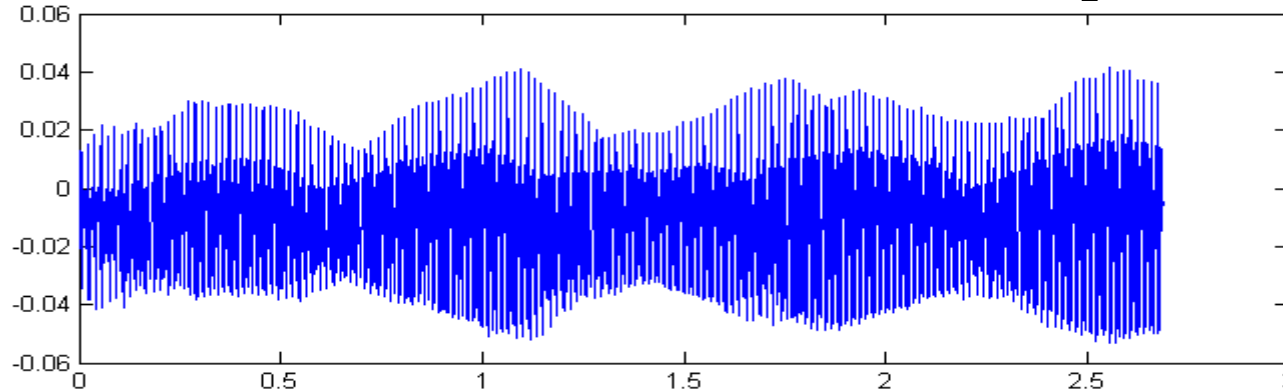
- Pro zjištění chování neuronové sítě používám jednočárkovanou oktávu (c_1-h_1) i s půltóny (12 různých frekvencí)
- Nástroj: flétna (má čistý a nekolísající zvuk)
- Trénovací Wav-soubory jsem nahrál pomocí midi
- Testovací přes mikrofon pomocí skutečné flétny

Trénovací množina dat

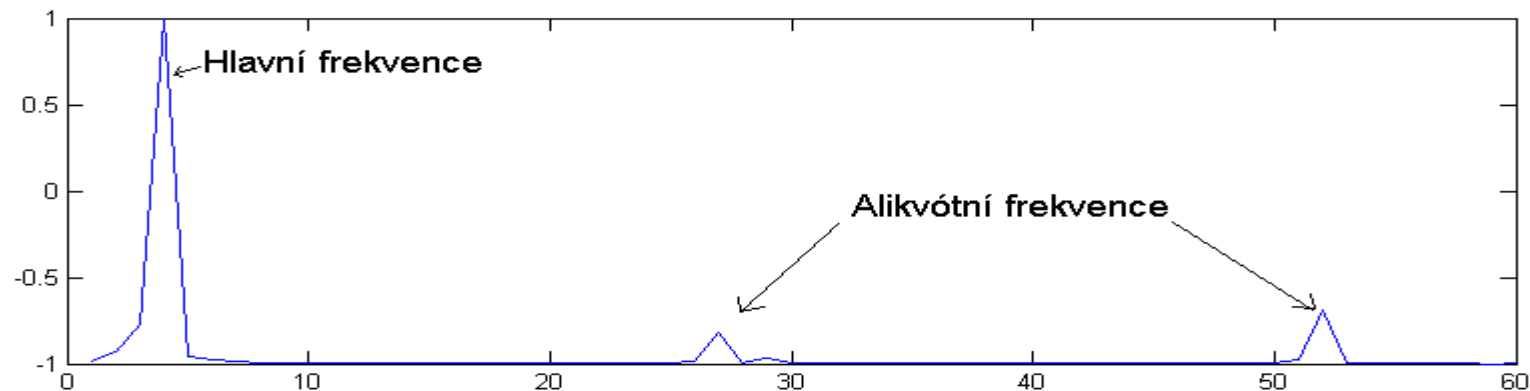
pokračování

- Akordy jsem zmixoval z jednotlivých tónů až na počítači
 - Nevýhody tónů nahrávaných přes mikrofon
 - Různá hlasitost jednotlivých tónů ve složených akordech (to je ale u reálných dat běžné)
 - Vyšší úroveň šumu (také běžné)
- => Testovací data jsou „Jako ze života“

Příprava trénovací množiny



- Převedení dat na spektrum intenzity jednotlivých frekvencí (pomocí fft)
- Nascalování na interval $\langle 0;1 \rangle$



Řešení

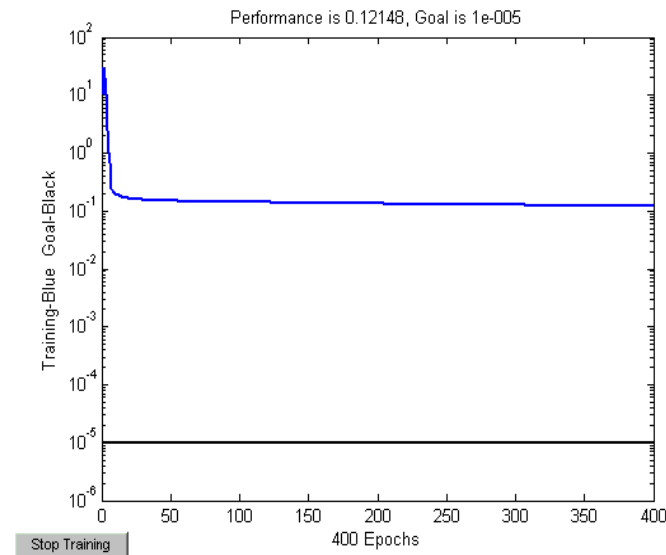
- Síť se zpětným šířením s 60ti vstupními a 12ti výstupními neurony

Parametry:

- Parametr učení – 0,1
- Počet epoch – 400

Výsledky – na datech z midi

- Vstupní data – všechny jednotóny a dvojtóny v oktávě nahrávané z midi = 78 vzorů
- Síť se zřejmě nepodařilo naučit správně
- Chyba při vybavování byla odpovídající, tzn. Toto bylo v praxi nepoužitelné



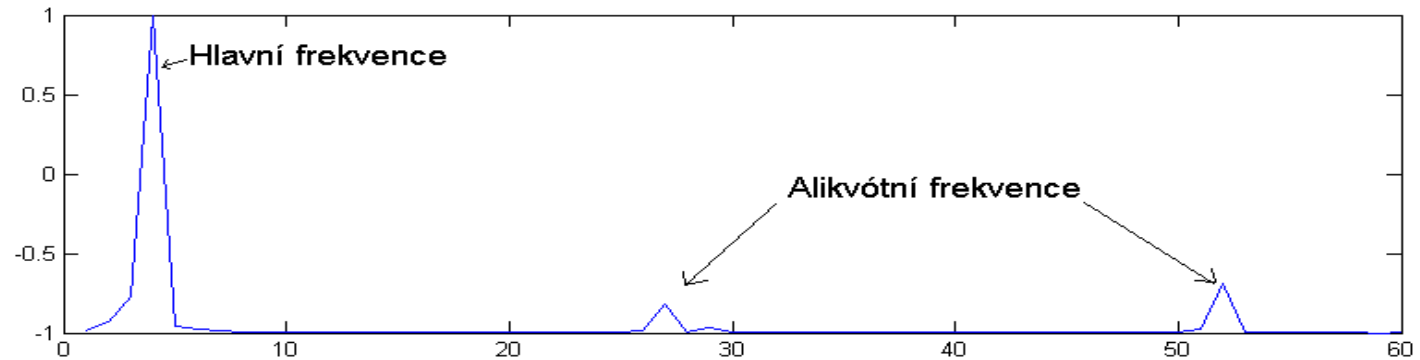
Výsledky – na datech ze skutečného nástroje

- Vstupní data – 12 tónů jednočárkové oktávy + c_2
- Učení proběhlo rychle a kvalitně – na trénovacích datech úspěšnost 100%
- Rozpoznávání tónů ve dvojakordech mělo úspěšnost 83%
 - Chyby byly způsobeny hlavně zanedbáním slabších tónů (většinou hlubších)
- Proč ten rozdíl oproti „*uvařeným*“ datům z midi

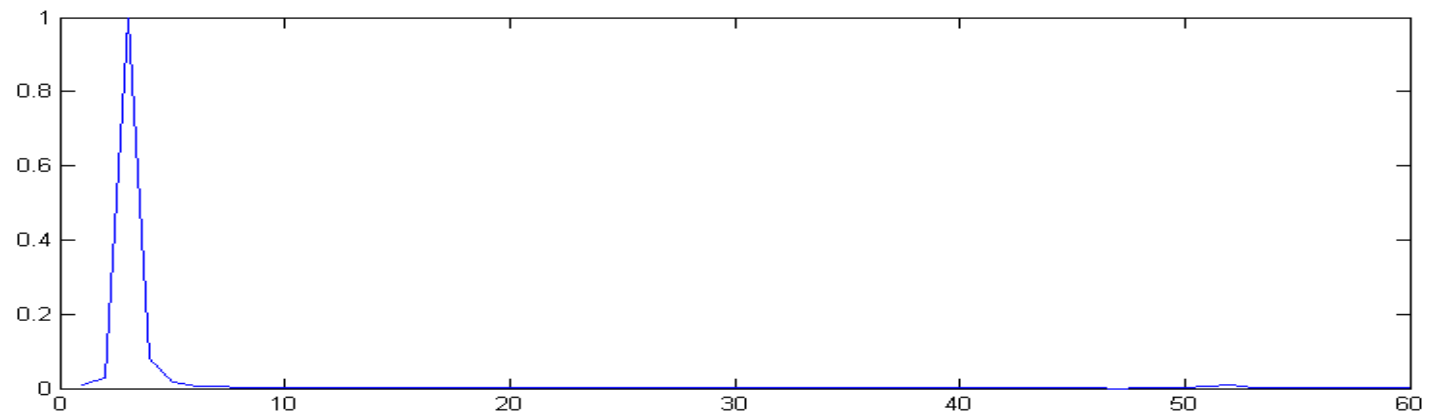
?

Alikvótní tóny

Tón C₁ z midi (po normalizaci)



Tón C₁ z mikrofonu (po normalizaci)



Jak zlepšit výsledky?

Další úpravou vstupních dat

- Lidské ucho slyší logaritmicky
 - Přesamplovat tóny, aby šly lineárně za sebou
 - Logaritmicky upravit hlasitosti jednotlivých frekvencí
 - Řešení na odstranění alikvótních tónů jsem nenašel. Neuronová síť mnou vytvořená zřejmě také ne..