

Statistická analýza hudebních signálů

O. Havelka*
M. Machalová**
Š. Jelínek***

Gymnázium Trutnov, Jiráskovo náměstí 325*
Wichterlovo gymnázium Ostrava, Čs. exilu 669**
Gymnázium Chomutov, Mostecká 3000***

ondra10ax@centrum.cz*
marketa.machalova314@gmail.com**
simon123456@seznam.cz***

Abstrakt:

Cílem našeho projektu bylo zanalyzovat hudební nahrávky různých hudebních stylů (rock, elektro house, jazz, eurodance, house, progressive, vážná hudba) o délce 10 až 20 vteřin se vzorkovací frekvencí 44 100 Hz. Poté pomocí metody průchodů nulou jsme určili náhodnost či pravidelnost daných hudebních nahrávek a dále ji statisticky kvantifikovali pomocí programu MATLAB.

1 Úvod

Samotné téma analyzování predikovatelnosti a závislosti hudebních signálů není ještě moc prozkoumáno. My jsme použili klasických statistických charakteristik, jako je průměr, směrodatná odchylka a autokorelační koeficient. Zvukový signál jsme modelovali pomocí Gaussova zlomkového šumu. Statistické výpočty jsme provedli v samotném programu MATLAB.

2 Teorie

K úspěšnému rozboru byly skladby (časová řada, diskrétní signál) rozděleny na segmenty. Samotná vzorkovací frekvence (počet členů časové řady za 1s) byla 44 100 Hz (parametry nekomprimovaného *.wav souboru), čímž splňuje podmínky Nyquistovi frekvence, při níž může dojít k přesné rekonstrukci spojitého, frekvenčně omezeného signálu z jeho vzorků. Tato vzorkovací frekvence musí převyšovat dvojnásobek maximální frekvence signálu a lidské ucho reaguje na nejvyšší frekvenci 20 kHz.

Hudební signály, ačkoliv se zapisují jako časové řady, mají fraktální dimenzi a to konkrétně hodnotu 2. Jedná se však o Hausdorffovu dimenzi (tzv. zobecněnou), nikoliv o topologickou.

Při zjišťování předvídatelnosti ukázek jsme využili různých statistických charakteristik, například směrodatné odchylky, která určuje vzájemnou odlišnost dat ve

vzorku, a autokorelačního koeficientu k -tého řádu, který určuje závislost dat od sebe vzdálených o k bodů.

Výše zmíněných údajů jsme využili při definování Gaussova zlomkového šumu a při nalezení spojitosti mezi zvukovým signálem a tímto šumem. Zlomkový Gaussův šum je spojitý náhodný proces s nulovou střední hodnotou, s jednotkovým rozptylem (což je kvadrát směrodatné odchylky) a s korelační funkcí

$$\rho_k = \frac{1}{2}(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}); H \in (0; 1). \quad (1)$$

Jeho Hausdorffova dimenze je 2, stejně jako u zvukového signálu, a proto považujeme zvukový signál za Gaussův zlomkový šum. Hurstův exponent (dále H) určuje předvídatelnost Gaussova šumu. Pokud se blíží 1, šum je silně předvídatelný, pokud 0,5, je chaotický a nepředvídatelný. Pokud se blíží k 0, je nechaotický, ale nepředvídatelný.

Pravděpodobnost p odhadneme podle vzorce

$$p = \frac{r}{n}, \quad (2)$$

kde r znázorňuje počet průchodů nulou, n je rovno počtu vzorků, a využijeme ji pro odhadnutí Hurstova koeficientu

$$H \approx 1 + \log_2 \left(\cos \frac{\pi \cdot p}{2} \right). \quad (3)$$

3 Metody

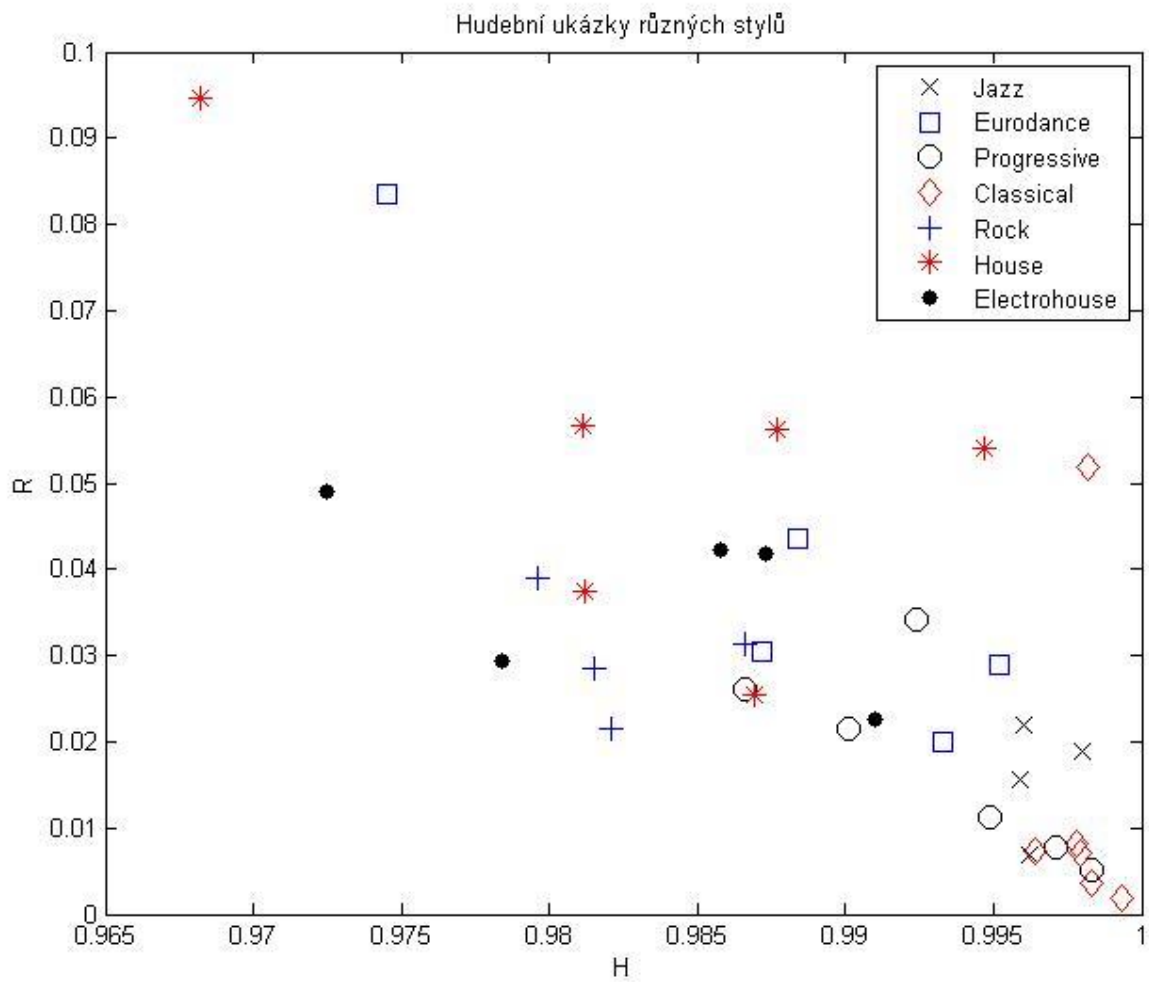
Využili jsme prostředí MATLAB, ve kterém jsme ukázky rozdělili na přibližně 250 ms dlouhé intervaly, u kterých jsme určili H a následně vykreslili jeho průběh v čase. Z výsledných Hurstových exponentů jsme spočítali průměr, směrodatnou odchylku, rozsah funkčních hodnot a rozdíl 80% kvantilu a 20% kvantilu. Tento postup byl využit pro všechny ukázky. Následně jsme skladby stejného žánru označili stejným symbolem a vynesli do grafu, kde osa x znázorňovala průměr H a osa y znázorňovala směrodatnou odchylku, rozsah funkčních hodnot nebo rozdíl 80% kvantilu a 20% kvantilu.

Snažili jsme se najít charakteristiku, s pomocí které by jednotlivé žánry byly odděleny ideálně tak, abychom po rozboru ukázky dokázali zjistit její žánr. Nejlépe se osvědčilo použít průměr a rozsah H .

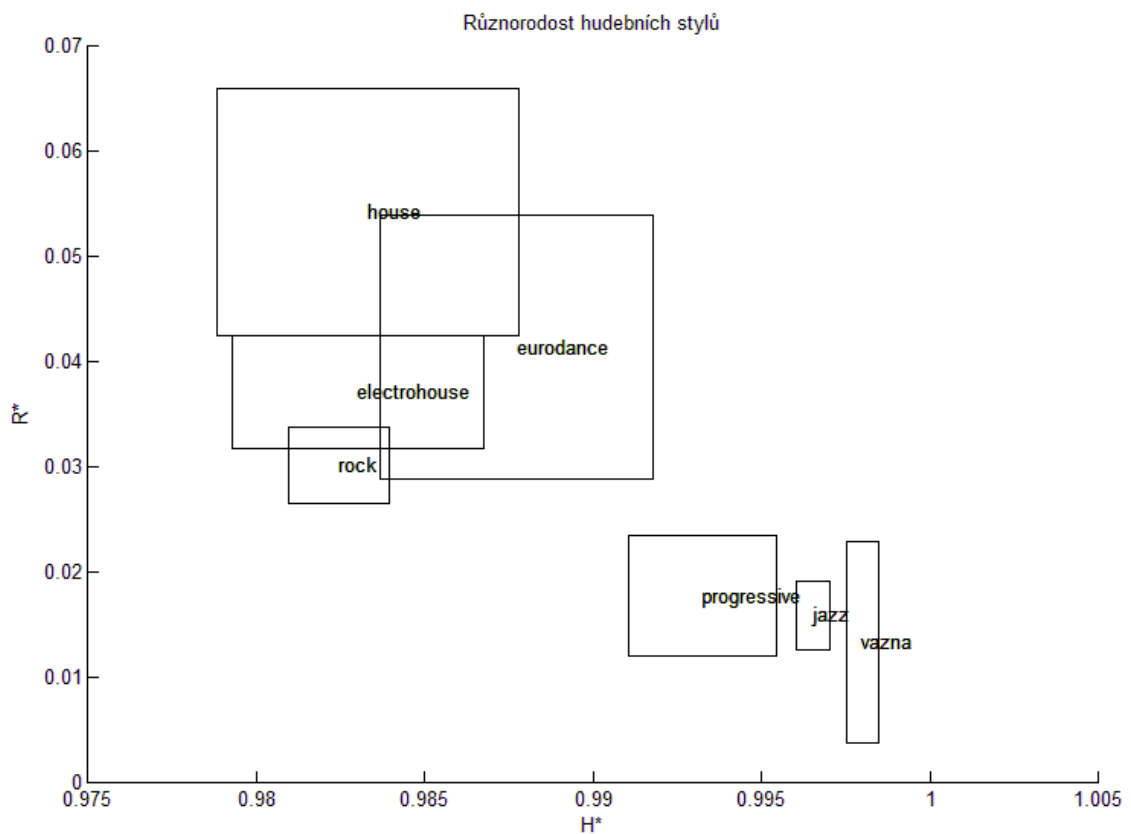
4 Výsledky

Na obr. 1 vidíme rozdělení jednotlivé ukázky zařazené do hudebních stylů podle rozsahu funkčních hodnot R a průměru H . Předvídatelnost je nejnižší v levém horním rohu a nejvyšší v pravém dolním. Na obr. 2 jsou oblasti typické pro každý hudební žánr, které byly vypočteny na základě znalosti průměru a směrodatné odchylky rozsahu funkčních hodnot. H^* je rovno průměru Hurstových exponentů pro daný žánr a R^* průměru rozsahů funkčních hodnot.

Podle našeho měření je vážná hudba nejpředvídatelnější, následuje jazz a progressive. Analýza jednotlivých ukázek těchto tří stylů a rocku se liší málo. Analýzy eurodance, electrohouse a house jsou rozdílné a tyto tři styly jsou si podobné (obr. 2).



Obr. 1 Hudební ukázky různých stylů



Obr. 2 Různorodost hudebních stylů

5 Závěr

Výsledky u vážné hudby a jazzu ukazují, že tato metoda má jistou vypovídací hodnotu v rozdělování žánrů hudby. S větším množstvím času, ukázek, zapálení a hudebních znalců by bylo možné zpřesnit výsledky a zaměřit se např. na závislost předvídatelnosti na použitých hudebních nástrojích. Metoda určování průchodů nulou se osvědčila.

6 Poděkování

Rádi bychom v první řadě poděkovali organizátorům Týdne vědy na Jaderce a vedení Fakulty jaderné a fyzikálně inženýrské ČVUT za uspořádání této jedinečné akce.

Velké poděkování dále patří hlavnímu organizátorovi, Vojtěchu Svobodovi a zbytku organizačního týmu - Ondřeji Groverovi, Kateřině Jirákové, Zbyňku Nguyenovi. Tento miniprojekt by však nebylo možné uskutečnit bez našeho supervizora Martina Dlaska, který nás zahrnoval hromadou úžasných nápadů.

7 Reference

- [1] DLASK, Martin, KUKAL, Jaromir, TRAN Quang Van. Revisited Zero-Crossing Method for Hurst Exponent Estimation in Time Series. In: *Mathematical Methods in Economics Conference Proceedings 2015*. Cheb: University of West Bohemia, 2015, s. 115-120. ISBN 978-80-261-0539-8.
- [2] JOHN W. LEIS. *Digital signal processing using matlab for students and researchers*. Hoboken, N. J.: Wiley, 2013. ISBN 978-11-180-3380-7.