

KONVOLUCIJSKI REKURZIVNI MODELI ZA OZNAČAVANJE GLAZBE

KRISTIJAN BARTOL

ZADATAK - OZNAČAVANJE GLAZBENIM ŽANROVIMA

- primjeri žanrova: rock, pop, jazz, instrumental, electronic, spoken...
- jedna pjesma može pripadati u više žanrova - *multi-output classification*
- problem s glazbom: **ispravne oznake?** (*ground truth*)



INTERPRETACIJA ZVUKA

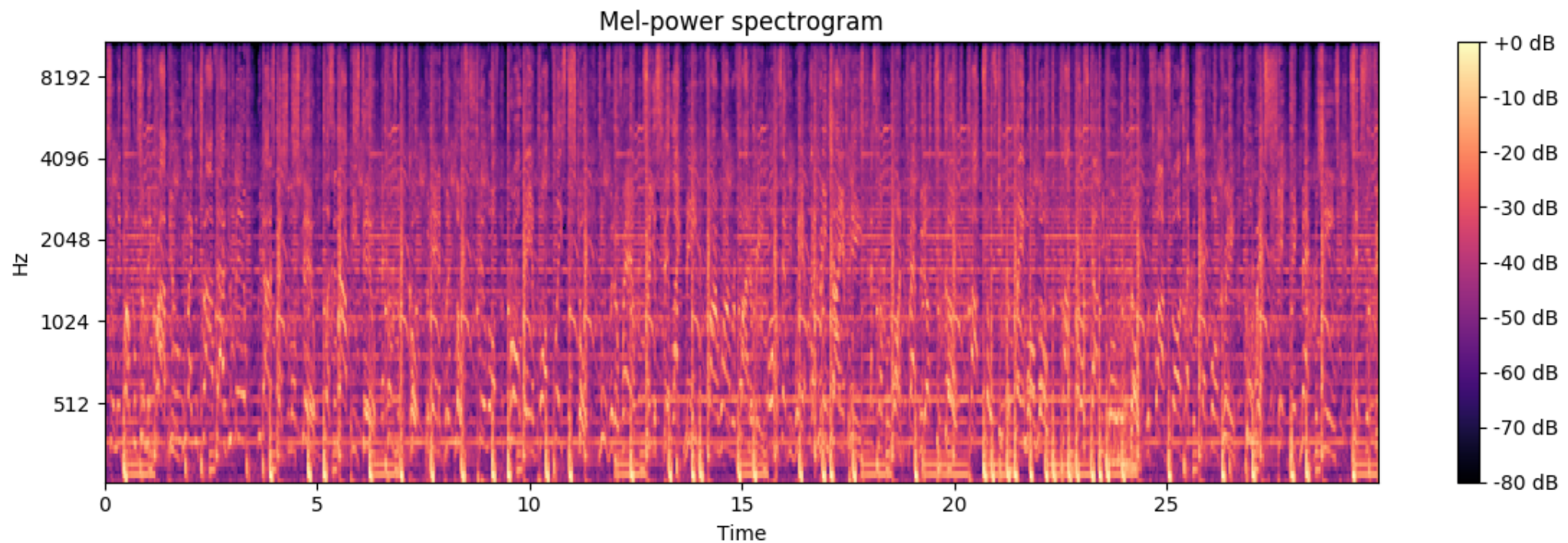


1 Second

<https://deepmind.com/blog/wavenet-generative-model-raw-audio/>

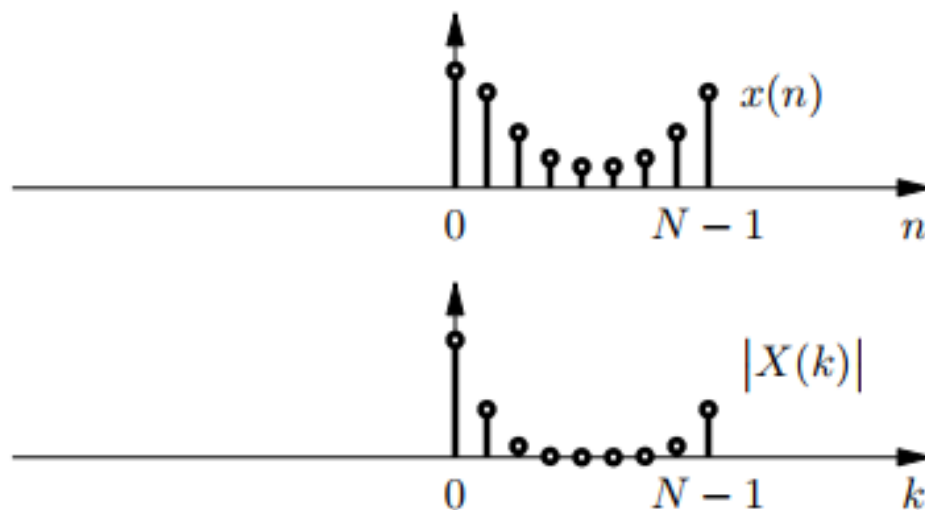


EKSTRAKCIJA (BITNIH) ZNAČAJKI - MEL-SPEKTROGRAMI



EKSTRAKCIJA BITNIH ZNAČAJKI - MEL-SPEKTROGRAMI

DFT



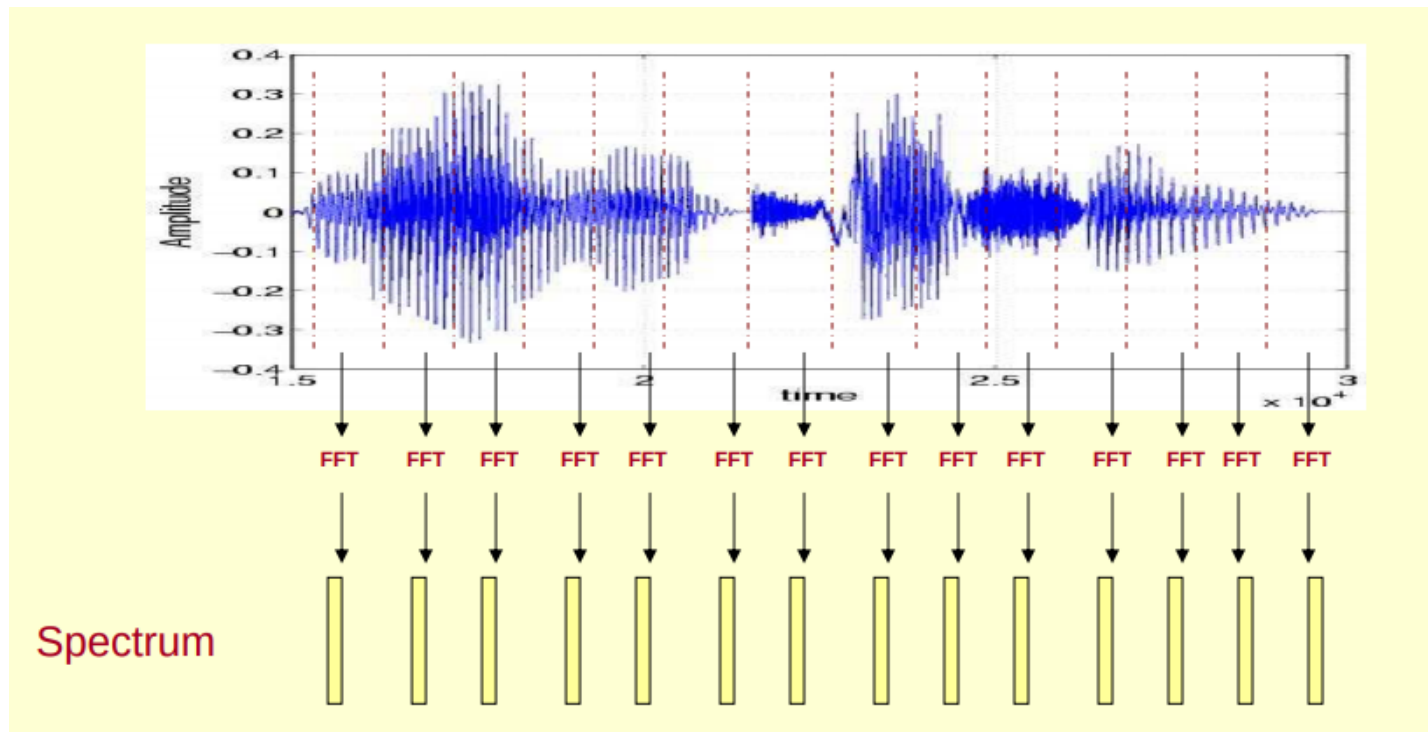
$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-kn}, \quad 0 \leq n \leq N-1$$

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{kn}, \quad 0 \leq k \leq N-1$$

$$X_k = x_0 e^{-jW_N^{k0}} + x_1 e^{-jW_N^{k1}} + \dots + x_n e^{-jW_N^{kn}}$$

$$W_N^{nk} = e^{-2\pi jnk/N}$$





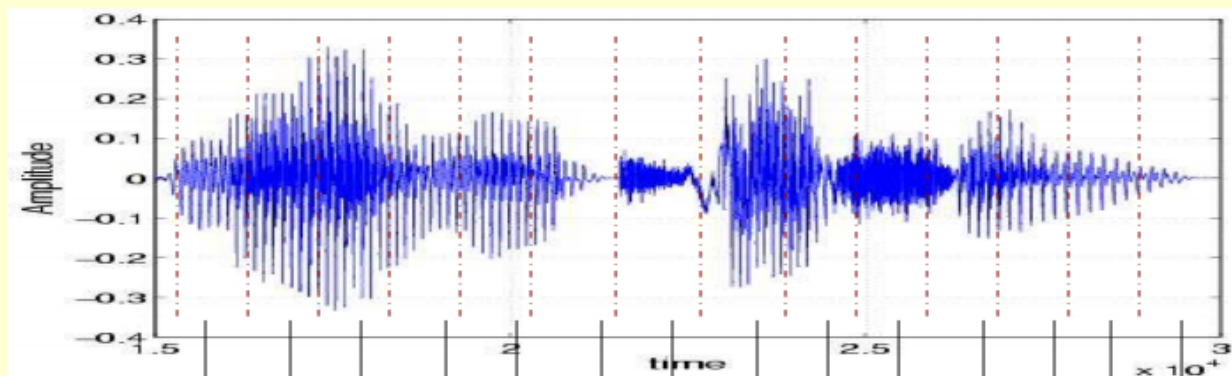
FFT = optimizirani DFT

EKSTRAKCIJA
(BITNIH)
ZNAČAJKI

-

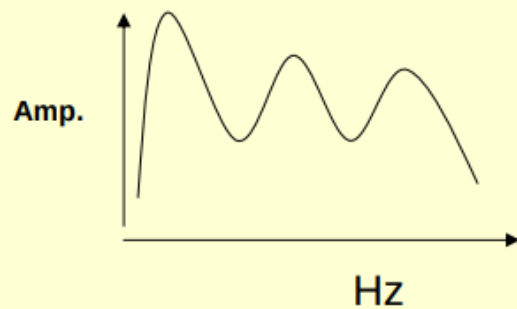
MEL-
SPEKTROGRAMI





FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT

Spectrum

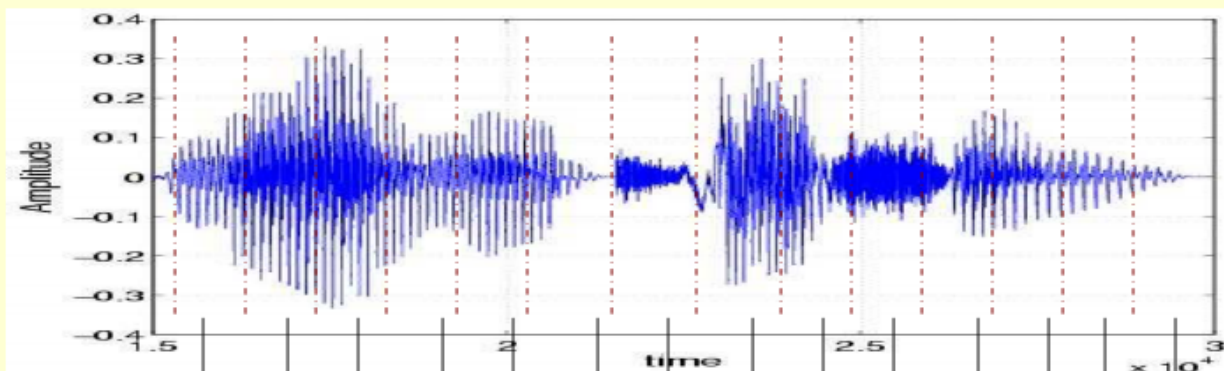


EKSTRAKCIJA
(BITNIH)
ZNAČAJKI

-

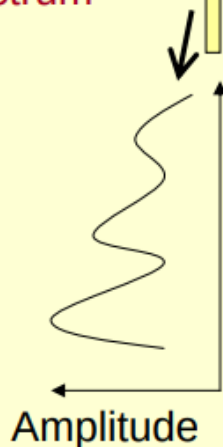
MEL-
SPEKTROGRAMI





FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT FFT

Spectrum



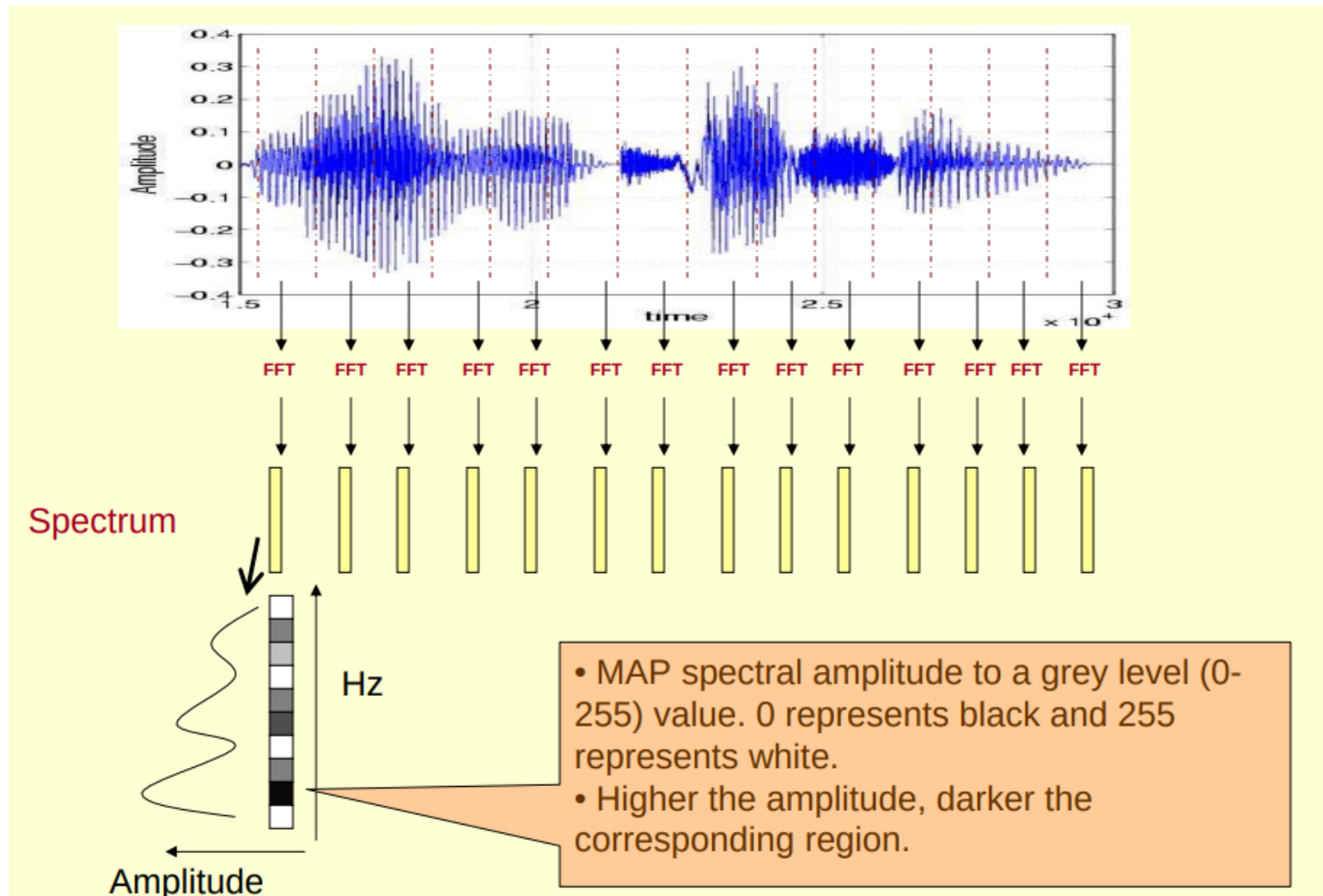
Rotate it by 90 degrees

EKSTRAKCIJA
(BITNIH)
ZNAČAJKI

-

MEL-
SPEKTROGRAMI



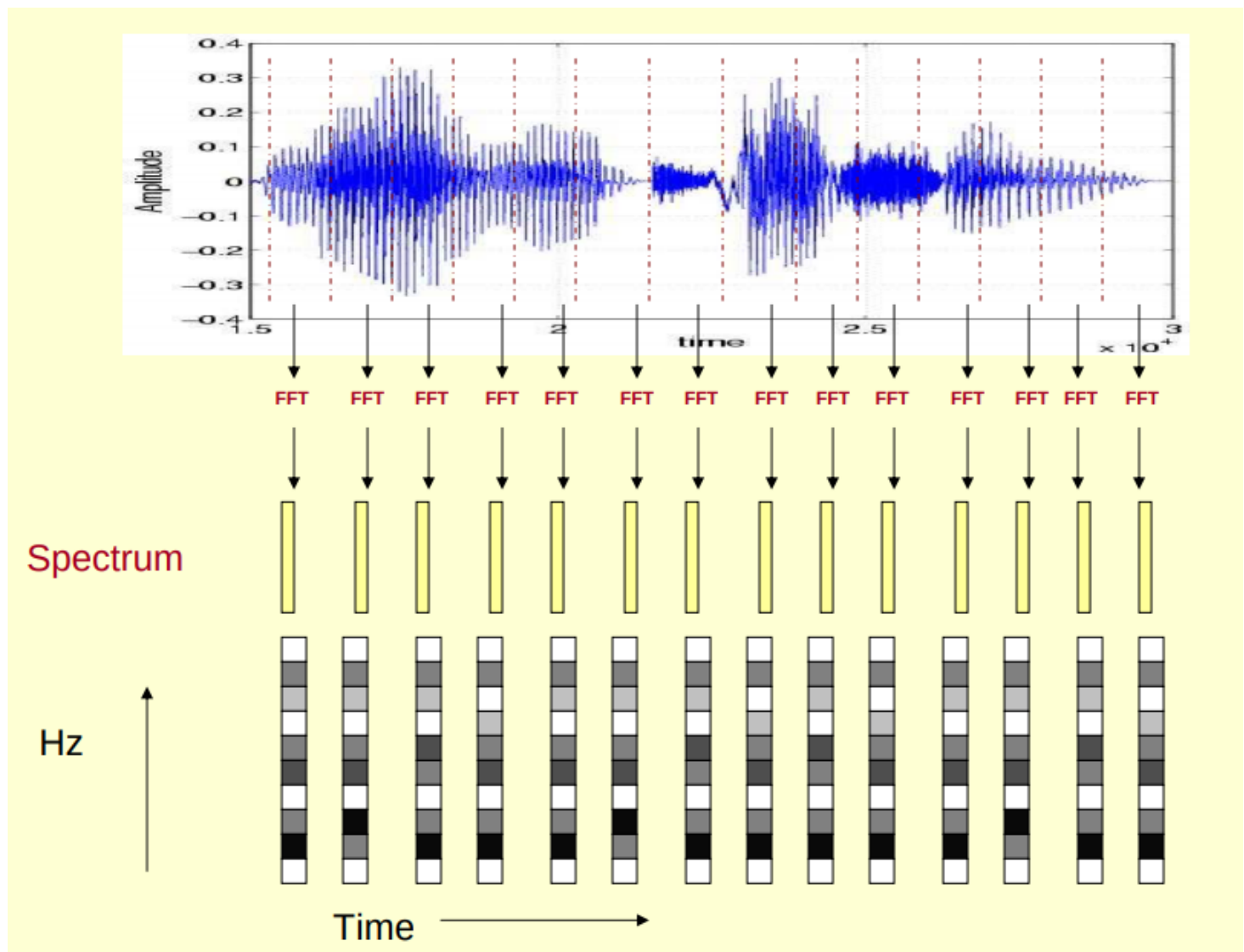


EKSTRAKCIJA (BITNIH) ZNAČAJKI

-

MEL- SPEKTROGRAMI





EKSTRAKCIJA (BITNIH) ZNAČAJKI

MEL- SPEKTROGRAMI



ZAŠTO "MEL"?

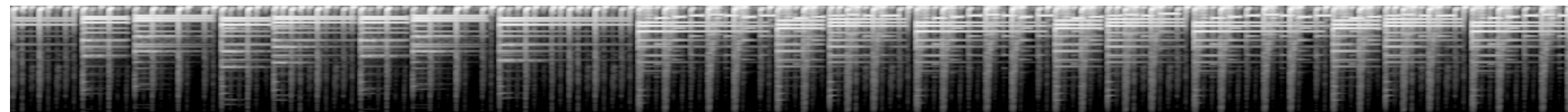
- melodijska skala – ljudima prirodnija
- nelinearna skala - moguća pretvorba u točnu vrijednost frekvencije i obrnuto

$$M(f) = 1125 \ln(1 + f/700)$$

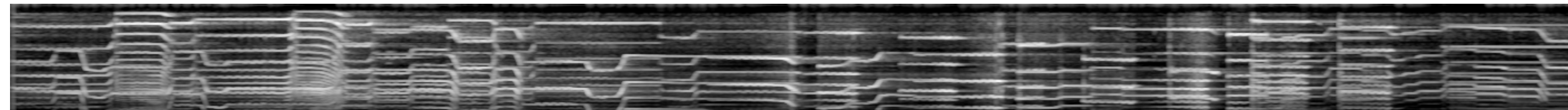
$$M^{-1}(m) = 700(\exp(m/1125) - 1)$$



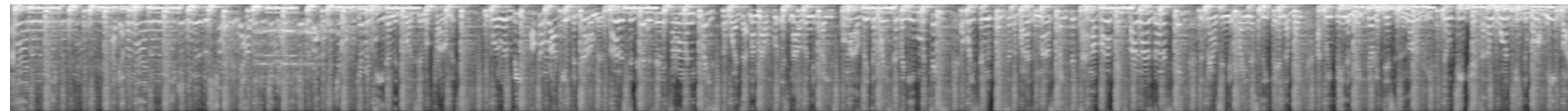
PRIMJERI IZRAČUNATIH SPEKTROGRAMA



Electronic



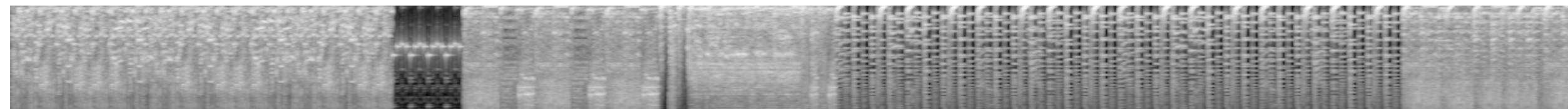
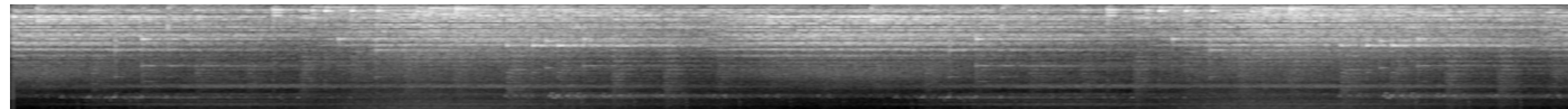
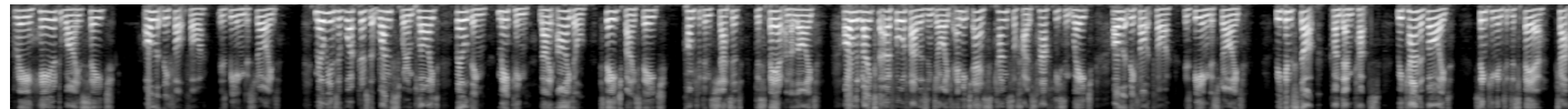
Experimental



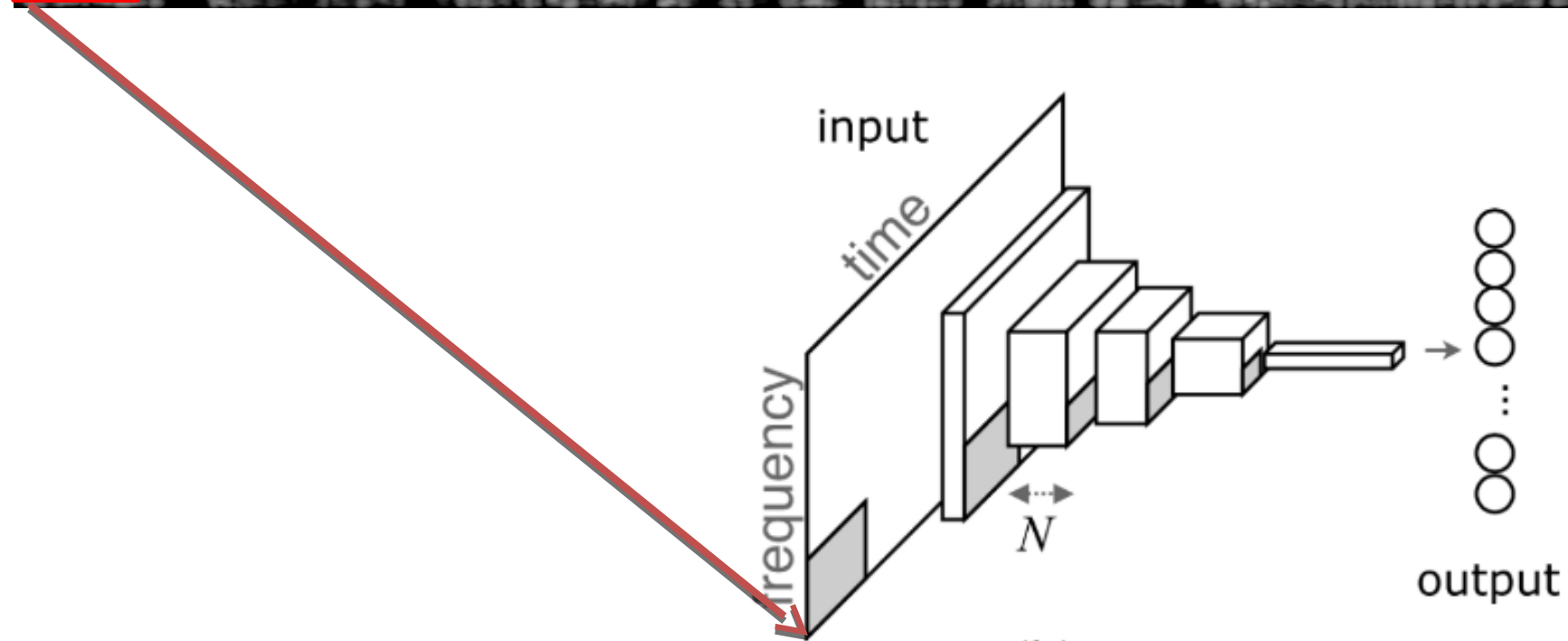
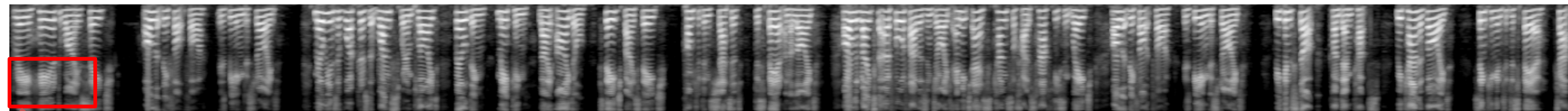
Hip-Hop



PRIMJERI IZRAČUNATIH SPEKTROGRAMA



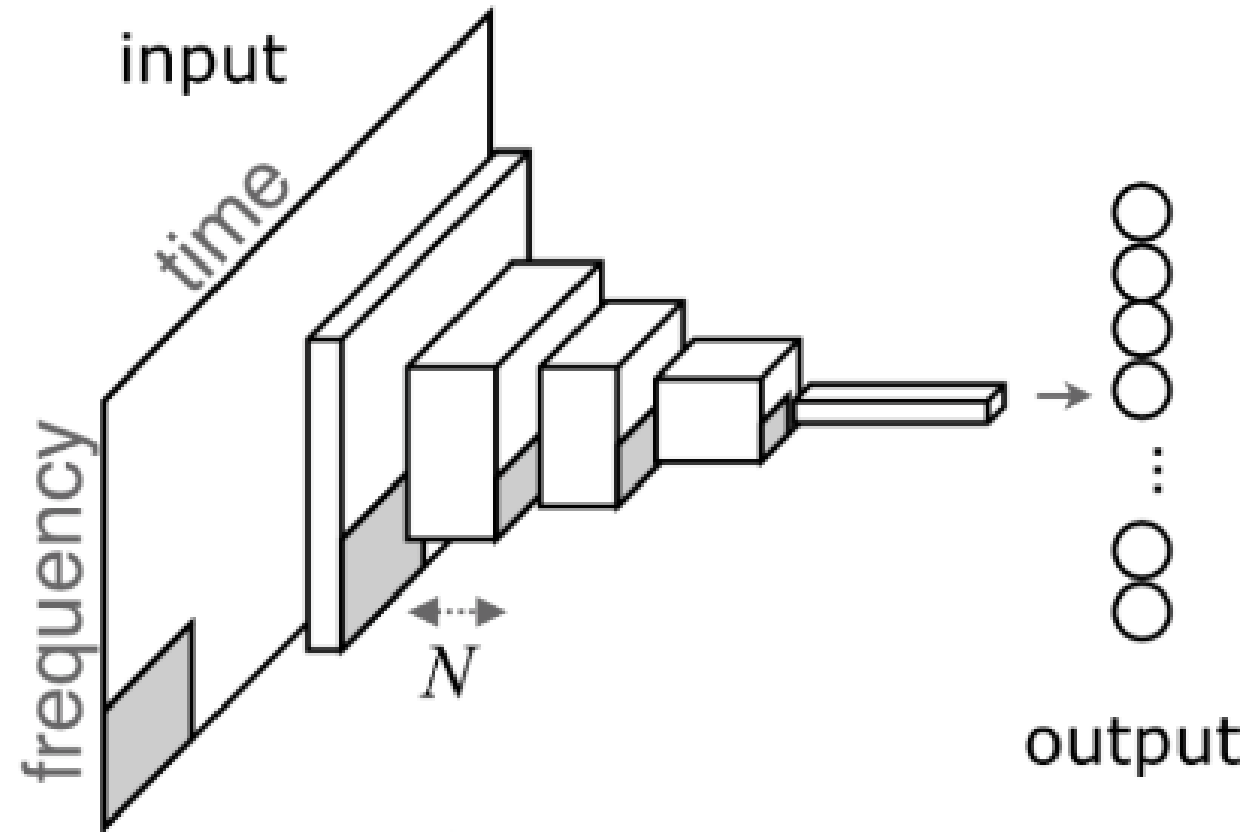
KONVOLUCIJSKA ARHITEKTURA - INSPIRACIJA IZ PODRUČJA SLIKE



KONVOLUCIJSKA ARHITEKTURA

- sakuplja lokalne značajke spektrograma (!)
- spektrogrami često imaju harmonijske korelacije; ne toliko lokalne [1]

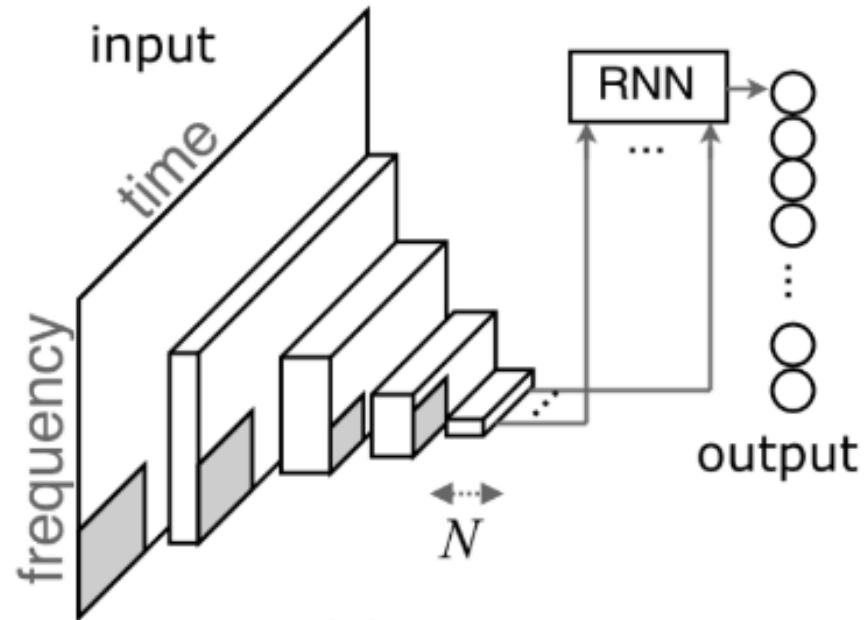
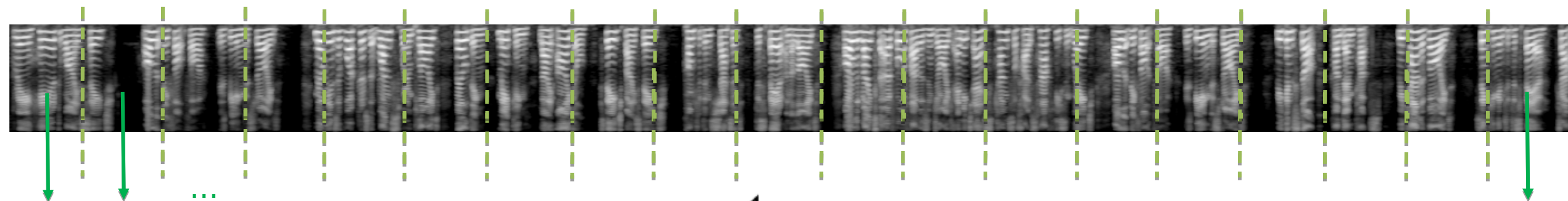
[1] <https://arxiv.org/pdf/1709.04396.pdf>



https://github.com/keunwoochoi/music-auto_tagging-keras

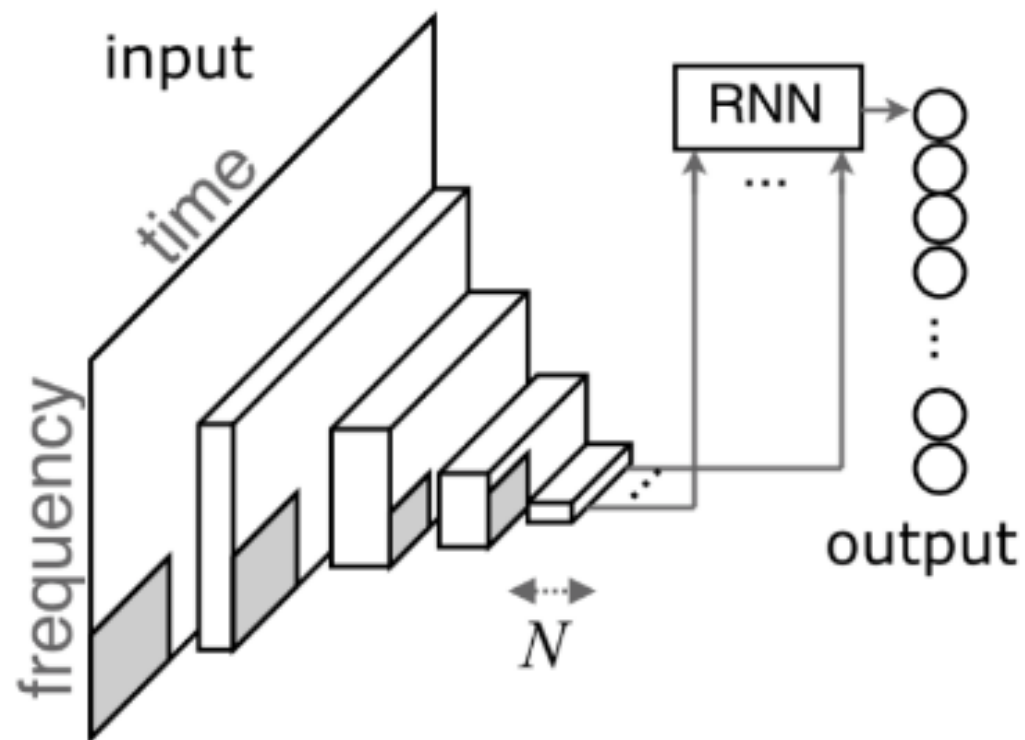
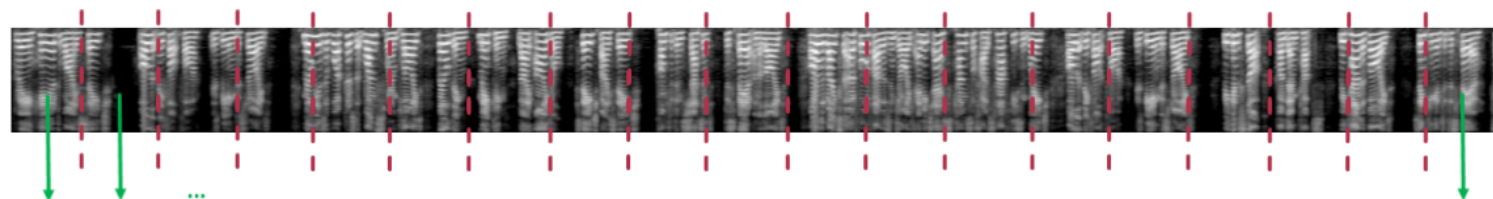


KONVOLUCIJSKA REKURZIVNA ARHITEKTURA



KONVOLUCIJSKA REKURZIVNA ARHITEKTURA

- razdvaja uzorak u sekvence što intuitivno više odgovara prirodi glazbe
- sakuplja globalne značajke spektrograma nakon konvolucijskih slojeva



https://github.com/keunwoochoi/music-auto_tagging-keras

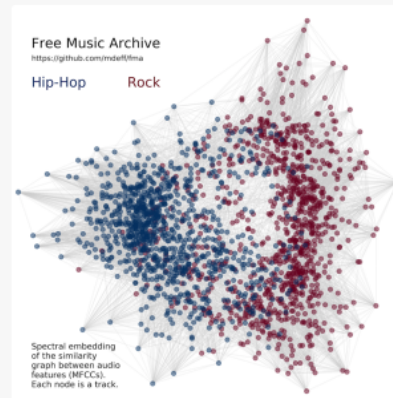


RADI LI U PRAKSI?

- proradilo jučer!
- treba ponovno pokušati trenirati nad čitavim skupom za učenje i minimizirati gubitak unakrsne entropije



[Challenges](#) [Knowledge Base](#) [Job Board](#) [Sign up](#) [Log in](#)



WWW 2018 Challenge: Learning to Recognize Musical Genre

Learning to Recognize Musical Genre from Audio on the Web

By **EPFL**

25 days left

7223 207 439
Views Participants Submissions

50

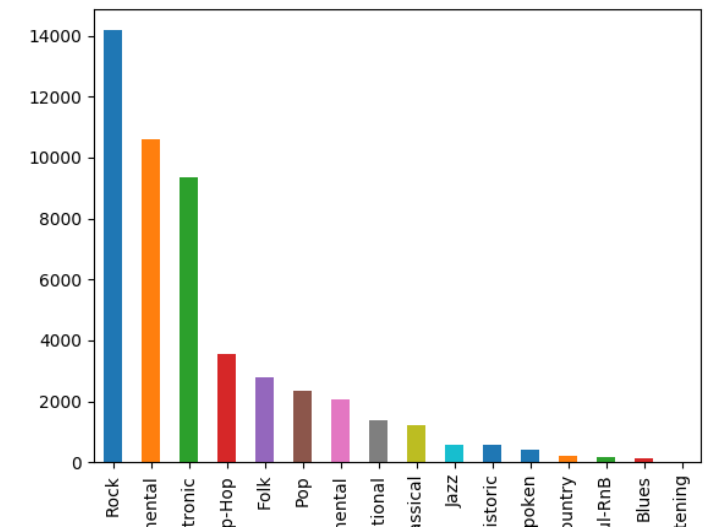
FOLLOW

<https://www.crowdai.org/challenges/www-2018-challenge-learning-to-recognize-musical-genre>



KORIŠTENI DATASET – FMA

- *Free Music Archive*, FMA ([Github](#))
- većina pjesama pod CC licencom
- ukupno 161 klasa



	Broj pjesama	Trajanje	Veličina
Mali	8000	30s	7.2GiB
Srednji	25.000	30s	22GiB
Veliki	105.574	30s	93GiB
Potpuni	105.574	Cijela pjesma	879GiB



OSTALI GLAZBENI DATASETI

- glavni problemi: licenca, mali broj primjera, skupovi bez audio primjera

	Broj pjesama	Različitih oznaka	Dostupan audio
GTZAN	1000	10	da
MSD	1000000	522.366	ne
Vlastiti dataset?	neograničeno	npr. 161	da*

<http://www.audiocontentanalysis.org/data-sets/>



"DEBUGGIRANJE" MREŽE

- odakle početi? - npr. *37 Reasons why your Neural Network is not working*
- mogući razni problemi s podacima (Učitavaju li se dobro? Pravilno pridružene izlazne oznake? Skup dovoljno velik?)
- moguće implementacijske pogreške (Uči li se dobra funkcija pogreške? Buggovi?)
- model nedovoljnog kapaciteta, nedovoljno vremena za učenje, itd...



ZAKLJUČAK

- problem je **težak**
- uspješna *multi-output* klasifikacija značila bi uspješnu **ekstrakciju bitnih značajki**



DALJNJE MOGUĆNOSTI

- nenadzirano grupiranje prema stilu / razdoblju / drugim apstraktnijim značajkama
- generiranje nove glazbe - mogućnost specifikacije žanra / stila / razdoblja
- umjetna kreativnost u glazbi

