

Opracowanie algorytmu generacji grafu DSP do rozwiązania problemu syntezy dźwięku

Design of a DSP graph generation algorithm for solving the sound
synthesis problem

Autor pracy: Mateusz Bączek

Opiekun pracy: Dr Inż. Maciej Hojda

Seminarium Dyplomowe – prezentacja 1

2023

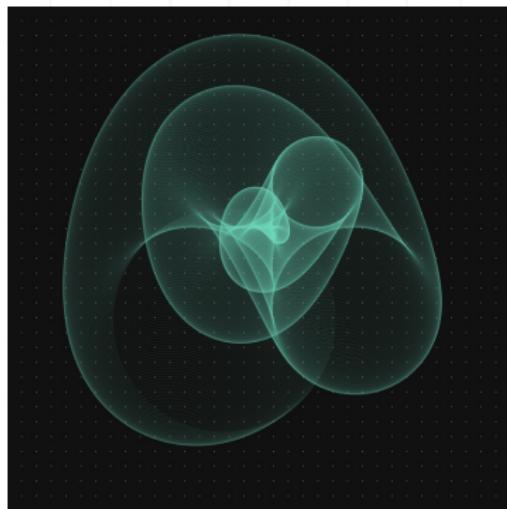


Politechnika Wrocławskiego

Sztuka generatywna

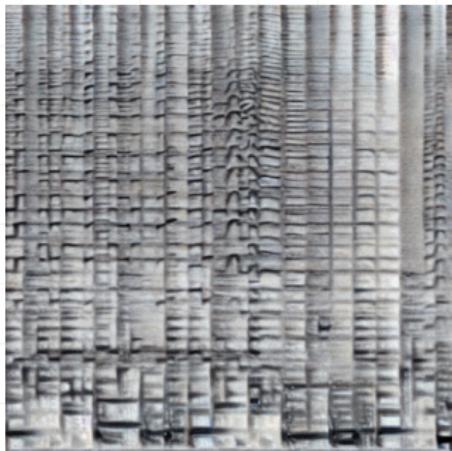


Rysunek: Obraz wygenerowany za pomocą algorytmu *Stable Diffusion*.



Rysunek: Rekursywna spirala wygenerowana w programie *Ronin*.

Algorytmiczna kompozycja muzyki



Rysunek: Spektrogram
wygenerowany za pomocą
algorytmu *Stable Riffusion*.

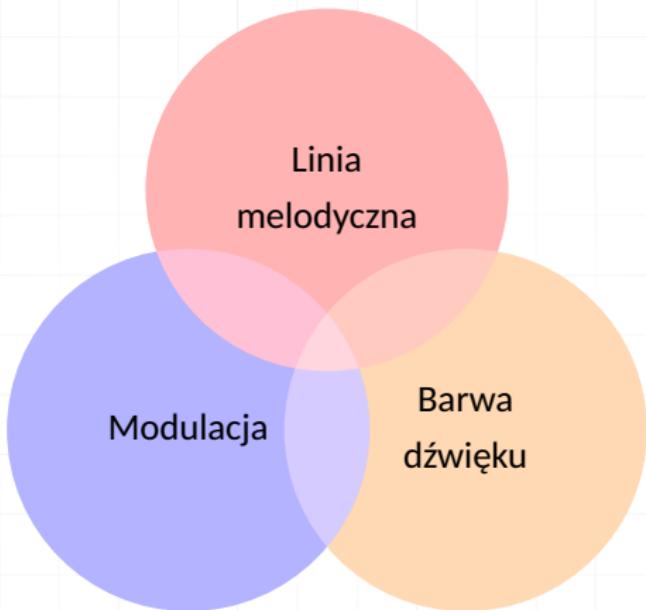


Rysunek: Zapis nutowy utworu
wygenerowany przez komputer
Lamus.

Popularne podejścia do kompozycji algorytmicznej

1. Generowanie zapisu nutowego [Zhang, 2023]:
 - ▶ wykorzystuje reguły matematyczne oparte o teorię muzyki,
 - ▶ generuje zapis symboliczny, który łatwo później wykorzystać.
2. Generowanie pełnego pliku audio [Forsgren, 2023]:
 - ▶ teoretycznie najbardziej imponujący efekt działania,
 - ▶ małe możliwości wykorzystania rezultatu algorytmu.
3. Symulacja instrumentów za pomocą sieci neuronowych [Engel, 2017]:
 - ▶ potencjalnie bardzo szeroka gama generowanych brzmień,
 - ▶ małe możliwości dostosowania brzmienia,
 - ▶ słaby wgląd we „wnętrze” wygenerowanego algorytmu DSP.

Elementy kompozycji muzycznej (w uproszczeniu)

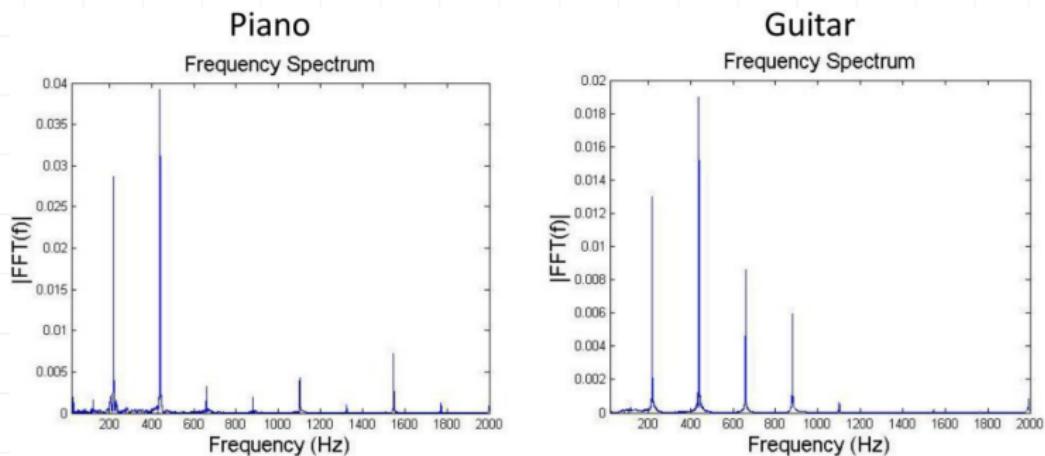


Rysunek: Elementy opisujące pojedynczy głos/instrument w kompozycji muzycznej.

Cel pracy

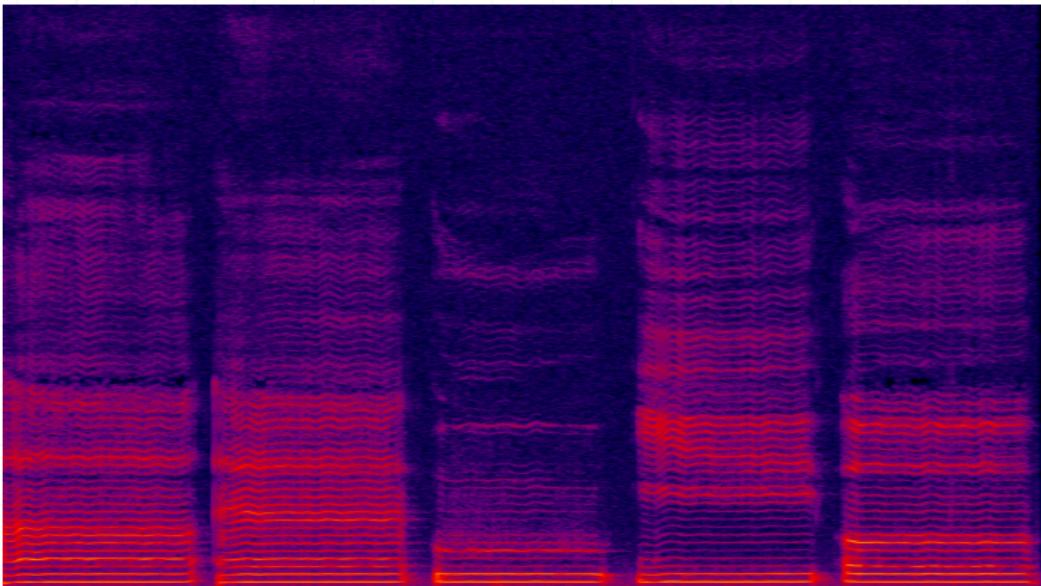
Wytwarzanie grafu przetwarzania sygnałów,
generującego dźwięk o określonej **barwie**.

Czym jest barwa dźwięku?



Rysunek: Porównanie transformaty Fouriera dla tej samej nuty granej na pianinie i na gitarze.

Czym jest **barwa** dźwięku?

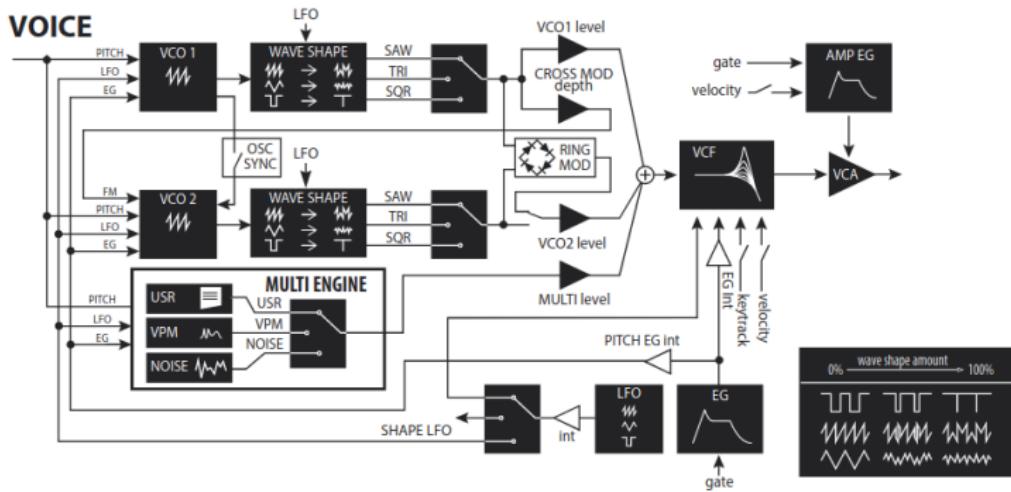


Rysunek: Spektrogram (transformata Fouriera w czasie) dla nagrania samogłosek „a, e, u, i, o”.

Skąd bierze się barwa dźwięku?

- ▶ Materiał, z którego wykonane są struny,
- ▶ pudło rezonansowe (lub jego brak),
- ▶ kształt pudła rezonansowego,
- ▶ materiał, z którego wykonany jest instrument,
- ▶ materiał, który uderza w struny,
- ▶ ... i wiele innych cech fizycznych instrumentu.

Jak kontroluje się barwę w syntezatorach dźwięku



Rysunek: Diagram blokowy pojedynczego głosu w syntezatorze
Korg Minilogue xd [Korg, 2019].

Generalizacja tradycyjnego procesu syntezy dźwięku

1. Bloki generujące sygnał:

- 1.1 synteza prostych sygnałów (sinus, trójkąt, prostokątny),
- 1.2 odtwarzanie sampli dźwiękowych,
- 1.3 sygnał modulujący (rytm, ADSR).

2. Bloki przetwarzające sygnał:

- 2.1 filtry (górnoprzepustowy, dolnoprzepustowy, itd.),
- 2.2 efekty (pogłos, echo).

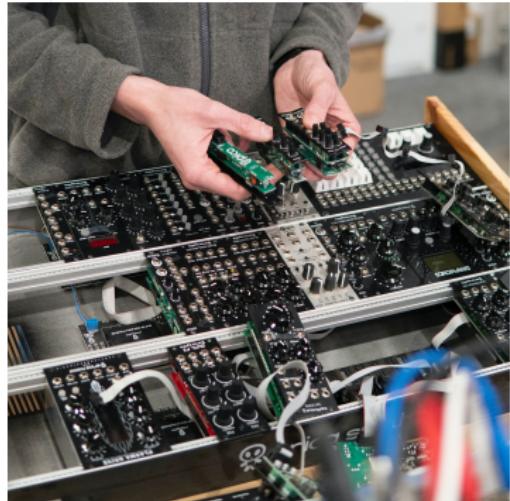
3. Połączenia między blokami:

- 3.1 modulowanie parametrów generacji/przetwarzania sygnału.

Synteza oparta o moduły

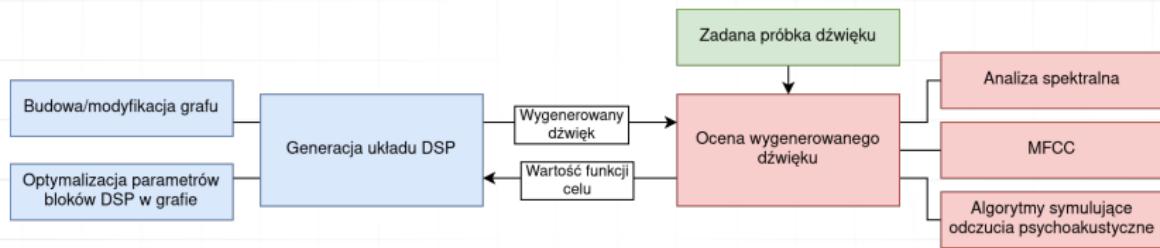


Rysunek: Oprogramowanie Bespoke.



Rysunek: Syntezator modułowy.

Jak algorytmicznie wytworzyć układ DSP generujący zadany dźwięk?



Rysunek: Diagram algorytmu realizowanego w ramach pracy.

Aspekt badawczy pracy

Do końca maja:

- ▶ reprezentacja grafu przetwarzania sygnałów [**Macret**]:
 - ▶ pozwalająca na automatyczne modyfikacje grafu,
 - ▶ zachowującą „sens” w domenie syntezy dźwięku.

Do końca kwietnia:

- ▶ analiza dostępnych algorytmów porównujących barwę dźwięku [**Yan**] [**Jacobsen**] [**Caspe**],
- ▶ wybór funkcji celu spośród przeanalizowanych algorytmów,
- ▶ wybór algorytmu optymalizującego parametry wewnętrz grafu.

Aspekt inżynierski pracy

Już gotowe:

- ▶ Implementacja bazowych algorytmów DSP w wydajnym, komplikowanym języku .

Do końca maja:

- ▶ generyczny algorytm syntezy dźwięku na podstawie grafu DSP,
- ▶ interfejs graficzny, wizualizacja grafu.

Aspekt inżynierski pracy

Do końca kwietnia:

- ▶ integracja algorytmu syntezy z API języka Python,
- ▶ implementacja/integracja algorytmów „oceniających” bliskość dźwięku do zadanej próbki,
- ▶ wykorzystanie narzędzi optymalizacyjnych z pakietów obliczeniowych języka Python do optymalizacji grafu DSP.

Podobne prace

Automatic design of sound synthesizers (...) [Macret]

- ▶ Niewielki zbiór testowanych operacji DSP w grafie,
- ▶ wykorzystanie tylko MFCC jako funkcji celu,
- ▶ słabe wyniki dla dźwięków o spektrum zmiennym w czasie.

DDX7: Differentiable FM Synthesis (...) [Caspe]

- ▶ Wykorzystanie tylko jednego algorytmu syntezy (FM),
- ▶ zadana z góry struktura grafu DSP,
- ▶ sieć neuronowa jako źródło modulacji.



Zhang, Li and Callison-Burch, Chris

Language Models are Drummers: Drum Composition with
Natural Language Pre-Training

<https://arxiv.org/abs/2301.01162>



Forsgren, Seth* and Martiros, Hayk*

Riffusion - Stable diffusion for real-time music generation

<https://riffusion.com/about>,



Jesse and Resnick, Cinjon and Roberts, Adam and Dieleman,
Sander and Eck, Douglas and Simonyan, Karen and Norouzi,
Mohammad

Neural Audio Synthesis of Musical Notes with WaveNet
Autoencoders

<https://arxiv.org/abs/1704.01279>

-  Korg Inc,
Korg Minilogue xd User Manual
<https://www.korg.com/us/support/download/manual/0/811/4277/>
-  Yan Ke and Hoiem, D. and Sukthankar, R.
Computer vision for music identification
10.1109/CVPR.2005.105
-  M. Pasquier, P.
Automatic Design of Sound Synthesizers as Pure Data Patches
using Coevolutionary Mixed-typed Cartesian Genetic
Programming
<https://metacreation.net/wp-content/uploads/2015/08/p309-macret.pdf>

- ❑ Caspe, Franco and McPherson, Andrew and Sandler, Mark
DDX7: Differentiable FM Synthesis of Musical Instrument Sounds
<https://arxiv.org/abs/2208.06169>
- ❑ Jacobsen, E. and Lyons, R.
The Sliding Fourier
10.1109/MSP.2003.1184347



Rysunek: Obraz wygenerowany przez DALL-E dla promptu „futuristic musical instrument”.

Dziękuję za uwagę!