香港中文大學
The Chinese University of Hong Kong

A Very Quick Overview to AI Generation of Music

Dr Chuck-jee CHAU 周卓之

Dept of Computer Science and Engineering,
The Chinese University of Hong Kong

Let’s have some music?

The Lazy Composers

- *Musikalisches Würfelspiel* in the 18th century
  - *"To compose without the least knowledge of Music so much Countrydances as one pleases, by throwing a certain Number with two Dice."*
  - To be exact, $2 \times 11^4 =$ 759,499,667,166,482 waltzes *(according to Wikipedia)*
  - http://www.playonlinedicegames.com/mozart
- "***Chance music***" / "Aleatoric music"

Image: https://i.redd.it/deh5m6taf0s51.png

WOLFGANG AMADEUS MOZART

Musikalisches Würfelspiel

Table of Measure Numbers

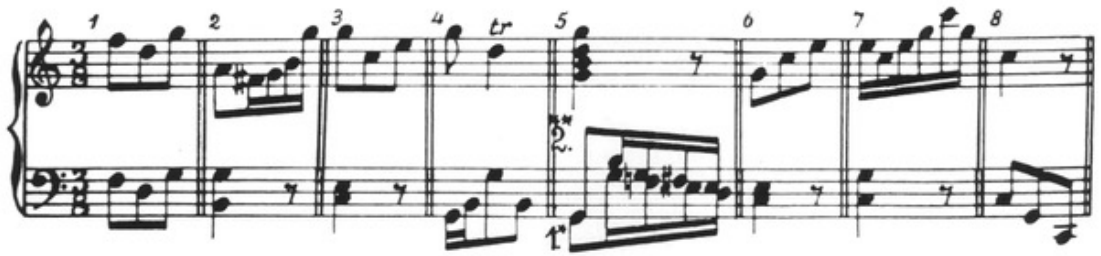
Part One

	I	II	III	IV	V	VI	VII	VIII
2	96	22	141	41	105	122	11	30
3	32	6	128	63	146	46	134	81
4	69	95	158	13	153	55	110	24
5	40	17	113	85	161	2	159	100
6	148	74	163	45	80	97	36	107
7	104	157	27	167	154	68	118	91
8	152	60	171	53	99	133	21	127
9	119	84	114	50	140	86	169	94
10	98	142	42	156	75	129	62	123
11	3	87	165	61	135	47	147	33
12	54	130	10	103	28	37	106	5

Part Two

	I	II	III	IV	V	VI	VII	VIII
2	70	121	26	9	112	49	109	14
3	117	39	126	56	174	18	116	83
4	66	139	15	132	73	58	145	79
5	90	176	7	34	67	160	52	170
6	25	143	64	125	76	136	1	93
7	138	71	150	29	101	162	23	151
8	16	155	57	175	43	168	89	172
9	120	88	48	166	51	115	72	111
10	65	77	19	82	137	38	149	8
11	102	4	31	164	144	59	173	78
12	35	20	108	92	12	124	44	131

Table of Measures



The Computer as a Composer

- ***Algorithmic composition*** — using algorithms to create music
  - Mathematical models
  - Knowledge based systems
  - Grammars
  - Evolutionary methods

- Music by neural networks
  - First trained music NN in 1989 at Stanford
  - *Big question: What is the next music note?*

Image: https://en.wikipedia.org/wiki/AIVA

AIVA (2016)

https://www.aiva.ai

The first virtual composer acknowledged by a music society (SACEM)

Google’s Magenta: AI Duet (2017)

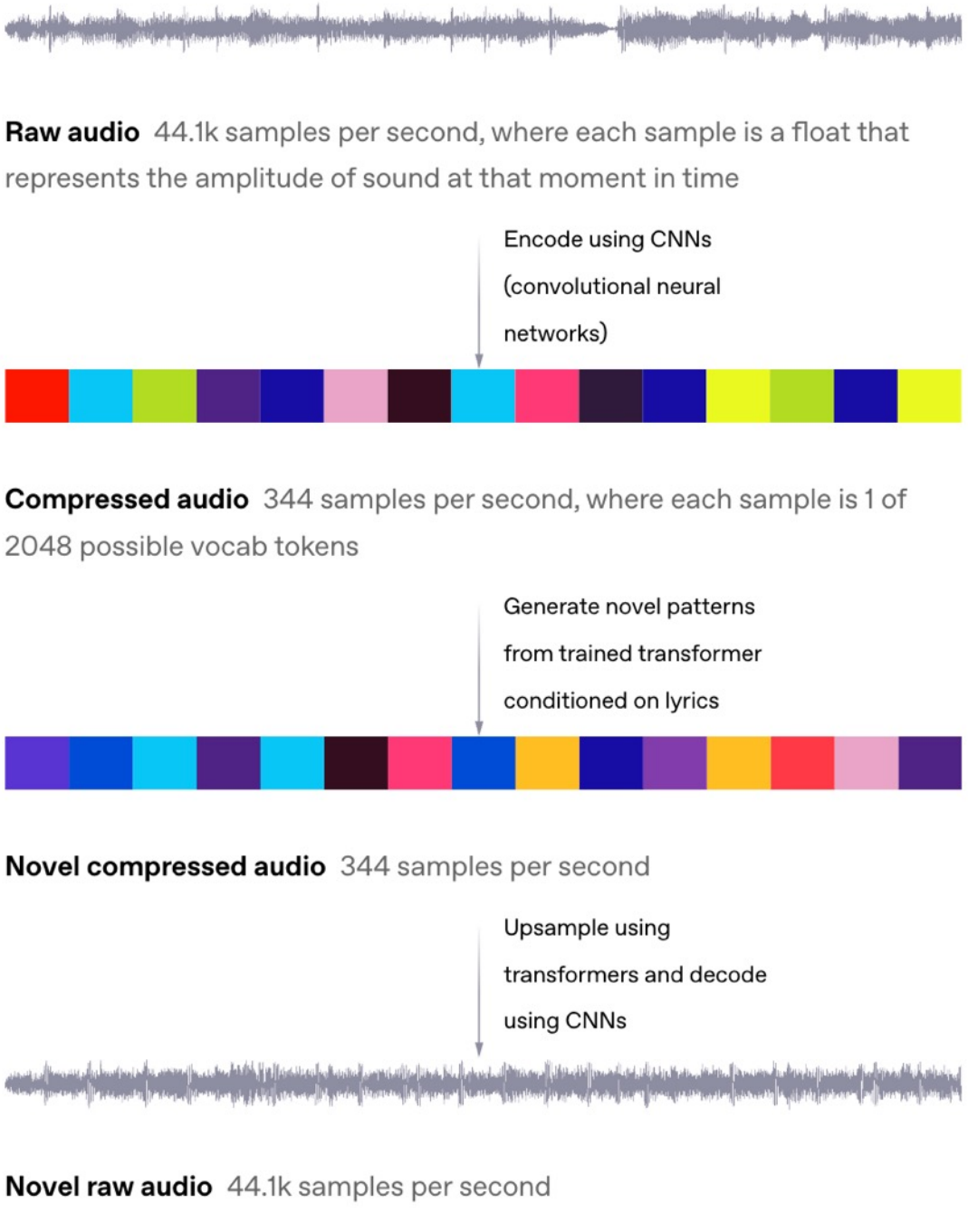
https://experiments.withgoogle.com/ai/ai-duet/view/

Google’s developer toolbox to algorithmic art

OpenAI's Jukebox (2020)

https://openai.com/index/jukebox/

Compressing and decompressing



Riffusion and Dance Diffusion (2022)

https://wandb.ai/wandb_gen/audio/reports/A-Gentle-Introduction-to-Dance-Diffusion--VmlldzoyNjg1Mzky

Generating audio as images from text prompts

Google’s MusicLM (2023)

https://google-research.github.io/seanet/musiclm/examples/

From text prompt to music clips

Suno and Bark (2024)

https://suno.com

One big commercial success

Imitation...

hkbirdcall2025 and 2 others

hkbirdcall2025 “🎶鳴就明 — 鳥聲模仿大賽2025 即日起正式接受報名!🐓” 升Key雀嘅叫聲，相信香港人都唔陌生。聽咗咁多次噪鵑叫，我哋又有幾了解佢哋？就算你發夢都聽到佢，清醒嘅你又有冇信心叫到佢咁？

一班朝朝比噪鵑叫醒嘅香港人，特別舉辦「鳴就明」鳥聲模仿大賽，俾個機會全世界一齊學噪鵑咁歇斯底里地叫返次。即刻參賽，爭奪成為全港第一位「鳥鳴之王」，過程入面更加可以學習到本地鳥類鳴聲背後嘅深義。

【參賽資格】 僅接受個人報名。歡迎十二歲或以上人士參加。

【賽制介紹】
⭐報名+初賽：參賽者在報名時須參考我們提供的音源（YouTube Link in

Liked by joeywanc and 2,509 others
June 11

Add a comment... Post

Video: https://youtu.be/mSvHU6hz-dI

Imitation and AI

Given large amount of data, it can “create”

Music Data Labelling

Where do we obtain data?

- The Internet!
  - Database of CD Records
  - Lyrics
  - Music reviews or analysis
  - Social tagging
  - “Metadata”
- ***Rough understanding*** only
- Can we regenerate music from these ideas?

Datasets

- Music captioning
  - MusicCaps (from MusicLM)
  - Song Describer https://github.com/mulab-mir/song-describer-dataset
    - *“A retro-futurist drum machine groove drenched in bubbly synthetic sound effects and a hint of an acid bassline.”*
    - *“Elegant and sophisticated Latin jazz piece with a Cuban base and a whispered melodic female voice.”*
    - *“Calm sitar and Indian tabla with dramatic synthetic strings background.”*

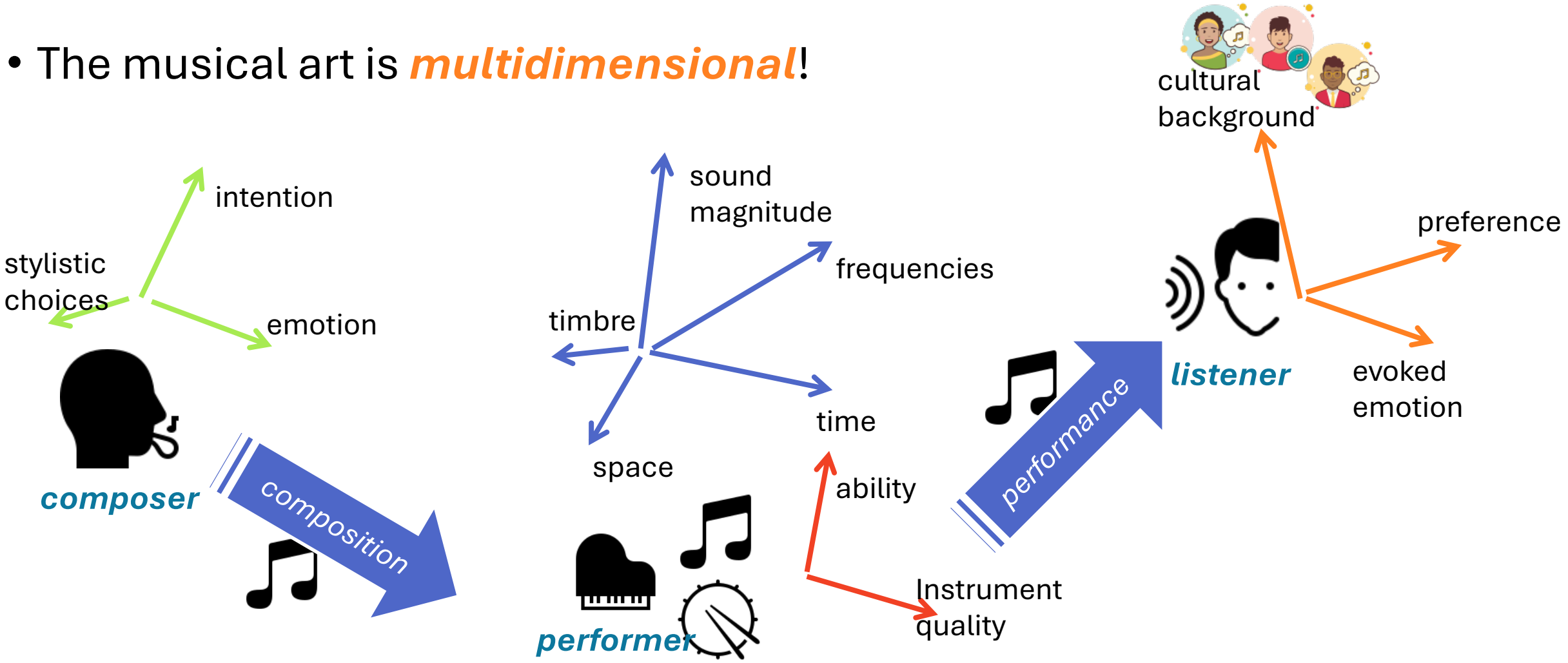
- More ***direct descriptions*** provided by human (expert) labellers
- Multi-modal language modal

Captions ⇄ Music

Multi-modal language models

Music as an Art...

- The musical art is ***multidimensional***!



Music Information Retrieval

Machine Learning vs. “Machine *Listening*”

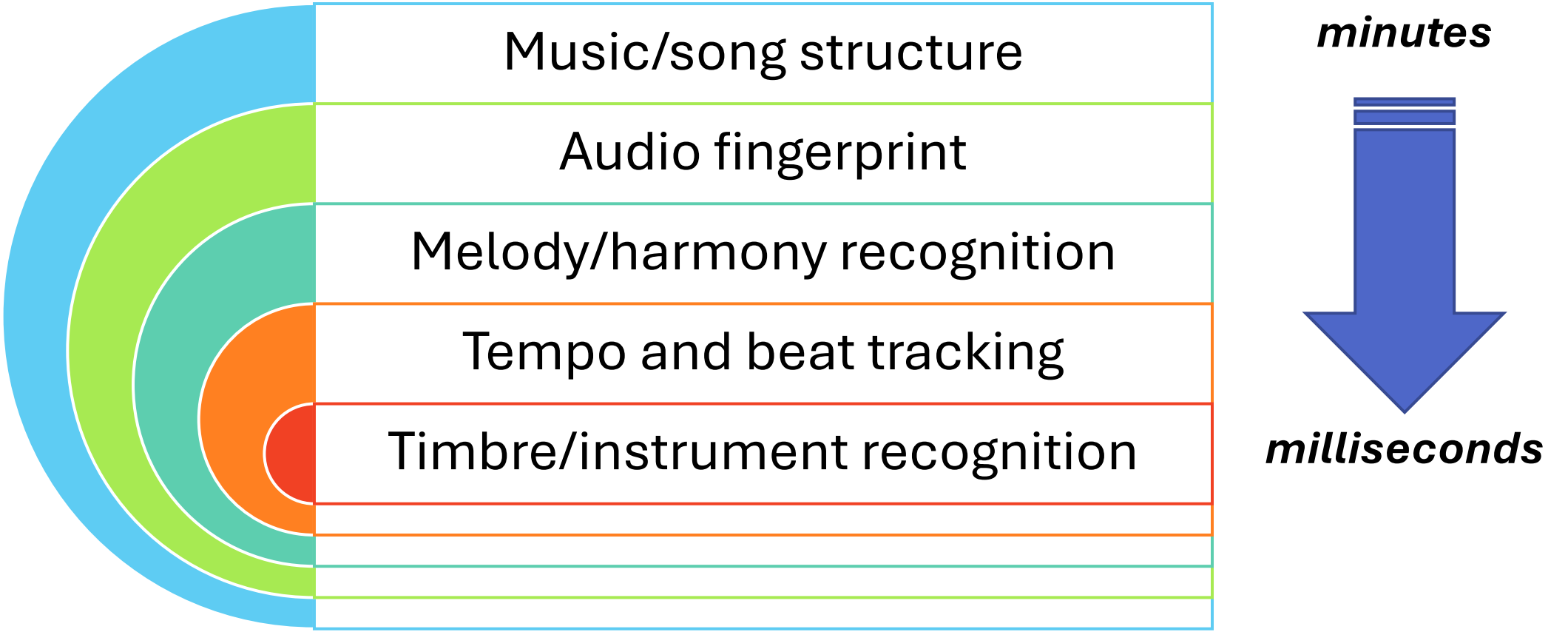
Music Information Retrieval

- MIR is an ***interdisciplinary*** field
  - Musical studies
  - Psychoacoustics and acoustics
  - Signal processing and analysis
  - Statistics
  - Machine learning and AI
  - Cognitive science and psychology
  - and could be more!

A dedicated research area

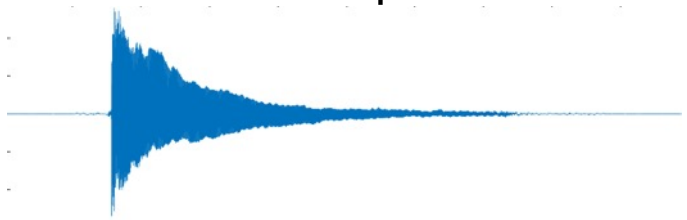
- International Society of Music Information Retrieval (***ISMIR***) conference (since 2000): *https://www.ismir.net/conferences*

From Global to Local Characteristics

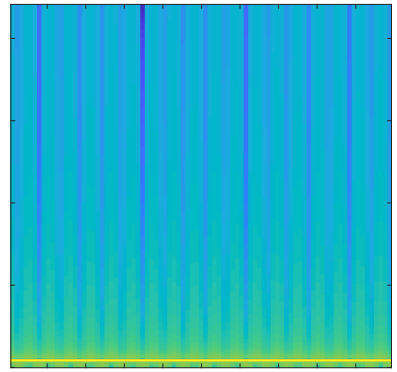


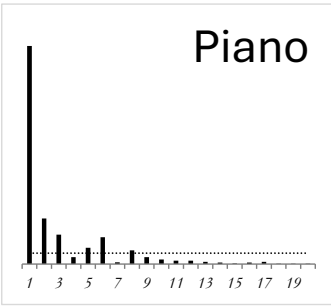
Feature Extraction

- Music analysis aims to produce ***acoustics features*** for further processing and compact representation, e.g.,
  - ***Time domain features***
    - Dynamic changes in music
    - Envelope of individual



- ***Frequency domain features***
  - Fundamental frequency
  - Spectrum characteristics
- ***Harmonic domain features***
  - Relationship or ratio between harmonics
  - Music instruments
- ***Music features***
  - Tonal centre, pitch and melody
  - Rhythm/pulse clarity, beat tracking





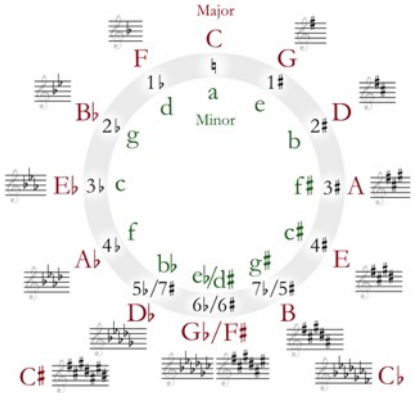
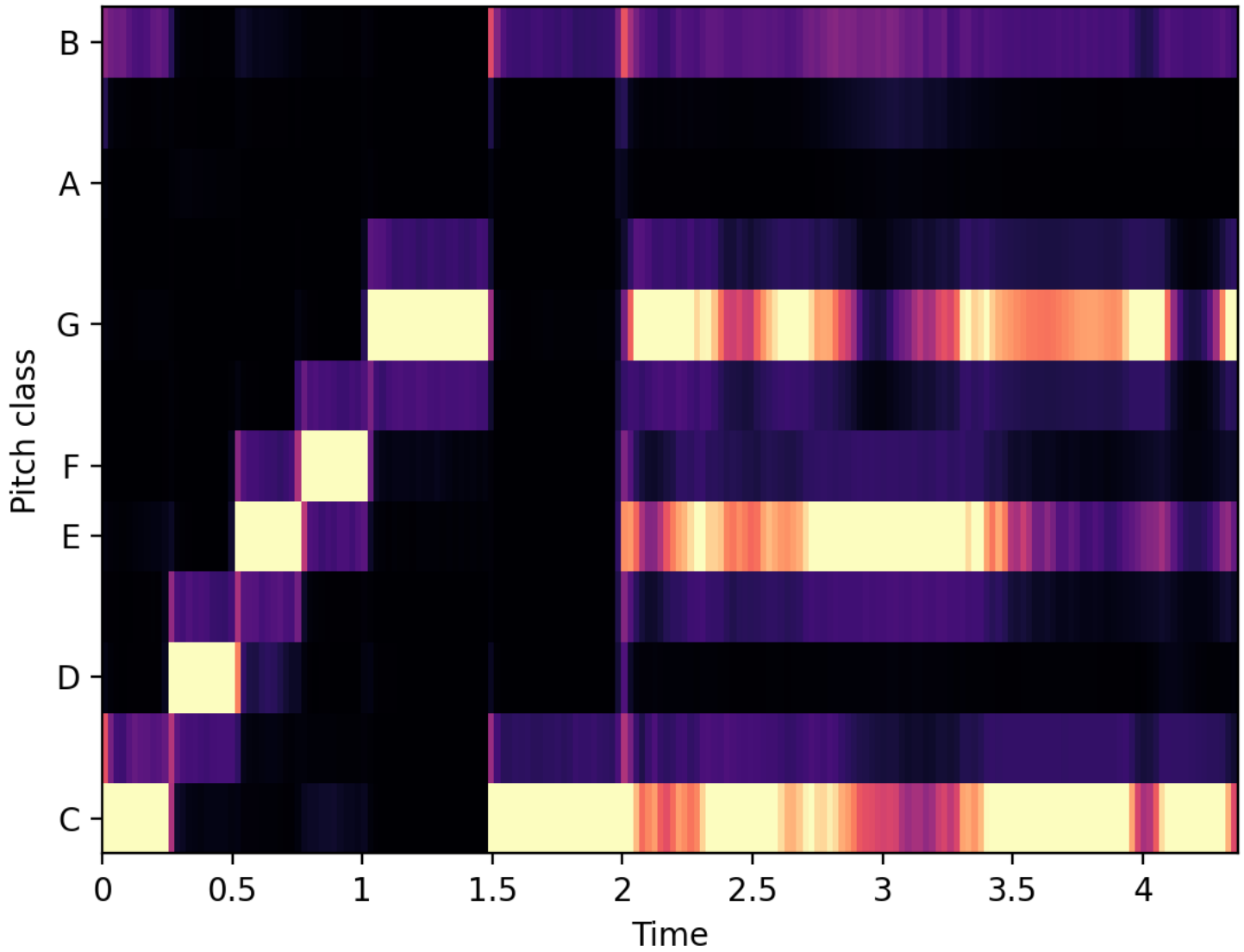


Image from: https://en.wikipedia.org/wiki/Key_signature

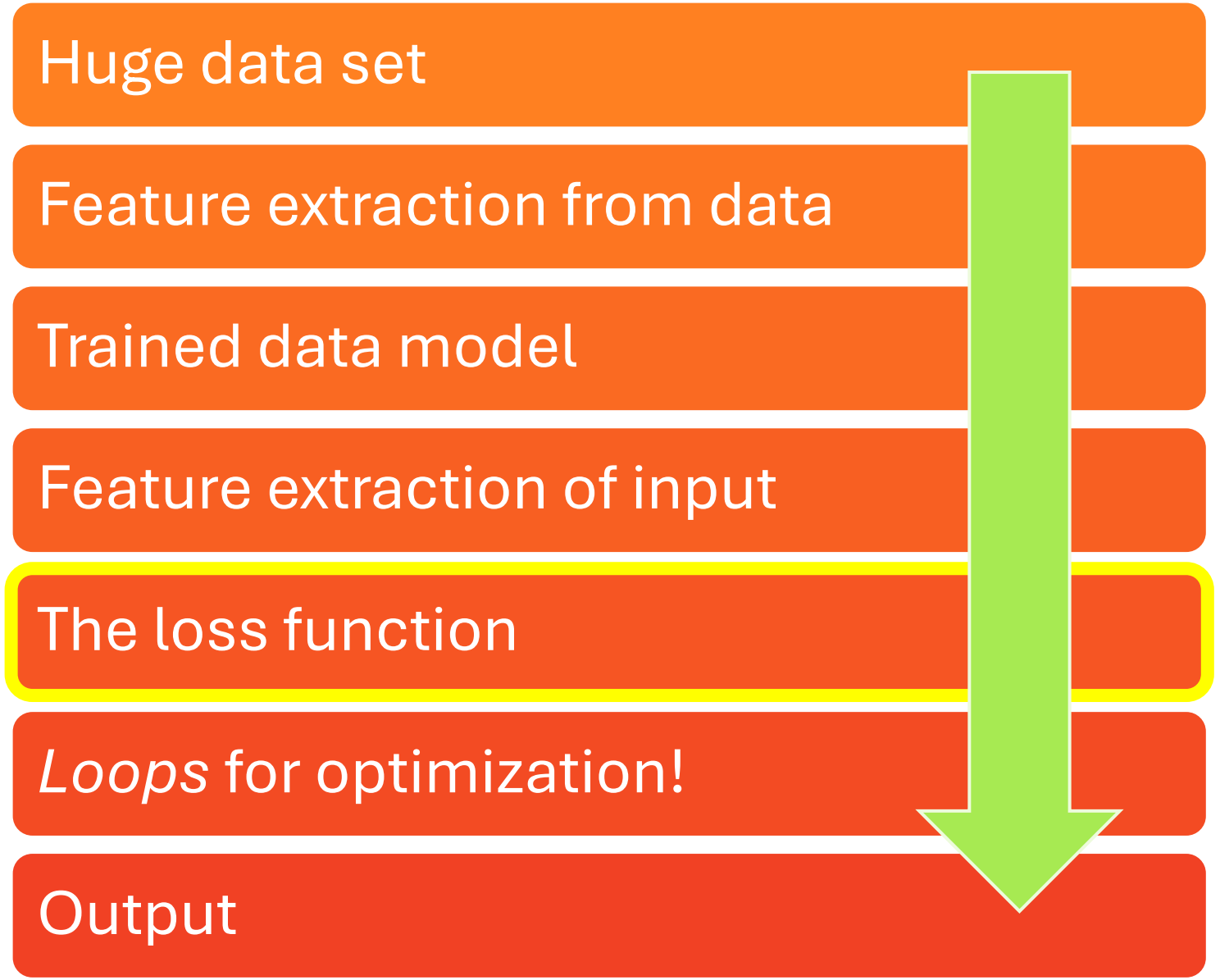

B
A
G
F
E
D
C
Pitch class
0
0.5
1
1.5
2
2.5
3
3.5
4
Time

https://deepdreamgenerator.com/ddream/dge1021q38b

Style Transfer

New creation while keeping certain features

The process of machine learning



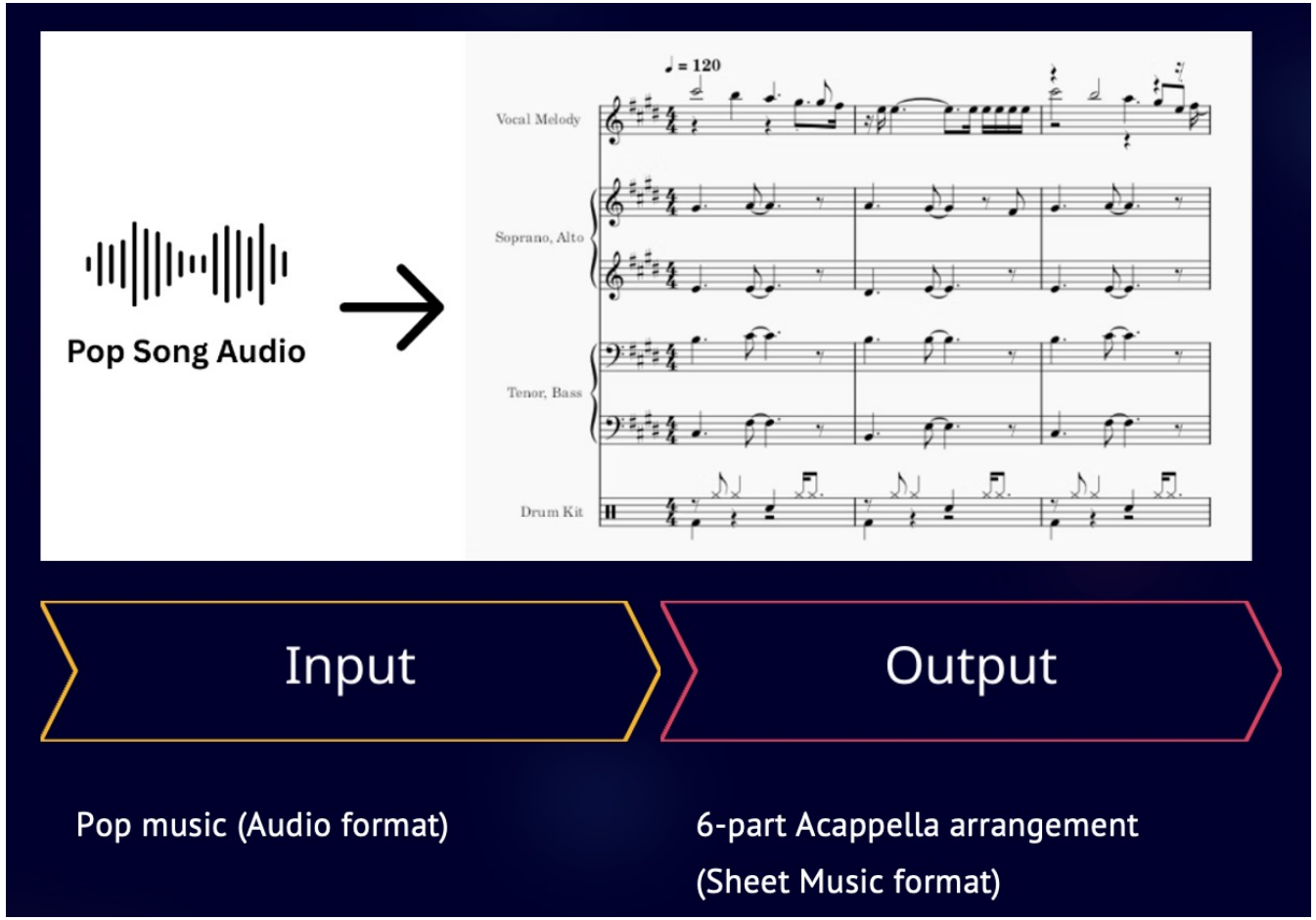
Machine Listening / Computer Audition

- Granting the ability for a computer to understand audio
  - Studying how ***human*** perceives, understands, and produces sound and music
  - How can the computer ***mimic such behaviours***?
  - Ensure transformation in audio editing to be musically meaningful

- Machine musicianship
  - A computer will be able to make music if it can hear and understand “music” like human being!

“Pop Music to Acappella Sheet Music Generation”

- By undergraduate students as a Final Year Project at our CUHK AIST programme
- Accepted for presentation at The 6th Conference on ***AI Music Creativity***, 10-12 September 2025, Brussels, Belgium



AI-generated Music

- While AI music sounds wonderful, there are still potential issues
  - Creativity and dependency
  - Data bias
  - Legal and ***ethical*** issues

- Applicable directions
  - Supporting composers for music creation and arrangement
  - Automatic accompaniment and education
  - Personalized entertainment

Thank you!

Dr Chuck-jee CHAU 周卓之
chuckjee@cse.cuhk.edu.hk
http://linkedin.com/in/chuckjee

Dept of Computer Science and Engineering,
The Chinese University of Hong Kong