

手書き入力で盛り上がりを  
コントロールするループシーケンサ：  
スペクトログラムからの盛り上がり度の自動割り振り

北原研究室

飯島孔右 鶴岡亜也佳

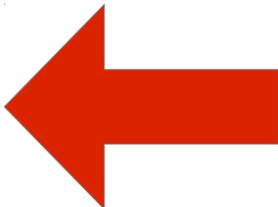
# 目次

- 研究の背景
- 先行研究
- 今回の研究
- 実験
- 考察
- 今後の課題

# 研究の背景

DTM (DeskTop Music) は発達し、  
作曲は身近なものになっている。

DTMとは、デスクトップ上で作曲をすること。

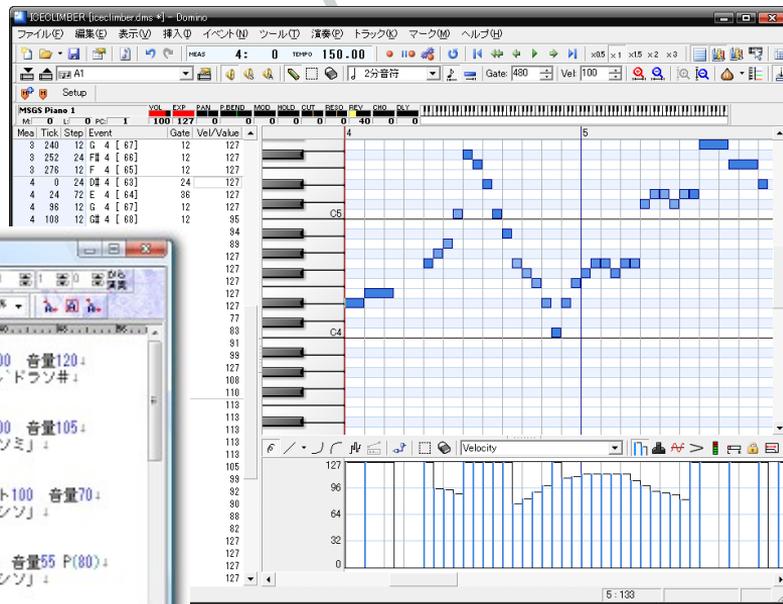
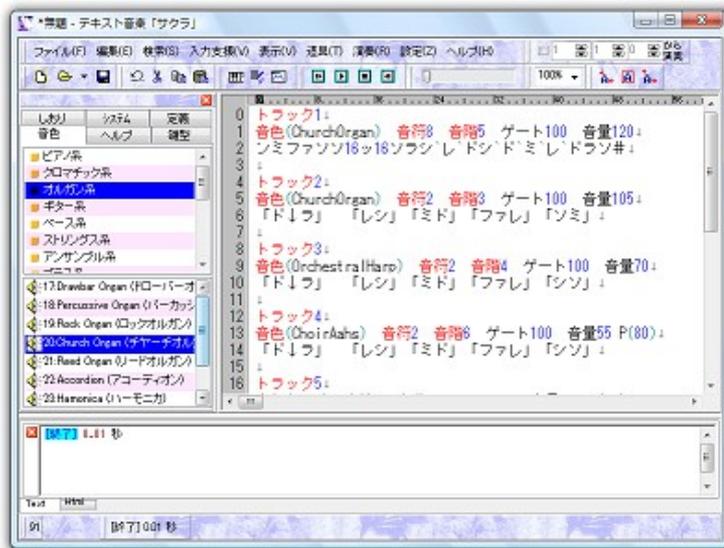


# DTMの例

DTMのソフトには  
いくつかの手法がある。

ピアノロール  
例

- ・DOMINO
- 楽譜を書く  
例
- ・サクラ

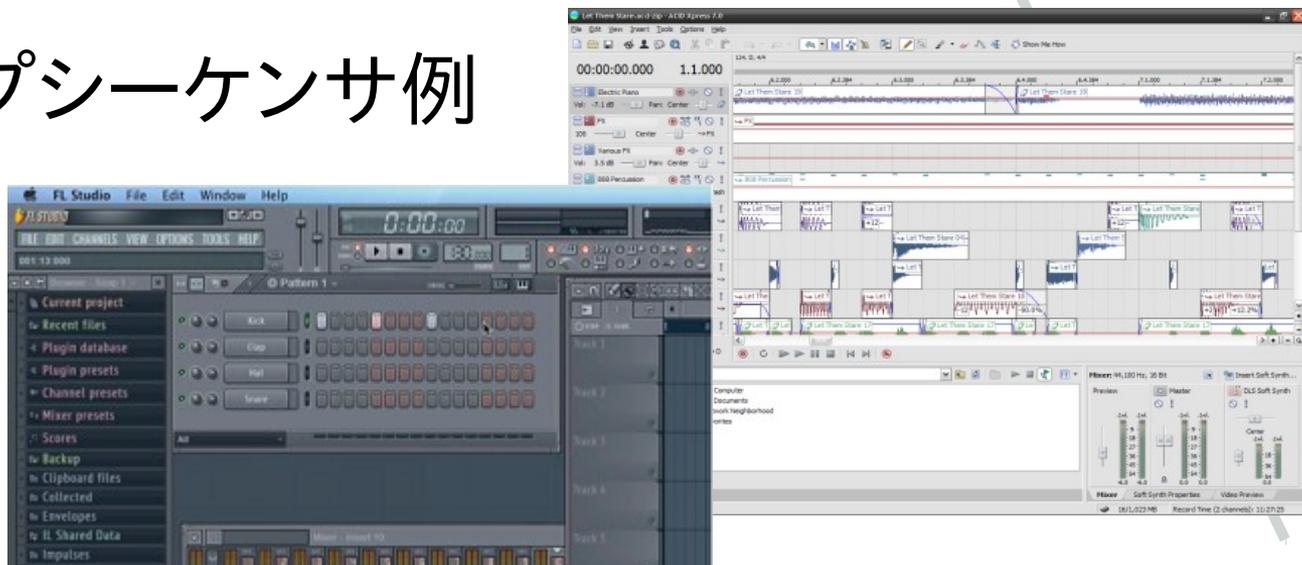


# ループシーケンサとは

DTMの中、ループシーケンサという手法がある。  
これは音素材を配置するだけで作曲できる。

## 代表的ループシーケンサ例

- ・MusicShake
- ・FL Studio
- ・ACID



# 音素材とは

数小節単位のフレーズ

素材1

素材2

素材3

試聴

The screenshot shows the musicshake software interface. At the top, it displays the time 00:09:34 / 03:11:99 and various playback controls. Below that, there are controls for song title, style (BnB), tempo (90), key (G), and effects (FADE, REPEAT, DSP). The main area is a table with columns for TRACK, Tone, Volume, and seven measure columns (1-7). A vertical line is positioned under measure 1.

TRACK	Tone	Volume	1	2	3	4	5	6	7
1 [Slow] The River Of Obl	1	100	Standard Drum	Standard Drum	Another drum	-	-	-	-
2 [Arpeggio] Warm Acous	1	70	-	1	1	-	-	-	-
3 [Chord Voicing] Beautif	3	20	1	-	-	-	-	-	-
4 [Simple Long Tone] Moc	5	30	-	-	-	-	-	-	-
5 [Melody] Whispering So	4	20	-	-	-	-	-	-	-

# music shake

The screenshot displays the musicshake software interface. At the top left, the logo "musicshake" is visible. The top center shows a timer at "00:00:00 / 02:07:99" and playback controls: REW, PLAY, STOP, FF, REC, and a volume meter labeled "vol:100". The top right has "My Songs" and "Login" buttons. Below the timer is a text input field "CLICK to add a song title (up to 40 characters)". The interface includes controls for "STYLE" (set to "Dance"), "TEMPO" (90), "KEY" (G), and buttons for "FADE", "REPEAT", and "DSP".

The main area features a track list and a mixing table. The track list has columns for "TRACK", "Tone", and "Volume". The mixing table has columns numbered 1 to 6, each with a waveform and a volume slider.

TRACK	Tone	Volume	1	2	3	4	5	6
1 [Effect] Increasing Hear	1	100	Another drum	—	—	—	—	—
2 [Complicated] Attack Mc	3	42	—	—	—	—	—	—
3 [Arpeggiator] Mysteriou	1	31	—	—	—	—	—	—
4 [Chord Harmony] Marvel	1	18	—	—	—	—	—	—
5 [Rhythm] Sensitive High	1	33	—	—	—	—	—	—

At the bottom, there are navigation buttons: "menu", "TRACK", "EFFECTS", and "PART".

# music shake

カクッカ!

The screenshot shows the musicshake software interface. At the top, the title 'musicshake' is displayed in a colorful font. Below it, a digital display shows '00:00:00 / 02:07:99'. To the right of the display are playback controls: a play button, a stop button, a fast forward button, a record button, and a volume indicator labeled 'vol100'. Further right, there is a 'My Songs' section with a 'Login' button and a window control icon. Below the playback controls is a search bar with the placeholder text 'CLICK to add song title (up to 40 characters)'. To the right of the search bar are several control buttons: 'STYLE' (set to 'Dance'), 'TEMPO' (set to '90'), 'KEY' (set to 'G'), 'FADE', 'REPEAT', and 'DSP'. The main area of the interface is a track list with columns for 'TRACK', 'Tone', and 'Volume'. The tracks are numbered 1 through 5. Track 1 is '[Effect] Increasing Hear' with a volume of 100. Track 2 is '[Complicated] Attack M...' with a volume of 42. Track 3 is '[Arpeggiator] Mysteriou...' with a volume of 31. Track 4 is '[Chord Harmony] Marvel' with a volume of 18. Track 5 is '[Rhythm] Sensitive Higt' with a volume of 33. To the right of the track list is a piano roll area with six lanes labeled 1 through 6. A hand is pointing at the second track, which has a yellow highlight and the word 'Another' written in orange. At the bottom of the interface, there is a navigation bar with icons for 'menu', 'TRACK', 'EFFECTS', and 'PART'. The bottom right corner has window control icons.

# music shake

The screenshot displays the musicshake software interface. At the top, the logo "musicshake" is on the left, and "My Songs" and "Login" are on the right. The central control bar includes a time display "00:00:00 / 02:07:99", buttons for REW, PLAY, STOP, FF, REC, and a volume meter labeled "vol100". Below this is a search bar with the text "CLICK to add a song title (up to 40 characters)", and controls for STYLE (set to "Dance"), TEMPO (90), KEY (G), and buttons for FADE, REPEAT, and DSP.

The main area is a multi-track mixer with six tracks. The tracks are:

TRACK	Tone	Volume	1	2	3	4	5	6
1 [Effect] Increasing Hear	1	100	Another drum	—	—	—	—	—
2 [Complicated] Attack M	3	42	1	—	—	—	—	—
3 [Arpeggiator] Mysteriou	1	31	1	—	—	—	—	—
4 [Chord Harmony] Marvel	1	18	—	—	—	—	—	—
5 [Rhythm] Sensitive High	1	33	—	—	—	—	—	—

At the bottom, there is a navigation bar with "menu", "TRACK", "EFFECTS", and "PART" buttons, along with a volume icon and a "PART" button.

# music shake

試聽

The screenshot shows the musicshake software interface. At the top, the logo "musicshake" is on the left, and "My Songs" and "Login" are on the right. Below the logo is a timer showing "00:00:00 / 02:07:99". To the right of the timer are playback controls: REW, PLAY, STOP, FF, REC, and a volume meter labeled "vol100".

Below the playback controls is a control bar with a text input field "CLICK to add a song title (up to 40 characters)", a "STYLE" dropdown menu set to "Dance", a "TEMPO" field set to "90", a "KEY" field set to "G", and buttons for "FADE", "REPEAT", and "DSP".

The main area is a piano roll with a table of tracks and notes. The tracks are:

TRACK	Tone	Volume	1	2	3	4	5	6
1 [Effect] Increasing Hear	1	100	Another drum	—	Another drum	—	—	—
2 [Complicated] Attack M	3	42	1	—	1	—	—	—
3 [Arpeggiator] Mysteriou	1	31	1	—	1	—	—	—
4 [Chord Harmony] Marvel	1	18	—	1	—	—	—	—
5 [Rhythm] Sensitive Higt	1	33	—	1	1	—	—	—

At the bottom of the interface, there is a navigation bar with "f menu", "TRACK", "EFFECTS", and "PART" buttons, along with a volume icon and a window management icon.

# ループシーケンサとは

つまり!

音素材を**組み合わせるだけ**なので  
比較的**簡単**に曲を作ることが**可能!**



# 研究の背景

先ほどのMusicShake  
素材の数は500000を超える



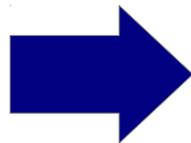
選べる素材の数が**多い**



ユーザーにとっての選択肢が**多い**



希望のものを  
探し出すのが困難



**負担増!**

# 研究の背景

ユーザーにとって**難しくない**操作で  
自動で音素材を挿入してくれる...



そんなループシーケンサに

**需要!**

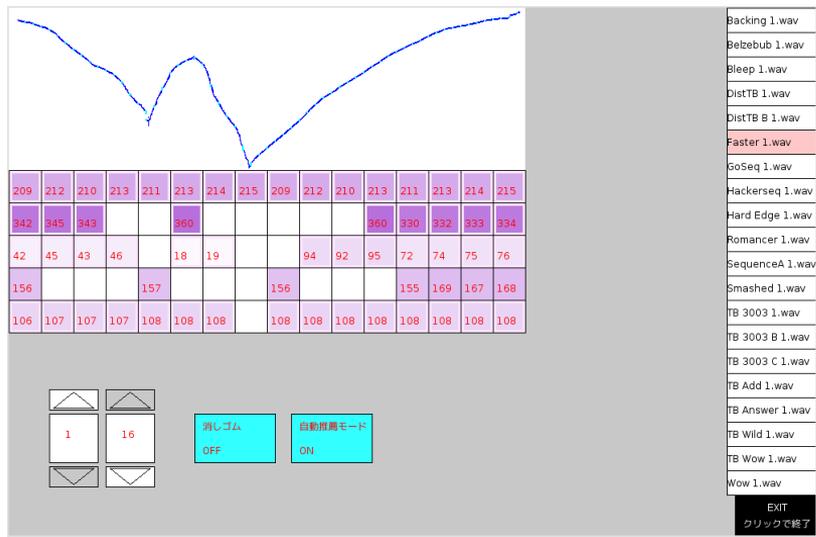
線を描くだけで音素材が選べたら理想的!

# 先行研究

“手書き入力によって盛り上がりをコントロールするループシーケンサ”

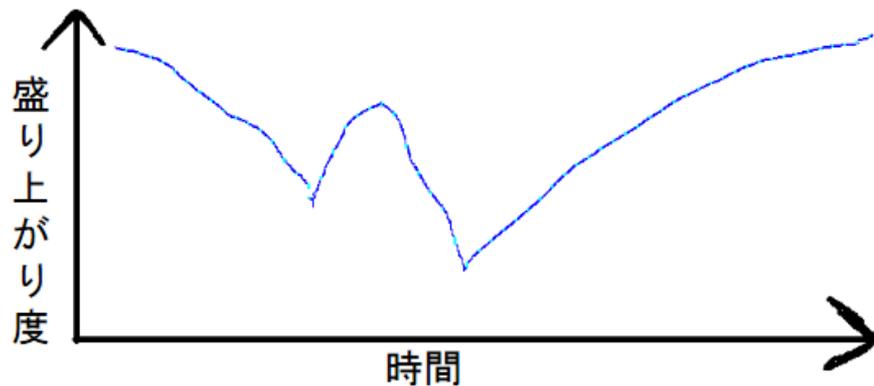
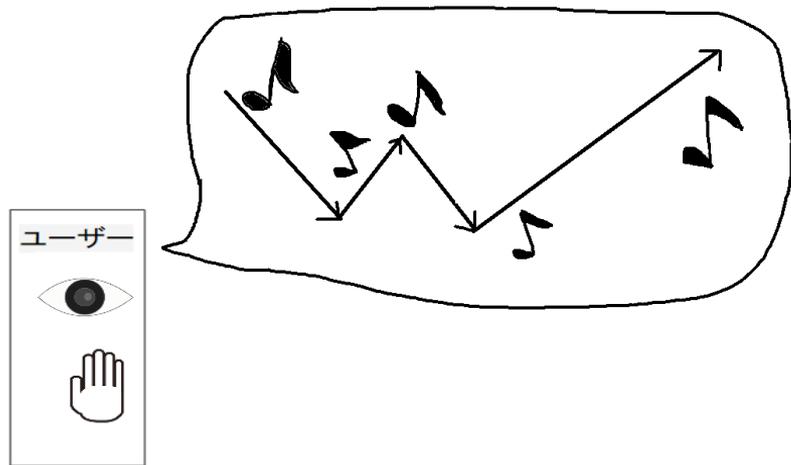
山下雄史, 岡田美咲, 北原鉄朗

情報処理学研究報告, Vol.2013-MUS-100, No.36, pp.1-6, 2013



# 山下らのループシーケンサのしくみ

ユーザーは曲の盛り上がり方を想像



ユーザーは線を描く

# 素材選択のしくみ

高い曲線を描いたら

楽器数多

より盛り上がる素材

低い曲線を描いたら

楽器数少

より静かな素材

素材

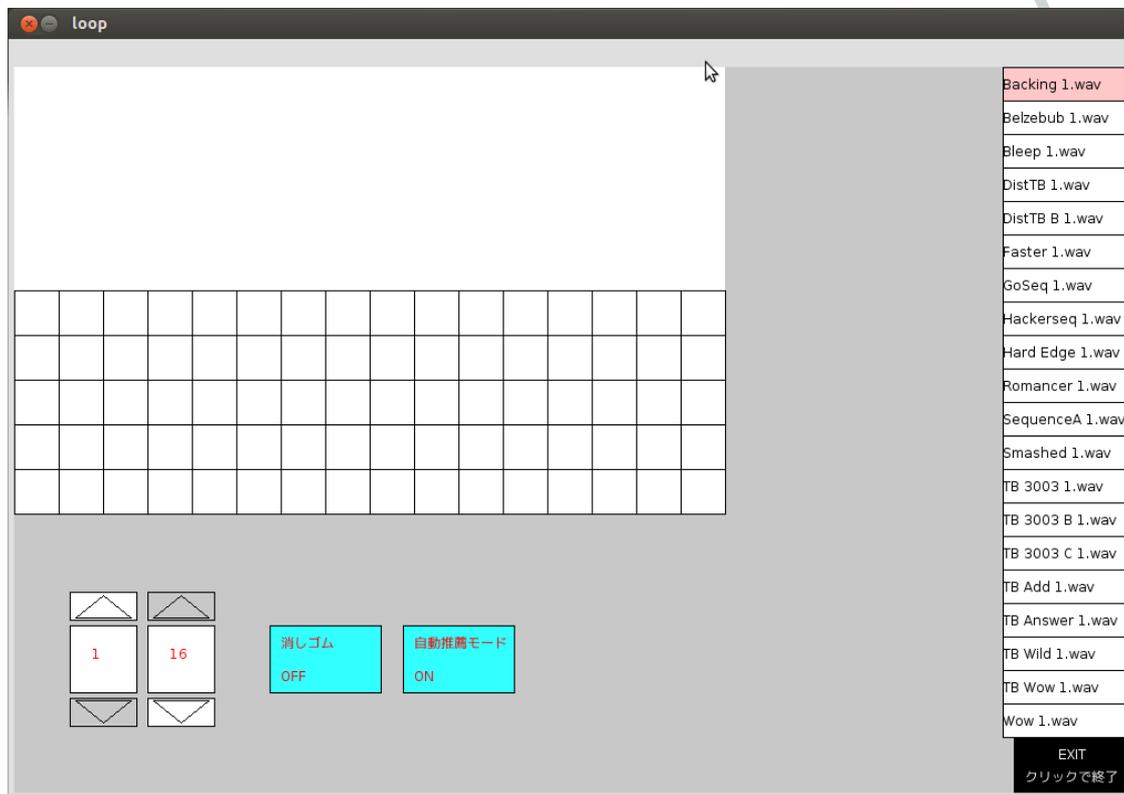
楽器数

盛り上がり具合

考慮

ユーザーの意図した  
盛り上がりを再現!

# 山下らのループシーケンサ



# 山下らのループシーケンサ

線を描く



The screenshot shows a software window titled "loop". At the top, there is a waveform plot with a blue line. Below the plot is a grid of audio samples. The grid has 4 rows and 16 columns. The top row contains sample numbers: 174, 177, 175, 178, 176, 178, 179, 180, 174, 177, 175, 178, 176, 178, 179, 180. The second row contains: 367, 330, 353, 323, 325, 326. The third row contains: 78, 81, 86, 88, 98, 101, 99, 102, 65, 67, 68. The fourth row contains: 154, 158, 154, 158. The bottom row contains: 109, 109, 109, 108, 108, 108, 108, 108, 108, 108, 108, 108, 108, 108, 108. Below the grid are four navigation buttons (up, down, left, right) and two buttons labeled "1" and "16". To the right of the grid are two buttons: "消しゴム OFF" and "自動推薦モード ON". On the right side of the window is a list of audio files: Backing 1.wav, Belzebub 1.wav, Bleep 1.wav, DistTB 1.wav, DistTB B 1.wav, Faster 1.wav, GoSeq 1.wav, Hackerseq 1.wav, Hard Edge 1.wav, Romancer 1.wav, SequenceA 1.wav, Smashed 1.wav, TB 3003 1.wav, TB 3003 B 1.wav, TB 3003 C 1.wav, TB Add 1.wav, TB Answer 1.wav, TB Wild 1.wav, TB Wow 1.wav, Wow 1.wav. At the bottom right, there is a black button labeled "EXIT" with the text "クリックで終了" below it.

音素材  
が挿入



試聴

# 山下らのループシーケンサ

The interface displays a waveform and a grid of 16 numbered cells (1-16) with various colors (yellow, green, red). Below the grid are four arrow buttons and two buttons labeled '消しゴム OFF' and '自動推薦モード ON'. On the right, a list of audio files is shown, with 'Backing 1.wav' highlighted in red. At the bottom right, there is an 'EXIT' button with the text 'クリックで終了'.

素材の選択も出来る!

174	177	175	178	176	178	179	180	174	177	175	178	176	178	179	180
			367	330				AlarmLead 1.wav			353	323	325	326	
		78	81	86	88			BigLead High 1.wav							
								BigLead Low 1.wav			102	65	67	68	
154				158				Lead 1.wav							
								Leader 1.wav				158			
	109	109	109	108	108	108	108	Shrabnelle 1.wav			108	108	108	108	

- Backing 1.wav
- Belzebub 1.wav
- Bleep 1.wav
- DistTB 1.wav
- DistTB B 1.wav
- Faster 1.wav
- GoSeq 1.wav
- Hackerseq 1.wav
- Hard Edge 1.wav
- Romancer 1.wav
- SequenceA 1.wav
- Smashed 1.wav
- TB 3003 1.wav
- TB 3003 B 1.wav
- TB 3003 C 1.wav
- TB Add 1.wav
- TB Answer 1.wav
- TB Wild 1.wav
- TB Wow 1.wav
- Wow 1.wav

EXIT  
クリックで終了

# 今回の研究：山下らの研究の問題点

1. 各素材の盛り上げりの手動付与

2. 拡張が困難

3. 評価実験の未実施

# 音素材の盛り上がり度

曲線の盛り上がり度が5

曲線の盛り上がり度が1



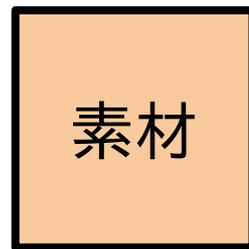
選択される!

盛り上がり度

1

盛り上がり度

5



# 音素材の盛り上がり度

つまり

音素材にも盛り上がり度の  
割り振りは**必要!**

しかし

素材数が**多い** → 割り振るの**大変!**

なぜ?

全部山下らが聴いて  
決めていた!



用意した素材全部聴く**必要**がある!



# 音素材の盛り上がり度

そこで我々は

## 盛り上がり度の 自動割り振りをする手法

を提案する!



# 盛り上がり度の自動推定

試聴

山下らが決めた盛り上がり度

盛り上がり度 **低** ↓ 素材

DRUMS

BASS SEQUENCE

盛り上がり度 **高** ↑ 素材

DRUMS

BASS

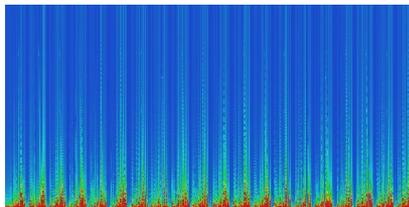
SEQUENCE



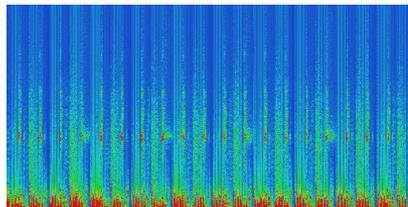
# 音素材のスペクトログラム

盛り上がり度 高

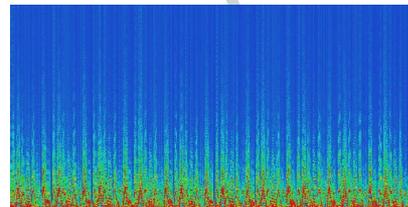
BASS



DRUMS

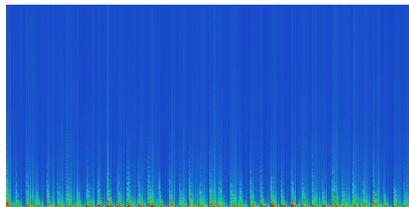


SEQUENCE

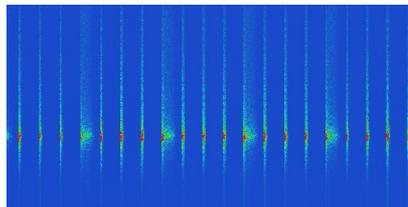


盛り上がり度 低

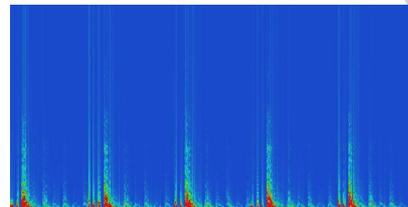
BASS



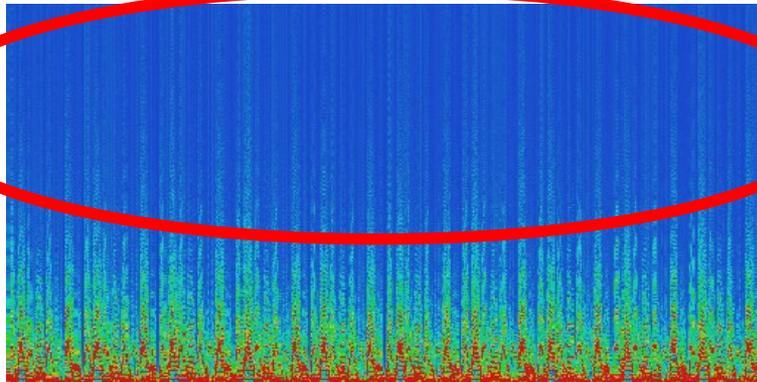
DRUMS



SEQUENCE

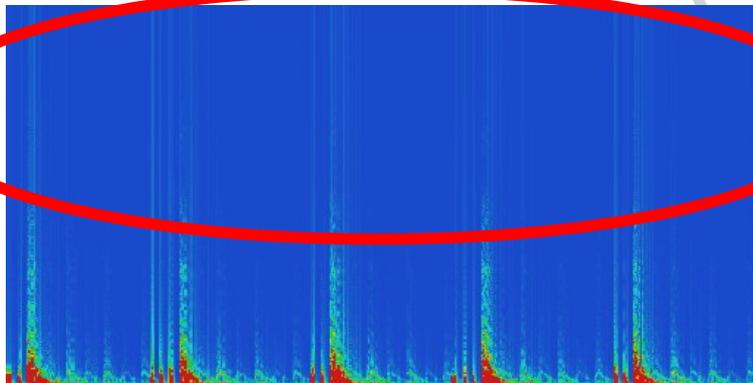


# 音素材のスペクトログラム：SEQUENCE



← 盛り上がり度 **高**

盛り上がり度 **低** →  
見比べた時  
とても振幅が違う!



# 仮説

各時刻、各周波数において振幅が強い割合が多い  
これは

盛り上がりが高い



←閑散とした観客席

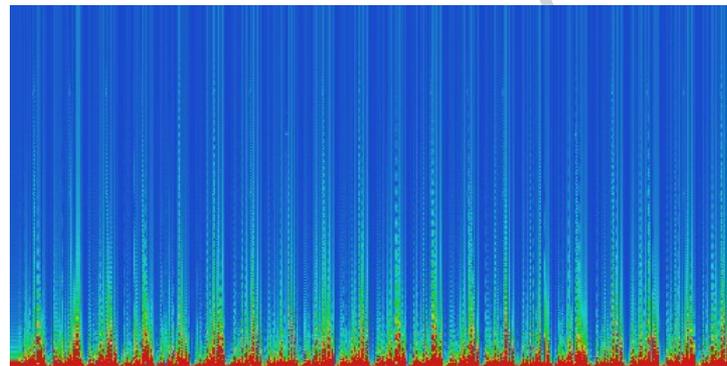
沸き立つ観客席→



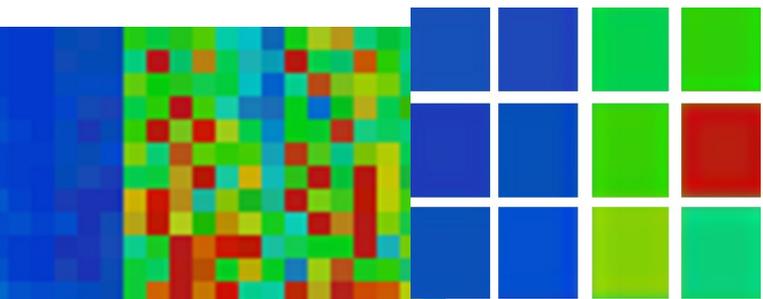
# 盛り上がり度を得る1

さきほどの仮説を計算で形にする。

音素材に対してフーリエ変換をする。  
→スペクトログラムを得る。



# 盛り上がり度を得る2



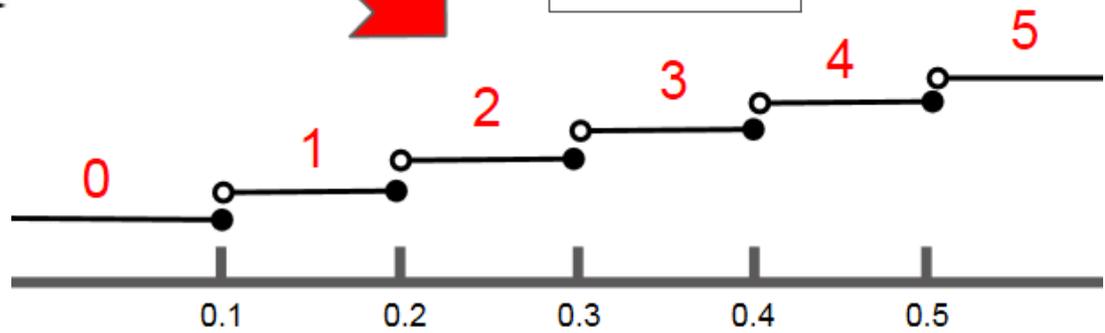
振幅の数値



変換



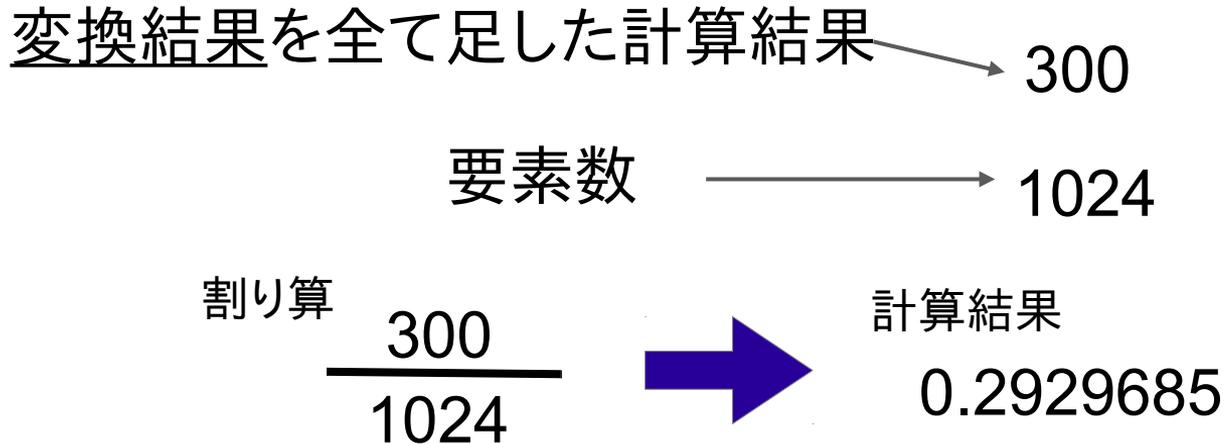
0~5の数値へ!



# 盛り上がり度を得る3

さきほどの変換結果を全て足して、要素の総数で割る。

例

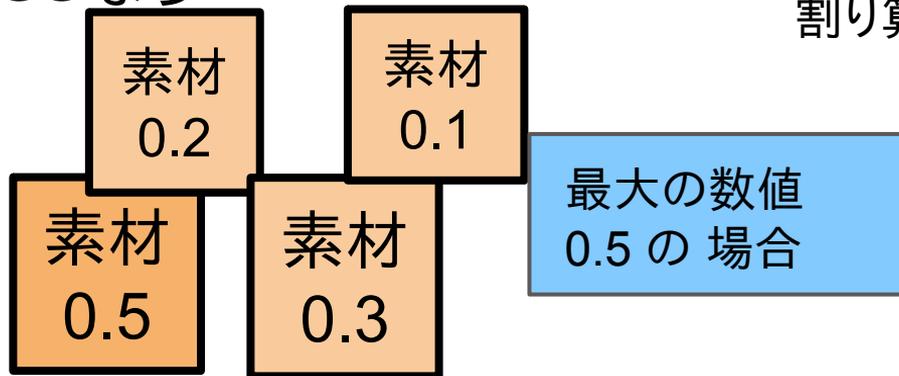


2, 3にて  
振幅の成分のある  
面積の割り合いを  
計る事ができた。

# 盛り上がり度を得る4

楽器毎に最大の数値を一つ求める。  
楽器毎にそれぞれの素材をその数値で割る。  
これによって0~1の数値になる。

例 :BASSなら



割り算

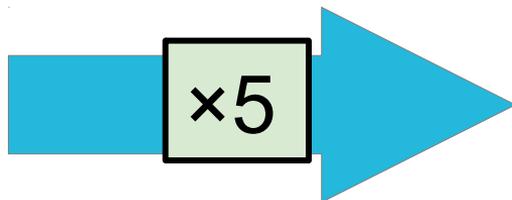
$$0.5 / 0.5 = 1.0$$
$$0.2 / 0.5 = 0.4$$
$$0.1 / 0.5 = 0.2$$
$$0.3 / 0.5 = 0.6$$

# 盛り上がり度を得る5

5を掛ける  $\boxed{\times 5}$   
(小数点以下は繰り上げる)

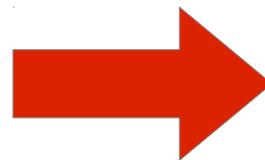
例

1.0  
0.4  
0.2  
0.6



盛り上がり度 5  
盛り上がり度 2  
盛り上がり度 1  
盛り上がり度 3

これで5つの数字に割り振られる



これこそ  
**盛り上がり度!**

# 盛り上がり度を得た

盛り上がり度を機械的に決められる

これで...

量も怖くない!

主観的な判断から客観的判断に!



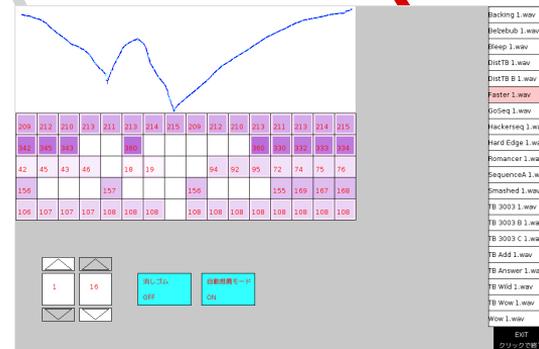
# 拡張が困難

山下らの研究を我々は受け継いだ。

しかし

表示や処理の記述がソースコード内に点在してしまっていた。

それぞれの関係性や機能の内容が把握しづらく、拡張が困難。



# 決断

基本の動作は出来る限り変えぬよう

全部作り直す!

ことに決定

そこで

- ・ 構造の改善
- MVCモデルを採用

# 構造の改善

MVCモデルとは？

Model

モデル

View

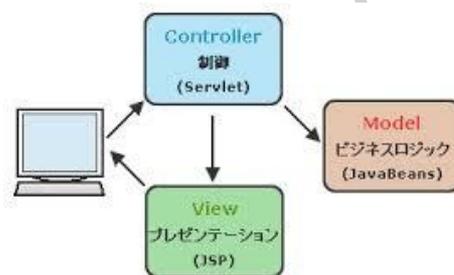
ビュー

Controller

コントローラー

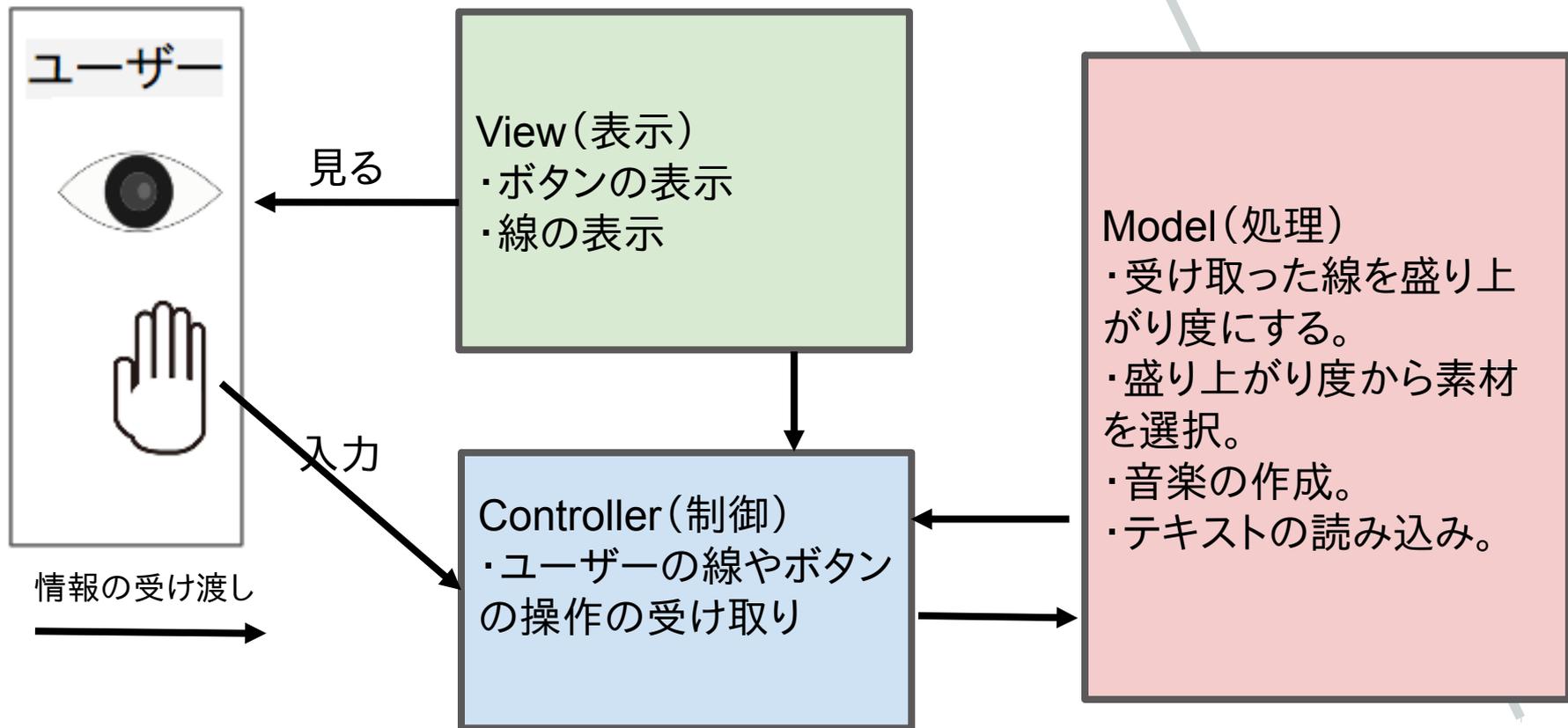
処理  
表示  
制御

システムの要素を三つに  
分離させるプログラミングモデル。



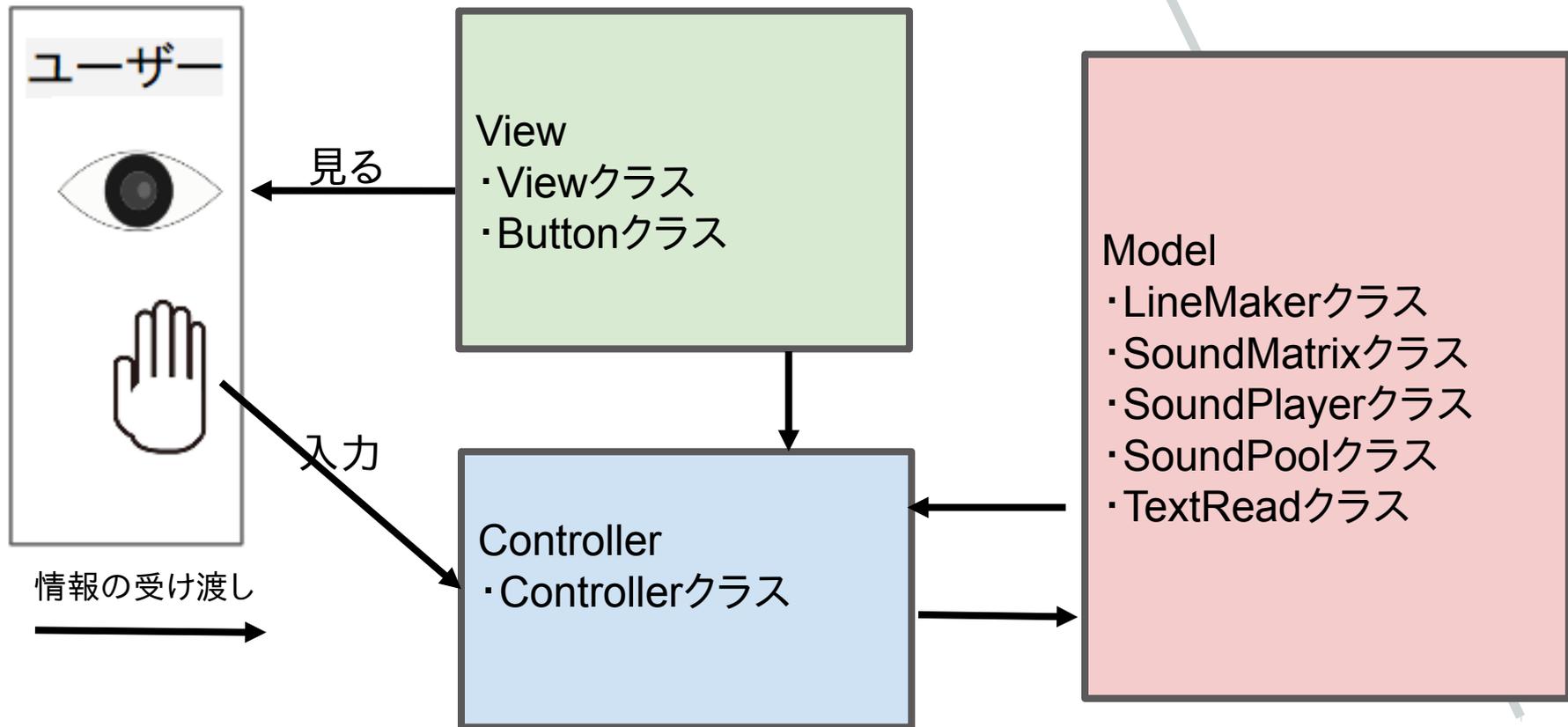
# MVCモデルを基礎とした

# システムの説明

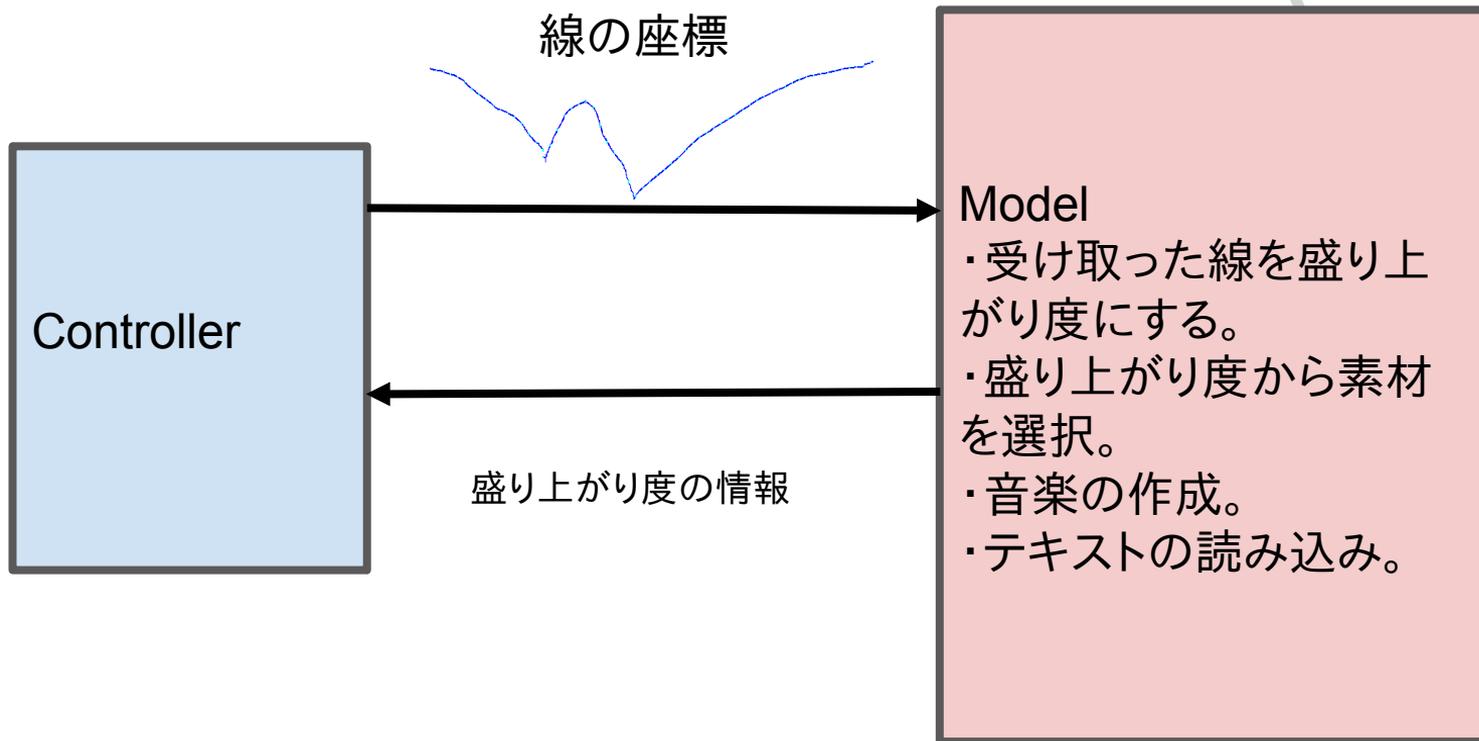


# MVCモデルを基礎とした

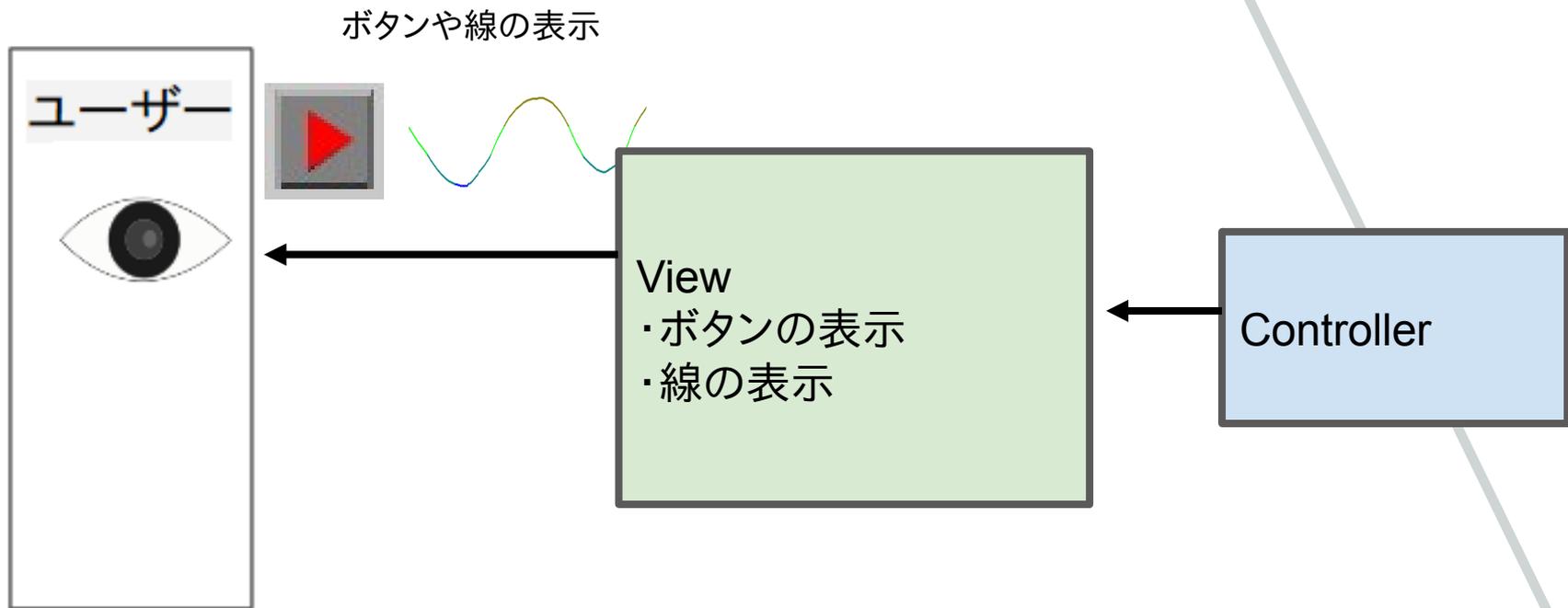
# システムの説明



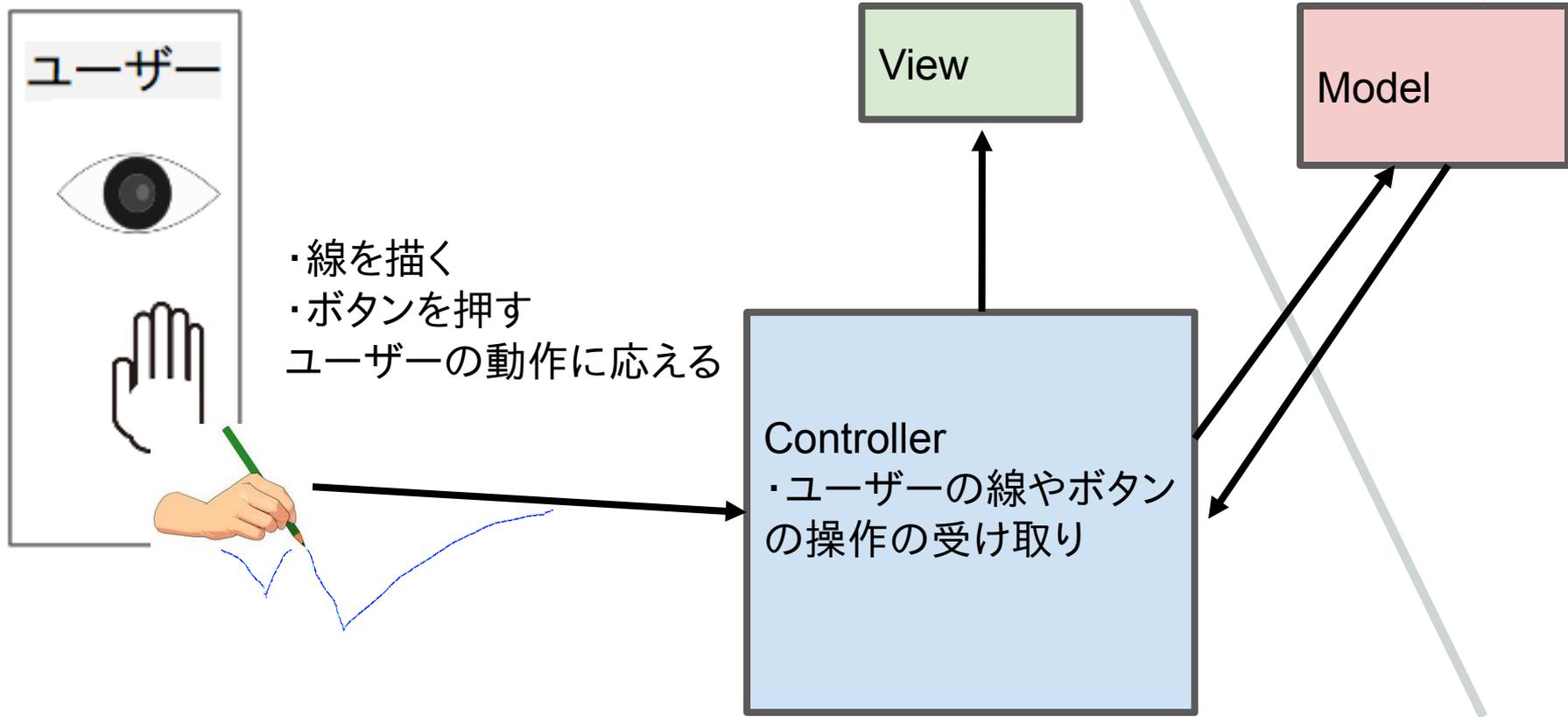
# 本システムにおける Model



# 本システムにおけるView



# 本システムにおけるController



# GUIの変更

Old GUI interface showing a waveform plot and a data table.

209	212	210	213	211	213	214	215	209	212	210	213	211	213	214	215
342	345	343			360					360	330	332	333	334	
42	45	43	46		18	19			94	92	95	72	74	75	76
156				157				156				155	169	167	168
106	107	107	107	108	108	108		108	108	108	108	108	108	108	108

Controls: 1, 16, 消しゴム OFF, 自動消音モード ON

←改良前

改良後→

New GUI interface showing a waveform plot and a detailed data table.

0	4	0	0	0	0	6	7	1	4	2	0	0	5	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	81	79	82	80	82	83	84	22	25	23	26	24	26	0	0
3	0	0	0	3	0	0	0	6	0	0	0	4	0	0	0
30	0	27	27	27	25	15	15	34	11	0	25	29	15	31	31

Timeline: 0:00 to 0:54. Status: 0:00 / 0:54

**DEMO**



# 実験(目的)

- 1.自動割り振りされた音素材の盛り上がり度はあっているか
- 2.音楽知識がなくても簡単に作曲をすることができるか



# 実験（被験者）

実験人数：10人（大学生）

楽器の演奏経験がある：7人  
（ピアノ、ギター、ベース、他）  
シーケンサを知ってる人：2人



# 実験（内容1）

試聴

一対比較法を用いてスペクトログラムから盛り上がり度の自動割り振りされた音素材を聞く  
盛り上がっていると感じた方を答えてもらう  
「Sequence」「Synth」「Bass」「Drums」「Percussion」  
を4問ずつの計20問



# 実験（結果1）

一対比較法の結果

20問中16.6問の正解率  
どの楽器でも高い正解率であった。



**約83%**

# 実験（内容2）

本システムと比較用システムを使用

（比較用システムは本システムから自動で音素材を挿入する機能をなくしたもの）

盛り上がりをイメージしやすいようにスポーツカーのプロモーション映像に合った曲の作成をしてもらった

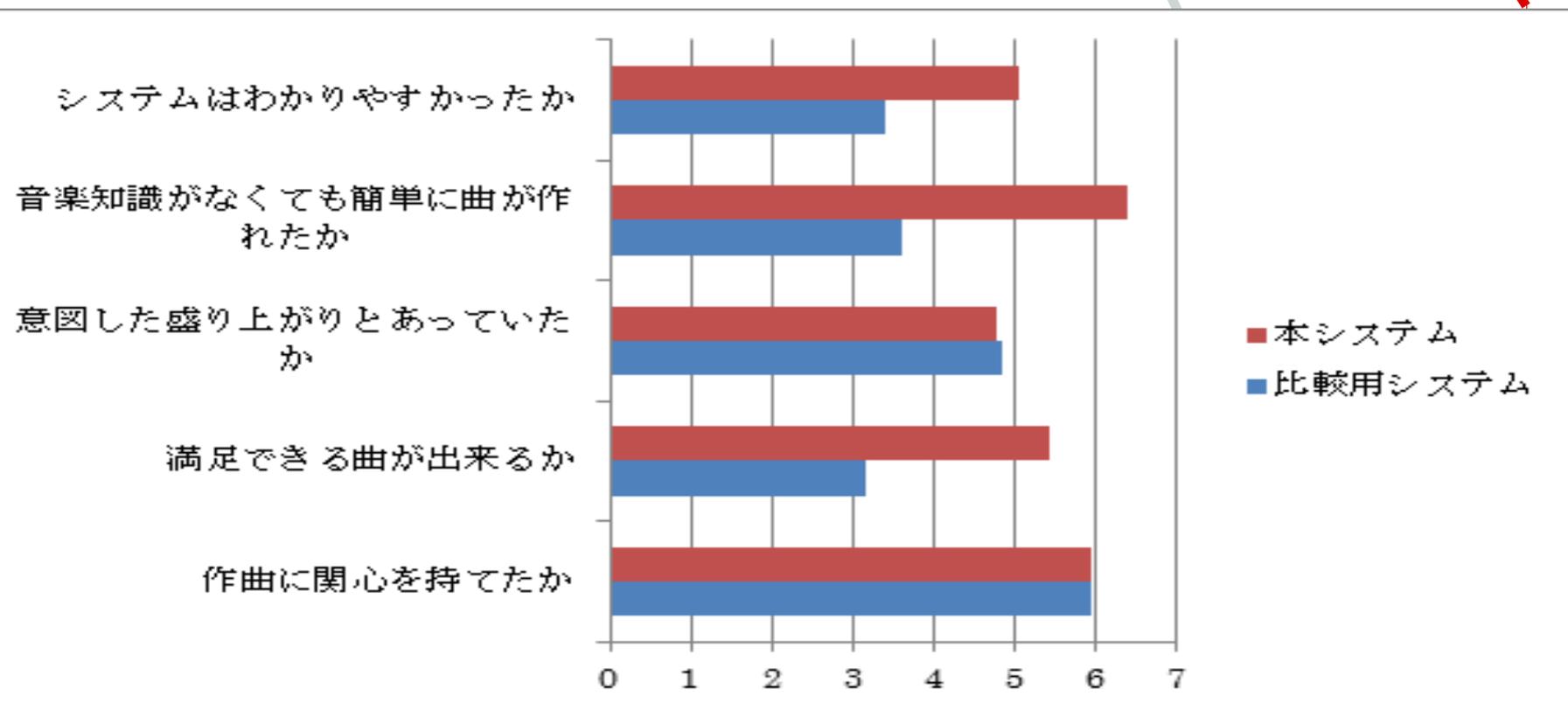
制限時間は30分



# 実験(内容2)



## 実験（結果2）



## 実験（結果2）

### 比較用システム

- ・音の試聴があるといい
- ・途中再生出来ないから動画と合わせるのが大変

### 本システム

- ・線を引くだけで簡単に作曲が出来て良かった
- ・急に盛り上げたいときや盛り下げたいときに反映されていないように思えた

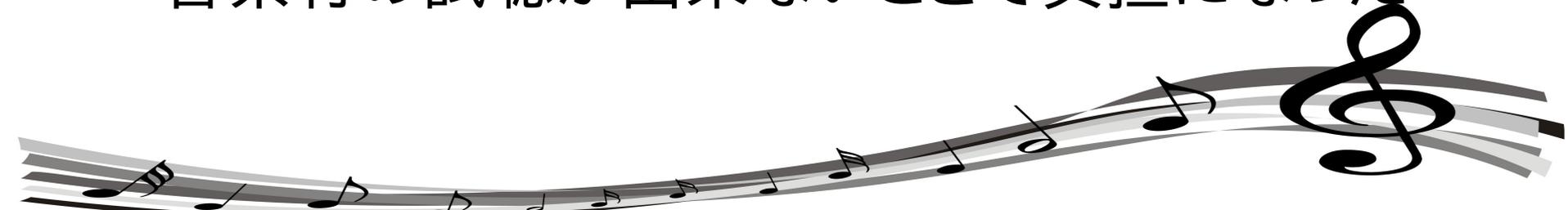
## 実験（考察）

- ・比較用システムを使用して曲を作った時、2人が曲を完成出来なかった
- ・本システムを使用して曲を作った時は、20分前後で完成させていた



# 実験（考察）

- ・事後アンケートで比較用システムより本システムの方が全体的に高い評価が得られた
  - 線を引くだけで簡単に作曲が出来る
- ・比較用システムでは、多くの素材の中から一つ一つ決めていくので大変だった
  - 音素材の試聴が出来ないことで負担になった



## まとめ

- ・スペクトログラムを用いて音素材の盛り上がり度の自動割り振りを実装した→約83%の正解率
- ・被験者実験を行った
- ・先行研究からの目的でもあった、簡単に作曲をすることができるかどうか検証をした→音楽知識がなくても線を引くだけで曲が出来ると認識された

# 今後の課題

- ・音素材の選択時に音素材の試聴が出来ること
- ・途中再生が出来ること



ご清聴ありがとうございました。