Professur
Allgemeine Nachrichtentechnik

Prof. Dr.-Ing. Udo Zölzer

HELMUT SCHMIDT
UNIVERSITÄT

Parameters Automation of Dynamic Range Compressor

Literature Review - Project Work

Aritra Mazumdar

20 December 2018

Outline

HELMUT SCHMIDT UNIVERSITÄT

Outline

- Introduction
- Theory
- Automation
- Approach 1
- Approach 2
- Discussions
- Conclusion

Introduction

Dynamic Range Compressor

Dynamic range of a signal is defined as the logarithmic ratio of maximum to minimum amplitude of a signal and is expressed in dB.

HELMUT SCHMIDT UNIVERSITÄT

Dynamic range of a signal is defined as the logarithmic ratio of maximum to minimum amplitude of a signal and is expressed in dB.

Dynamic Range of a signal can be controlled using an process named as **Dynamic Range Compressor**.

Dynamic range of a signal is defined as the logarithmic ratio of maximum to minimum amplitude of a signal and is expressed in dB.

Dynamic Range of a signal can be controlled using an process named as **Dynamic Range Compressor**.

Motivation:

HELMUT SCHMIDT UNIVERSITÄT

Dynamic range of a signal is defined as the logarithmic ratio of maximum to minimum amplitude of a signal and is expressed in dB.

Dynamic Range of a signal can be controlled using an process named as **Dynamic Range Compressor**.

Motivation:

- Project work involved building DRC application

HELMUT SCHMIDT UNIVERSITÄT

Dynamic range of a signal is defined as the logarithmic ratio of maximum to minimum amplitude of a signal and is expressed in dB.

Dynamic Range of a signal can be controlled using an process named as **Dynamic Range Compressor**.

Motivation:

- Project work involved building DRC application
- Application accepts static set of parameters

Dynamic range of a signal is defined as the logarithmic ratio of maximum to minimum amplitude of a signal and is expressed in dB.

Dynamic Range of a signal can be controlled using an process named as **Dynamic Range Compressor**.

Motivation:

- Project work involved building DRC application
- Application accepts static set of parameters
- Adapting parameters to signal dynamics gives better results

HELMUT SCHMIDT UNIVERSITÄT

Dynamic range of a signal is defined as the logarithmic ratio of maximum to minimum amplitude of a signal and is expressed in dB.

Dynamic Range of a signal can be controlled using an process named as **Dynamic Range Compressor**.

Motivation:

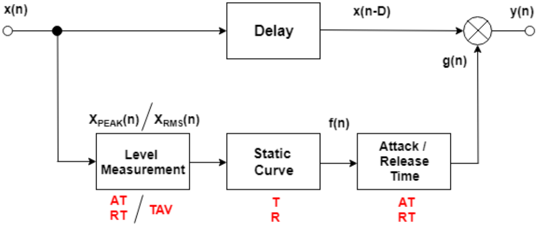
- Project work involved building DRC application
- Application accepts static set of parameters
- Adapting parameters to signal dynamics gives better results
- Compressor with automated parameters is better approximation

Outline

- Introduction
- Theory
- Automation
- Approach 1
- Approach 2
- Discussions
- Conclusion

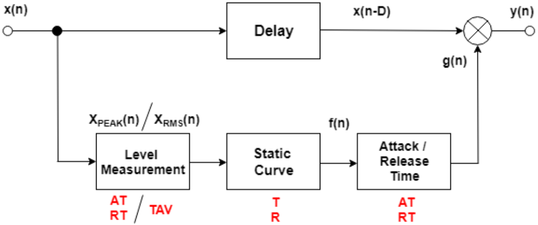
Theory
Block Diagram

HELMUT SCHMIDT UNIVERSITÄT



The figure above represents a basic Dynamic Range Compressor system.

Theory
Block Diagram



The figure above represents a basic Dynamic Range Compressor system.

$$y(n) = x(n - D) \cdot g(n) \quad (1)$$

Theory
Parameters for Automation

- Threshold (T)

Parameters for Automation

- Threshold (T)
- Ratio (R)

Theory

Parameters for Automation

- Threshold (T)
- Ratio (R)
- Knee Width (W)

Theory
Parameters for Automation

- Threshold (T)
- Ratio (R)
- Knee Width (W)
- Attack Time (τ_A)

Parameters for Automation

- Threshold (T)
- Ratio (R)
- Knee Width (W)
- Attack Time (τ_A)
- Release Time (τ_R)

Theory

Parameters for Automation

- Threshold (T)
- Ratio (R)
- Knee Width (W)
- Attack Time (τ_A)
- Release Time (τ_R)
- Make-up Gain (g)

Outline

- Introduction
- Theory
- **Automation**
- Approach 1
- Approach 2
- Discussions
- Conclusion

Automation

Goals and Approaches

- Based on signal

Automation
Goals and Approaches

HELMUT SCHMIDT UNIVERSITÄT

- Based on signal
  - Calculate signal metrics

Automation

Goals and Approaches

- Based on signal
  - Calculate signal metrics
  - Find a relation between calculated metrics and Compressor parameters

Automation

Goals and Approaches

- Based on signal
  - Calculate signal metrics
  - Find a relation between calculated metrics and Compressor parameters
- Based on Supervised Learning

Automation

Goals and Approaches

- Based on signal
  - Calculate signal metrics
  - Find a relation between calculated metrics and Compressor parameters
- Based on Supervised Learning
  - Extract audio features from reference audio

Automation

Goals and Approaches

HELMUT SCHMIDT UNIVERSITÄT

- Based on signal
  - Calculate signal metrics
  - Find a relation between calculated metrics and Compressor parameters
- Based on Supervised Learning
  - Extract audio features from reference audio
  - Train Regression Model based on dataset

Automation

Goals and Approaches

- Based on signal
  - Calculate signal metrics
  - Find a relation between calculated metrics and Compressor parameters
- Based on Supervised Learning
  - Extract audio features from reference audio
  - Train Regression Model based on dataset
  - Map features to effect Compressor parameters

Automation

Goals and Approaches

- Based on signal
  - Calculate signal metrics
  - Find a relation between calculated metrics and Compressor parameters
- Based on Supervised Learning
  - Extract audio features from reference audio
  - Train Regression Model based on dataset
  - Map features to effect Compressor parameters
  - Process target audio with estimated parameters for similar effect

Automation

Goals and Approaches

HELMUT SCHMIDT UNIVERSITÄT

- Based on signal
  - Calculate signal metrics
  - Find a relation between calculated metrics and Compressor parameters
- Based on Supervised Learning
  - Extract audio features from reference audio
  - Train Regression Model based on dataset
  - Map features to effect Compressor parameters
  - Process target audio with estimated parameters for similar effect
  - Music similarity measure between reference and processed audio

Outline

HELMUT SCHMIDT UNIVERSITÄT

Approach 1

Attack and Release Times

- **Crest Factor Method:**

Approach 1
Attack and Release Times

- **Crest Factor Method:**
  - Ratio of Peak to RMS signal level

Approach 1

Attack and Release Times

- **Crest Factor Method:**
  - Ratio of Peak to RMS signal level
  - Low for steady state signal

Approach 1

Attack and Release Times

- **Crest Factor Method:**
  - Ratio of Peak to RMS signal level
  - Low for steady state signal
  - High for signal with transients

Approach 1
Attack and Release Times

- **Crest Factor Method:**
  - Ratio of Peak to RMS signal level
  - Low for steady state signal
  - High for signal with transients
  - Can locate transient parts

Approach 1
Attack and Release Times

- **Crest Factor Method:**
  - Ratio of Peak to RMS signal level
  - Low for steady state signal
  - High for signal with transients
  - Can locate transient parts
  - **Calculation:**

Approach 1

Attack and Release Times

- **Crest Factor Method:**
  - Ratio of Peak to RMS signal level
  - Low for steady state signal
  - High for signal with transients
  - Can locate transient parts
  - **Calculation:**

$$y_C[n] = y_{PEAK}[n]/y_{RMS}[n] \tag{2}$$

Approach 1

Attack and Release Times

- **Crest Factor Method:**
  - Ratio of Peak to RMS signal level
  - Low for steady state signal
  - High for signal with transients
  - Can locate transient parts
  - **Calculation:**

$$y_C[n] = y_{PEAK}[n]/y_{RMS}[n] \tag{2}$$

$$\tau_A[n] = 2 \cdot \tau_{Amax}[n]/y_C^2[n] \tag{3}$$

- **Crest Factor Method:**
  - Ratio of Peak to RMS signal level
  - Low for steady state signal
  - High for signal with transients
  - Can locate transient parts
  - **Calculation:**

$$y_C[n] = y_{PEAK}[n]/y_{RMS}[n] \tag{2}$$

$$\tau_A[n] = 2 \cdot \tau_{Amax}[n]/y_C^2[n] \tag{3}$$

$$\tau_R[n] = 2 \cdot \tau_{Rmax}[n]/y_C^2[n] - \tau_A[n] \tag{4}$$

Approach 1

Attack and Release Times (contd..)

- **Spectral Flux Method:**

Approach 1

Attack and Release Times (contd..)

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal

Approach 1
Attack and Release Times (contd..)

HELMUT SCHMIDT UNIVERSITÄT

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal
  - Low for steady state signal

Approach 1

Attack and Release Times (contd..)

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal
  - Low for steady state signal
  - High for signal with transients

Approach 1

Attack and Release Times (contd..)

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal
  - Low for steady state signal
  - High for signal with transients
  - More sensitive than Crest Factor

Approach 1
Attack and Release Times (contd..)

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal
  - Low for steady state signal
  - High for signal with transients
  - More sensitive than Crest Factor
  - **Calculation:**

HELMUT SCHMIDT UNIVERSITÄT

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal
  - Low for steady state signal
  - High for signal with transients
  - More sensitive than Crest Factor
  - **Calculation:**

$$SF(n) = \frac{\sum_{k=-N/2}^{N/2-1} H(|X(n,k)| - |X(n-1,k)|)}{\sum_{k=-N/2}^{N/2-1} |X(n,k)|} \tag{5}$$

Approach 1

Attack and Release Times (contd..)

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal
  - Low for steady state signal
  - High for signal with transients
  - More sensitive than Crest Factor
  - **Calculation:**

$$SF(n) = \frac{\sum_{k=-N/2}^{N/2-1} H(|X(n,k)| - |X(n-1,k)|)}{\sum_{k=-N/2}^{N/2-1} |X(n,k)|} \quad (5)$$

$$SF_{smooth}[n] = max(x[n], \alpha \cdot SF_{smooth}[n-1] + (1-\alpha) \cdot SF[n]) \quad (6)$$

Approach 1
Attack and Release Times (contd..)

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal
  - Low for steady state signal
  - High for signal with transients
  - More sensitive than Crest Factor
  - **Calculation:**

$$SF(n) = \frac{\sum_{k=-N/2}^{N/2-1} H(|X(n,k)| - |X(n-1,k)|)}{\sum_{k=-N/2}^{N/2-1} |X(n,k)|} \quad (5)$$

$$SF_{smooth}[n] = max(x[n], \alpha \cdot SF_{smooth}[n-1] + (1-\alpha) \cdot SF[n]) \quad (6)$$

$$\tau_A[n] = 2 \cdot \tau_{Amax}[n] / SF_{smooth}[n] \quad (7)$$

- **Spectral Flux Method:**
  - Rapidity in change of power spectrum of signal
  - Low for steady state signal
  - High for signal with transients
  - More sensitive than Crest Factor
  - **Calculation:**

$$SF(n) = \frac{\sum_{k=-N/2}^{N/2-1} H(|X(n,k)| - |X(n-1,k)|)}{\sum_{k=-N/2}^{N/2-1} |X(n,k)|} \quad (5)$$

$$SF_{smooth}[n] = max(x[n], \alpha \cdot SF_{smooth}[n-1] + (1-\alpha) \cdot SF[n]) \quad (6)$$

$$\tau_A[n] = 2 \cdot \tau_{Amax}[n] / SF_{smooth}[n] \quad (7)$$

$$\tau_R[n] = 2 \cdot \tau_{Rmax}[n] / SF_{smooth}^{\gamma}[n] - \tau_A[n] \quad (8)$$

Approach 1

Attack and Release Times - Evaluation

- Subjective evaluation based on human preference of parameters

Approach 1

Attack and Release Times - Evaluation

HELMUT SCHMIDT UNIVERSITÄT

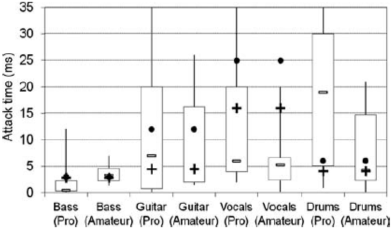
- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters

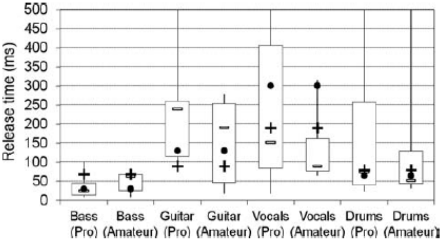
Approach 1

Attack and Release Times - Evaluation

HELMUT SCHMIDT UNIVERSITÄT

- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters



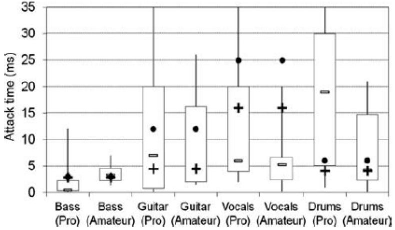


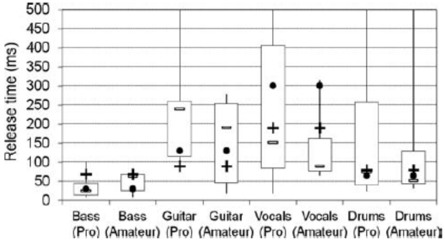
Approach 1

Attack and Release Times - Evaluation

HELMUT SCHMIDT UNIVERSITÄT

- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters





(−) Median value of response
(•) Crest Factor Automation
(+) Spectral Flux Automation

Approach 1
Threshold and Ratio

- **Threshold:**

Approach 1
Threshold and Ratio

- **Threshold:**
  - Defines desired compression

Approach 1
Threshold and Ratio

- **Threshold:**
  - Defines desired compression
  - To be controlled by user

Approach 1
Threshold and Ratio

- **Threshold:**
  - Defines desired compression
  - To be controlled by user
- **Ratio:**

Approach 1
Threshold and Ratio

- **Threshold:**
  - Defines desired compression
  - To be controlled by user
- **Ratio:**
  - To be set to infinity

Approach 1
Threshold and Ratio

- **Threshold:**
  - Defines desired compression
  - To be controlled by user
- **Ratio:**
  - To be set to infinity
  - Soft Knee with variable width equivalent to automatic ratio

Approach 1
Knee Width

- **Basic Method:** Bases on gain reduction of input signal

Approach 1
Knee Width

- **Basic Method:** Bases on gain reduction of input signal

$$g_{Dev}[n] = \alpha \cdot g_{Dev}[n-1] + (1-\alpha) \cdot (g[n] - g_{Est}) \tag{9}$$

Approach 1
Knee Width

- **Basic Method:** Bases on gain reduction of input signal

$$g_{Dev}[n] = \alpha \cdot g_{Dev}[n-1] + (1-\alpha) \cdot (g[n] - g_{Est}) \tag{9}$$

$$W[n] = 2.5 \cdot (g_{Dev}[n] + g_{Est}) \tag{10}$$

Approach 1
Knee Width

- **Basic Method:** Bases on gain reduction of input signal

$$g_{Dev}[n] = \alpha \cdot g_{Dev}[n-1] + (1-\alpha) \cdot (g[n] - g_{Est}) \tag{9}$$

$$W[n] = 2.5 \cdot (g_{Dev}[n] + g_{Est}) \tag{10}$$

$$g_{Est} = T \cdot (1 - 1/R)/2$$

Approach 1
Knee Width

HELMUT SCHMIDT UNIVERSITÄT

- **Basic Method:** Bases on gain reduction of input signal

$$g_{Dev}[n] = \alpha \cdot g_{Dev}[n-1] + (1-\alpha) \cdot (g[n] - g_{Est}) \tag{9}$$

$$W[n] = 2.5 \cdot (g_{Dev}[n] + g_{Est}) \tag{10}$$

$$g_{Est} = T \cdot (1 - 1/R)/2$$

- **Optimized Method:** Bases on SF information of input signal

Approach 1
Knee Width

- **Basic Method:** Bases on gain reduction of input signal

$$g_{Dev}[n] = \alpha \cdot g_{Dev}[n-1] + (1-\alpha) \cdot (g[n] - g_{Est}) \tag{9}$$

$$W[n] = 2.5 \cdot (g_{Dev}[n] + g_{Est}) \tag{10}$$

$$g_{Est} = T \cdot (1 - 1/R)/2$$

- **Optimized Method:** Bases on SF information of input signal

$$SF_{min}[n] = min(|SF[n]|, \alpha \cdot SF_{min}[n-1] + (1-\alpha) \cdot SF[n]) \tag{11}$$

Approach 1
Knee Width

- **Basic Method:** Bases on gain reduction of input signal

$$g_{Dev}[n] = \alpha \cdot g_{Dev}[n-1] + (1-\alpha)\cdot(g[n] - g_{Est}) \tag{9}$$

$$W[n] = 2.5\cdot(g_{Dev}[n] + g_{Est}) \tag{10}$$

$$g_{Est} = T\cdot(1 - 1/R)/2$$

- **Optimized Method:** Bases on SF information of input signal

$$SF_{min}[n] = min(|SF[n]|, \alpha\cdot SF_{min}[n-1] + (1-\alpha)\cdot SF[n]) \tag{11}$$

$$SF_{min,avg}[n] = (1-\alpha_2)\cdot SF_{min}[n] + \alpha_2\cdot SF_{min,avg}[n-1]) \tag{12}$$

Approach 1

Knee Width

HELMUT SCHMIDT UNIVERSITÄT

- **Basic Method:** Bases on gain reduction of input signal

$$g_{Dev}[n] = \alpha \cdot g_{Dev}[n-1] + (1-\alpha) \cdot (g[n] - g_{Est}) \quad (9)$$

$$W[n] = 2.5 \cdot (g_{Dev}[n] + g_{Est}) \quad (10)$$

$$g_{Est} = T \cdot (1 - 1/R)/2$$

- **Optimized Method:** Bases on SF information of input signal

$$SF_{min}[n] = min(|SF[n]|, \alpha \cdot SF_{min}[n-1] + (1-\alpha) \cdot SF[n]) \quad (11)$$

$$SF_{min,avg}[n] = (1-\alpha_2) \cdot SF_{min}[n] + \alpha_2 \cdot SF_{min,avg}[n-1]) \quad (12)$$

$$W[n] = 2.5 \cdot g^k_{Avg}[n] \quad (13)$$

Approach 1
Knee Width

- **Basic Method:** Bases on gain reduction of input signal

$$g_{Dev}[n] = \alpha \cdot g_{Dev}[n-1] + (1-\alpha)\cdot(g[n] - g_{Est}) \tag{9}$$

$$W[n] = 2.5\cdot(g_{Dev}[n] + g_{Est}) \tag{10}$$

$$g_{Est} = T\cdot(1 - 1/R)/2$$

- **Optimized Method:** Bases on SF information of input signal

$$SF_{min}[n] = min(|SF[n]|, \alpha\cdot SF_{min}[n-1] + (1-\alpha)\cdot SF[n]) \tag{11}$$

$$SF_{min,avg}[n] = (1-\alpha_2)\cdot SF_{min}[n] + \alpha_2\cdot SF_{min,avg}[n-1]) \tag{12}$$

$$W[n] = 2.5\cdot g_{Avg}^{k}[n] \tag{13}$$

$$k = \begin{cases} 0.6, & SF_{min,avg} > 0.1 \\ 0.05, & SF_{min,avg} <= 0.1 \end{cases} \tag{14}$$

Approach 1

Knee Width - Evaluation

- Subjective evaluation based on human preference of parameters

Approach 1

Knee Width - Evaluation

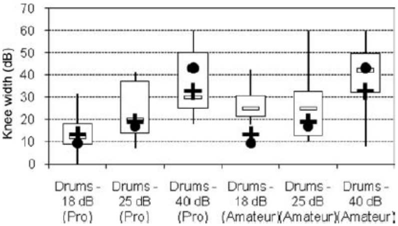
HELMUT SCHMIDT UNIVERSITÄT

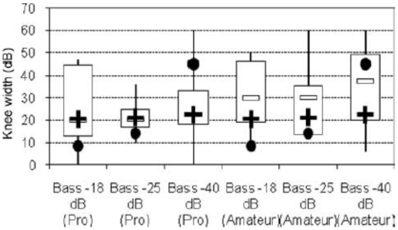
- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters

Approach 1

Knee Width - Evaluation

- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters



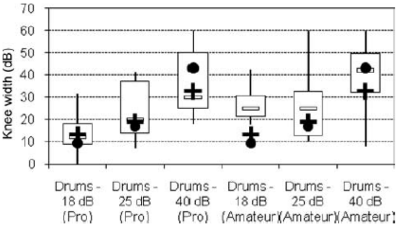


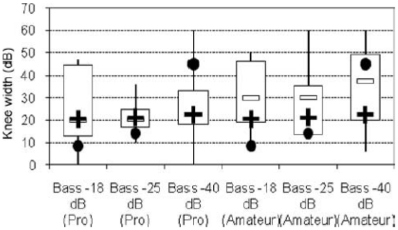
Approach 1

Knee Width - Evaluation

HELMUT SCHMIDT UNIVERSITÄT

- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters





(−) Median value of response
(•) Gain Reduction Dependent Automation
(+) Signal Information Dependent Automation

Approach 1
Make-up Gain

HELMUT SCHMIDT UNIVERSITÄT

- **Compression based Method:**

Approach 1
Make-up Gain

- **Compression based Method:**

$$g_{make-up}[n] = -(g_{Dev}[n] + g_{Est}) \tag{15}$$

Approach 1
Make-up Gain

- **Compression based Method:**

$$g_{make-up}[n] = -(g_{Dev}[n] + g_{Est}) \quad (15)$$

- **Loudness based Method:**

Approach 1
Make-up Gain

HELMUT SCHMIDT UNIVERSITÄT

- **Compression based Method:**

$$g_{make-up}[n] = -(g_{Dev}[n] + g_{Est}) \tag{15}$$

- **Loudness based Method:**
  - Compares perceived loudness before and after compression

Approach 1
Make-up Gain

HELMUT SCHMIDT UNIVERSITÄT

- **Compression based Method:**

$$g_{make-up}[n] = -(g_{Dev}[n] + g_{Est}) \tag{15}$$

- **Loudness based Method:**
  - Compares perceived loudness before and after compression
  - Calculates make-up gain from loudness difference

Approach 1
Make-up Gain - Evaluation

HELMUT SCHMIDT UNIVERSITÄT

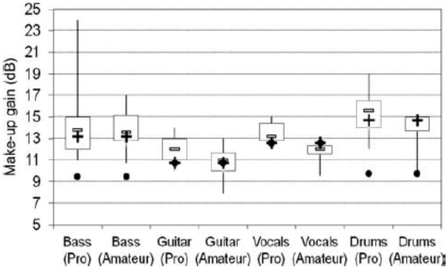
- Subjective evaluation based on human preference of parameters

Approach 1
Make-up Gain - Evaluation

- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters

Approach 1
Make-up Gain - Evaluation

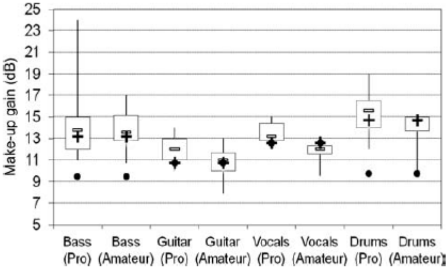
- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters



Approach 1

Make-up Gain - Evaluation

- Subjective evaluation based on human preference of parameters
- Evaluation results compared with automated parameters



(–) Median value of response
(•) Compression based Automation
(+) Loudness based Automation

Outline

- Introduction
- Theory
- Automation
- Approach 1
- Approach 2
- Discussions
- Conclusion

Approach 2

Audio Features Extraction

$$SC_{mean} = E[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}], Y(n,k) = |X(n,k)| \tag{16}$$

Approach 2
Audio Features Extraction

$$SC_{mean} = E[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}], Y(n,k) = |X(n,k)| \tag{16}$$

$$SC_{var} = Var[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}] \tag{17}$$

Approach 2
Audio Features Extraction

$$SC_{mean} = E[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}], Y(n,k) = |X(n,k)| \quad (16)$$

$$SC_{var} = Var[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}] \quad (17)$$

$$SV_{mean} = E[\sqrt{E[Y(n,k)^2] - (E[Y(n,k)])^2}] \quad (18)$$

Approach 2
Audio Features Extraction

HELMUT SCHMIDT UNIVERSITÄT

$$SC_{mean} = E[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}], Y(n,k) = |X(n,k)| \tag{16}$$

$$SC_{var} = Var[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}] \tag{17}$$

$$SV_{mean} = E[\sqrt{E[Y(n,k)^2] - (E[Y(n,k)])^2}] \tag{18}$$

$$SV_{var} = Var[\sqrt{E[Y(n,k)^2] - (E[Y(n,k)])^2}] \tag{19}$$

Approach 2

Audio Features Extraction

HELMUT SCHMIDT UNIVERSITÄT

$$SC_{mean} = E[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}], Y(n,k) = |X(n,k)| \tag{16}$$

$$SC_{var} = Var[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}] \tag{17}$$

$$SV_{mean} = E[\sqrt{E[Y(n,k)^2] - (E[Y(n,k)])^2}] \tag{18}$$

$$SV_{var} = Var[\sqrt{E[Y(n,k)^2] - (E[Y(n,k)])^2}] \tag{19}$$

$$RMS_{mean} = E[\sqrt{\frac{1}{M} \cdot \sum_{m=0}^{M-1} x(m)^2}] \tag{20}$$

Approach 2
Audio Features Extraction

$$SC_{mean} = E[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}], Y(n,k) = |X(n,k)| \tag{16}$$

$$SC_{var} = Var[\frac{\sum_{k=0}^{K-1} k \cdot Y(n,k)}{\sum_{k=0}^{K-1} Y(n,k)}] \tag{17}$$

$$SV_{mean} = E[\sqrt{E[Y(n,k)^2] - (E[Y(n,k)])^2}] \tag{18}$$

$$SV_{var} = Var[\sqrt{E[Y(n,k)^2] - (E[Y(n,k)])^2}] \tag{19}$$

$$RMS_{mean} = E[\sqrt{\frac{1}{M} \cdot \sum_{m=0}^{M-1} x(m)^2}] \tag{20}$$

$$RMS_{var} = Var[\sqrt{\frac{1}{M} \cdot \sum_{m=0}^{M-1} x(m)^2}] \tag{21}$$

Approach 2

Regression Model - Training

- **Training Data:** Violin Samples (RWC isolated note database)

Approach 2

Regression Model - Training

- **Training Data:** Violin Samples (RWC isolated note database)
- **Training Set:**

Training sets (size)	Conditions			
	Thr(dB)	Ratio	Att(ms)	Rel(ms)
A (60*50)	0:1:49	2	5	200
B (60*50)	37.5	1:0.4:20	5	200
C (60*100)	37.5	2	1:1:100	200
D (60*100)	37.5	2	5	50:10:1000

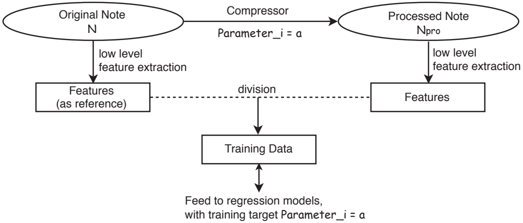
Approach 2

Regression Model - Training

- **Training Data:** Violin Samples (RWC isolated note database)
- **Training Set:**

Training sets (size)	Conditions			
	Thr(dB)	Ratio	Att(ms)	Rel(ms)
A (60*50)	0:1:49	2	5	200
B (60*50)	37.5	1:0.4:20	5	200
C (60*100)	37.5	2	1:1:100	200
D (60*100)	37.5	2	5	50:10:1000

- **Training Data Generation:**



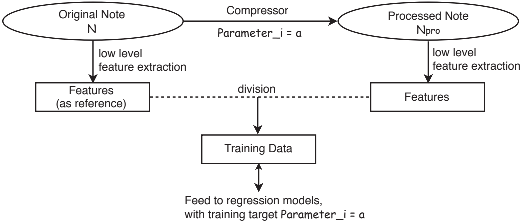
Approach 2

Regression Model - Training

- **Training Data:** Violin Samples (RWC isolated note database)
- **Training Set:**

Training sets (size)	Conditions			
	Thr(dB)	Ratio	Att(ms)	Rel(ms)
A (60*50)	0:1:49	2	5	200
B (60*50)	37.5	1:0.4:20	5	200
C (60*100)	37.5	2	1:1:100	200
D (60*100)	37.5	2	5	50:10:1000

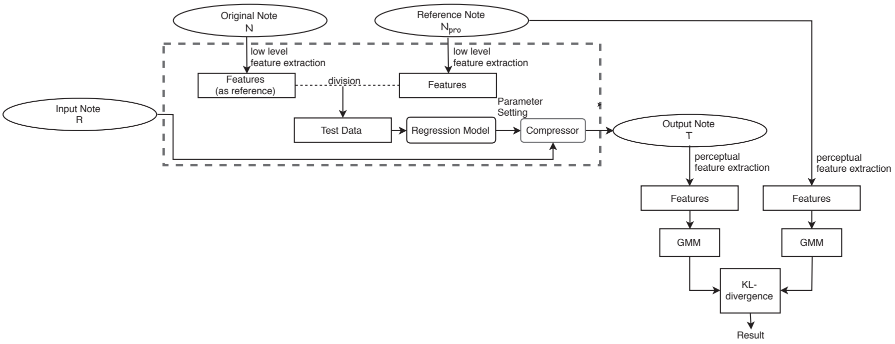
- **Training Data Generation:**



- **Regression Models:** Linear and Random Forest Regression

Approach 2
Automation Process

HELMUT SCHMIDT UNIVERSITÄT



Approach 2
Similarity Measure

- **Crest Factor Difference:**

Approach 2

Similarity Measure

HELMUT SCHMIDT UNIVERSITÄT

- **Crest Factor Difference:**

$$D_{Crest}(A, B) = mean(|Crest(A) - Crest(B)|)$$

Approach 2
Similarity Measure

- **Crest Factor Difference:**

$$D_{Crest}(A, B) = mean(|Crest(A) - Crest(B)|)$$

Violin	Threshold	Ratio	Attack	Release
$D_{Crest}(N_{pro}, R)$	60.31	94.13	104.93	85.31
$D_{Crest}(N_{pro}, T)_{LR}$	**12.53**	39.72	46.76	48.62
$D_{Crest}(N_{pro}, T)_{RF}$	15.27	**38.24**	**45.23**	**47.19**

Approach 2
Similarity Measure

HELMUT SCHMIDT UNIVERSITÄT

- **Crest Factor Difference:**

$$D_{Crest}(A, B) = mean(|Crest(A) - Crest(B)|)$$

Violin	Threshold	Ratio	Attack	Release
$D_{Crest}(N_{pro}, R)$	60.31	94.13	104.93	85.31
$D_{Crest}(N_{pro}, T)_{LR}$	**12.53**	39.72	46.76	48.62
$D_{Crest}(N_{pro}, T)_{RF}$	15.27	**38.24**	**45.23**	**47.19**

- **Kullback-Leibler Divergence:**

Approach 2
Similarity Measure

HELMUT SCHMIDT UNIVERSITÄT

- **Crest Factor Difference:**

  $$D_{Crest}(A, B) = mean(|Crest(A) - Crest(B)|)$$

Violin	Threshold	Ratio	Attack	Release
$D_{Crest}(N_{pro}, R)$	60.31	94.13	104.93	85.31
$D_{Crest}(N_{pro}, T)_{LR}$	**12.53**	39.72	46.76	48.62
$D_{Crest}(N_{pro}, T)_{RF}$	15.27	**38.24**	**45.23**	**47.19**

- **Kullback-Leibler Divergence:**

  Between GMM distributions based on MFCC coefficients

HELMUT SCHMIDT UNIVERSITÄT

- **Crest Factor Difference:**

$$D_{Crest}(A,B) = mean(|Crest(A) - Crest(B)|)$$

Violin	Threshold	Ratio	Attack	Release
$D_{Crest}(N_{pro}, R)$	60.31	94.13	104.93	85.31
$D_{Crest}(N_{pro}, T)_{LR}$	**12.53**	39.72	46.76	48.62
$D_{Crest}(N_{pro}, T)_{RF}$	15.27	**38.24**	**45.23**	**47.19**

- **Kullback-Leibler Divergence:**

Between GMM distributions based on MFCC coefficients

Violin	Threshold	Ratio	Attack	Release
$D(N_{pro}, R)$	38.122	53.187	44.018	55.206
$D(N_{pro}, T)_{LR}$	19.799	20.911	22.852	20.994
$D(N_{pro}, T)_{RF}$	**19.742**	**20.856**	**22.213**	**20.807**

Outline

HELMUT SCHMIDT UNIVERSITÄT

Discussions

- **Music Samples:** 4 drum samples (EDM, Hip-Hop, Rock, Jazz)

Discussions

- **Music Samples:** 4 drum samples (EDM, Hip-Hop, Rock, Jazz)
- **Subjective evaluation**: Human preference of Attack and Release times across different styles of music

Discussions

- **Music Samples:** 4 drum samples (EDM, Hip-Hop, Rock, Jazz)
- **Subjective evaluation**: Human preference of Attack and Release times across different styles of music
- **Preference Chart:**

Discussions

- **Music Samples:** 4 drum samples (EDM, Hip-Hop, Rock, Jazz)
- **Subjective evaluation**: Human preference of Attack and Release times across different styles of music
- **Preference Chart:**

Comp	Fast Attack	Slow Attack	Fast Release	Slow Release
EDM	0.38	0.62	0.65	0.35
Hip-hop	0.48	0.52	0.46	0.55
Rock	0.29	0.71	0.51	0.54
Jazz	0.25	0.78	0.25	0.77

Discussions

HELMUT SCHMIDT UNIVERSITÄT

- **Music Samples:** 4 drum samples (EDM, Hip-Hop, Rock, Jazz)
- **Subjective evaluation**: Human preference of Attack and Release times across different styles of music
- **Preference Chart:**

Comp	Fast Attack	Slow Attack	Fast Release	Slow Release
EDM	0.38	0.62	0.65	0.35
Hip-hop	0.48	0.52	0.46	0.55
Rock	0.29	0.71	0.51	0.54
Jazz	0.25	0.78	0.25	0.77

- **Deduction:**

Discussions

HELMUT SCHMIDT UNIVERSITÄT

- **Music Samples:** 4 drum samples (EDM, Hip-Hop, Rock, Jazz)
- **Subjective evaluation**: Human preference of Attack and Release times across different styles of music
- **Preference Chart:**

Comp	Fast Attack	Slow Attack	Fast Release	Slow Release
EDM	0.38	0.62	0.65	0.35
Hip-hop	0.48	0.52	0.46	0.55
Rock	0.29	0.71	0.51	0.54
Jazz	0.25	0.78	0.25	0.77

- **Deduction:**
  - Different preference of Attack and Release across different styles

Discussions

HELMUT SCHMIDT UNIVERSITÄT

- **Music Samples:** 4 drum samples (EDM, Hip-Hop, Rock, Jazz)
- **Subjective evaluation**: Human preference of Attack and Release times across different styles of music
- **Preference Chart:**

Comp	Fast Attack	Slow Attack	Fast Release	Slow Release
EDM	0.38	0.62	0.65	0.35
Hip-hop	0.48	0.52	0.46	0.55
Rock	0.29	0.71	0.51	0.54
Jazz	0.25	0.78	0.25	0.77

- **Deduction:**
  - Different preference of Attack and Release across different styles
  - Machine Learning approach can make use of this information

Outline

HELMUT SCHMIDT UNIVERSITÄT

HELMUT SCHMIDT UNIVERSITÄT

- Better results with approaches using Spectral Flux, SF based Optimization and Loudness

HELMUT SCHMIDT UNIVERSITÄT

- Better results with approaches using Spectral Flux, SF based Optimization and Loudness
- Better results with Random Forest Regression

Conclusion

Summary

- Better results with approaches using Spectral Flux, SF based Optimization and Loudness
- Better results with Random Forest Regression
- Variation of preference of ballistics across styles

Conclusion
Future Work

- Include additional audio features mentioned in Approach 1 to train Regression Model in Approach 2 for better results

Conclusion
Future Work

- Include additional audio features mentioned in Approach 1 to train Regression Model in Approach 2 for better results
- Include human preference information for parameters across different styles while training Regression Model

Conclusion
Future Work

- Include additional audio features mentioned in Approach 1 to train Regression Model in Approach 2 for better results
- Include human preference information for parameters across different styles while training Regression Model
- Evaluation on more complex tracks

1. Dimitrios Giannoulis, Michael Massberg, Joshua D. Reiss, "Parameter automation in a dynamic range compressor", J. Audio Eng. Soc., Vol. 61, No. 10, October 2013
2. Di Sheng, György Fazekas ,"Automatic control of the dynamic range compressor using a regression model and a reference sound", Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17), Edinburgh, UK, September 5-9, 2017
3. Gary Bromham, David Moffat, Mathieu Barthet, György Fazekas ,"The impact of compressor ballistics on the perceived style of music", Audio Engineering Society, Convention Paper 10080, presented at the 145th Convention New York, NY, USA, October 17-20, 2018

HELMUT SCHMIDT UNIVERSITÄT

Thanks for your time

Any Question or Suggestion?