EXPERIMENTS IN AUTOMATIC SAMPLE DETECTION IN HIP HOP

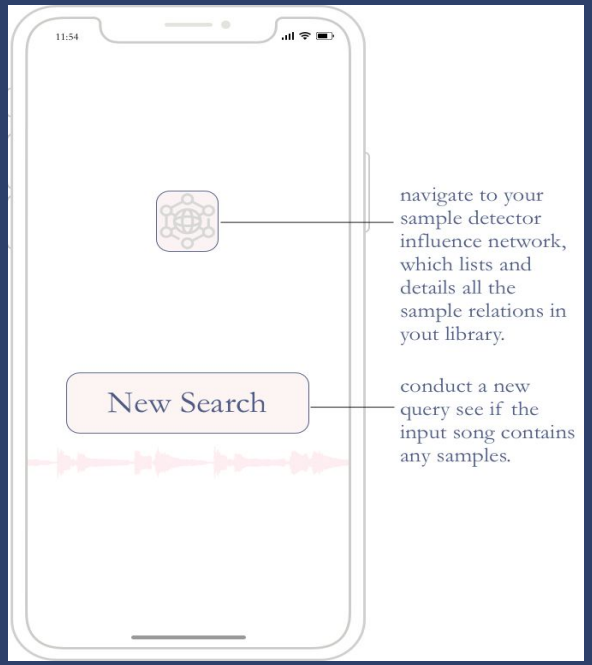
M.M. Music Technology Thesis Presentation

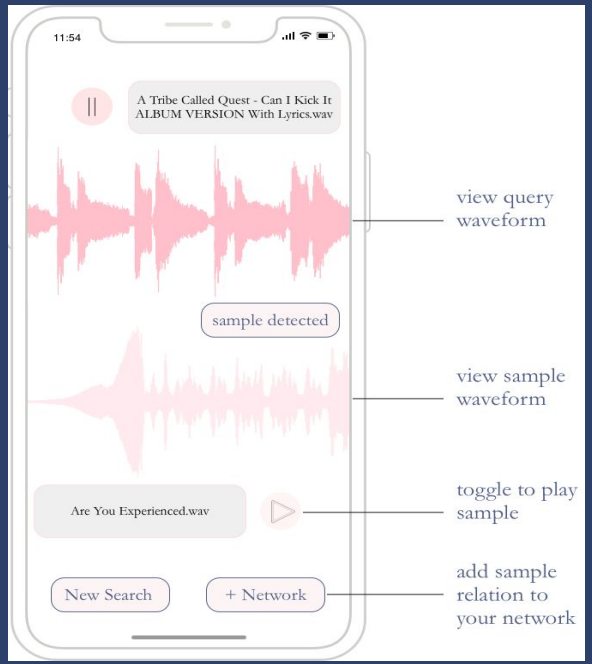
TREY BRADLEY

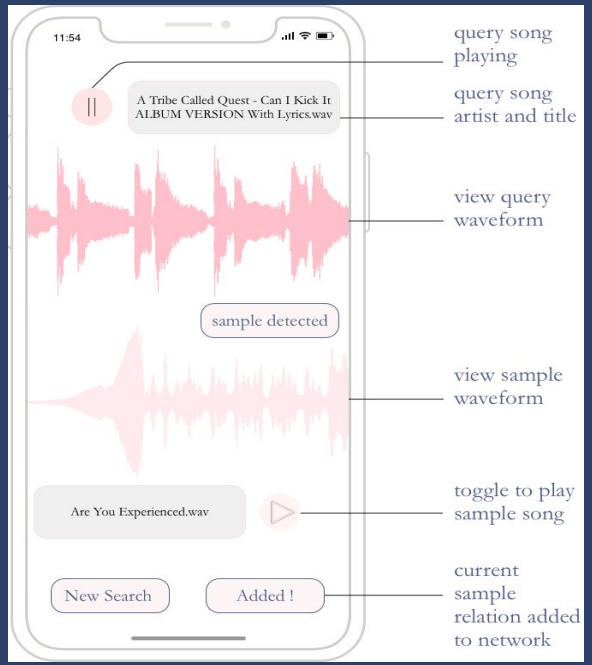
WIREFRAMES FOR THE PROJECT

I illustrated a GUI for an music discovery app that goes along with my this thesis topic. To learn more about this audio analysis that went behind this sample detection app, please read the attached presentation, or visit the github repo at **AutomaticSampleDetection**

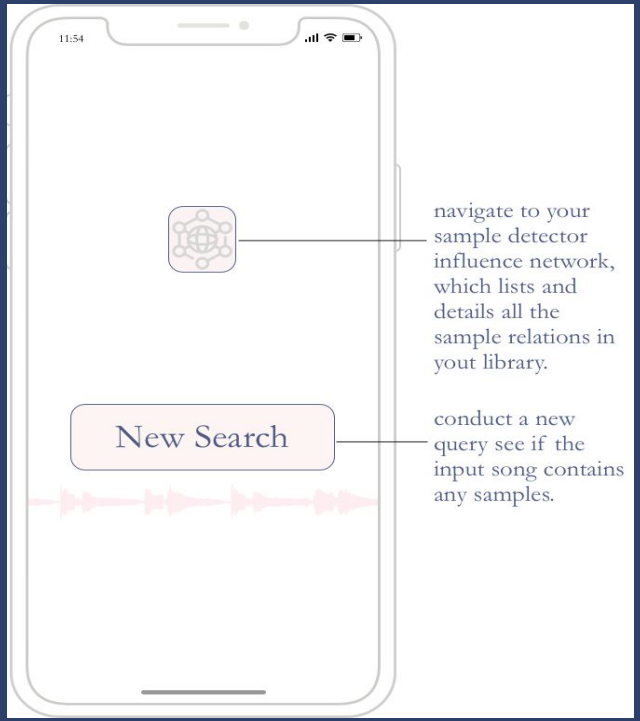
WIREFRAMES_NEW SEARCH

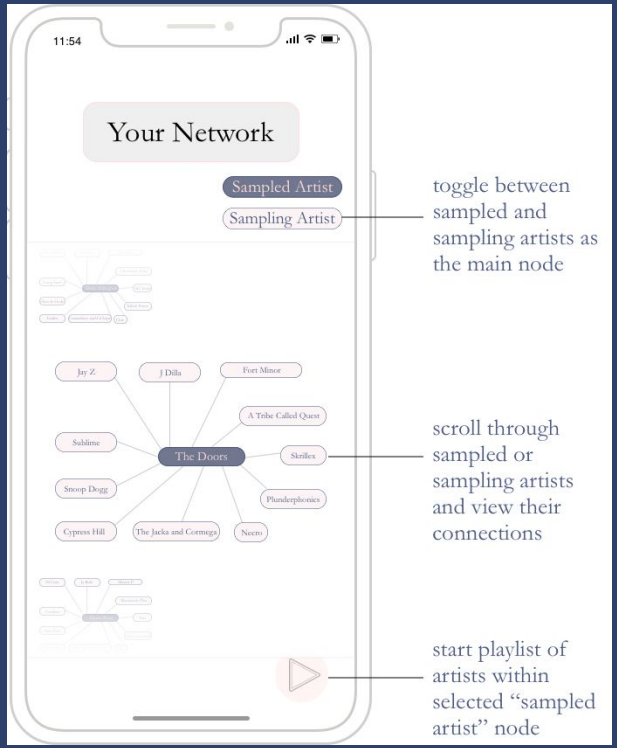






WIREFRAME_YOUR NETWORK





PRESENTATION LAYOUT

I INTRODUCTION

Goals, Background on Sampling and Music Information Retrieval

II METHODOLOGY

Data, Experiments and Results

III ANALYSIS

Interpreting the results

IV CONCLUSION

Immediate takeaways, future work

DEFINITIONS

Sampling is the use of pre recorded audio in a new, complex, audio mixture. Examples include:

- Music Concrete
- Electronic Music
- Hip Hop Music

Example of a Sample Relation (pair):

Sample Song	Sampling (Query) Song
Saudade vem Correndo by Stan Getz and Luis Bonfa	Runnin by The Pharcyde (prod. J. Dilla)

MOTIVATIONS

1. Data for Musicology and music discovery

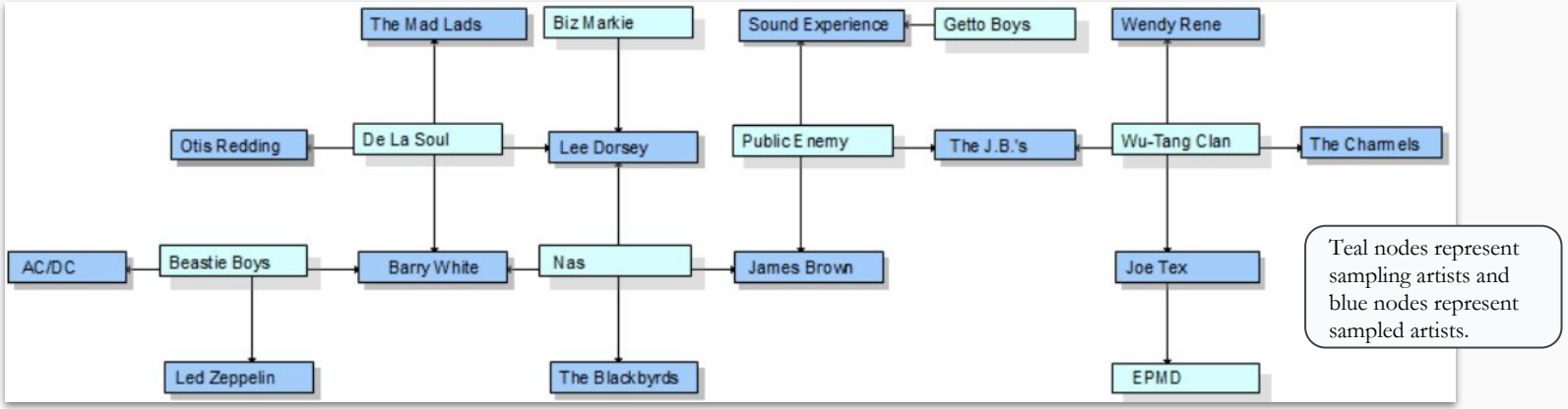


Figure 1 Sampling network in van Balen, J. (2011). Automatic Recognition of Samples in Musical Audio. *Master Thesis for Department of Information and Communication Technologies*. p. 2.

2. Groundtruth database for legal decision

GOAL + SYSTEM REQUIREMENTS

Design and prototype an algorithm that can detect samples in query songs.

Sample Types:
- Chopped
- Pitch shifted
- Time stretched
- Reversed
- Signal processed
- Looped
- One shots

- Accurate
- Fast
- Robust

whosampled.com

RELATED AREAS OF RESEARCH

Audio Identification

- Identify samples within query songs
- Landmark fingerprints

Source Separation

- Extract samples from query songs

van Balen, J. (2011). Automatic Recognition of Samples in Musical Audio. *Master Thesis for Department of Information and Communication Technologies*. pp. 1-88.

Whitney, J. L. (2013). Automatic recognition of samples in hip-hip music through non-negative matrix factorization. Master's thesis, University of Miami.

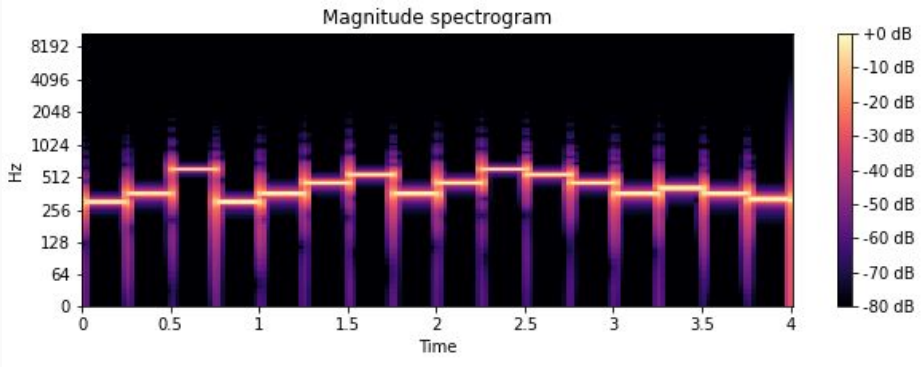
Gururani, S., Lerch, A. (2017). Automatic Sample Detection in Polyphonic Music. *18th International Society for Music Information Retrieval Conference.*

de Carvalho, L. L. (2019). Processamento digital de a´udio aplicado a` detecc ̧a ̃o de samples musicais. Master's thesis, Universidade Federal do Rio de Janeiro.

AUDIO IN THE FREQUENCY DOMAIN

Short Time Fourier Transform (STFT)

- An algorithm that conducts a fast fourier transform on each window of a signal, to reveal the frequency content of the signal over time.



Energy (color) in frequency bands (Hz) over time (sec)

BACKGROUND - MATRIX DECOMPOSITION

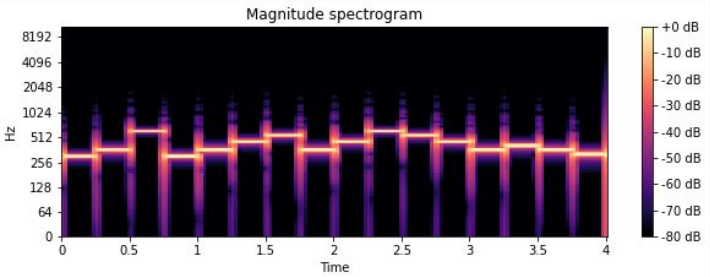
Non-negative Matrix Factorization **(NMF)**

- Decompose a matrix V into matrices W and H, to reveal the matrix' (signal's) components (# of ranks).

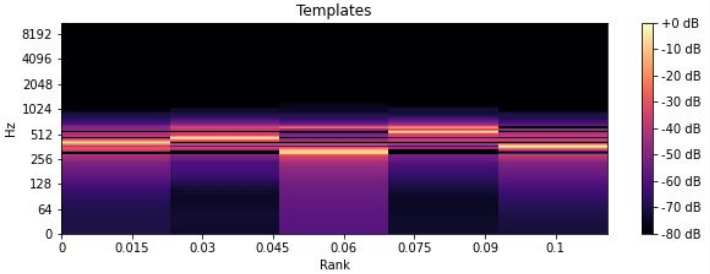
$$V \approx WH$$

- Initialize
- Rank
- Update rule
- Loss function
- # iterations

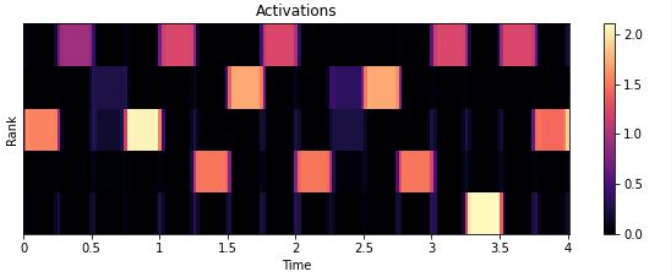
V



W



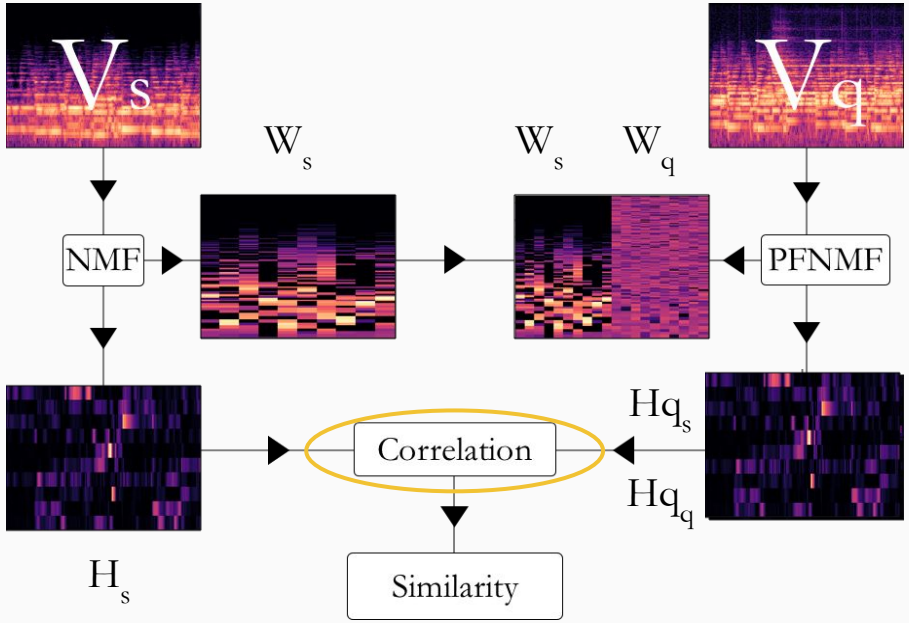
H



INITIALIZATION - PFNMF

Knowledge-based constraint

- An **NMF** using musical information about the sample song to decompose the query song.
  - In the diagram, it is carried out via a Ws **initialization** during the PFNMF of Vq. The Wq templates are randomly intialized and updated.



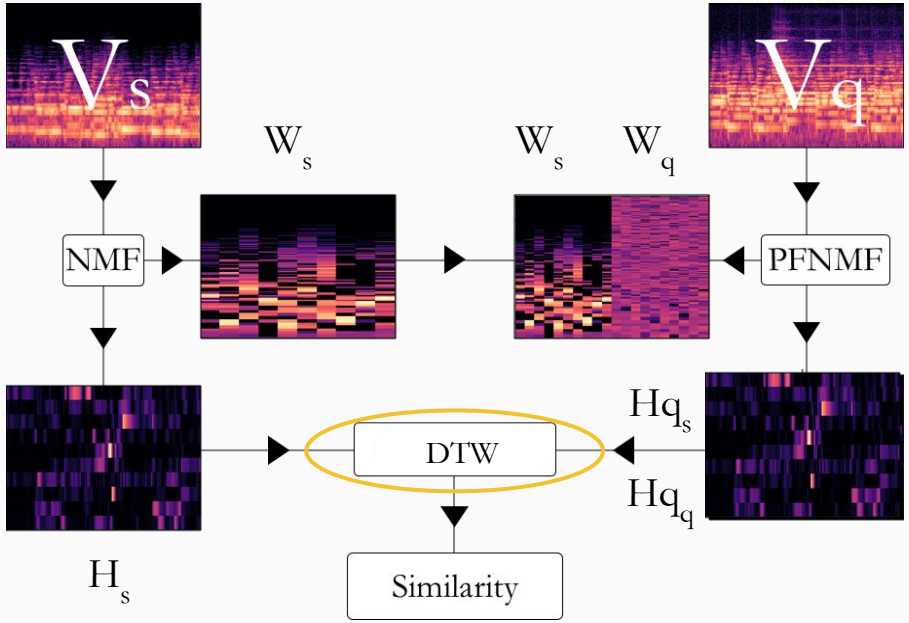
Müller, Meinard. Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications, Springer International Publishing AG, 2015

Dittmar, C. et. al (2012). Audio forensics meets music information retrieval: a toolbox for inspection of music plagiarism. *20th European Signal Processing Conference.*

BASELINE SYSTEM

Dynamic Time Warping **(DTW)**

- Sequence alignment
- Align H_s and H_q activations, for time-stretched samples



Gururani, S., Lerch, A. (2017). Automatic Sample Detection in Polyphonic Music. *18th International Society for Music Information Retrieval Conference.*

EVALUATION METRICS

Grountruth Label	Predicted Label	Metric
1	1	True Positive (TP)
0	0	True Negative (TN)
0	1	False Positive (FP)
1	0	False Negative (FN)

	+ Sampling	- Sampling
+ Sampling (prediction)	TP	FP
- Sampling (prediction)	FN	TN

Accuracy	Precision (P)	Recall (R)	**F-1**	Fp Rate
TP+FP / Total	TP/TP+FP	TP/TP+FN	$\frac{2*P*R}{P+R}$	FP/TN+FP

DATASET

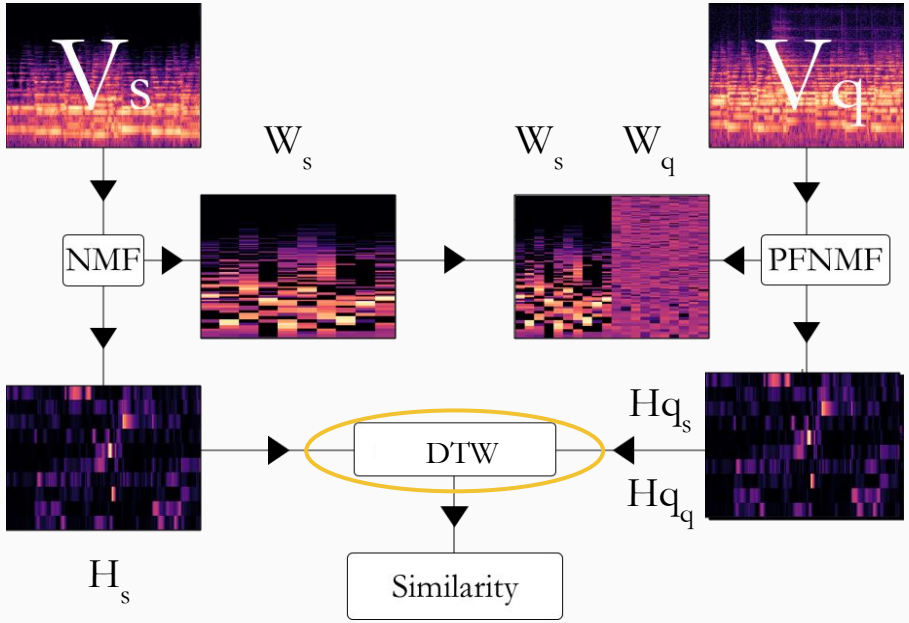
		Query Songs									
		Missy Elliott -	Kanye West -	A Tribe Calle	VI Seconds-	Biggie Smalls	Missy Elliot f	50 Cent - Hu	Aaliyah - Mo	Missy Elliott -	A Tribe Calle
Sample Songs	Ann Peebles	1	0	0	0	0	0	0	0	0	0
	Move On Up	0	1	0	0	0	0	0	0	0	0
	Milt Jackson	0	0	1	0	0	0	0	0	0	0
	One Directio	0	0	0	1	0	0	0	0	0	0
	Im Coming C	0	0	0	0	1	0	0	0	0	0
	HOT STREA	0	0	0	0	0	1	0	0	0	0
	Maze ft Franl	0	0	0	0	0	0	1	0	0	0
	Mayada EL H	0	0	0	0	0	0	0	1	0	0
	RUN-DMC -	0	0	0	0	0	0	0	0	1	0
	WE GETTIN	0	0	0	0	0	0	0	0	0	1

- 80 training examples
- 40 testing examples
- 10 x 10 batches

BASELINE EXPERIMENT - PFNMF

	+ Sampling	- Sampling
+ Sampling (prediction)	23	151
- Sampling (prediction)	17	209

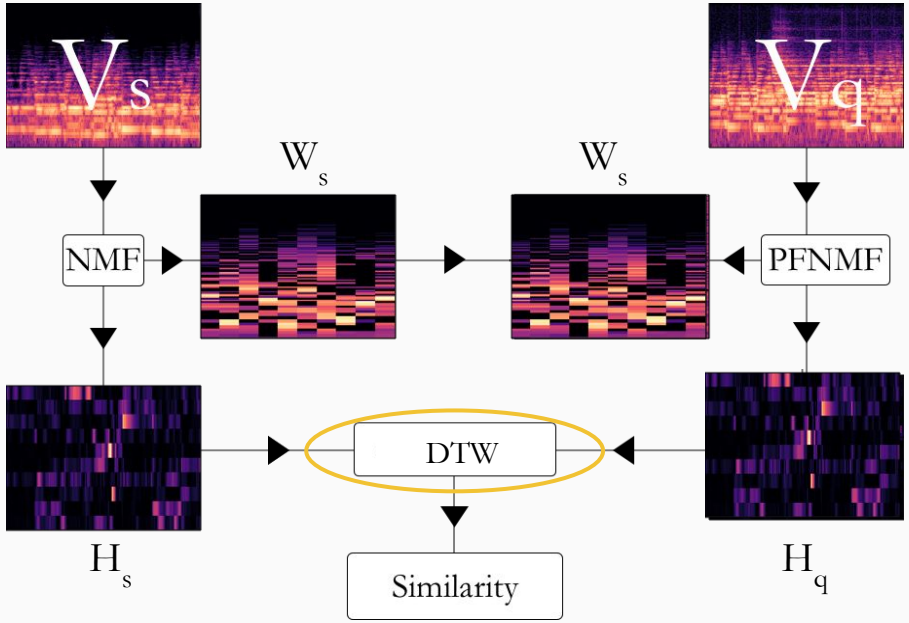
Accuracy	Precision	Recall	F-1	Fp Rate
58%	13%	57%	21%	41%



EXPERIMENT I - FULLY FIXED NMF

	+ Sampling	- Sampling
+ Sampling (prediction)	25	93
- Sampling (prediction)	15	267

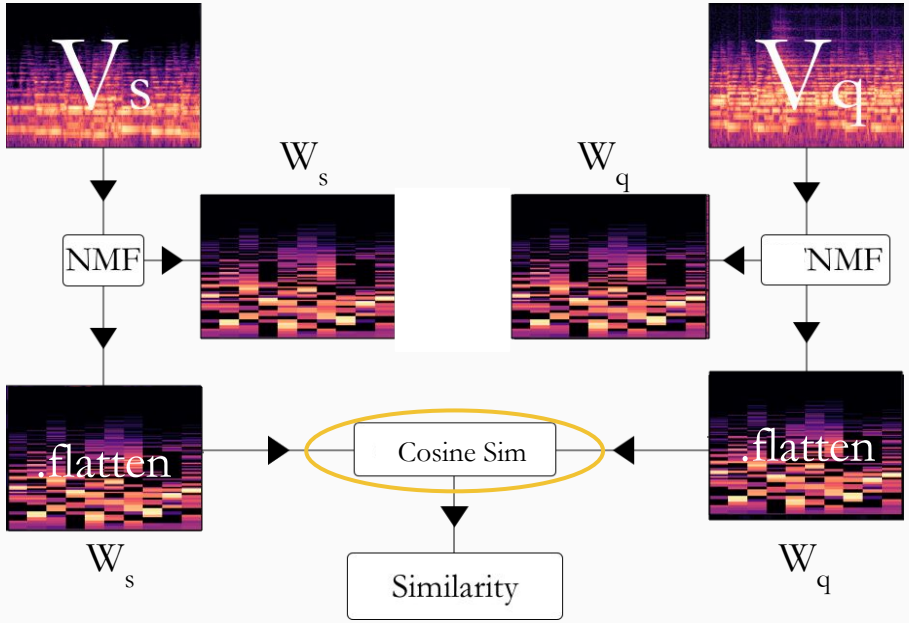
Accuracy	Precision	Recall	F-1	Fp Rate
73%	21%	62%	31%	25%



EXPERIMENT II - TRADITIONAL NMF

	+ Sampling	- Sampling
+ Sampling (prediction)	20	156
- Sampling (prediction)	20	204

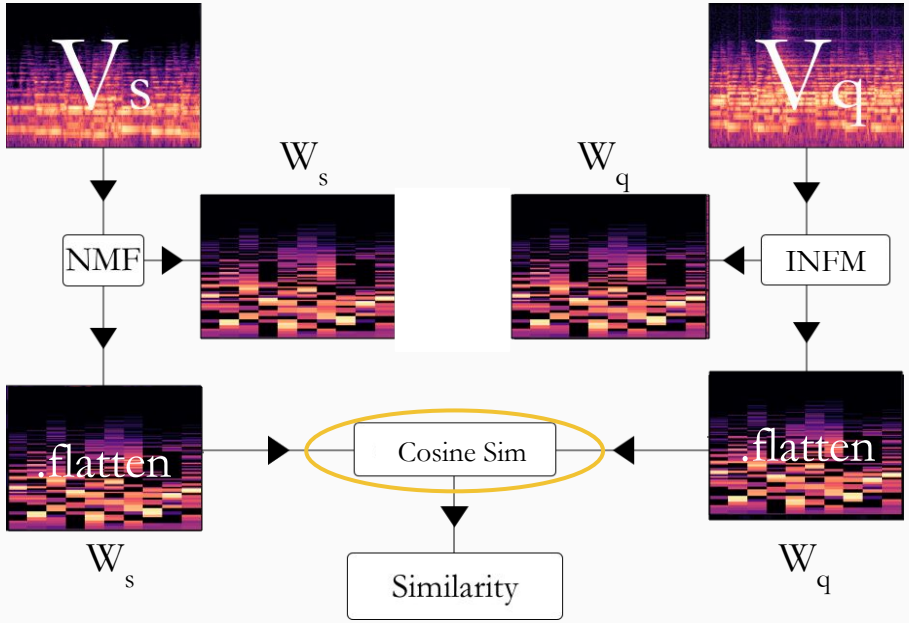
Accuracy	Precision	Recall	F-1	Fp Rate
56%	11%	50%	18%	43%



EXPERIMENT III - INITIALIZED NMF (INMF)

	+ Sampling	- Sampling
+ Sampling (prediction)	20	94
- Sampling (prediction)	20	266

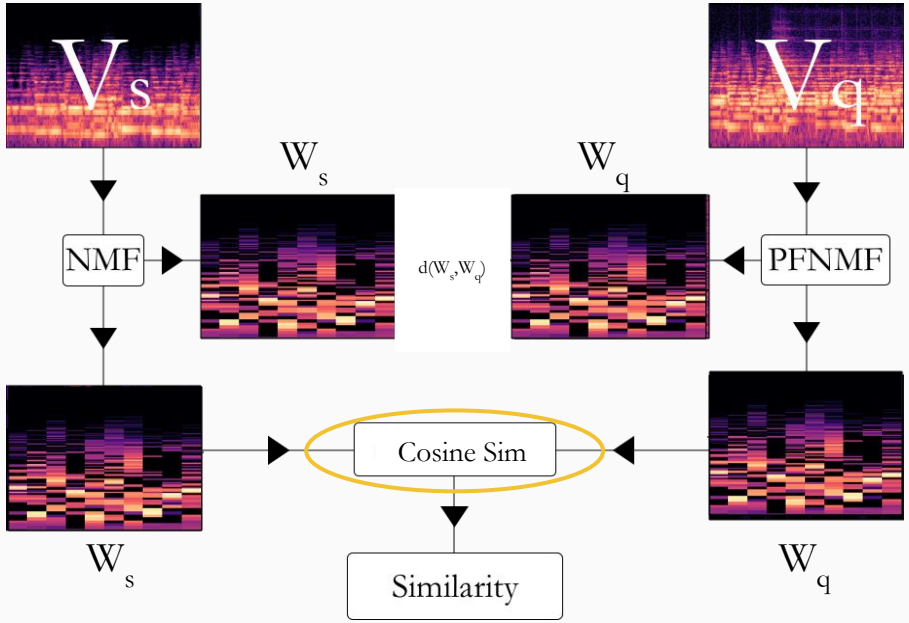
Accuracy	Precision	Recall	F-1	Fp Rate
71%	17%	50%	25%	26%



EXPERIMENT IV - INMF + SELECTION

	+ Sampling	- Sampling
+ Sampling (prediction)	21	93
- Sampling (prediction)	19	267

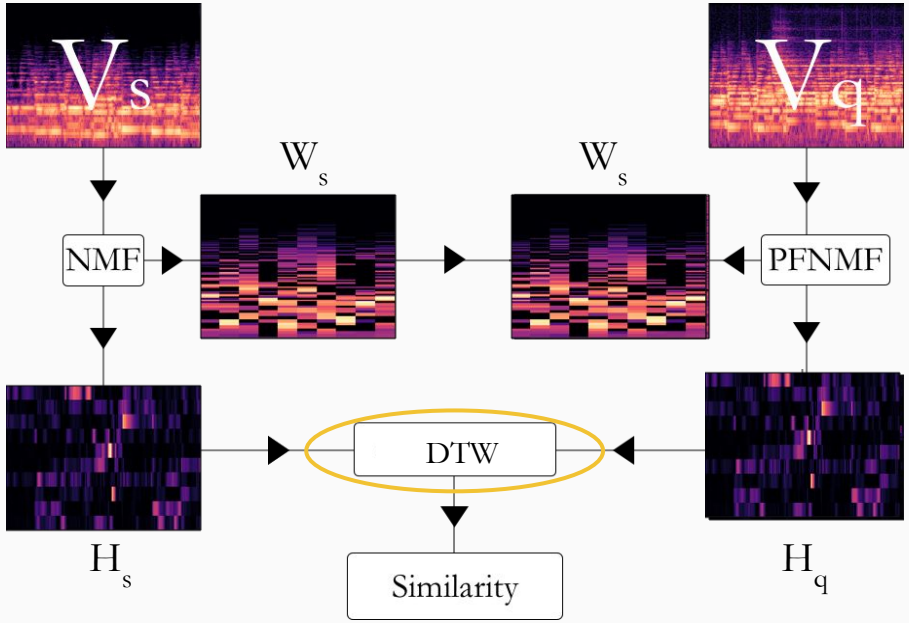
Accuracy	Precision	Recall	F-1	Fp Rate
72%	18%	52%	27%	25%



EXPERIMENT V - HPSS + FFNMF

	+ Sampling	- Sampling
+ Sampling (prediction)	19	118
- Sampling (prediction)	21	242

Accuracy	Precision	Recall	F-1	Fp Rate
65%	13%	47%	21%	32%



ANALYSIS - EXPERIMENT I - FFNMF

	+ Sampling	- Sampling
+ Sampling (prediction)	25	93
- Sampling (prediction)	15	267

Accuracy	Precision	Recall	F-1	Fp Rate
73%	21%	62%	31%	25%

TRUE POSITIVE

vocal, copy
vocal, perc, copy
vocal, pitched
tonal, copy
vocal, copy
tonal, pitched
tonal, copy
tonal, copy
vocal, copy
vocal, copy
vocal, pitched
vocal, copy
perc, pitched
tonal, pitched
tonal, perc, pitched
vocal, perc, copy
tonal, perc, pitched
tonal, copy
tonal, perc, pitched
tonal, pitched
vocal, copy
tonal, perc, pitched
tonal, perc, pitched
vocal, copy
vocal, copy

FALSE NEGATIVE

vocal, perc, stretched
tonal, pitched, buried
vocal, short, buried
vocal, filtered
perc, pitched
tonal, pitched, burred
perc, shifted
tonal, pitched, buried
vocal, pitched, buried
perc, tonal, pitched, shifted
perc, pitched, buried
vocal, copy
tonal, pitched, shifted, buried
perc, tonal, shifted
vocal, chopped, buried

CONCLUSIONS

TAKEAWAYS

- FFNMF can detected samples in query songs
  - song-song level
  - various types of samples
- A more refined system for pitched shifted samples
- Templates contain some information worth exploiting
- More work needs to be done in order to reduce the false-positive rate of the system and gain in robustness

FUTURE WORK

- Song-song analysis
  - Select candidate regions of the sample and query, using frame-wise reconstruction error of templates, to undergo DTW alignment.

BIBLIOGRAPHY

de Carvalho, L. L. (2019). Processamento digital de a´udio aplicado a` detecc¸a˜o de samples musicais. Master's thesis, Universidade Federal do Rio de Janeiro.

Dittmar, C., Hildebrand, K. F., Gaertner, D., Winges, M., Mu¨ller, F., and Aichroth, P. (2012). Audio forensics meets music information retrieval - a toolbox for in- spection of music plagiarism. In *20th European Signal Processing Conference*.

Gururani, S., Lerch, A. (2017). Automatic Sample Detection in Polyphonic Music. *18th International Society for Music Information Retrieval Conference.*

Müller, Meinard. Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications, Springer International Publishing AG, 2015

van Balen, J. (2011). Automatic Recognition of Samples in Musical Audio. *Master Thesis for Department of Information and Communication Technologies.* pp. 1-88.

Whitney, J. L. (2013). Automatic recognition of samples in hip-hip music through non-negative matrix factorization. Master's thesis, University of Miami.