

AUTOMATIC DETECTION OF OUTLIERS IN WORLD MUSIC RECORDINGS

Maria Panteli, Emmanouil Benetos, Simon Dixon
Centre for Digital Music
Queen Mary University of London

CONTENTS

- Dataset
- Descriptors
- Application

OUR GOAL

- Given a (small) world music corpus:
 - can we detect recordings that stand out from the rest...,
 - with respect to their music style,
 - by analysing the audio signal with computational tools?

OUR GOAL

- Why are we interested in this:
 - filter out noisy or irrelevant recordings
 - identify recordings with unique music characteristics

CONTENTS

- Dataset
- Descriptors
- Application

WHAT IS WORLD MUSIC

- ambiguous [1]
 - folk, art, popular music
 - amateur, professional musicians
 - sacred, secular, commercial purpose
 - Western, non-Western culture

[1] P.V. Bohlman. *World Music: A Very Short Introduction*. Oxford University Press, 2002.

WHAT IS WORLD MUSIC

- for our study it means:
 - mainly folk and traditional
 - from as many countries as possible
 - as old as possible
 - recorded and digitised

DATASET

- Many large digital archives of world music available today:
 - **Smithsonian Folkways Recordings**
 - **British Library**
 - **Centre de Recherche en Ethnomusicologie**
 - ...

DATASET



The screenshot shows the Smithsonian Folkways website. At the top, there is a navigation bar with links for 'FIND RECORDINGS', 'EXPLORE FOLKWAYS', 'TOOLS FOR TEACHING', 'ABOUT US', 'CUSTOMER SERVICE', and a 'VIEW CART' button. The main header features the 'SMITHSONIAN FOLKWAYS' logo. Below this, a section titled 'Smithsonian Folkways Recordings' provides a mission statement: 'Smithsonian Folkways Recordings is the nonprofit record label of the Smithsonian Institution. We are dedicated to supporting cultural diversity and increased understanding among peoples through the documentation, preservation, and dissemination of sound. We believe that musical and cultural diversity contributes to the vitality and quality of life throughout the world. Through the dissemination of audio recordings and educational materials we seek to strengthen people's engagement with their own cultural heritage and to enhance their awareness and appreciation of the cultural heritage of others. Our mission is the legacy of Moses Asch, who founded Folkways Records in 1948 to document "people's music," spoken word, instruction, and sounds from around the world. The Smithsonian acquired Folkways from the Asch estate in 1987, and Smithsonian Folkways Recordings has continued the Folkways commitment to cultural diversity, education, increased understanding, and lively engagement with the world of sound. Smithsonian Folkways has expanded on Asch's legacy, adding several other record labels to the collections and releasing over 300 new recordings that document and celebrate the sounds of the world around us.' Below the text, there are two featured recordings: 'Music of Central Asia vol. 4: Bardic Divas: Women's Voices in Central Asia' (SFW40523, 2007) and 'Classic Old-Time Music from Smithsonian Folkways' (SFW40093, 2003). A 'RELATED CONTENT' box lists 'The Folkways Collection podcast', 'Sound Sessions from Smithsonian Folkways podcast', 'Worlds of Sound book', and 'Worlds of Sound film'. There are also social media share buttons and a 'JOIN OUR EMAIL' button at the bottom left.

- Smithsonian Folkways Recordings
- metadata: Country, Year, Culture, Genre, Language, ...
- 30-second audio previews online

DATASET

COUNTRY

- united states (3,681)
- united kingdom (1,411)
- ussr (former) (1,098)
- canada (926)
- russia (923)

see more [+](#)

CULTURE GROUP

- anglo-american (721)
- african caribbean (659)
- jewish (583)
- scottish (481)
- african american (347)

see more [+](#)

INSTRUMENT

- vocals (3,675)
- guitar (2,813)
- drum (1,020)
- accordion (802)
- violin (508)



**Music of the World's
4** →
ARTIST Various Artists
YEAR 1958
CATALOG NO. FW04507
BUY Custom CD: \$16.98
Digital Download: \$9.99

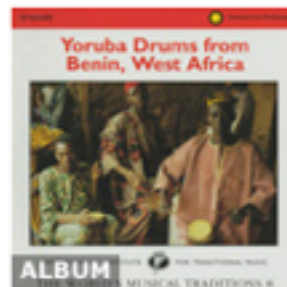
LANGUAGE

- english (2,640)
- french (569)
- spanish (552)
- hebrew (modern) (262)
- yiddish (183)

see more [+](#)



The World in My Mail Box
ARTIST Tony Schwartz
YEAR 1958
CATALOG NO. FW05562
BUY Custom CD: \$16.98
Digital Download: \$9.99



**The World's Musical Traditions
8:...** →
ARTIST Various Artists
YEAR 1996
CATALOG NO. SFW4044
BUY CD: \$14.98 |
Digital Download: \$9.99

DECADE

- 1950s (4,635)
- 1960s (3,561)
- 1970s (3,327)
- 1990s (2,074)
- 1980s (1,787)

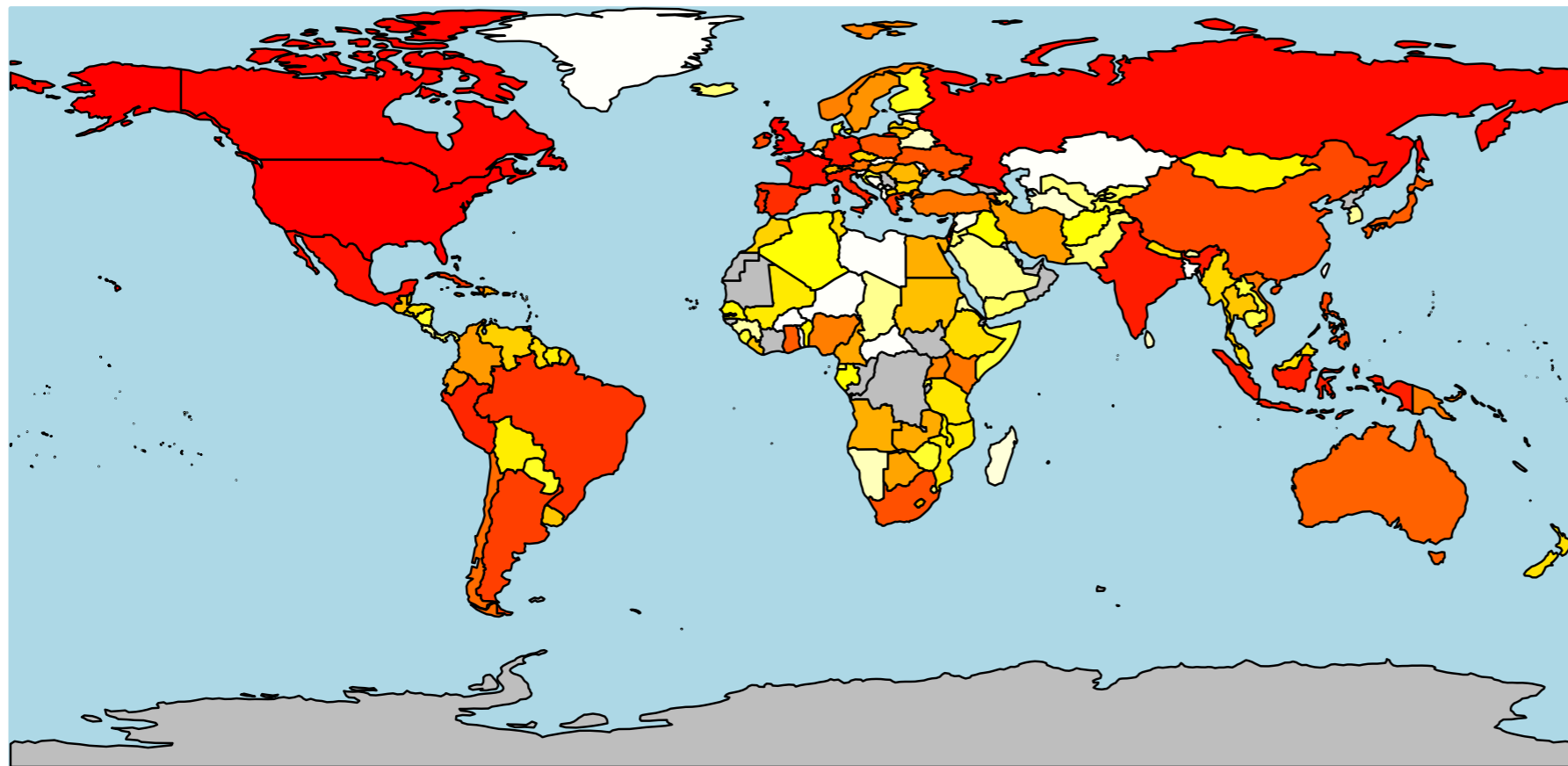
see more [+](#)



**The Dances of the World
Vol. 2:...** →
ARTIST Various Artists
YEAR 1958
CATALOG NO. FW06502
BUY Custom CD: \$16.98
Digital Download: \$9.99



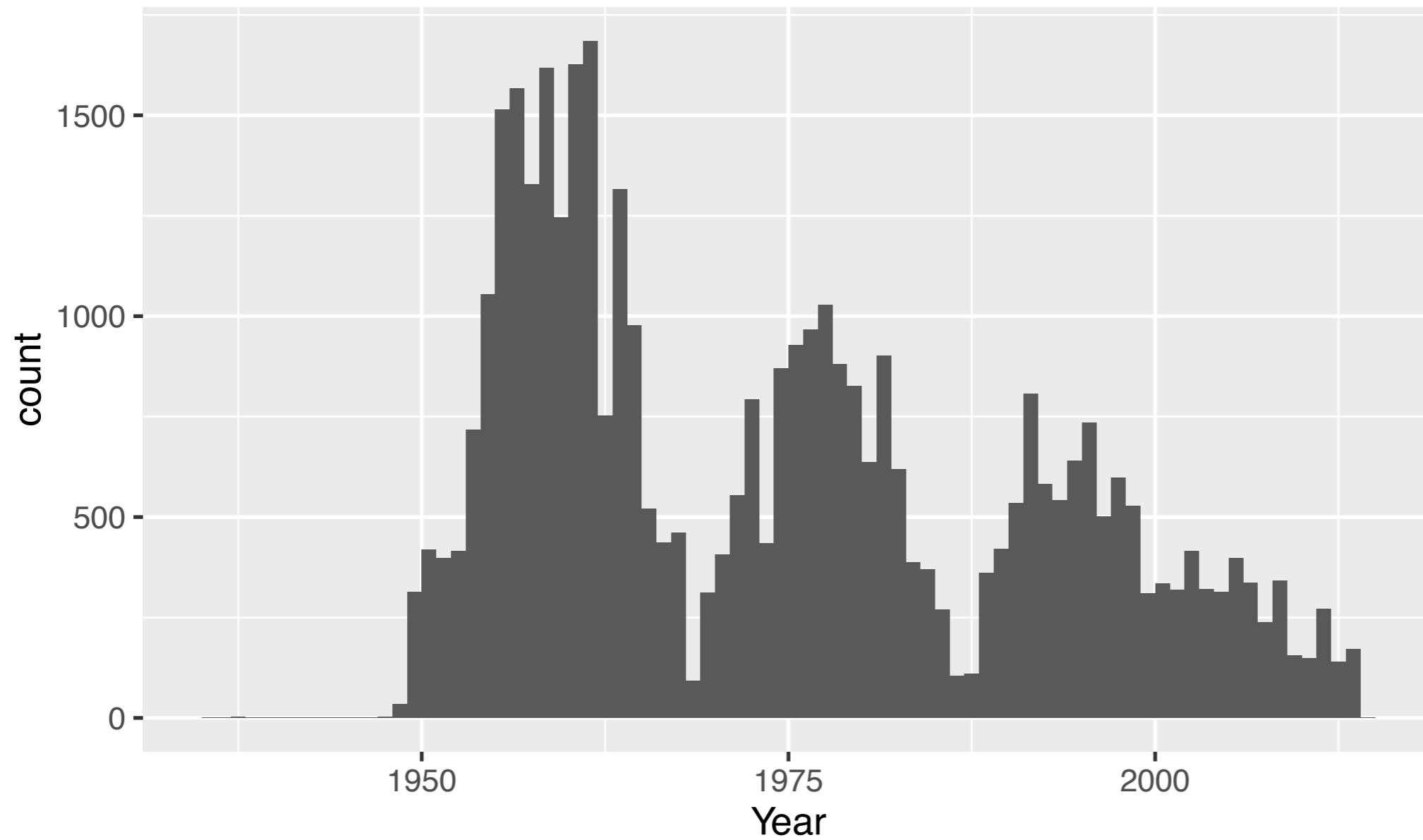
**Worlds of Sound: The
Smithsonian...** →
CATEGORY Recording Sp



1 8 17 28 43 61 84 155 246 437

DATASET

Figure: Number of recordings per country.



DATASET

Figure: Number of recordings per year.

DATASET

- 820 recordings
- 82 countries x 10 recordings
- as old as possible
- chosen at random

CONTENTS

- Dataset
- Descriptors
- Application

WHAT IS MUSIC STYLE

- style can be recognised by characteristic uses of form, texture, harmony, melody, and rhythm [2]

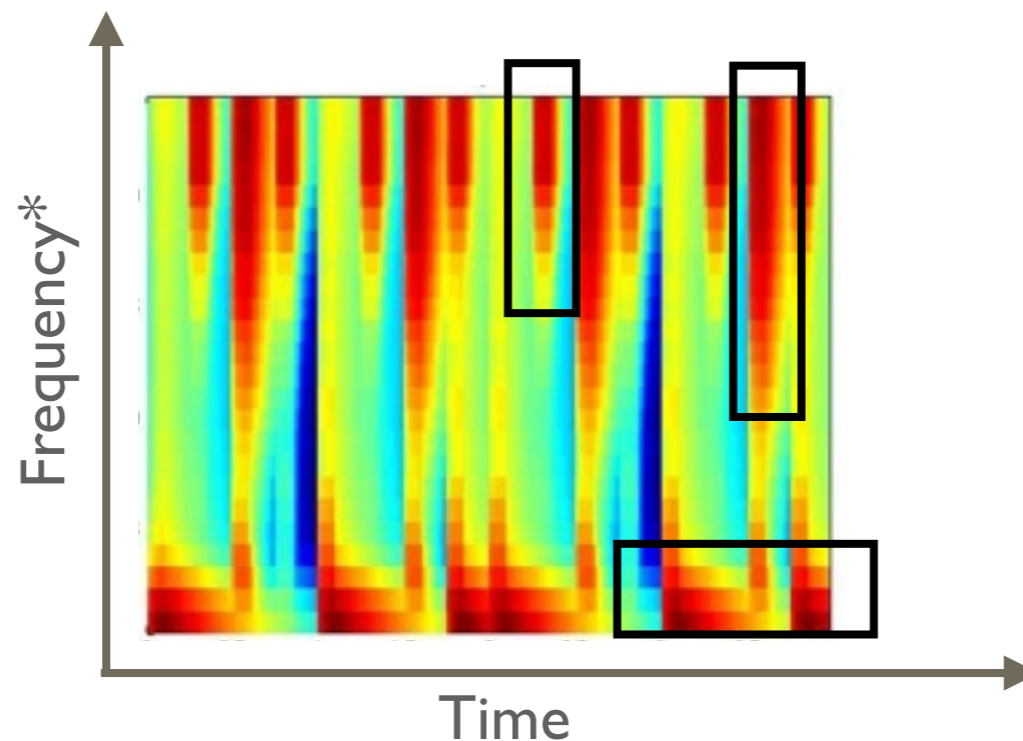
[2] S. Sadie, J. Tyrrell, and M. Levy. *The New Grove Dictionary of Music and Musicians*. Oxford University Press, 2001.

STYLE DESCRIPTORS

- Timbre, Rhythm, Melody, Harmony, (Form: ignored for now)
- from the field of Music Information Retrieval (MIR)
 - Low-level: more abstract representations but likely robust to the diversity of music style characteristics

STYLE DESCRIPTORS

- Timbre: Mel Frequency Cepstrum Coefficients (MFCC) [3]

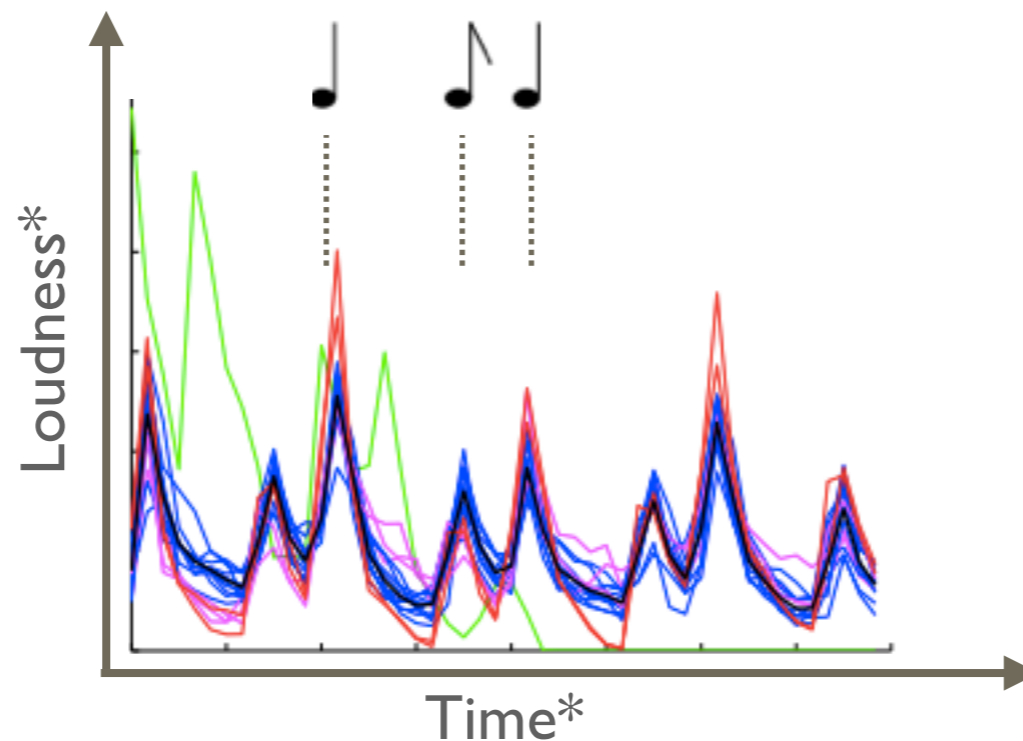


spectrogram:
bass, snare, hi-hat pattern

[3] J.J. Aucouturier, F. Pachet, and M. Sandler, “The way it sounds’: Timbre models for analysis and retrieval of music signals,” *IEEE Transactions on Multimedia*, vol. 7, no. 6, pp. 1028–1035, 2005.

STYLE DESCRIPTORS

- Rhythm: scale transform [4]

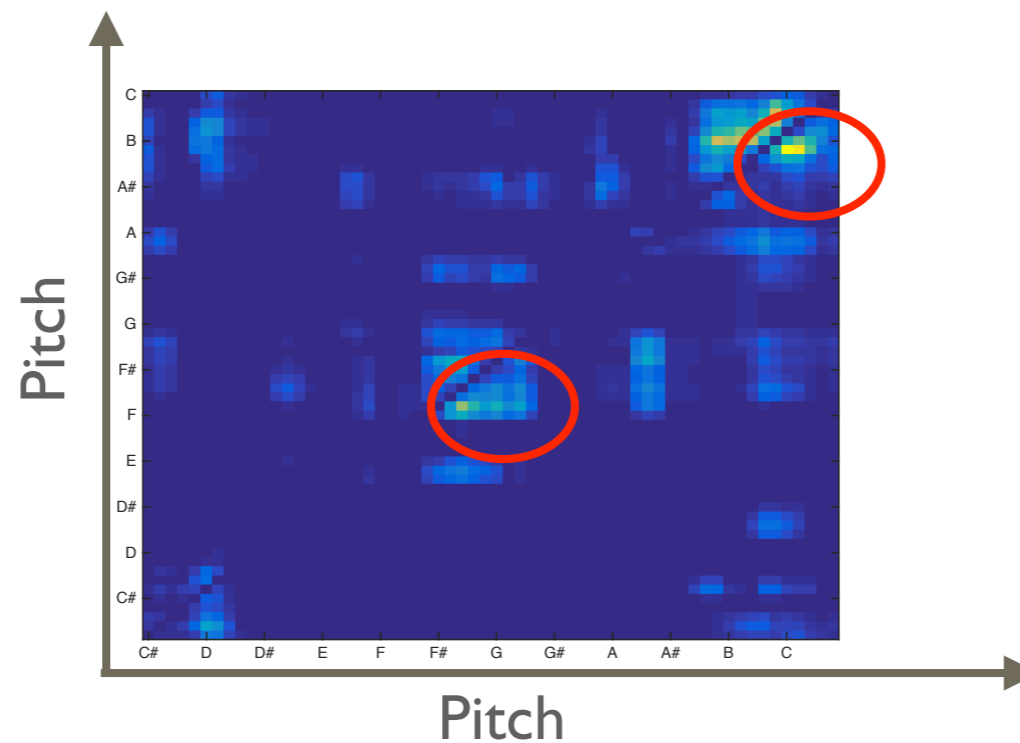


onset pattern:
swing eighth note

[4] A. Holzapfel and Y. Stylianou, “Scale Transform in Rhythmic Similarity of Music,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 1, pp. 176–185, 2011.

STYLE DESCRIPTORS

- Melody: pitch bigrams [5]

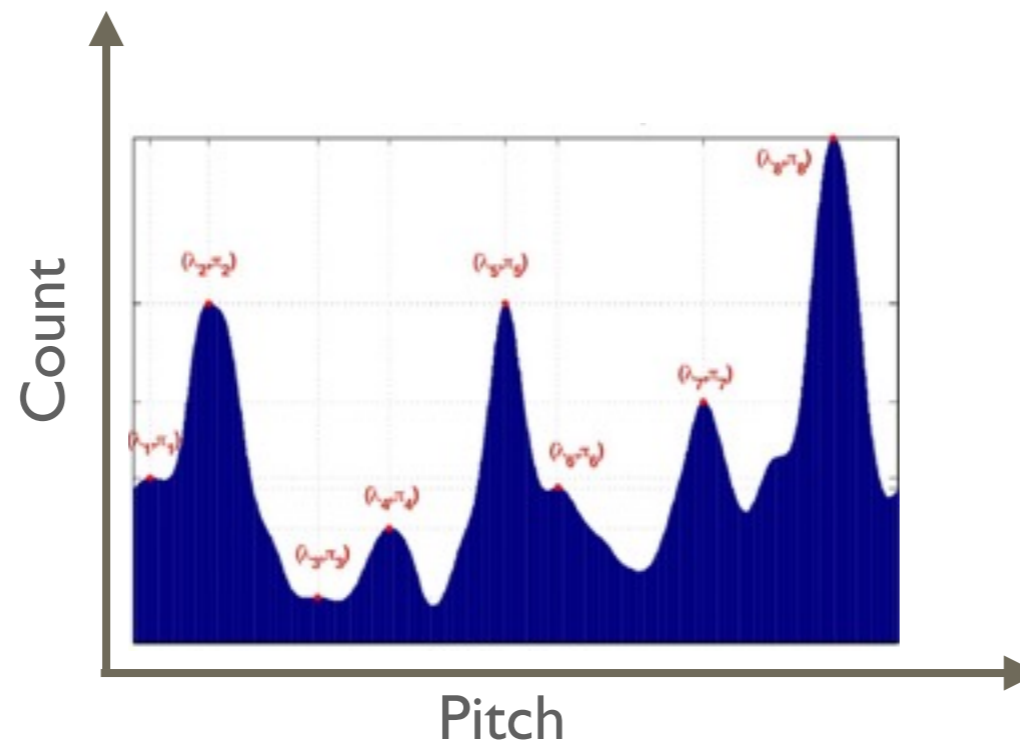


pitch-bigram matrix:
from B to C, from F to G

[5] J. Van Balen, D. Bountouridis, F. Wiering, and R. Veltkamp, "Cognition-inspired Descriptors for Scalable Cover Song Retrieval," in International Society for Music Information Retrieval Conference, pp. 379–384, 2014.

STYLE DESCRIPTORS

- Harmony: chroma averaged over time [6]



pitch histogram:
8? scale notes

[6] M. Bartsch and G. Wakefield, "Audio thumbnailing of popular music using chroma-based representations," *IEEE Transactions on Multimedia*, vol. 7, no. 1, pp. 96–104, 2005.

STYLE DESCRIPTORS

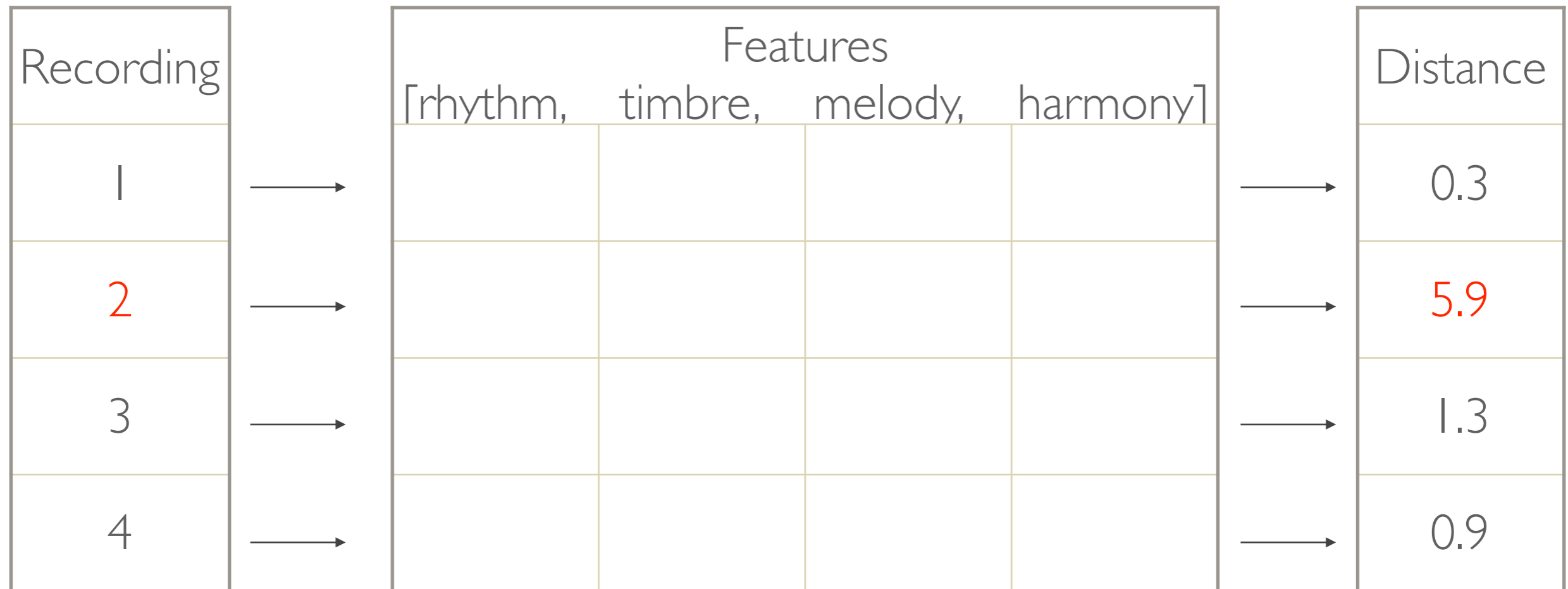


- dimensionality reduction:
 - Principal Component Analysis (PCA)
 - from 1000+ dimensions to 30

CONTENTS

- Dataset
- Descriptors
- Application

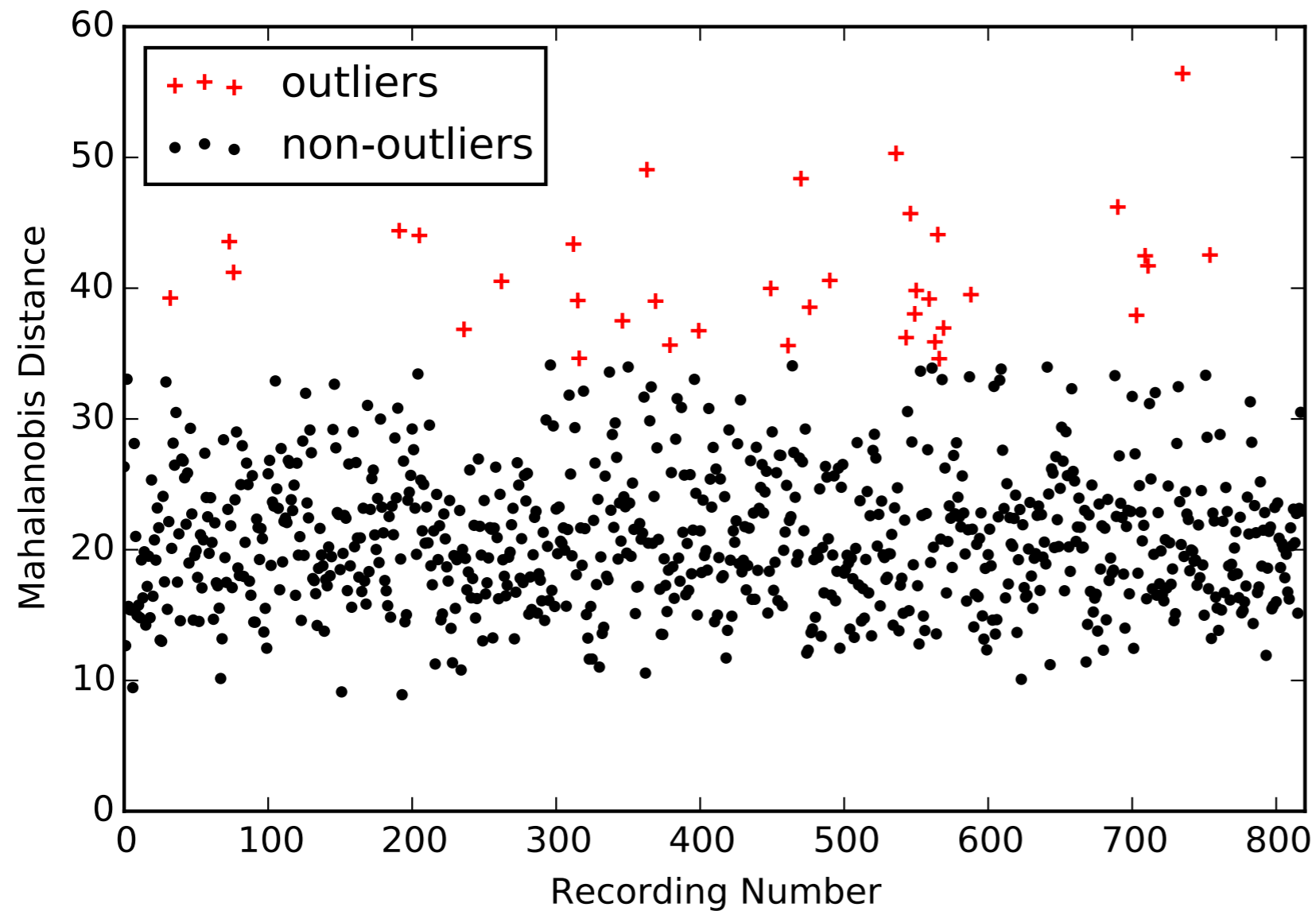
OUTLIER DETECTION



OUTLIER DETECTION

- Mahalanobis distance [7] from the set of recordings: distance to the mean of the distribution in standard deviation units
- (Minimum Covariance Determinant for robustness)

[7] C. C. Aggarwal and P. S. Yu, “Outlier detection for high dimensional data,” International Conference on Management of Data (ACM SIGMOD), pp. 37–46, 2001.



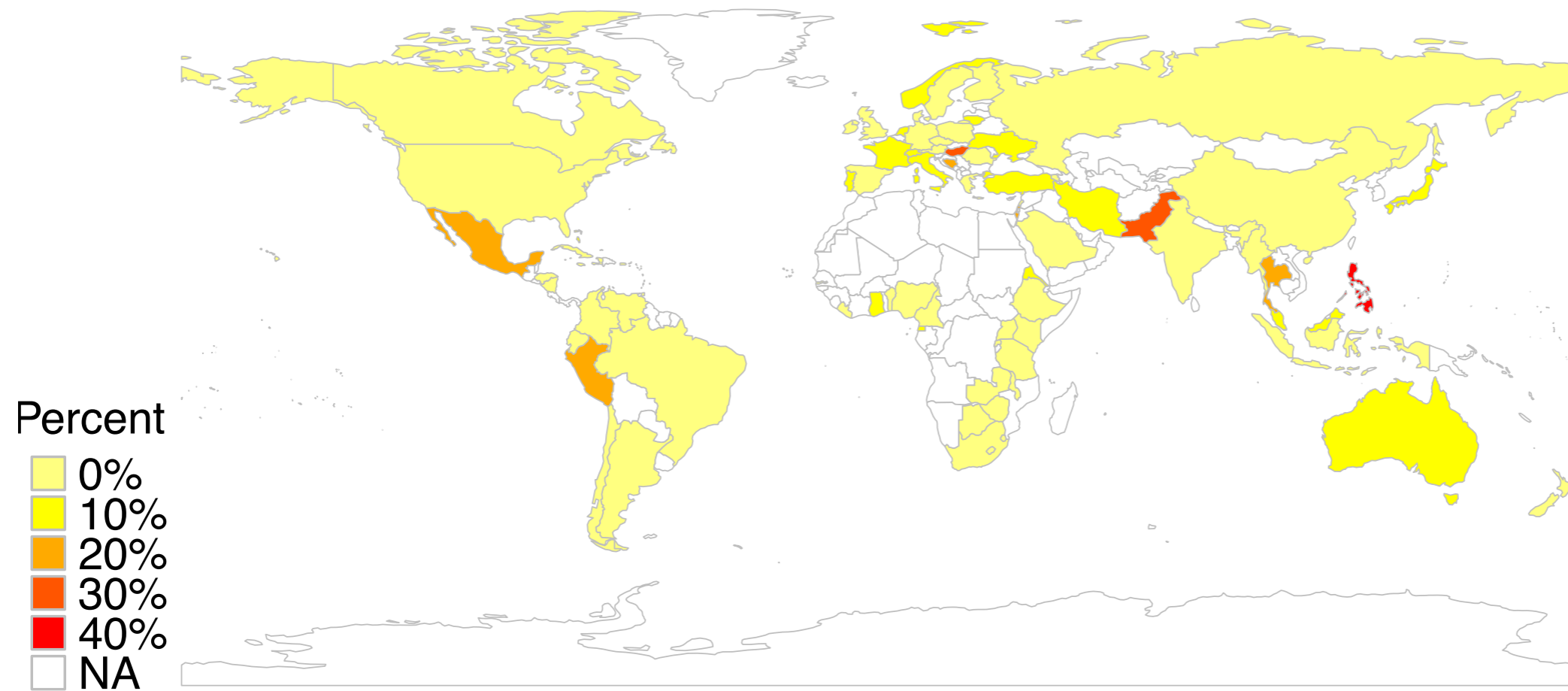
[Audio 1, Turkey](#)

[Audio2, Vietnam](#)

[Audio 3, Greece](#)

RESULTS

Figure: Mahalanobis distances and outliers at the 97.5% quantile of chi-square distribution.



RESULTS

Figure: Outliers per country for a total of 820 recordings (10 from each of 82 countries) in our sample collection.

EVALUATION

- odd one out experiment [8]

Select the odd one out

non-outlier



non-outlier



outlier



[8] D.Wolff and T.Weyde, “Adapting Metrics for Music Similarity Using Comparative Ratings,” in International Society for Music Information Retrieval Conference, pp. 73–78, 2011.

EVALUATION

- 23 participants
- 690 triad ratings
- each outlier rated min 9 and max 14 times
- 53% agreement (random baseline 33%)

CRITICAL REMARKS

- subjective judgement
- enough ratings
- reliable descriptors
- representative dataset

CONCLUSION

- characterise ‘style’ from audio recordings with MIR tools
- detect outliers in highdimensional data
- outlier detection in a small world music corpus
 - noisy examples, countries with most outliers
- moderate agreement in perceptual evaluation of outliers

Thank you!