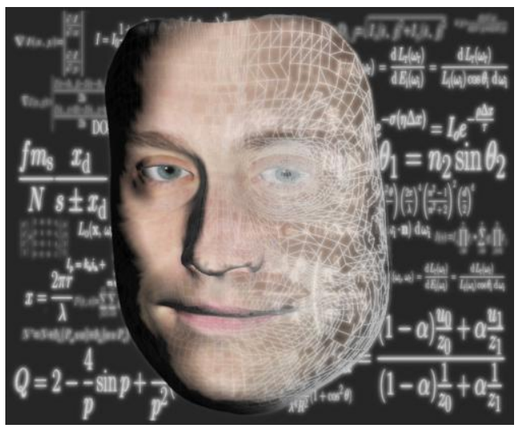


Music Instrument Classification using Traditional Methods

Natasha Shereen Benita,

Masters in Artificial Intelligence , Univ. BTU Cottbus-Senftenberg, Cottbus, Germany

1. INTRODUCTION

This study adapts a GMM-SVM method from speaker to instrument classification. Originally for speakers, it uses GMM supervectors in SVMs, showing high accuracy. For instruments, we modify it by extracting MFCCs, training GMMs per instrument, and using log-likelihood scores in SVMs. We optimize features and parameters to enhance classification accuracy.

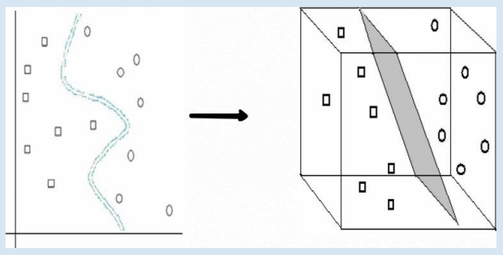
2. RESEARCH QUESTION

How can the GMM-SVM hybrid model, originally designed for speaker identification, be effectively adapted for musical instrument classification, and what improvements can be made to enhance its performance?

3. Dataset

Input:10 instrument classes
(e.g., piano, violin, guitar).
30 audio samples per instrument

4. Background



SVM discriminative classifier

$$p(\vec{x}_t|\lambda_j) = \sum_{i=1}^{M} g_i N(\vec{x}_t; \vec{\mu}_i, \sum_l)$$

GMM probabilistic model

5. METHODOLOGY

Base Paper Approach (Speaker Identification)
Extracts MFCCs as primary features. Trains GMMs per speaker to capture spectral characteristics. Constructs GMM supervectors representing the probability distribution. Uses SVM for final classification based on GMM supervectors.

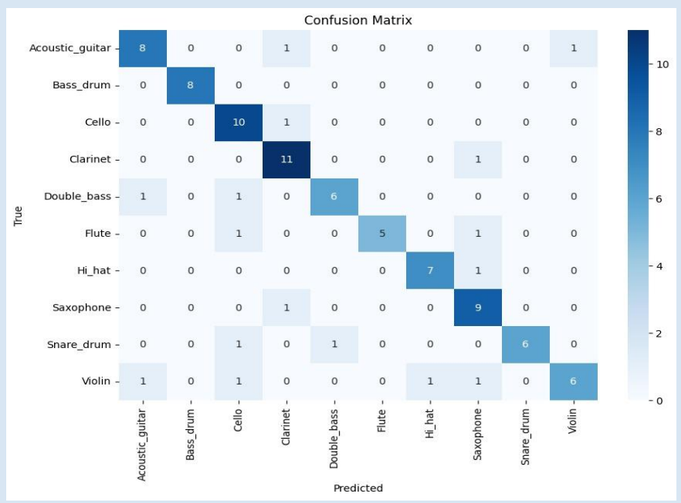
Our Approach (Instrument Classification)

✓ Similarities to Base Paper:
MFCCs as core feature representation.
GMMs trained per instrument to model spectral variations.
SVM classifier trained on GMM-derived features.

✓ Modifications & Enhancements:

- **Feature Representation**: Instead of GMM supervectors, we use log-likelihood scores as inputs to SVM. Applied dimensionality reduction (PCA/LDA) to improve feature efficiency.
- **Domain Adaptation:** Base paper focuses on speaker timbre, whereas we apply it to instrument timbre. Adjusted GMM parameters (number of components, covariance type) for instrument classification.
- **Dataset Handling:** The base paper uses a speech dataset, while we trained on a custom instrument dataset with 10 classes.

6. Results



Double Bass & Cello: feature similarities
Violin & Cello: may share tonal features

MODEL	ACCURACY	DIFFERENCE FROM BASE PAPER
BASE PAPER (GMM-SVM for Speakers)	100 %	Speaker Identification
Our GMM-SVM Approach	83 %	Instrument Classification

7. Summary

- **Cello, Violin, and Double Bass** show the most mutual confusion, likely because they are all string instruments.
- GMM-SVM is a robust approach for timbre-based classification tasks, extending beyond speech recognition into the realm of music analysis.

8. Metadata Analysis

Overview of Models and Performance
(GMM-SVM, MFCCs only) – Accuracy: 83%
(SVM with MFCC, Delta, Delta-Delta) – Accuracy: 82%
(LSTM) – Accuracy: 91%

9. Future Scope

Feature Engineering Enhancements: Adding delta, and delta-delta coefficients, spectral contrast, chroma features, or Zero Crossing Rate (ZCR).
Hybrid Models: Using an ensemble of GMM-SVM and LSTM for robust classification.
Dataset Expansion: More diverse datasets with real-world variations

References *: [1] A hybrid system based on GMM-SVM for speaker identification https://ieeexplore.ieee.org/document/7489195*

Contact :
natashas@b-tu.de