Category-level Identification of Non-registered Musical Instrument Sounds

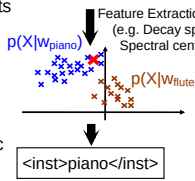
Tetsuro Kitahara, Masataka Goto** and Hiroshi G. Okuno**

*(*Graduate School of Informatics, Kyoto University, Japan, **National Institute of Advanced Industrial Science and Technology, Japan)*

1. Introduction

Musical Instrument Identification

- To obtain the names of musical instruments from sounds
- Useful for many applications including
  - Automatic Music Transcription
  - MPEG-7 based Music Annotation
  - Music Information Retrieval
  - Human-robot Communication via Music



Problem: Non-registered Instruments

Identifying musical instruments usually requires their template sounds.
Then, ***can we prepare template sounds of all instruments used in actual musical pieces?***

Because
- There are numerical kinds of musical instruments in the world

No!
- Recent development of digital audio technology has made it possible to create infinite kinds of original sounds

The mechanism for dealing with instruments that are ***not contained*** in the template sounds

non-registered instruments

Our solution: Category-level Identification

Distinguish between registered and non-registered insruments and identify the category names of the non-registered ones

- If a given sound is registered
  - ⇨ "It's a violin"
- If the sound is not registered
  - ⇨ "I don't know this, but it's a kind of strings"

(This approach would be similar to humans' feelings toward sounds that they have heard for the first time)

2. Musical Instrument Categorization for Category-level Identification

What categorization is appropriate for the category-level identification?

⇨ One that reflects the similarity of timbres

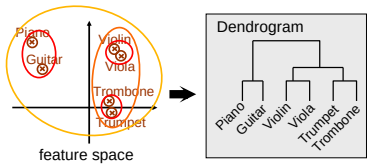
- The conventional categorization, which is based on sounding mechanisms, is not applicable, because
  - It does not satisfy the above requirement
  - It does not consider electric instrument sounds
- The categorization reflecting the timbre similarity has not been reported yet
  - ⇨ We make it from a large musical sound database

The conventional categorization

Higher-level	Middle-level	Lower-level	Instruments
Strings	-----	Struck	PF
		Plucked	CG, UK, AG
		Bowed	VN, VL, VC
Winds	Wood winds	Air reeds	PC, FL, RC
		Single reeds	SS, AS, TS, BS, CL
		Double reeds	OB, FG
	Brasses	-----	TR, TB
Percuss.	(omitted)	(omitted)	(omitted)

Basic Idea: Hierarchical Clustering

1. Let each instrument be a cluster
2. Merge the closet pair of the clusters into a single cluster
3. Repeat Step 2 until all of the instruments are merged into a single cluster



Problems of Hierarchical Clustering

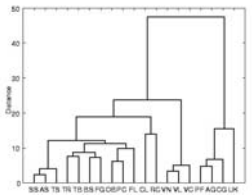
1. Results depend on a feature space
2. If one sound is used for each instrument, the result will not be robust

Our solution

1. Use the same feature space for both idetification and clustering
2. Perform clustering on multivariate normal distributions of instruments, which are obtained from a large number of sounds

Actual Acquisition of Categorization

- **Database:** An excerpt from RWC-MDB-I-2001 (19 instrs., 6,247 solo tones, Normal artic. only)
- **Feature space:** What we proposed in ICASSP '03 (18 dim. obtained with PCA&LDA from 129 features incl. Spectral Centroid and Decay Speed)



2 Major Differences
- Decayed and Sustained
- Sax. and Clarinet (conical, cylindrical)

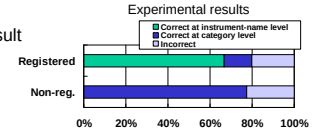
The Categorization obtained by our method

Higher-level	Middle-level	Lower-level	Instruments
Decayed	-----	Ukulele	UK
		Others	PF, CG, AG
Sustained	Strings	-----	VN, VL, VC
	Woods	Saxophones	SS, AS, TS
		Clarinet	CL
		Recorder	RC
		Brasses, etc.	TR, TB, BS, FG
		Others	OB, PC, FL

3. Identification of Non-registered Instruments

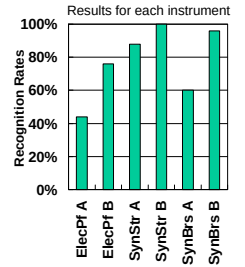
Method

1. Identify a given sound at instrument-name level
2. Calculate the Mahalanobis distance from the sound to the distribution of the above result
3. If the distance is less than a threshold, output the instrument-name result
4. If the distance is not, re-identify the sound at category level



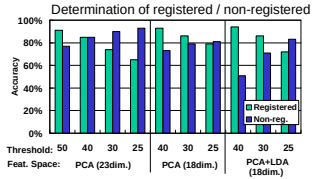
Database

- **Training data: Real instruments** (Half of 6,247 solo tones of 19 instrs.)
- **Testset for registered instruments: Real instruments** (The rest of 6,247 tones)
- **Testset for non-registered instruments: Electric sounds** (Elec. Piano, Synth Strings, Synth Brass)



Results

- Success rate of category identification: **92%**
- Success rate of reg./non-reg. determ.: **85%**
- Succcess rate of both process: **77%**
- Recognition rate using the conventional categorization: **43%**
  - ⇦ Unsuitable for electric sounds, which do not have sounding mechanisms
- Recognition rate for ElecPf A was low
  - ⇦ It was recognized as a registered instr.



4. Conclusions

- We pointed out a new problem in identifying instruments: ***non-registered instruments***
- We solved this problem by identifying the **categories** of non-registered instruments
- We automatically constructed musical instrument categorization for this identification
- Experimental results show that **77%** of non-registered sounds were identified
- Future work will include evaluation on mixtures of sounds and real musical pieces

The 2004 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (17th-21st May 2004 in Montreal, Canada)