Acoustic-similarity-based Musical Instrument Hierarchy and Its Application to Musical Instrument Identification

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

***Grad. Sch'l of Informatics, Kyoto Univ.**
****Nat'l Inst. Adv. Ind. Sci. & Tech.**

ISMA 2004 (March/April 2004 in Nara)

Motivation and Aim

Timbre:

One of the basic parameters of sounds, but it is not known well

What physical features are corresponded to timbres?
How can we feel the (un)similarity of timbres?

Taxonomy:

Study on making a hierarchy of something for understanding it

Making a hierarchy of timbres will help us to understand timbres

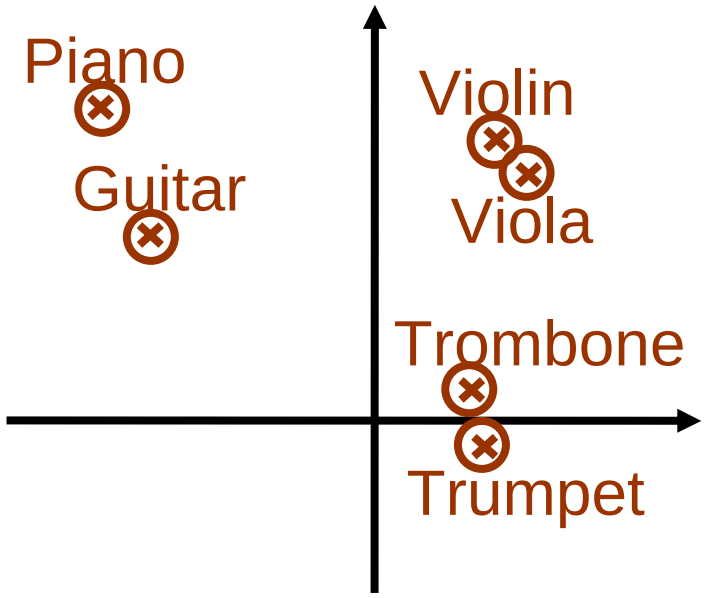
Two Strategies for Making a Hierarchy of Timbres

- **Human-perception-based**
  - Studied in the field of psychoacoustics
  - but few reports of large-scale experiments
    ⇦ the burden on human subjects
- **Acoustical-similarity-based**
  - Facilitates a large-scale one
  - but no reports
    ⇦ the lack of large-scale DB

We adopt the second strategy
We report a large-scale one using RWC-MDB

Basic Idea for Making a Hierarchy: Hierarchical Clustering

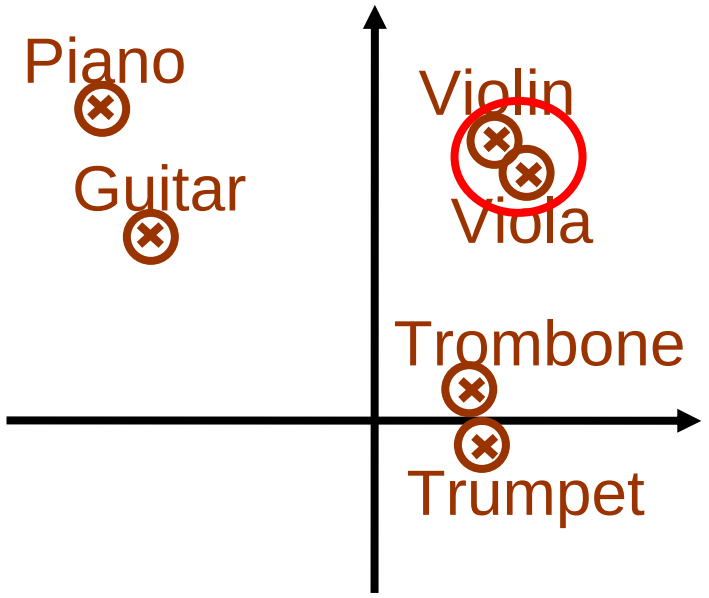
1. Let each instrument be a cluster
2. Merge the closet pair of clusters into a single cluster
3. Repeat step 2 until all of the instruments are merged into a single cluster



feature space

Basic Idea for Making a Hierarchy: Hierarchical Clustering

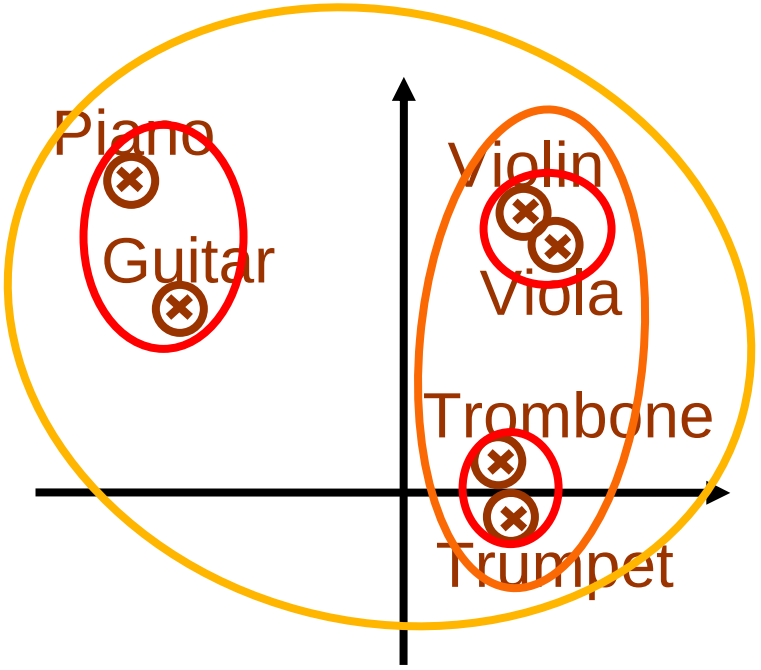
1. Let each instrument be a cluster
2. Merge the closet pair of clusters into a single cluster
3. Repeat step 2 until all of the instruments are merged into a single cluster



feature space

Basic Idea for Making a Hierarchy: Hierarchical Clustering

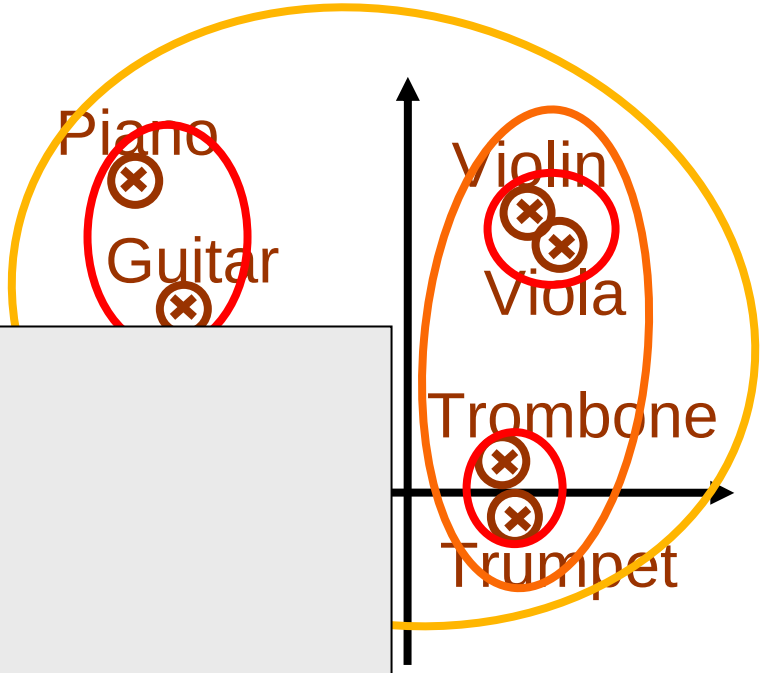
1. Let each instrument be a cluster
2. Merge the closet pair of clusters into a single cluster
3. Repeat step 2 until all of the instruments are merged into a single cluster

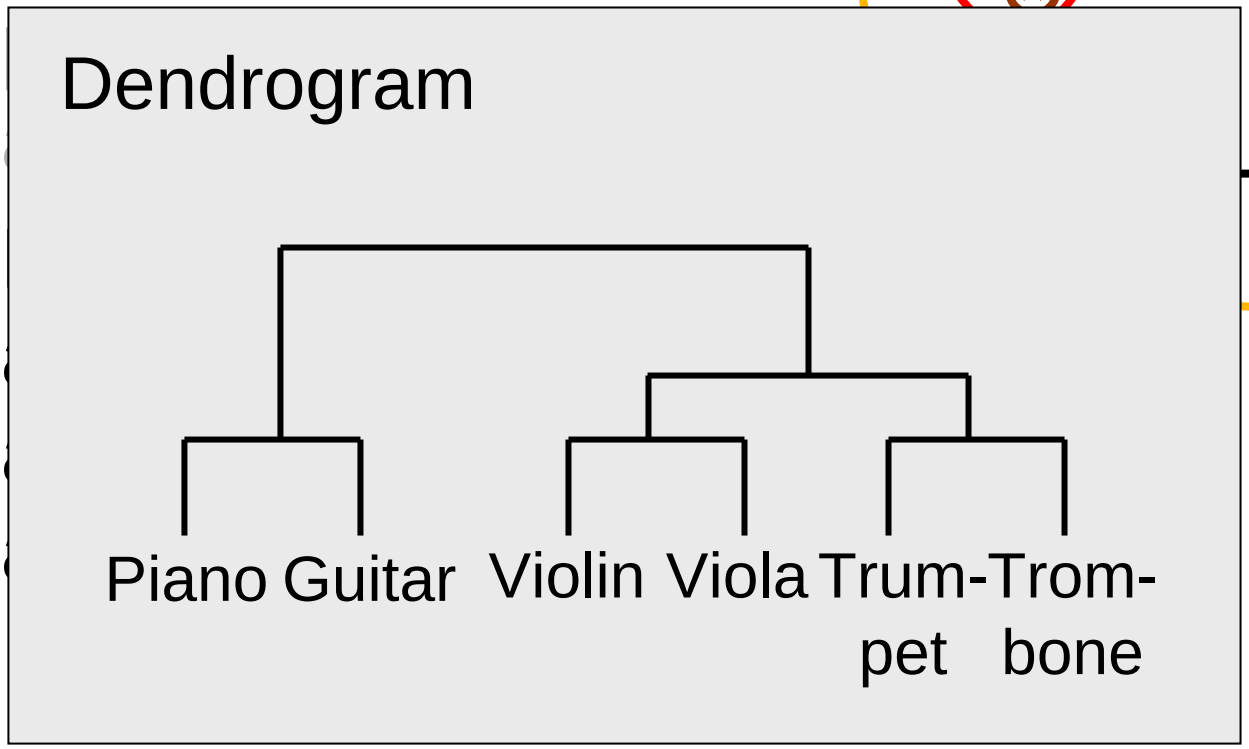


feature space

Basic Idea for Making a Hierarchy: Hierarchical Clustering

1. Let each instrument be a cluster
2. Merge the closet
3.





What is the Issue?

- **What feature space is used?**

- **How to obtain positions in the feature space**
  (They vary according to pitch, etc.)

What is the Issue?

- **What feature space is used?**
  The feature space that we previously proposed
  (Accuracy of identifying 19 instrs: **80%**)
  ⇨ It must capture the timbres of instrs well
- **How to obtain positions in the feature space**
  (They vary according to pitch, etc.)

What is the Issue?

- **What feature space is used?**

  The feature space that we previously proposed
  (Accuracy of identifying 19 instrs: **80%**)

  ⇨ It must capture the timbres of instrs well

- **How to obtain positions in the feature space**
  (They vary according to pitch, etc.)

  Approximate the distribution of each instr
  from a **large number** of sounds

  ⇨ More robust than using single sound

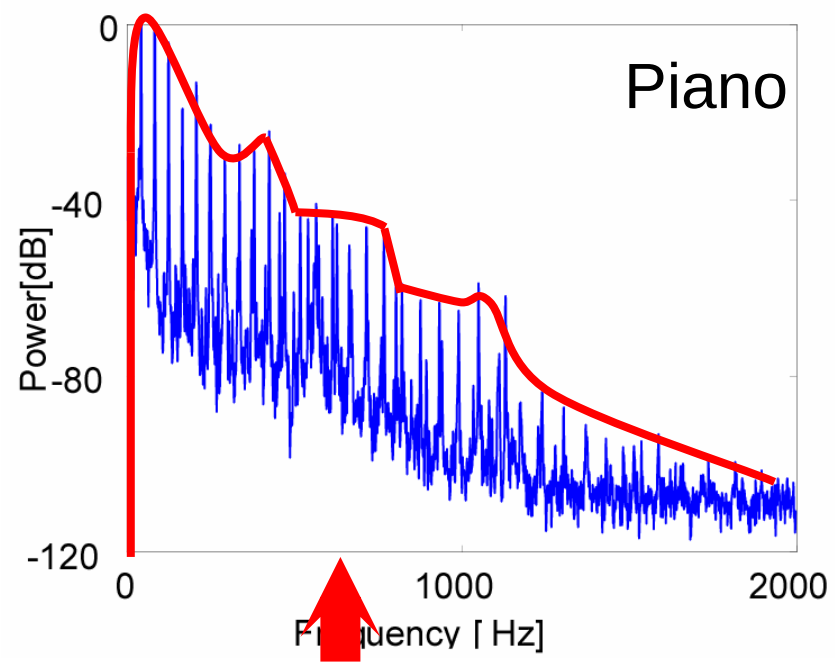
Details of the Method of Making a Hierarchy

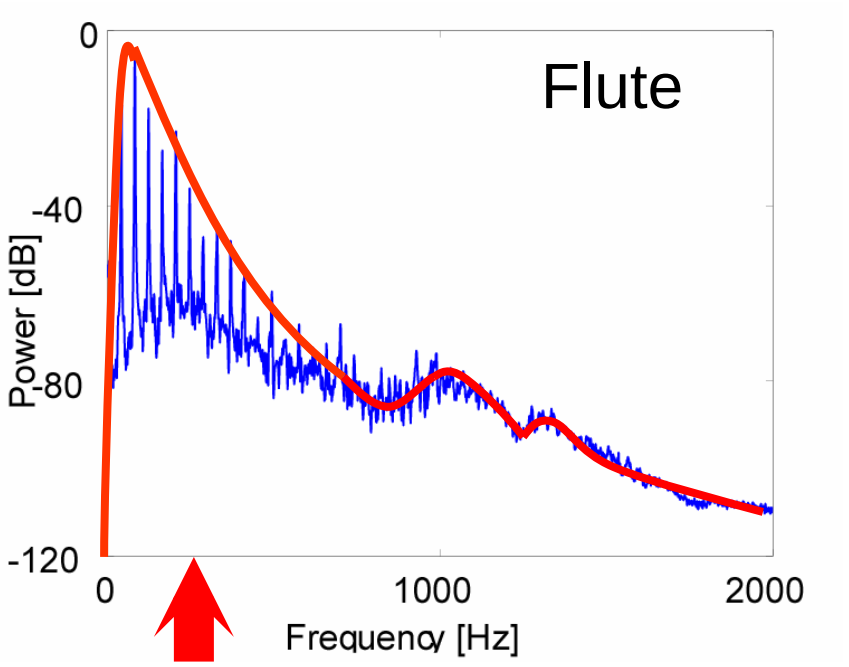
1. Feature Extraction
2. Dimensionality Reduction
3. Calculation of the Mahalanobis Distances
4. Hierarchical Clustering

Details of the Method

1. **Feature Extraction**
   The 129 features we previously proposed are extracted

e.g. Spectral centroid

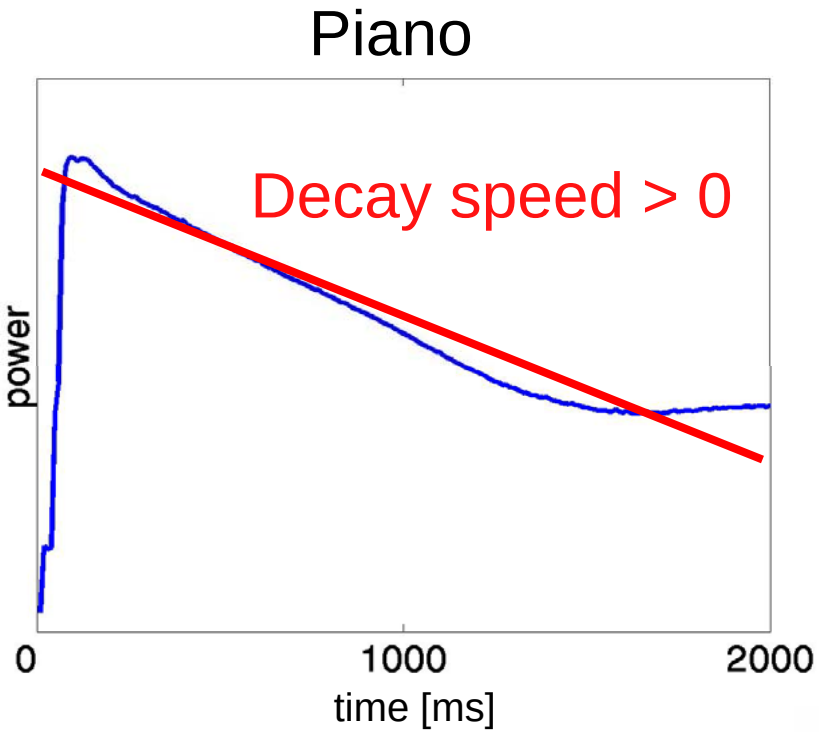


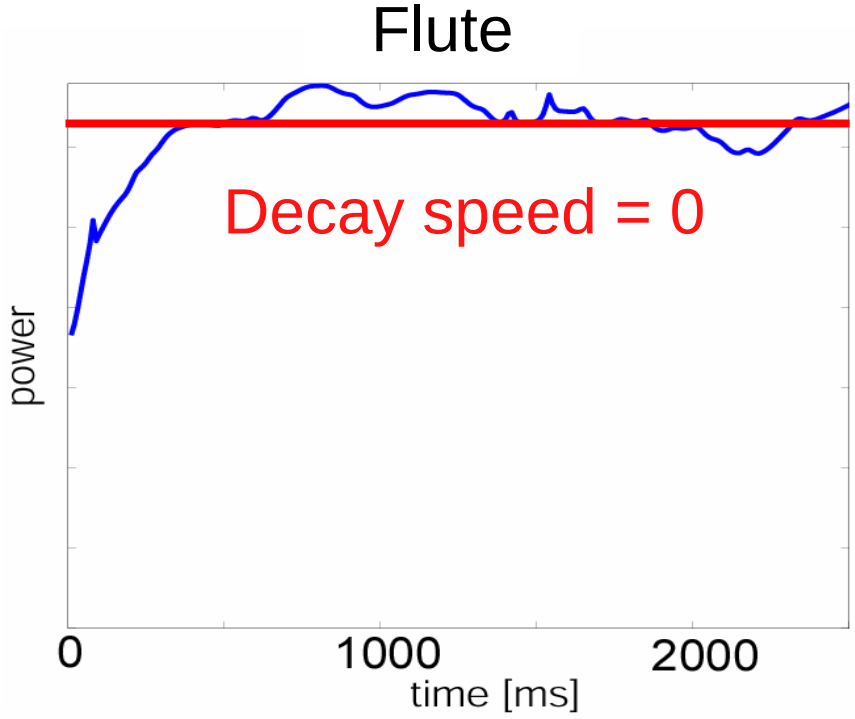


Details of the Method

1. **Feature Extraction**
   The 129 features we previously proposed are extracted

e.g. Decay speed of power





Details of the Method

2. Dimensionality Reduction

129 dim.
79 dim. } Principal Component Analysis
18 dim. } Linear Discriminant Analysis

3. Calculation of Mahalanobis Distances

(1) Calculate mean μ_i and covariance Σ_i of each instrument

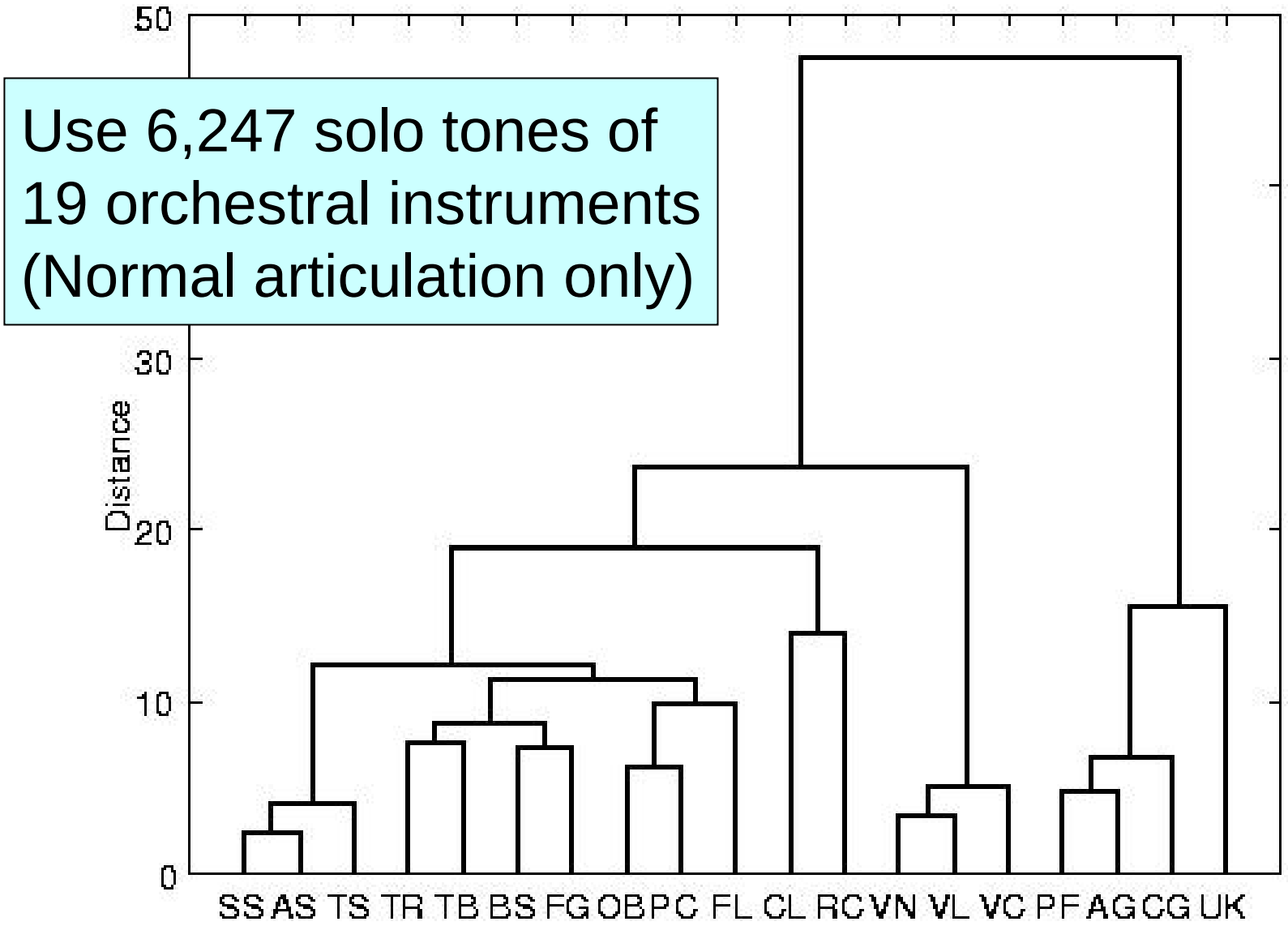
(2) Calculate the following equation:

$$D(i, j) = (\mu_j - \mu_j)' \, \Sigma^{-1} (\mu_j - \mu_j)$$

$$(\Sigma = (\Sigma_i + \Sigma_j) / 2)$$

4. Hierarchical Clustering

Results of Making a Hierarchy



Categories based on Our Hierarchy

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

Categories based on Our Hierarchy

Higher level	Middle level	Lower level	Instruments
Decayed	———	Ukulele	PF, CG, AG
		Others	UK
Sustained	Strings	———	VN, VL, VC
	Winds	Saxophones	SS, AS, TS
		Clarinet	CL
		Recorder	RC

Different from conventional ones

Matched to psychoacoustic exps

Categories based on Our Hierarchy

Higher level	Middle level	Lower level	Instruments
Decayed		Ukulele	PF, CG, AG
Sustained	Strings		VN, VL, VC
		Saxophones	SS, AS, TS
		Clarinet	CL
			FG

conical

cylindrical

Different categories although both are single reeds

Because of the timbre difference caused by the shape difference

Reed-based classification doesn't consider this difference

Application: Category-level Identification of Non-registered Instruments

Non-registered Instruments:

instrs that are *not included* in training data

- Numerous kinds of instruments
- Recent technology of making novel sounds

⇨ Impossible to prepare all of the sounds

Treating non-registered instrs is needed

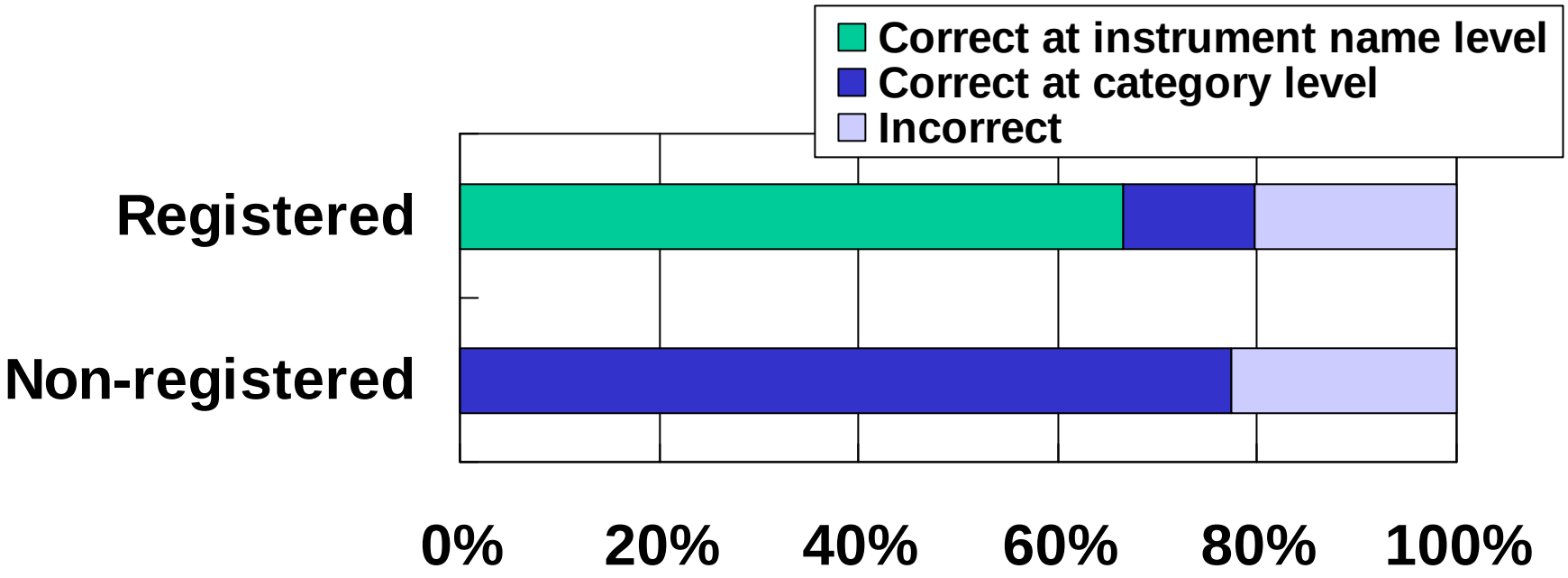
Our solution:
Identify them at category level

Application: Category-level Identification of Non-registered Instruments

- If a sound is registered
  ⇨ “It’s a violin”
- If a sound is not registered
  ⇨ “I don’t know this, but it’s a kind of strings”

For this identification,
our instrument categorization is applied

Application: Category-level Identification of Non-registered Instruments



- Training & registered test sets: **Real Instrs**
  (6,247 solo tones of 19 instrs from RWC-MDB-I-2001)
- Non-registered test set: **Electric sounds**
  (from MU2000 (Yamaha))

Conclusions

Acoustical-similarity-based Musical Instrument Hierarchy:

- Made from a large-scale musical sound DB
- Compared with psychoacoustic exps
  ⇨ It refrects timbres better

Non-registered Musical Instruments:

- Essential in musical instrument identification
- We proposed identifying categories of them
  ⇨ Our instr hierarchy was applied

Piano	Piano	
Guitars	Classical Guitar	Acoustic Guitar
	Ukulele	
Strings	Violin	Cello
	Viola	
Brass	Trumpet	Trombone
Saxophones	Soprano Sax	Tenor Sax
	Alto Sax	Baritone Sax
Double Reeds	Oboe	Faggoto
Clarinet	Clarinet	
Air Reeds	Piccolo	Recorder
	Flute	