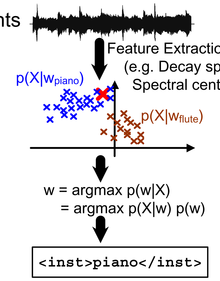
Musical Instrument Identification based on F0-dependent Multivariate Normal Distribution

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

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

1. What is musical instrument identification?

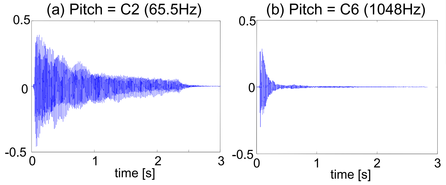
- It is to obtain the names of musical instruments from sounds (acoustical signals).
- It is a kind of pattern recognition.
- It is useful for various applications.
  e.g. automatic music transcription,
  music information retrieval,
  MPEG-7 annotation,
  human-robot interaction via music,
  and many entertainment applications
- Its research began recently (since 1990s).



2. What is difficult in musical instrument identification?

The *pitch dependency* of timbre

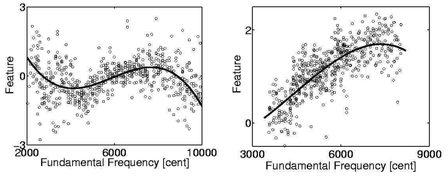
e.g. Low-pitch piano sound = Slow decay
High-pitch piano sound = Fast decay



In previous studies…
The pitch dependency of timbre was pointed out, but was NOT dealt with explicitly.

3. How is the pitch dependency coped with?

1. Approximate the pitch dependency of each feature as ***a function of fundamental frequency*** (F0).
2. Estimate feature distributions of each F0 using this function.
   ⇨F0-dependent multivariate normal distribution



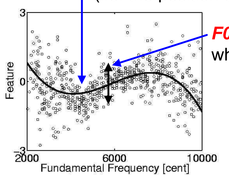
⇦ The pitch dependency of timbre and its function approximation

4. F0-dependent multivariate normal distribution

- It is a distribution for representing musical sound features depending on the pitch.
- It has following two parameters:
  - ***F0-dependent mean function***: obtained by function approximation of the pitch dependency of each feature.
  - ***F0-normalized covariance***: obtained by normalizing the F0-dependent mean.
- The ***pitch dependency*** and the ***non-pitch dependency*** of timbre can be separated by estimating these parameters.

F0-dependent mean function
which captures the pitch dependency
(i.e. the position of distributions of each F0)

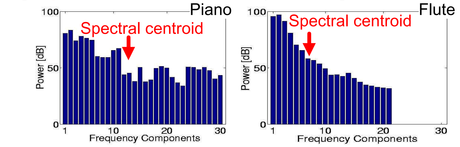
F0-normalized covariance
which captures the non-pitch dependency



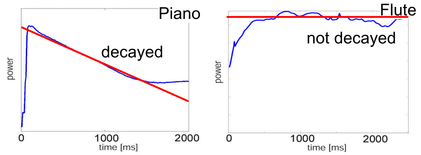
5. A musical instrument identification method using the F0-dependent multivariate normal distribution

1st step: **Feature extraction**
129 features defined based on consulting literatures are extracted.

e.g. **Spectral centroid** (which captures brightness of tones)

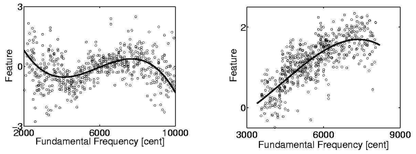


Decay speed of power

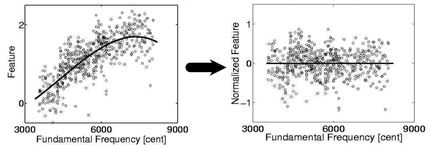


2nd step: **Dimensionality reduction**
First: **PCA** (principal component analysis)
129-dimension ⇨ 79-dimension
(with the proportion value of 99%)
Second: **LDA** (linear discriminant analysis)
79-dimension ⇨ 18-dimension

3rd step: **Parameter estimation of the F0-dependent multivariate normal distribution**
First: the **F0-dependent mean function** is approximated as a cubic polynomial.



Second: the **F0-normalized covariance** is obtained by normalizing the F0-dependent mean.
⇩
eliminating the pitch dependency



Final step: **Applying the Bayes decision rule**
The instrument w satisfying
w = argmax [log p(X|w; f) + log p(w; f)]
is determined as the result.

6. Experiments

Experimental conditions:

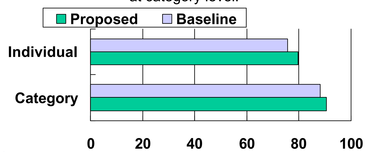
- Database: A subset of **RWC-MDB-I-2001**
  - Consists of solo tones of **19** real instruments with all pitch range.
  - Contains **3** individuals and **3** intensities for each instrument.
  - Contains normal articulation only.
  - The number of all sounds is **6,247**.
- Using the **10-fold cross validation**.
- Evaluate the performance both at individual-instrument level and at category level.

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

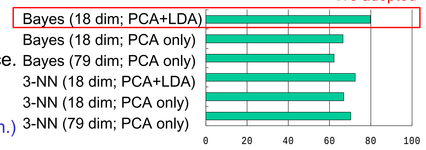
The above categorization is adopted for evaluating the performance at category level.

Experimental results (Recognition rates):

- The proposed method improved recognition rates:
  75.73%⇨**79.73%** (at individual level)
  (Error reduction rate: **16.48%**)
  88.20%⇨**90.65%** (at category level)
  (Error reduction rate: **20.67%**)
- Recognition rates of **6** instruments were improved by **more than 7%**.
- Recognition rates of the piano were best improved. (74.21% ⇨ 83.27%)
  ⇨Because the piano has the wide pitch range.
- The Bayes decision rule vs. k-NN rule
  - **PCA+LDA+Bayes** achieved the best performance.
  - **LDA** improved the performance.
  - Bayes with 79 dim. showed poor performance.
  (⇨ # of training data is not enough.)



We adopted



7. Conclusions

- To cope with the ***pitch dependency*** of timbre in musical instrument identification, the ***F0-dependent multivariate normal distribution*** is proposed.
- Experimental results of identifying 6,247 solo tones of 19 instruments show that the proposed method improved the recognition rate (75.73%⇨79.73%).
- Future works include evaluation against mixture of sounds and development of application systems using the proposed method.

The 4th IEEE Int'l Conf. on Multimedia & Expo (6th-9th July 2003 in Baltimore, MD, USA)