# Musical Instrument Identification Based on Unreliable Features

○Jun Wu, Stanisław Andrzej Raczyński, Takuya Nishimoto,
Nobutaka Ono and Shigeki Sagayama (The University of Tokyo)

## 1 Introduction

Music Instrument Identification is an important and difficult research topic in Music Information Retrieval (MIR). It involves both F0 estimation and the identification of each estimated F0 to a specific instrument. An efficient multipitch estimation algorithm was proposed in [1] and some efficient features were derived in our previous research [2] for instrument identification. However, there are often some cases that parts of the learning data are unreliable in real applications while no algorithm was proposed to solve this problem. In this paper, we applied a weighted PCA approach to deal with the unreliable features and build a consistent representation to cope with certain training data which contain noise, overlapped effect or other undesirable effects in musical instrument identification.

The paper organization is as follows. In Section 2, the proposed model is introduced. In section 3, the experimental results are demonstrated and compared with baseline. Finally, the conclusion is given in section 4.

## 2 Proposed model

In our previous research, we have proposed a harmonic model capable of both estimating multipitch and deriving timbre features. [2] For constructing classifier for musical instrument identification, we use the spectral envelope features, temporal features, Harmonic Temporal Timbre Energy Ratio and Harmonic Temporal Timbre Envelop Similarity after estimating multipitch from the input music signal. [2] We use the following criterion to decide whether the features are reliable or unreliable:

Denoting by k as the source index and n as the partial index, the features derived from a partial (k,n) is considered to be unreliable at a given time y if there exists another partial (k',n') such that $\min(\tau_k + (y+1)*\phi_k, \tau_{k'} + Y*\phi_{k'}) - \max(\tau_k + (y-1)*\phi_k, \tau_{k'} - \phi_{k'}) > 0$ and

$\min(\mu_k + \log n + \sigma_k, \mu_{k'} + \log n' + \sigma_{k'}) - \max(\mu_k + \log n - \sigma_k, \mu_{k'} + \log n' - \sigma_{k'}) > 0$.

Then we denote by $o_{kny}$ the weight of partial (k,n) at time y (0 if overlapped and 1 otherwise), weight each feature by the mean weight of the corresponding partial power coefficients.

After the features with weights are derived, weighted PCA is used as classifier. The goal of using weighted PCA is to the approximate the [original feature matrix] by the product of a [basis matrix] and a [reduced dimension matrix]. The Principal component analysis (PCA) is often used for model building. However, original PCA has several shortcomings such as not able to deal with unreliable features appropriately. In PCA, the basis vectors of the principal subspace, i.e., the principal directions in the input space, can be estimated by minimizing the reconstruction error of all reconstructed input vectors. EM algorithm to introduce weights in order to perform weighted learning or learning from incomplete data in [3]. In the weighted PCA for musical instrument identification, the weighted Euclidean criterion is followed:

$$\xi = \sum_{i=1}^{M} \sum_{j=1}^{N} \omega_{ij} (\hat{x}_{ij} - \sum_{p=1}^{k} u_{ip} a_{pj})^2 \qquad (1)$$

In the training stage we use all the single notes and build for each instrument category an x-dimensional subspace. In the testing stage we project every test note into each x-dimensional eigenspace. For every testing note the smallest reconstruction error is obtained when the note is projected into the eigenspace representing the correct category.

* 不確定特徴に基づく楽器認識、呉軍、ラチンスキ・スタニスワフ、西本卓也、小野順貴、嵯峨山茂樹(東大情報理工).

Input: data matrix X, weight matrix W, number of principal axes to be estimated k.

Output: weighted mean vector μ, U spanning principal subspace.

E-step: $a_j = \left((.\sqrt{w_j} l_{1\times k}) \circ U\right)^{\dagger} \left(.\sqrt{w_j} \circ \widehat{x_j}\right), j = 1 \dots N.$ (2)

M-step: $u_{i:} = (.\sqrt{w_j} \circ \hat{x}_{i:})((l_{k\times 1}.\sqrt{w_{i:}}) \circ \left((U^T U)^{-1} U^T \widehat{X})^{\dagger}, i = 1 \dots M.\right.$ (3)

Subscript $x_i$ denotes the ith column vector in the matrix X, while $x_{i:}$ denotes the ith row vector in the matrix X. $.\sqrt{A}$ is an operator that calculates the square root of each element of the matrix A. $A \circ B$ denotes the Hadamard (entrywise) product between two matrices of equal dimension.

## 3 Experiment

To evaluate the proposed algorithm, we did the experiments with the music notes chosen from the RWC music database [4]. Since the RWC database also includes the MIDI files associated with each real-performed music signal data, we will evaluate the accuracy by comparing the estimated fundamental frequency and the MIDI files. The accuracy for instrument identification experiment is the multiplication of the accuracy for F0 estimation and the accuracy for identifying each pitch to corresponding instrument.

271 music signal pieces (including 6 instruments: 32 altosax pieces, 36 guitar pieces, 88 piano pieces, 45 violin pieces, 36 flute pieces and 34 oboe pieces) chosen from the RWC music database [4]. 70% of the signal pieces were selected randomly as the training data. Then the proposed model was applied to generate the training features. The testing data was selected randomly from the rest 30% music pieces and mixed randomly to generate new polyphonic signals. In Table 1, the proposed algorithm was compared with the algorithm in [2] as baseline. The proposed algorithm outperforms baseline by 3.7% for 2 instruments task, 6.2% for 3 instruments task and 7.6% for 4 instruments task.

Table 1. Accuracy of musical instrument identification algorithms

| | 2 Instruments (%) | 3 Instruments (%) | 4 instruments (%) |
|---|---|---|---|
| Baseline | 74.8 | 60 | 50.7 |
| Proposed | 78.5 | 66.2 | 58.3 |

## 4 Conclusion

Based on the previous research of using a parametric model for both estimating multipitch and identifying each pitch to a specific instrument, this paper applied a weighted PCA algorithm for musical instrument identification based on unreliable features. The proposed algorithm was intuitive and efficient for solving the musical instrument identification problem, which was proved by the experiments.

## References

[1] J. Wu, Y. Kitano, T. Nishimoto, N. Ono, S. Sagayama, "Flexible Harmonic Temporal Structure for modeling musical instrument," ICEC, Seoul, South Korea, Sep., 2010.

[2] J. Wu, Y. Kitano, S. Raczynski, S. Miyabe, T. Nishimoto, N. Ono, S. Sagayama, "Musical Instrumental Identification Based on Harmonic Temporal Timbre Features," in Proceedings of Workshop on Statistical and Perceptual Audition (SAPA), pp. 7-12, 2010.

[3] D. Skocaj , A. Leonardis, "Weighted and Robust Incremental Method for Subspace Learning", Proceedings of the Ninth IEEE International Conference on Computer Vision, p.1494, October 13-16, 2003.

[4] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, “RWC music database: Popular, classical, and jazz music database,” in Proc. ISMIR, pp. 287–288, Paris, Oct, 2002.