

Ligges, Uwe; Krey, Sebastian

Feature clustering for instrument classification. (English) Zbl 1304.65052
Comput. Stat. 26, No. 2, 279-291 (2011).

Summary: We propose a method that allows for instrument classification from a piece of sound. Features are derived from a pre-filtered time series divided into small windows. Afterwards, features from the (transformed) spectrum, Perceptive Linear Prediction (PLP), and Mel Frequency Cepstral Coefficients (MFCCs) as known from speech processing are selected. As a clustering method, k-means is applied yielding a reduced number of features for the classification task. A SVM classifier using a polynomial kernel yields good results. The accuracy is very convincing given a misclassification error of roughly 19% for 59 different classes of instruments. As expected, misclassification error is smaller for a problem with less classes. The `rastamat` library (Ellis in PLP and RASTA (and MFCC, and inversion) in Matlab. <http://www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/>, online web resource, 2005) functionality has been ported from Matlab to R. This means feature extraction as known from speech processing is now easily available from the statistical programming language R. This software has been used on a cluster of machines for the computer intensive evaluation of the proposed method.

MSC:

65C60 Computational problems in statistics (MSC2010)

Cited in 1 Document

Keywords:

feature clustering; SVM; classification; timbre; instrument; music

Software:

Auditory Toolbox; ElemStatLearn; FNN; Kernlab; LIBSVM; MASS (R); Matlab; mlr; R; randomForest; rastamat

Full Text: [DOI](#)

References:

- [1] Bischl B, Wornowizki M, Borg K (2009) The mlr package: machine learning in R. <http://www.algorithm-forge.com/bischl/mlr/>
- [2] Davis SB, Mermelstein P (1980) Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans Acous Speech Signal Process ASSP* 28(4): 357–366 · [doi:10.1109/TASSP.1980.1163420](https://doi.org/10.1109/TASSP.1980.1163420)
- [3] Ellis DPW (2005) PLP and RASTA (and MFCC, and inversion) in Matlab. <http://www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/>, online web resource
- [4] Halkidi M, Batistakis Y, Vazirgiannis M (2001) On clustering validation techniques. *J Intell Inf Syst* 17(2–3): 107–145 · [Zbl 0998.68154 · doi:10.1023/A:1012801612483](https://doi.org/10.1023/A:1012801612483)
- [5] Hastie TJ, Tibshirani RJ, Friedman J (2001) *The elements of statistical learning. Data mining inference and prediction.* Springer, New York · [Zbl 0973.62007](https://doi.org/10.1007/978-1-4939-9792-3)
- [6] Hermansky H (1990) Perceptual linear predictive (PLP) analysis of speech. *J Acoust Soc Am* 87(4): 1738–1752 · [doi:10.1121/1.399423](https://doi.org/10.1121/1.399423)
- [7] Hsu CW, Chang CC, Lin CJ (2009) A practical guide to support vector classification. National Taiwan University, Taipei, <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>
- [8] Karatzoglou A, Smola A, Hornik K, Zeileis A (2004) kernlab—an S4 package for kernel methods in R. *J Stat Softw* 11(9):1–20, <http://www.jstatsoft.org/v11/i09/>
- [9] Klapuri A, Davy M (2006) *Signal processing methods for music transcription.* Springer, New York
- [10] Krey S (2008) SVM basierte Klangklassifikation. Diplomarbeit, TU Dortmund, Dortmund
- [11] Li S (2010) FNN: Fast nearest neighbor search algorithms and applications. <http://CRAN.R-project.org/package=FNN>
- [12] Liaw A, Wiener M (2002) Classification and regression by randomforest. *R News* 2(3):18–22, <http://CRAN.R-project.org/doc/Rnews/>
- [13] Opolko F, Wapnick J (1987) McGill University master samples (CDs)
- [14] R Development Core Team (2009) *R: A language and environment for statistical computing.* Vienna, Austria, <http://www.r-project.org/>, ISBN 3-900051-07-0

- [15] Roever C (2003) Musikinstrumentenerkennung mit Hilfe der Hough-Transformation. Universität Dortmund, Fakultät Statistik, [http://www.aei.mpg.de/\\(\sim\\)chroev/publications/RoeverDiplom.pdf](http://www.aei.mpg.de/\(\sim\)chroev/publications/RoeverDiplom.pdf)
- [16] Slaney M (1998) Auditory toolbox: A MATLAB Toolbox for auditory modeling work version 2. Tech. Rep. 1998-010, [http://rvl4.ecn.purdue.edu/\\(\sim\\)malcolm/interval/1998-010/](http://rvl4.ecn.purdue.edu/\(\sim\)malcolm/interval/1998-010/)
- [17] Traummüller H (1990) Analytical expressions for the tonotopic sensory scale. *J Acoust Soc Am* 88: 97–100 · doi:10.1121/1.399849
- [18] Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer, New York, <http://www.stats.ox.ac.uk/pub/MASS4> · Zbl 1006.62003
- [19] Walker JS (1996) Fast fourier transforms, 2nd edn. CRC Press, Boca Raton
- [20] Weihs C, Reuter C, Ligges U (2005) Register classification by timbre. In: Weihs C, Gaul W (eds) *Classification: the ubiquitous challenge*. Springer, Berlin, pp 624–631
- [21] Weihs C, Szepannek G, Ligges U, Luebke K, Raabe N (2006) Local models in register classification by timbre. In: Batagelj V, Bock HH, Ferligoj A, Žiberna A (eds) *Data science and classification*. Springer, Berlin, pp 315–322
- [22] Weihs C, Ligges U, Mörchen F, Müllensiefen D (2007) Classification in music research. *Adv Data Anal Classif* 1(3): 255–291 · Zbl 1183.62109 · doi:10.1007/s11634-007-0016-x

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.