Summary « Automatic Genre Classification : Is it worth pursuing and how can it be improved ?"

- Intro : improvements are really minor recently, leading to the question of "is it still worth investigating ?" / Article says it is relevant, but more insights from the musicology and psychology fields should be taken into account in automatic genre classification research / genre classification is somewhat subjective, even amongst humans
- Qualitative aspects :
- Many other parameters are taken into account by human people, but are not strictly related to musical content : message, behavior, social context, economical and juridical context… These are not used by automatic genre classification
- It's interesting to better understand how humans classify genres, so as to build more relevant models to replicate this automatically
- It appears that humans to do not define necessary and sufficient conditions, but rather organize genres according to prototypical exemplars
- "Only limited agreement can be achieved among human annotators when classifying music by genre, and such limits impose an unavoidable ceiling on automatic genre classification performance"
- Many recordings belong to several genres to a certain degree, with some of them frequently overlapping
- Study : inter-participant genre agreement only of 76% with 30s long recordings (S. Lippens, J. P. Martens, M. Leman, B. Baets, H. Meyer, and G. Tzanetakis. "A Comparison of Human and Automatic Musical Genre Classification," in *Proceedings of the IEEE International Conference on Audio, Speech and Signal Processing,* 2004) → Is this an upper limit that cannot be surpassed by computers ?
- Genre classification is often made by album or by artist and not by individual recording
- New genres are often introduced, and previously existing genres evolve
- MIREX 2005 : 75% accuracy with 10 genres, dropping to 46% with 38 genres → Not yet practically viable
- Genre classification vs. similarity-based suggestions ? Customers tend to still prefer the first approach
- Genre allows to identify with a culture, and plays a major role in the appreciation of a track (sometimes more than the intrinsic characteristics of the record)
- More technical considerations :
- A lot of features commonly used were based on timbre, and were low-level features (!= high-level features, based on musical abstractions)
- Extract cultural features through text mining ?
- Possibility to add multiple labels to a recording ? Would be more relevant with the subjectivity of the annotator
- Music in traditionally used musical databases (such as Magnatune, Epitonic) is not representative of current commercial music
- Databases are too small to avg out "annotation noise", annotators are not always experts… As much work should be done on the algorithms as on the ground truth they rely on. + labeling by artist or album is problematic (some of them are incredibly diverse)

- Genres should be more representative, only the ~10 commonly used are insufficient, when taking into account that EDM itself can be subdivided into dozens of very different genres
- Structured classification strategies : hierarchical representation, with certain genres appearing in multiple categories. Allows for different levels of granularity (certain people may want to subdivide EDM into 30 genres, while others will label it all “techno” but need dozens of Rock subgenres)
- (Linked with hierarchical classification) certain misclassifications are less serious than others (Hard Rock classified as Heavy Metal is less serious than Hard Rock classified as Orchestral) → Trees allow to “weigh” the penalty in case of misclassification
- Certain recordings can be segmented in different genres (ex : Linkin Park can be rap or metal depending on the section of a song) → must not use only features averaged over an entire recording
- Some evolutions over time are also characteristic of certain genres (ex : 12 bar blues chords, EDM buildups…) → certain classifiers with memory can be effective (hidden Markov models, recurrent neural networks)
- Dimensionality reduction techniques, although effective in terms of pure performance, do not allow to retrieve which musical features are most useful in which context → might be preferable to use DRT that preserve the identity of the features (ex : forward-backward selection, generic algorithm-based selection)
- Humans identify genres very differently, and automatic genre classification should take that into account (a musicologist will have very different needs and appreciation of classification compared to a teenager for example)

Keywords/definitions :

- MIR : Music Information Retrieval
- Genre != style : genre is a kind of music, recognized by a community, whereas style is more linked with the personal touch of an artist/band/composer
- Ground truth : statements widely recognized and accepted as true by the community → often hard to find