THE MACHINE LEARNING AND INTELLIGENT MUSIC PROCESSING GROUP AT THE AUSTRIAN RESEARCH INSTITUTE FOR ARTIFICIAL INTELLIGENCE (ÖFAI), VIENNA

Gerhard Widmer[1,2], *Simon Dixon*[1], *Arthur Flexer*[1], *Werner Goebl*[1], *Peter Knees*[2],
Søren Tjagvad Madsen[1], *Elias Pampalk*[1], *Tim Pohle*[1], *Markus Schedl*[1,2], *Asmir Tobudic*[1]

[1]Austrian Research Institute for Artificial Intelligence
Freyung 6/6, A-1010 Vienna, Austria
[2]Department of Computational Perception,
Johannes Kepler University Linz, Austria

ABSTRACT

The report introduces our research group in Vienna and briefly sketches the major lines of research that are currently being pursued at our lab. Extensive references to publications are given, where the interested reader will find more detail.

1. ABOUT THE GROUP

The Machine Learning and Intelligent Music Processing Group at the Austrian Research Institute for Artificial Intelligence (ÖFAI) in Vienna focuses on applications of Artificial Intelligence (AI) techniques to (mostly) musical problems. Originally composed of a few researchers performing general machine learning and data mining research, it currently consists of 11 full-time researchers, 9 of whom work in projects specifically dedicated to musical topics. The group focuses on several lines of research in computational music processing, both at symbolic and audio levels. The uniting aspect is the central methodological role of Artificial Intelligence (AI), in particular, of machine learning, intelligent data analysis, and automatic machine discovery.

The group is involved in a number of national and multinational (EU) research projects (see Acknowledgments below) and performs cooperative research with many music research labs worldwide.

2. MAJOR RESEARCH LINES IN INTELLIGENT MUSIC PROCESSING

In the following, we briefly sketch the major lines of research that are currently being pursued at our lab, with extensive references to publications where the interested reader can find more detail. More information can be found at http://www.oefai.at/oefai/ml/music.

2.1. AI-Based Models of Expressive Music Performance

AI-based research on expressive performance in classical music has been a long-standing research topic at ÖFAI.

The principal goal of this work is to learn more about this complex artistic activity by measuring expressive aspects such as timing, dynamics, etc. in large numbers of recordings by famous musicians (currently: pianists) and using AI techniques — mainly from machine learning and data mining — to analyze these data and gain new insights for the field of performance research. This line of research covers a vast variety of different subtopics and research tasks including:

- data acquisition from MIDI and audio recordings (score extraction from expressive performances, score-to-performance matching, beat and tempo tracking in MIDI files and in audio data [7, 21]);

- detailed acoustic studies of the piano (analysis of the timing properties of piano actions, quality assessment of reproducing pianos like the Bösendorfer SE system or the Yamaha Disklavier) — see, e.g., [8, 10];

- automated structural music analysis (segmentation, clustering, and motivic analysis) — e.g., [1];

- studies on the perception of expressive aspects (e.g., perception of tempo [3], perception of note onset asynchronies – 'melody lead' [11], and similarity perception of expressive performances [17]);

- performance visualisation (e.g., in the form of animated two-dimensional tempo-loudness trajectories and real-time systems; a screenshot of the "Performance Worm" [6] is shown in Fig. 1);

- inductive building of computational models of expressive performance via Machine Learning (inducing rule-based peformance models from data [20], learning multi-level models of phrase-level and note-level performance [19]), and

- characterisation and automatic classification of great artists (e.g., discovering performance patterns characteristic of famous pianists [9, 13], learning to recognize performers from characteristics of their style [16]).
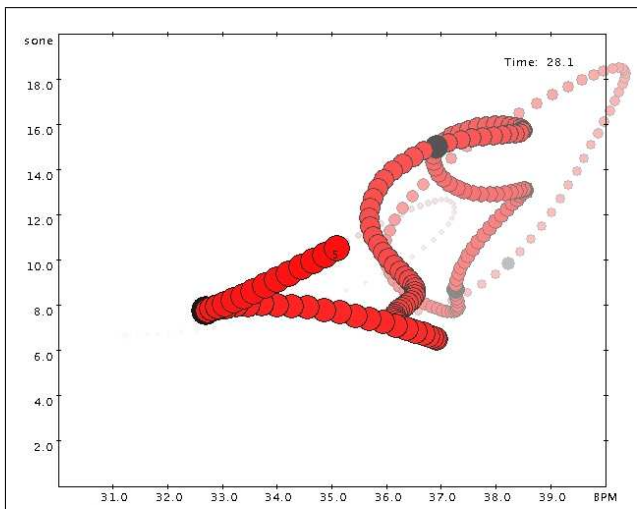
Figure 1. The 'Performance Worm'

A more comprehensive overview of this long-term basic research project can be found in [18].

2.2. Real-time Analysis and Visualisation of Musical Structure and Performance

A recent, more application-oriented research line that grew out of the above work is concerned with the development of new interfaces — in the most general sense of the word — that enable new ways of presenting, teaching, understanding, experiencing, and also shaping music in creative ways. These 'interfaces' will provide intuitive access to music through novel visualisation paradigms and new methods for interactive manipulation and control of music performances and recordings (for instance, by visualising aspects of the musical structure and performance via computer animations, or by permitting a user to interactively play with and modify a given recording according to his/her taste). That requires basic research on intelligent musical structure recognition algorithms, new visualisation paradigms, and methods for (real-time) computer-based interaction with music. These activities are currently funded by a sizeable national project. Fig. 2 shows a first result of this kind of work (see also [2]).

2.3. Music Information Retrieval

The recently established research field of Music Information Retrieval (MIR) is of growing interest for both the research community and "normal" music consumers (and, not least, for the music industry). MIR focuses on the development of computational methods for the intelligent automated handling of digital music and audio. That includes problems such as the automatic recognition and classification of musical styles, artists, instruments, etc.; intuitive interfaces for presenting and offering music via Internet; intelligent search engines that can find and propose new, interesting music based on an understanding of the taste of the individual user; and many more.

Figure 2. Gerhard Widmer controlling a piano performance via a MIDI theremin (cf. [2]).

In this area, we are currently engaged in a variety of research directions, especially in the context of the EU project SIMAC (www.semanticaudio.org). Current work comprises, among other things, the development of algorithms for extracting musically relevant high-level information and meta-data from low-level audio (e.g., [4]), methods for automatically structuring and visualising large digital music collections [14], tools for navigating in such music collections, algorithms for the automatic classification of musical genres and styles [5], and also 'personal DJs' that generate play-lists adapted to the musical taste of the individual user.

As a graphical example of our work in this area, Fig. 3 shows a screen shot of the program *Islands of Music v.2* [14], which is designed to permit a user to interactively explore the structure of music collections according to different musical criteria. The upper part of the screen shows the projection of a collection of mp3 files onto a two-dimensional visualisation plane, according to musical or sound similarity. Songs that are somehow similar should be located close to each other. Islands represent coherent areas containing similar music, snow-covered mountains are areas of particularly high density. In the lower part of the screen, various musically relevant properties of the music in different regions are shown — in this particular case, rhythmic patterns (left) and timbral aspects as described by MFCC's (Mel Frequency Cepstral Coefficients). The user can use a slider to interactively and continuously modify the relative weighting of these two factors in the similarity judgment. The resulting changes in the structure of the music collection then become immediately visible.
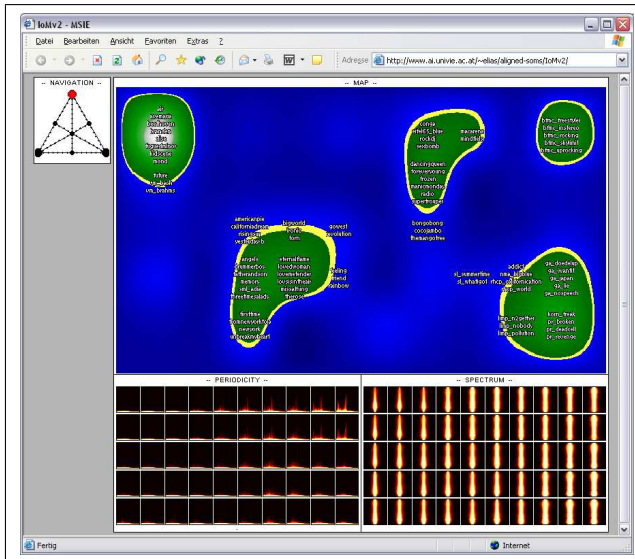
Figure 3. The 'Islands of Music' graphical metaphor for visualising the structure of digital music collections

3. RECENT HIGHLIGHTS AND FUTURE ACTIVITIES

Recent highlights include a computer program that learns to identify famous concert pianists from their playing style with high levels of accuracy (this was developed in cooperation with Prof. John Shawe-Tayler and his group at the University of Southampton, UK, and won the Best Paper Award at the European Conference on Machine Learning (ECML 2004) [15]) and the International Conference on Music Information Retrieval (ISMIR 2004) in Barcelona, where Elias Pampalk of our group won the *Genre Classification Contest*.

Also, we recently gave a public presentation to a concert audience at the *Konzerthaus*, one of Vienna's major concert halls, in the context of a concert featuring pianist Hélène Grimaud and the Cincinnati Symphony Orchestra, in which we used animations like the Performance Worm (see Fig. 1) to illustrate aspects of artistic music performance. In the future, we plan to extend these activities towards live multi-modal artistic presentations (such as performance visualisations of pianists performing live on stage). That is a definite goal within a new three-year project that is financed by the City of Vienna, which is aimed at artistic and didactic applications of music visualisation. Music visualisation will become a major research focus in the coming years, with applications both in artistic contexts and in Music Information Retrieval.

In the area of MIR, we have just started a new, nationally funded project that aims at the development of operational, multi-faceted metrics of music similarity, which will combine both music- and sound-related aspects (e.g., timbre, rhythm, etc.) and extra-musical information that can be discovered by the computer in the Internet (cultural information, lyrics, etc. [12]). We expect a large number of applications to come out of this research.

Recently, the group has also started to engage in a close cooperation with the newly founded *Department of Computational Perception* at the Johannes Kepler University of Linz, Austria (`www.cp.jku.at`). In this way, we can offer university credits and diploma (in the field of computer science) to guest students who want to join us for interesting research projects.

Acknowledgments

4. REFERENCES

[1] Cambouropoulos, E. and Widmer, G. (2001). Automatic Motivic Analysis via Melodic Clustering. *Journal of New Music Research* 29(4), 303-317.

[2] Dixon, S., Goebl, W., and Widmer, G. (2005). The "Air Worm": An Interface for Real-Time Manipulation of Expressive Music Performance. In *Proceedings of the International Computer Music Conference (ICMC'2005)*, Barcelona, Spain.

[3] Dixon, S., Goebl, W., and Cambouropoulos, E. (2005). Perceptual Smoothness of Tempo in Expressively Performed Music. *Music Perception* 23 (in press).

[4] Dixon, S., Gouyon, F., and Widmer, G. (2004). Towards Characterisation of Music via Rhyhmic Patterns. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR'04)*, Barcelona, Spain, October 10-14, 2004.

[5] Dixon, S., Pampalk, E., and Widmer, G. (2003). Classification of Dance Music by Periodicity Patterns. In *Proceedings of the Fourth International Conference on Music Information Retrieval (ISMIR 2003)*, pp. 159-165, Baltimore, MD.

[6] Dixon, S., Goebl, W., and Widmer G. (2002). The Performance Worm: Real Time Visualisation of Expression Based on Langner's Tempo-Loudness Animation. In *Proceedings of the 2002 International Computer Music Conference (ICMC'2002)*, Gothenburg, Sweden, pp. 361-364.

[7] Dixon, S. (2001). Automatic Extraction of Tempo and Beat from Expressive Performances. *Journal of New Music Research* 30(1), 39-58.

[8] Goebl, W., Bresin, R., and Galembo, A. (2004). Once Again: The Perception of Piano Touch and Tone. Can Touch Audibly Change Piano Sound Independently of Intensity? *Proceedings of the 2004 International Symposion on Music Acoustics (ISMA'04)*, Nara, Japan, pp. 332-335.

[9] Goebl, W., Pampalk, E., and Widmer, G. (2004). Exploring Expressive Performance Strajectories: Six Famous Pianists Play Six Chopin Pieces. In *Proceedings of the 8th International Conference on Music Perception and Cognition (ICMPC8)*, Evanston, IL. Causal Productions, Adelaide.

[10] Goebl, W. and Bresin, R. (2003). Measurement and Reproduction Accuracy of Computer-controlled Grand Pianos. *Journal of the Acoustical Society of America* 114(4), 2273-2283.

[11] Goebl W. (2001). Melody Lead in Piano Performance: Expressive Device or Artifact?, *Journal of the Acoustical Society of America* 110(1), 563-572.

[12] Knees, P., Pampalk, E., and Widmer, G. (2004). Artist Classification with Web-based Data. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR'04)*, Barcelona, Spain, October 10-14, 2004.

[13] Madsen, S.T. and Widmer, G. (2005). Exploring Similarities in Music Performances with an Evolutionary Algorithm. *Proceedings of the 18th International FLAIRS Conference*, Clearwater Beach, Florida.

[14] Pampalk, E., Dixon, S., and Widmer, G. (2004). Exploring Music Collections by Browsing Different Views. *Computer Music Journal* 28(2), 49-62.

[15] Saunders, C., Hardoon, D., Shawe-Taylor, J., and Widmer G. (2004). Using String Kernels to Identify Famous Performers from their Playing Style. In *Proceedings of the 15th European Conference on Machine Learning (ECML'2004)*, Pisa, Italy.

[16] Stamatatos, E. and Widmer, G. (2005). Automatic Identification of Music Performers with Learning Ensembles. *Artificial Intelligence* 165(1), 37-56.

[17] Timmers, R. (2005). Predicting the Similarity Between Expressive Performances of Music from Measurements of Tempo and Dynamics. *Journal of the Acoustical Society of America* 117(1), 391–399.

[18] Widmer, G., Dixon, S., Goebl, W., Pampalk, E., and Tobudic A. (2003). In Search of the Horowitz Factor. *AI Magazine* 24(3), 111-130.

[19] Widmer, G. and Tobudic, A. (2003). Playing Mozart by Analogy: Learning Multi-level Timing and Dynamics Strategies. *Journal of New Music Research* 32(3), 259-268.

[20] Widmer, G. (2002). Machine Discoveries: A Few Simple, Robust Local Expression Principles. *Journal of New Music Research* 31(1), 37-50.

[21] Widmer, G. (2001). Using AI and Machine Learning to Study Expressive Music Performance: Project Survey and First Report. *AI Communications* 14(3), 149-162.