Artificial Neural Networks Based Models for Automatic Performance of Musical Scores

Author: Roberto Bresin

Presented by Merrick Mosst
ISE 599, Spring 2006

Contents

Author's Background

- Roberto Bresin
  - Researcher in Music Acoustics at KTH
  - Earned his PhD in Music Acoustics
  - Current projects
    - Control models for sound synthesis (The Sounding Object - SOb)
    - Feedback Learning in Musical Expression (FEEL-ME)
    - Analysis and synthesis of emotional content in music (Multisensory Expressive Gesture Applications - MEGA)
    - * Articulation strategies in expressive music performance
    - * Emotional melodies for mobile phones
    - * Synthesis of Emotional Expression in Music Performance

Introduction

- Previous research
  - Seashore's piano performance research
  - Bruno Repp's quantitative analysis
  - Clynes' composer's pulses
  - De Poli, Irone, Vidolin's rule systems
  - Friberg, Fryden, Sundberg's rule system (KTH)
- KTH Rule System
  - Analysis-by-synthesis development
  - Additive rules
  - Tuned by professional musician
  - Based on "conscious decisions"

Approach

- Cognitive model of performer
  - Analyzes structure of composition
    - Structural (global) model
    - Symbolic rules logic appropriate
  - Focuses on local level performance
    - Local models
    - Non-symbolic rules more appropriate
  - → two time spaces, each of which involves conscious & unconscious deviations
- Hybrid system: rules for each time space
  - Symbolic
    - Decision rules
    - Performing rules (subset of KTH rules)
  - Sub-symbolic
    - Artificial Neural Networks (ANNs) controlled by decision rules

Approach

- Artificial Neural Networks
  - Properties
    - Can represent non-linear functions
    - Can be good non-linear interpolators (generalization)
    - Sensitive to architecture (no. of inputs, layers, outputs)
    - Sensitive to training (amount of data, prior knowledge, etc)
  - Applicability in hybrid model
    - Local model: non-linear function g() w/short context
      - $y_n = g(\underline{k}, \underline{x}_n)$
      - $\underline{x}_n$: parameters applied to n-th note
      - $\underline{k}$: vector of constants or variables that slowly evolve
    - With proper training, an ANN can achieve this
      - $y'_n = net(\underline{k}, \underline{x}'_n)$
      - $\underline{x}'_n$: x_n above & new parameters to teach rules
      - $\underline{k}$: allows user to weight ANN deviations

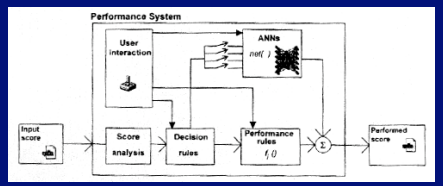
Approach

- KTH model
  - Summation of many rules
    - $y^N_n = \_i k_i * f_i(\underline{x}_n)$
    - $f_i()$: function for i-th rule
    - $\underline{x}_n$: vector of KTH rules parameters for n-th note
    - k_i: constant used to emphasize deviation
- Hybrid System
  - Deemed necessary after experimentation
  - Obtained by summing the deviations generated by a subset of S of the N KTH rules above

Approach

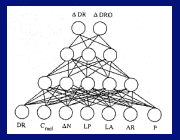
- Hybrid System

$$Y_n = y^s_n + y'_n = \sum_{i=1}^{J} k_i \cdot f_i(\underline{x}_n) + net(\underline{k}, \underline{x}'_n)$$



Experiment 1

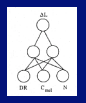
- Training with performance rules
  - ANN for Timing deviations
    - Input
      - Subset of KTH rules under consideration
      - AND additional inputs to aid ANN in recognizing context
    - Output: ΔDR, ΔDRO



Rule name	Input parameters	Output parameters
High loud	N	ΔL
Leap articulation	DR, ΔN, *LA*	ΔDRO
Leap tone duration	ΔN, *LP*	ΔDR
Articulation of repetition	*AR*	ΔDRO
Durational contrast	DR	ΔDR, ΔL
Melodic charge	C_{mel}	ΔDR, ΔL
Phrase	*P*	ΔDR, ΔDRO

Experiment 1

- ANN for loudness deviations
  - Input: DR, C_{mel}, N
  - Output: ΔL



- Training procedure
  - KTH rules were applied to W.A. Mozart's Piano Sonata K. 284 with "k" set to 1
  - Twenty deviations representing significant cases were used to train the ANN
- Evaluation
  - Twenty musicians volunteered
  - Disklavier piano recordings were evaluated for "musical quality"
  - K. 284 and K. 331 were "performed" using KTH rules and ANN model

Experiment 1

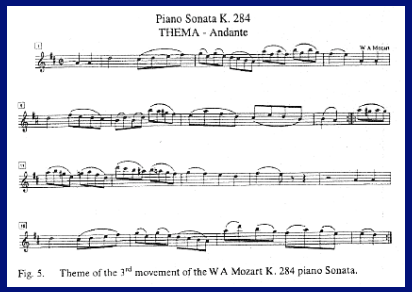
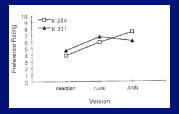


Fig. 5. Theme of the 3rd movement of the W A Mozart K. 284 piano Sonata.

Experiment 1

- Results
  - Mean ratings
    - ANN (6.73)
    - KTH rules (6.30)
    - Deadpan (4.28)
  - Perceived equivalence of ANN and KTH versions
    - → successful ANN emulation
  - Low ratings of deadpan version
- Discussion
  - Most subjects reported that performances were poor, possibly because of the lack of an accompaniment
  - Slight preference of ANNs may be due to interpolation properties of ANNs



Experiment 1

- User-controlled ANN
  - Additional input neuron “k” added for user control of performance style
  - ANNs were trained using same score played in different styles
    - Baroque (k=0)
    - Romantic (k=0.5)
    - Swing (k=0.9)
  - No results are given

Experiment 2

- Training with performing style of a pianist
  - Motivation: previous ANN structure found to be insufficient to describe pianist style
  - Training material: Robert Schumann’s Traumerei
    - Examined in-depth by Bruno Repp
    - Professional pianist recorded several MIDI performances
  - Processing
    - Melody was extracted
    - Mean metrical (nominal) quaver was computed

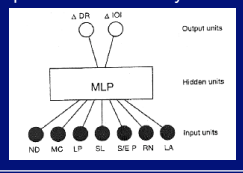
Experiment 2

- Pianist deviations are coded in terms of their “relative variations,” i.e. distance from nominal values
- Basic ANN structure
  - 3-layer MLP, 7 inputs, 2 outputs
  - Problem: notes with same input coding belonged to different contexts and were performed differently

$$\Delta DR = \frac{DR_{performance}}{DR_{nominal}}$$

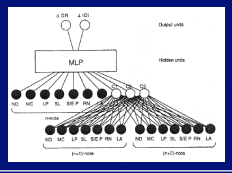
$$\Delta IOI = \frac{IOI_{performance}}{IOI_{nominal}}$$

$$\Delta L = \frac{L_{performance}}{L_{nominal}}$$



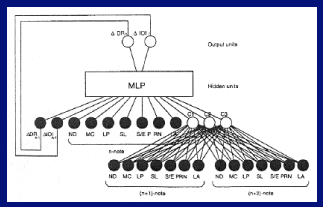
Experiment 2

- Solution: context nodes (CNs)
  - Three new nodes C1,C2,C3
  - Helped ANN classify identical patterns with differing outputs
  - Received information related to (n+1) note and (n+2) note
  - Results were more satisfactory



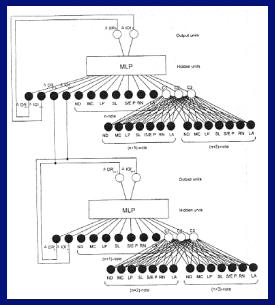
Experiment 2

- Ecological ANN
  - ANN output is fed back to the input
  - ANN virtual pianist "listens to himself"



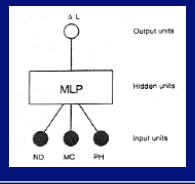
Experiment 2

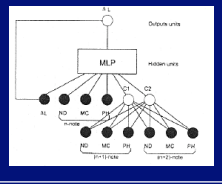
- Ecological-predictive ANN
  - Two more input nodes added
  - Lower ecological ANN computes DR & IOI for (n+1) note using (n+1), (n+2) and (n+3) notes
  - Upper ecological ANN uses the lower ANN inputs plus the n-th and (n-1) note



Experiment 2

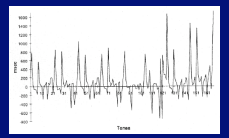
- Loudness deviations
  - Basic structure: 3 input, 1 output
  - Ecological and ecological-predictive versions produced loudness deviations very similar to pianist



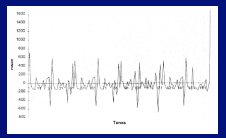


Experiment 2

- Results
  - In ANNs performance, there is a repetition of the deviations reflecting the piece structure, and small variations are made between structurally similar situations



Deadpan Pianist



Ecological-Predictive ANN

Experiment 2

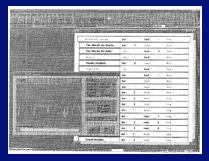
- Results (cont'd)
  - Informal listening tests resulted in "positive judgment of the ANN performance"
  - Langner and Kopiez (1995) found that the ANN's performance produced tempo variations very similar to those of a professional pianist
  - A comparison of ANN rules with KTH rules revealed that ANN learned very similar rules to those of the KTH system

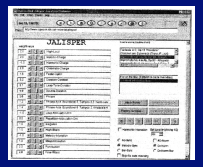
Conclusions

- Proposed ANN approach found to be successful in learning performing style of a real pianist
- ANNs allow real-time performance and offer possibility of controlling input parameters
- Future research
  - ANNs for punctuation of a score
  - ANNs for harmonic and structural analysis of music

Appendix

- Melodia: Software for testing the trained ANNs and KTH performance rules
- JALISPER (JAPER): Java (software/hardware independent) applet tool for performance





The End