A Distributed Model for Multiple Viewpoint Melodic Prediction

Srikanth Cherla[1,2], Tillman Weyde[1,2], Artur Garcez[2], Marcus Pearce[3]

[1]Music Informatics Research Group, City University London
[2]Machine Learning Group, City University London
[3]Centre for Digital Music, Queen Mary University of London

November 4, 2013

Outline

Introduction: Analysing sequences in symbolic music data

Background: Probabilistic modelling of melodic sequences

Approach: Modelling melodic sequences with RBMs

Results: Encouraging Prediction Performance

Next

Introduction: Analysing sequences in symbolic music data

Background: Probabilistic modelling of melodic sequences

Approach: Modelling melodic sequences with RBMs

Results: Encouraging Prediction Performance

Sequential Information in Notated Music

- A wealth of information in notated music.
- Increasingly available
  - in different formats (MIDI, Kern, GP4, etc.).
  - for different kinds of music (classical, rock, pop, etc.)
- Analysis of sequences key to extracting information.
- Melody — Good starting point for a broader analysis.

Relevance

Scientific:

- Computational musicology
- Organizing music data
- Generating musical stimuli
- Aiding acoustic models
- Music education

Creative:

- Automatic music generation
- Compositional assistance

Next

Introduction: Analysing sequences in symbolic music data

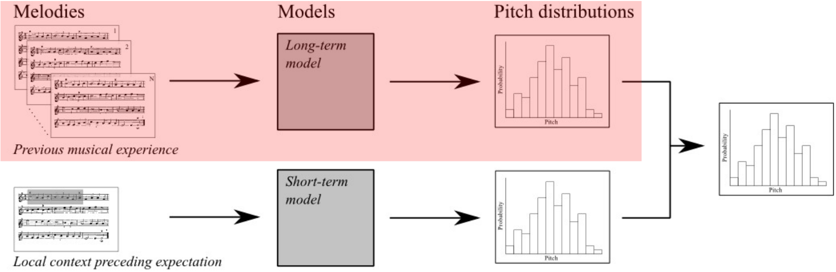
Background: Probabilistic modelling of melodic sequences

Approach: Modelling melodic sequences with RBMs

Results: Encouraging Prediction Performance

Information Dynamics of Music (IDyOM)

- **Predictive models of musical structure using probabilistic learning** (Pearce & Wiggins, 2004).
- Develop insights into the analysis of musical structure drawing on research in musicology (Whorley et al., 2013).
- Relate predictions to psychological and neural processing of music (Omigie et al., 2013).



Website: ***www.idyom.org***

Multiple Viewpoint Systems for Music Prediction (Conklin & Witten, 1995)

- Framework for analysis of symbolic music data.
- *Viewpoint type* (feature) sequences extracted from score.
- One Markov model per *type*.
- Mixture/product-of-experts to combine multiple models.

Viewpoint	Transformed sequence									
pitch	67	69	71	72	69	72	64	67	72	69
int	$\perp$	2	2	1	-3	3	-8	3	5	-3
onset	0	2	5	6	9	10	12	15	16	20
ioi	$\perp$	2	3	1	3	1	2	3	1	4
int $\otimes$ ioi	$\perp$	2, 2	2, 3	1, 1	-3, 3	3, 1	-8, 2	3, 3	5, 1	-3, 4

(Image Courtesy:Darrell Conklin)

Motivating a Distributed Model

At present...

1. A more scalable way to *link* viewpoint types.
2. An alternative approach to one relying directly on occurrence statistics.

In the future...

- Interest in knowledge extraction from neural networks.

Next

Introduction: Analysing sequences in symbolic music data

Background: Probabilistic modelling of melodic sequences

Approach: Modelling melodic sequences with RBMs

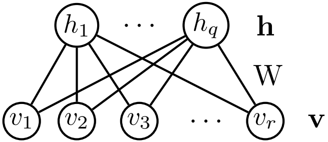
Results: Encouraging Prediction Performance

Goals

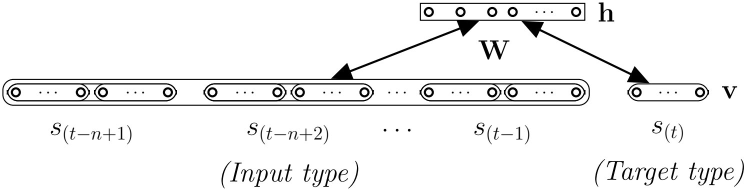
- Demonstrate the use of multiple-viewpoint systems with a distributed model - Restricted Boltzmann Machine.
- Compare the predictive performance of this model with the originally used Markov models on a melody corpus.

Restricted Boltzmann Machine (Smolensky, 1986)

- A bipartite network with binary stochastic units.
- Data in visible layer, features in hidden layer.
- Can model
  - joint distribution $p(v_1, \dots, v_r)$
  - **conditional distribution** $p(v_1, \dots, v_c | v_{c+1} \dots, v_r)$
- Can be stacked into a deep network and trained efficiently.



A Distributed Melodic Prediction Model



- Viewpoint subsequence $s_{(t-n+1)\ldots t}$ in visible layer.
- Models the conditional distribution $p(s_t | s_{(t-n+1)\ldots(t-1)})$.
- Generalized softmax visible units.
- Viewpoint types linked by vector-concatenation.
- Trained generatively using Contrastive Divergence.

Next

Introduction: Analysing sequences in symbolic music data

Background: Probabilistic modelling of melodic sequences

Approach: Modelling melodic sequences with RBMs

Results: Encouraging Prediction Performance

Evaluation Tasks

Predicting the next *pitch* with

1. a model that uses context of type *pitch*.
2. a model that uses context of type $pitch \otimes dur$.
3. a simple mixture-of-experts combination of 1 and 2.

Evaluation Setup

Corpus

- As used in Pearce et al., 2004.
- Subset of the Essen Folk Song Collection.
- A collection of 8 datasets of chorale and folk melodies.
- A total of $54,308$ musical events.

Evaluated models

- Context length $\in \{1, 2, 3, 4, 5, 6, 7, 8\}$
- Hidden units $\in \{100, 200, 400\}$
- Learning rate $\in \{0.01, 0.05\}$

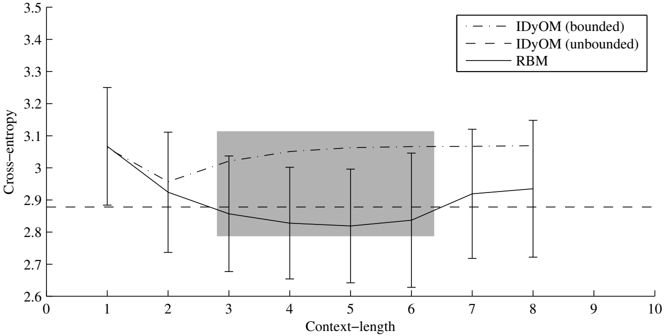
Evaluation criterion — cross-entropy (to be minimized)

$$H_c(p_{mod}, \mathcal{D}_{test}) = \frac{-\sum_{s_1^n \in \mathcal{D}_{test}} \log_2 p_{mod}(s_n | s_1^{(n-1)})}{|\mathcal{D}_{test}|}$$

Changing Context Length

- Dataset: Folk melodies of Nova-Scotia, Alsace, Yugoslavia, Switzerland, Austria, Germany; Chorale melodies
- Input: *pitch*, Target: *pitch*

Model Performance



Combining "Multiple Viewpoints"

Dataset: 185 chorale melodies

- Input: *pitch*, Target: *pitch*

context length	1	2	3	4
IDyOM	2.737	2.565	2.505	2.473
RBM	**2.698**	**2.530**	**2.490**	**2.470**

- Input: $pitch \otimes duration$, Target: *pitch*

context-length	1	2	3	4
IDyOM	2.761	2.562	2.522	**2.502**
RBM	**2.660**	**2.512**	**2.481**	2.519

- Input: $pitch \oplus (pitch \otimes duration)$, Target: *pitch*

context length	1	2	3	4
RBM (combined)	2.663	**2.486**	**2.462**	**2.413**

Conclusions & Future Work

We presented the following

- A distributed model for multiple-viewpoint melodic prediction using Restricted Boltzmann Machines.
- Improved prediction results in comparison to previously evaluated Markov models.

Some interesting directions for future work

- Deeper networks.
- Musical interpretation of hidden layers.
- A distributed Short-Term Model.
- Polyphonic music.
- Interesting MIR applications.

Acknowledgements

We would like to thank

Darrell Conklin (Universidad del Pais Vasco)
Son Tran (City University London)

Thank you!

Questions?