INTRODUCTION

Classifying melody in the brain

There is a broad consensus in the literature on music perception that Western musical phrases (“melodies”) can be described by several principal form bearing dimensions, i.e. dimensions that allow us to discriminate perceptually between different musical events. Standard among these dimensions are pitch, duration (“rhythm”/“timing”), amplitude (“loudness”/“dynamics”), and timbre [1-4]. Pitch is thought to have at least two sub-dimensions, namely register (the general frequency range or the octave of the pitch) and chroma (the octave cycle obtained when pitch is abstracted into octave independent note names, e.g. C, D, E, F, etc.) [4, 5].

Our aim was to correctly predict which of several melodies a subject was hearing on a trial-by-trial basis, using their scalp EEG. We then looked for invariances in stimuli representation across subjects and related these to stimuli structure.

Approach

• Use previously developed methods [6] to predict which of 4 melodic stimuli a subject hears on a given trial.
• Relate confusion matrices from classification to distance of stimuli in a mutual information-defined space

Goals:

• High single-trial classification rates (good fit)
• Relate structure in EEG to structure of melodic stimuli

METHODS: Basic Melodic Stimuli

10 subjects age 23-27 each completed 1200 trials involving hearing one of four melodies presented in a pseudorandom order.

We used a Monte Carlo simulation to generate a million potential groups of four stimuli that varied in pitch (chroma and register), duration and dynamics at three consecutive time points.

We then chose the group of four that was furthest apart both at each time point and across time points according to Levendesian distance and randomly assigned each a unique timbre.

METHODS: Analysis

Single-trial classification [6]

• Temporal ICA for artifact removal and to identify independent sources
• Spatial PCA to reduce problem dimensionality
• Discriminant Analysis applied to cleaned, dimension-reduced data
• Predict stimuli class trial-by-trial on held-out test data to generate confusion matrices

Relating classification to stimuli structure

• Mutual Information: for each dimension calculate within-dimension MI distances between stimuli to yield matrix

$D_{ij}^{Dim} = I(X_i^{Dim} ; X_j^{Dim})$

•BOOSTED GAM: flexible nonlinear fitting allows us to predict which of 4 melodic stimuli a subject is hearing on a given trial

RESULTS: Single-trial Classification

Confusion Matrix Experiment 1 (n = 30)

<table>
<thead>
<tr>
<th></th>
<th>True</th>
<th>Stimulate 1</th>
<th>Stimulate 2</th>
<th>Stimulate 3</th>
<th>Stimulate 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulate 1</td>
<td>60.5%±11.1%</td>
<td>11.1%±6.1%</td>
<td>6.6%±3.3%</td>
<td>12.6%±4.7%</td>
<td></td>
</tr>
<tr>
<td>Stimulate 2</td>
<td>11.8%±6.0%</td>
<td>70.1%±10.3%</td>
<td>10.6%±4.8%</td>
<td>7.7%±4.3%</td>
<td></td>
</tr>
<tr>
<td>Stimulate 3</td>
<td>17.3%±4.6%</td>
<td>10.6%±6.9%</td>
<td>6.0%±11.9%</td>
<td>13.1%±7.2%</td>
<td></td>
</tr>
<tr>
<td>Stimulate 4</td>
<td>13.2%±5.5%</td>
<td>7.1%±2.8%</td>
<td>13.5%±7.6%</td>
<td>66.1%±13.5%</td>
<td></td>
</tr>
</tbody>
</table>

$P$-value $6.7 \times 10^{-10}$ $1.8 \times 10^{-10}$ $9.7 \times 10^{-10}$ $1.6 \times 10^{-10}$

Confusion Matrix Experiment 2 (n = 7)

<table>
<thead>
<tr>
<th></th>
<th>True</th>
<th>Stimulate 1</th>
<th>Stimulate 2</th>
<th>Stimulate 3</th>
<th>Stimulate 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulate 1</td>
<td>59.9%±12.3%</td>
<td>14.4%±5.8%</td>
<td>15.4%±3.9%</td>
<td>13.7%±5.6%</td>
<td></td>
</tr>
<tr>
<td>Stimulate 2</td>
<td>12.3%±5.8%</td>
<td>67.2%±11.1%</td>
<td>10.9%±3.1%</td>
<td>16.3%±5.3%</td>
<td></td>
</tr>
<tr>
<td>Stimulate 3</td>
<td>16.0%±4.9%</td>
<td>11.4%±3.5%</td>
<td>66.1%±6.3%</td>
<td>11.9%±3.2%</td>
<td></td>
</tr>
<tr>
<td>Stimulate 4</td>
<td>15.6%±4.7%</td>
<td>9.0%±3.8%</td>
<td>14.4%±4.2%</td>
<td>66.5%±9.3%</td>
<td></td>
</tr>
</tbody>
</table>

$P$-value $2.1 \times 10^{-3}$ $3.0 \times 10^{-15}$ $4.8 \times 10^{-32}$ $1.2 \times 10^{-34}$

RESULTS: Relating Classification Rates to Stimuli Structure

Mutual information (MI) between certain stimuli dimensions predicts confusion rate

$\text{Confusion} = \beta_0 + \beta_1 \text{D}^{\text{Chroma}} + \beta_2 \text{D}^{\text{Dynamics}} + \beta_3 \text{D}^{\text{Register}} + \epsilon$

Actual Confusion Matrix

Experiment 1 (n = 10)

Estimated Confusion Matrix using GAM model from Experiment 1 and Experiment 2 (n = 7)

Finding:

• Excellent single-trial classification rates of melodies using EEG data
• Strong invariance in representation across subjects
• Structural model consistent with previous findings that show modularity of chroma and register [4]
• Affine mapping between information in stimulus space and ambiguity of representation in the EEG

REFERENCES