Automatic Synthesizer Preset Generation with PresetGen

Kıvanç Tatar, Matthieu Macret, Philippe Pasquier
The *preset generation* problem

- Modern synthesizers are very powerful and have many parameters resulting in a vast and complex search space.
- The possibilities of a given synthesizer are unknowns and the search space is beyond human grasp.
- Preset search is time-consuming and tedious.
  - Musicians and sound designers spend time tuning parameters instead of making music.
  - The solution found might not be optimal

**We want to automate preset generation:** Given a target sound, and a synthesizer, give me a preset for that sound.
The OP-1 is a commercial synthesizer that has a very large presets search space:

- 7 synthesis engines
- 3 types of LFO (Low frequency oscillators)
- 4 types of special effects
- 120 keys

The total number of distinct presets is of the order of $10^{76}$

Added challenges: The space is highly discontinuous and the synthesis engines are non-deterministic (adding warmth to the sound). Each with 4 parameters with 32767 possible values each.
Our solution: PresetGen

• We use evolutionary algorithms to locate multiple distinct OP-1 presets to replicate a given target sound

• We minimize the 3 objectives distances (Envelope, FFT, STFT) using a multi-objective genetic algorithm: the Non-dominated Sorting Genetic Algorithm-II (NSGA-II)
1\textsuperscript{st} Objective: Temporal envelope distance

Target sound

OP-1 generated sound

Euclidian distance
$2^{nd}$ objective: FFT distance for spectral signature
3rd objective: STFT distance for spectral content dynamic
Results

1. We analyse the target sound

2. We evolve presets

3. We cluster the Pareto front

4. We return a variety of presets that approximate the target sound using various synthesis methods!
Examples

<table>
<thead>
<tr>
<th>Engine</th>
<th>FX</th>
<th>LFO</th>
<th>Key</th>
<th>Octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster</td>
<td>Inactive</td>
<td>Inactive</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Engine</th>
<th>FX</th>
<th>LFO</th>
<th>Key</th>
<th>Octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster</td>
<td>Punch</td>
<td>Inactive</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Engine</th>
<th>FX</th>
<th>LFO</th>
<th>Key</th>
<th>Octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM</td>
<td>Grid</td>
<td>Element</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Examples of instruments

**Piano**

<table>
<thead>
<tr>
<th>Engine</th>
<th>FX</th>
<th>LFO</th>
<th>Key</th>
<th>Octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>String</td>
<td>Delay</td>
<td>Tremolo</td>
<td>44</td>
<td>1</td>
</tr>
</tbody>
</table>

**Clarinet**

<table>
<thead>
<tr>
<th>Engine</th>
<th>FX</th>
<th>LFO</th>
<th>Key</th>
<th>Octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital</td>
<td>Delay</td>
<td>Tremolo</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>
Empirical Evaluation

- PresetGen compared to human sound designers.
  - 8 target sounds:
  - 3 human sound designer
  - 14 auditors judge similarity across dimensions.

- Results:
  - PresetGen sounds rated more similar to target (avg 17%)
  - PresetGen outperform humans at the task both in competency and efficiency.
In Conclusion: **PresetGen** automates a creative task to human competitive levels and would fit well at a computer-assisted creativity tools in many synthesizers.