Music Search Engine

Li Cao, Jason Chang, & Tiffany Yeh

Advisor: Alex Spektor
Contents

• Inspiration
• Features
• System Details
• Testing
• Conclusion
Inspiration

• “What’s the name of that song!”
• “I can’t understand a word Enya is singing…”
• Music is a universal language
Features

• Search song by:
  – Virtual keyboard
  – Plug in audio signal
  – Microphone input

• USB Communications
**System Details**

*Software Components*

Building Database

User Input Preparation → Realtime Pitch Extraction* → Serial-to-USB Conversion → Song Search*
System Details - Building Database (Overview)

- **Building Melody Database**
  - **File Preprocessing**
    - Convert to *.wav
    - Bandpass Voice Filter (0.1kHz - 3kHz)
    - Convert to ASCII Data
  - **Pitch Extraction**
    - Break Up Data
    - File Done Reading? (no)
    - Pitch Extraction Algorithm
    - yes
  - **File Postprocessing**
    - Combine Data
    - Convert Lags to Frequencies
    - Clean Data (Median Filter & other non-linear de-noising techniques)
    - Code Frequencies
    - output
System Details – Building Database (Algorithm)

- Implemented and Researched by Yipeng Li and DeLiang Wang\(^1\)
- Extracts pitch perfectly at SNR = 10dB
- Typical music has SNR < 0dB
- Restricted to short input (~3 seconds)

\[\text{[1]} \text{ Li, Yipeng and DeLiang Wang. “Extracting Pitch of Singing Voice in Polyphonic Audio.” 2005.}\]
System Details – Building Database (Coding)

Database Song Format:

Mary Had a Little Lamb Coding Example

<table>
<thead>
<tr>
<th>Song – Freq</th>
<th>329.6</th>
<th>293.7</th>
<th>261.6</th>
<th>293.7</th>
<th>329.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Song – Note</td>
<td>E4</td>
<td>D4</td>
<td>C4</td>
<td>D4</td>
<td>E4</td>
</tr>
<tr>
<td>Song – Freq. Index</td>
<td>40</td>
<td>42</td>
<td>44</td>
<td>45</td>
<td>44</td>
</tr>
<tr>
<td>Song – Freq. Length (10ms)</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>150</td>
</tr>
</tbody>
</table>
System Details - Building Database

4PM – Sukiyaki
BBMak – More than Words
Billy Joel – The Longest Time
Brown Eyes – I Believe
Christina Aguilera – Beautiful
Deep Blue Something – Breakfast at Tiffany’s
Dido – White Flag
Little Mermaid – Part of Your World
Eagles – Hotel California
Everclear – I Will Buy You a New Life
G.O.D. – Trip
Goo Goo Dolls – Black Balloon
Howie Day – Collide
Jackson 5 – Rockin’ Robin
James Blunt – You’re Beautiful
John Denver – Leaving on a Jet Plane
John Denver – Take Me Home Country Road
John Mayer – My Stupid Mouth
Josh Kelley – Perfect 10
Kelly Clarkson – Beautiful Disaster

Kelly Clarkson – Behind These Hazel Eyes
Kelly Clarkson – Breakaway
Kelly Clarkson – Since You’ve Been Gone
Mamas and Papas – Puff the Magic Dragon
Mariah Carey – Hero
Mariah Carey – Can’t Live if Living is Without You
Marvin Gaye – Ain’t No Mountain High Enough
Michelle Branch – Are You Happy Now
Nat King Cole – L-O-V-E
Norah Jones – Come Away With Me
Norah Jones – Don’t Know Why
N’Sync – How Deep is Your Love
Panic! At The Disco – I Write Sins Not Tragedies
Paul McCartney – Yesterday
Paulina Rubio – Ni Una Sola Palabra
Red Hot Chili Peppers – Otherside
Take Me Out to the Ballgame
Twinkle Twinkle Little Star
Temptations – My Girl

*Members of this group obtained legal copies of these songs
System Details - User Input

- Microphone Input
- Keyboard (Line-in) Input
- Pre-amplifier
System Details - User Input

Preamplifier

[Diagram of a preamplifier circuit with various components such as resistors and capacitors labeled with values like 1.2m, 10u, 4k, 100k, 6.2k, and 15k.]

[Graph showing an oscilloscope trace with vertical and horizontal settings such as 20% and 1.00V/200ns.]

Agilent Technologies

[Options for measuring Amp1: 7.7mV, Amp2: 1.75V, and Frequency: 259Hz.]

[Options for Selecting and Clearing data.]
System Details - Pitch Extraction (FFT)

- Sampling rate 44,100 Hz
- Frequency resolution 10.766 Hz
- Decimation by 8
- 4096 point FFT
- Frequency resolution 1.346 Hz
**System Details - Pitch Extraction**

- Normal singing voice 150 – 1000 Hz
- Aliasing from harmonics
- Lowpass filter with cutoff 1000 Hz
System Details - Pitch Extraction

Prefilter

- OPAMP
- 1.2M ohm
- 0.02uF
- 0.01uF
- Vin
- 0V
4096 Point FFT of Tiffany’s Middle C

- Frequency (Hz)
- Magnitude

- Middle C
- 2nd Harmonic
- 3rd Harmonic
- 4th Harmonic
System Details - Pitch Extraction (Harmonics)

- Harmonics occur at 2x, 3x, etc., of fundamental frequency
- Harmonics of low frequency notes may fall within filtered range
- Find if strongest frequency is a harmonic of some other fundamental frequency
System Details - PC Communication

Why serial to USB?
System Details - PC Communication

- MAX232
- USB Breakout
System Details - PC Communication
System Details - Search Algorithm

• Hard to implement due to inaccurate database
• >5 Search Algorithms Implemented
• Optimized for quick search times and accurate results
• Search parameters effect results a great deal
System Details - Search Algorithm

- Search window for each note
- Search window for entire string
- Possible skipping of notes
### System Details - Search Algorithm

**Search String (differences)**

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>+2</th>
<th>+2</th>
<th>+2</th>
<th>+3</th>
<th>+9</th>
</tr>
</thead>
</table>

**Song – Freq. Index**

<table>
<thead>
<tr>
<th></th>
<th>40</th>
<th>42</th>
<th>43</th>
<th>45</th>
<th>46</th>
<th>48</th>
<th>49</th>
<th>48</th>
<th>49</th>
<th>55</th>
</tr>
</thead>
</table>

**Song – Freq. Length (10ms)**

<table>
<thead>
<tr>
<th></th>
<th>20</th>
<th>20</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>10</th>
</tr>
</thead>
</table>

*No 44! Skip 3 notes*
Testing & Results

• Perfect database & perfect input = perfect
• Bad database & perfect input = pretty good
• Bad database & bad input = not good

• Perfect database – Hard coded database
• Perfect input – Virtual keyboard
## Testing & Results

<table>
<thead>
<tr>
<th>Search</th>
<th>Language</th>
<th>&lt;Search Time/ N&gt;</th>
<th>&lt;Percent of finding song&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Split Block Search</td>
<td>VB6</td>
<td>10.4ms</td>
<td>56.0%</td>
</tr>
<tr>
<td>Split Block Search (with Error Factor)</td>
<td>VB6</td>
<td>12.5ms</td>
<td>54.0%</td>
</tr>
<tr>
<td>Complete Window Search</td>
<td>VB6</td>
<td>152.6ms</td>
<td>48.0%</td>
</tr>
<tr>
<td>Complete Window Search (with Error Factor)</td>
<td>VB6</td>
<td>165.4ms</td>
<td>70.7%</td>
</tr>
<tr>
<td>Complete Window Search</td>
<td>C++ &amp; VB6</td>
<td>16.4ms</td>
<td>68.6%</td>
</tr>
<tr>
<td>Complete Window Search (with Skip)</td>
<td>C++ &amp; VB6</td>
<td>20.4ms</td>
<td>83.6%</td>
</tr>
</tbody>
</table>

N: input string length

### Songs Tested

- **Deep Blue Something** – Breakfast at Tiffany’s
- **Mamas and Papas** – Puff the Magic Dragon
- **Paulina Rubio** – Ni Una Sola Palabra
- **Dido** – White Flag
- **Norah Jones** – Come Away with Me
- **Red Hot Chili Peppers** – Otherside
- **Josh Kelley** – Perfect 10
- **Paul McCartney** – Yesterday
Conclusions

Advantages
– Can search vast database
– Potentially retrieve similar music

Disadvantages
– Never as perfect as human
– Each song takes a long time to process
Conclusions

Future improvement
• Improve database algorithm
• Recognize and stabilize wavering from untrained singers
• Improve search algorithm – leniency for imperfect input
Thank You…

• Professor Swenson
• Alex Spektor & ECE445 TAs
• ECE Shop Technicians
• Yipeng Li & DeLiang Wang
• Professor Jones
• TI Support
References


Thank you for coming!

Questions?