Intelligent Communication Systems

Frontiers
2/22/10
Tim Carella
Outline

- Background on aphasia
- Problems faced by aphasics
- Objective of the system

The system:
- 6-sensor data glove
- Finger language recognition software
- Coordinate-based virtual keyboard

Conclusions
Background Information

- Aphasia is a speech handicap where a person is suddenly unable to speak clearly (or at all) and there is a lack of capacity to comprehend speech.

- Three key features of this system:
  - 6-Sensor data glove (hardware)
  - A finger language recognition subsystem (software)
  - A coordinate-based virtual keyboard
The AAC (Augmentative and Alternative Communication) systems out today are too complex for some users

- AAC devices have low text typing speed
- Some users lack the fine motor control required to work these devices
- Most of the AAC devices are imported and expensive
Objective of the System

- Create a new intelligent system allowing handicapped aphasia victims to perform basic communication
- To provide quick output from simple input methods (i.e. hand movements)
The user will wear a 6-sensor data glove made up of LEDs and Photodiodes.

By means of hand gestures (much like sign language) the user will be able to communicate.

Every hand gesture will be recognized by the developed software and will correlate to a neural network-based classification.

The classifications specify a value on the virtual keyboard.

Once recognized by the subsystem software text will be output digitally on a screen.

Source: [1]
Sign language is popularly used among the deaf-mute

- The complicated hand/body motions make sign language very difficult for handicapped aphasiacs

At the Southwest Research Institute a mechanical hand was developed for the deaf/blind
The Stanford Mechanical Department (Dr. Gilden) developed the DEXTER series (1993)

This series is comprised of DEXTER models I-IV

The series is capable of converting the 26 letter English alphabet into hand gestures for keyboards, computers, and other similar devices

This was the first piece used in creating this new intelligent communication system

Source: [2]
A new data glove was designed specifically for this system.

There are other gloves out in the market capable of similar results.

Those gloves are overly complicated and can cost up to US $5,000.

The goal is to design a reliable but low cost data glove (~$150).

Source: [3]
The Data Glove

- Uses only LEDs and photo-detectors (PD)
- The data glove is open source and can be used with MATLAB and C++ programs
- This simple design is low cost (~$400) and easily programmable

Source: [4]
The problems faced when creating any data glove is system reliability/accuracy and the differences between individual patients.

To bypass this problem a software system capable of noticing very small movements is necessary.
Finger Language Recognition Subsystem

Source: [4]
Feature Values Calculation

Oi : Original optical signal strength of each finger (scale 0 - 255)

Fi : Bending degree of each finger

\[
F_i(O_i) = \begin{cases} 
1 & \text{if } O_i \geq b_i \\
\frac{(O_i - a_i)}{(b_i - a_i)} & \text{if } a_i < O_i < b_i \\
0 & \text{if } O_i \leq a_i 
\end{cases}
\]

ai and bi are the signal values for the least bent and least straight events of finger i

\[ N_i = O_i / 255 \]

Ni : Normalized value of original optical signal strength of each finger
The feature selection equation creates an output $Y$ using all the features previously described features ($O_i$, $F_i$, and $N_i$)

$$Y = C_0 + (C_1 \times O_1 + C_2 \times O_2 + C_3 \times O_3 + C_4 \times O_4 + C_5 \times O_5)$$
$$+ (C_6 \times F_1 + C_7 \times F_2 + C_8 \times F_3 + C_9 \times F_4 + C_{10} \times F_5)$$
$$+ (C_{11} \times N_1 + C_{12} \times N_2 + C_{13} \times N_3 + C_{14} \times N_4 + C_{15} \times N_5)$$

Through testing it was discovered that $O_i$ and $N_i$ could be neglected and the simplified formula is now:

$$Y = C_0 + C_6 \times F_1 + C_7 \times F_2 + C_8 \times F_3 + C_9 \times F_4 + C_{10} \times F_5$$

The bending degree ($F_i$) is required for every individual user.

With this equation almost all users can achieve a success rate of 99%
Image shown is a simple model of the radial basis function neural network (RBFNN).

The feature selection values become the input values for the input layer.

The values then progress through the 2 remaining stages of the network and are given a single output value of either 1 or 0.

Source: [4]
## Neural Network Classifications

<table>
<thead>
<tr>
<th>Language Component Y</th>
<th>Finger gesture</th>
<th>Classification by neural network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>y1</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Source: [4]
Speed in communication systems is highly based around text entry.

Many approaches have been proposed:

- Chord keyboard
- Line-of-gaze technique
- Listening keyboard as a speech interface
- Japanese text input method for 12-button phone keypad

The problem with these approaches is that they are based off the standard keyboard.
Virtual Keyboard

Source: [4]

<table>
<thead>
<tr>
<th>Step</th>
<th>Meaning</th>
<th>Gesture Annotation</th>
<th>Finger Gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>English Input</td>
<td>Gesture: 0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>X coordinate: 5</td>
<td>First Gesture: 5</td>
<td>Thumb Up</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delimiter Gesture</td>
<td>Wrist Bent</td>
</tr>
<tr>
<td>3</td>
<td>Y coordinate: 1</td>
<td>Second Gesture: 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delimiter Gesture</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wrist Bent</td>
</tr>
</tbody>
</table>
Only components 1 - 7 are used to navigate the virtual keyboard.

Additional uses:

- Components 0 and 9 – Select English or Chinese
- Component 8 – erase an incorrect input
- Component 10 – “page down button”
- Component 11 – confirms the choice
Top Level System Summary

- **Phase 1:**
  - Glove adjustment (PDs and LEDs)

- **Phase 2:**
  - Adjusting the max and min values of each individual finger

- **Phase 3:**
  - Training (User will make each gesture 100 times)

- **Phase 4:**
  - Regression analysis is done on the training data

- **Phase 5:**
  - Phase 4 results are fed into the neural network
# Error Calculations

<table>
<thead>
<tr>
<th>Neural network</th>
<th>Features</th>
<th>Error in each gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td>User1</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;, N&lt;sub&gt;1&lt;/sub&gt;, N&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User2</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;, N&lt;sub&gt;2&lt;/sub&gt;, N&lt;sub&gt;3&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User3</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;, N&lt;sub&gt;1&lt;/sub&gt;, N&lt;sub&gt;2&lt;/sub&gt;, N&lt;sub&gt;4&lt;/sub&gt;, N&lt;sub&gt;5&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User4</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User5</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;, N&lt;sub&gt;1&lt;/sub&gt;– N&lt;sub&gt;5&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User6</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User7</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User8</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User9</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;4&lt;/sub&gt;, N&lt;sub&gt;4&lt;/sub&gt;, N&lt;sub&gt;5&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>User10</td>
<td>RBFNN F&lt;sub&gt;1&lt;/sub&gt; – F&lt;sub&gt;5&lt;/sub&gt;, N&lt;sub&gt;2&lt;/sub&gt;, N&lt;sub&gt;4&lt;/sub&gt;, N&lt;sub&gt;5&lt;/sub&gt;</td>
<td>0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

Source: [4]
Action 1:
• The user is trained for use with the data glove and virtual keyboard

Action 2:
• The recognition system finds the correct values and the virtual keyboard displays the designated letters / sentences

Action 3:
• The intended text is sent to a text-to-speech device
Things to remember:

• The data glove is fitted specifically to its user

• The bending degrees of each finger are calculated separately during the training process

• The recognition model uses only the best test data to allow for close to 100% success rate

• The coordinate-indexed virtual keyboard has continuous input for fast communication speed
Conclusions

- The system has been tested by various beta users and received positive results commenting at its ease of use and communication speed.
- The beta testers have commented positively saying the system is suitable for typing letters.
- Some future work may include other similar applications of the system and an improved customization process for the user allowing for the creation of hand gestures.
Major Engineering Concepts

- Mean Squared Error
- Classification by Neural Network (RBFNN)
- Electronics
  - Light Emitted Diode
  - Photodiode
- MATLAB
References


Questions?