Audience

The audience for my project is Skype users. This can be extended to people interested in using VoIP as a means for communication. Right now the prototype is being used with Skype; however, it might be possible to port the code over for other VoIP clients.

Motivation

Typing can take a long time. Some people can only have text conversations for a short period of time. Or during a text conversation, one may leave out details or simply not be very responsive to avoid typing. However, the text conversations allow people to easily log their conversations. One can search old conversations just to remember a fun conversation, or to find some piece of information.

VoIP is a convenient way to have a conversation when not face to face. Many people find it easier to express ideas vocally instead of through text. With larger hard drive sizes, people can now log their audio conversations as well. However, not many people have the patience to listen to an entire conversation to find a piece of information. What if we could visualize the conversation though? If we could find a way to summarize the conversation in an image, users would be able to just look at the image and gain information about what was talked about. This image also becomes an artifact of the conversation. Users will be able to create a gallery of their past conversation, and maybe even notice more patterns based on the set of images.

Idea

My idea is to create a tag cloud of sorts. Using a speech recognition system the audio captured is analyzed. The content is extract and placed into the visualization in the order the words are said. The words appear around a circle. Each circle represents a minute of the conversation. They start appearing towards the center of the screen and then as time passes, the words move outwards, away from the center to make room for the new words being spoken. When a word is moused-over, the word will increase in size to allow people to see what a word is, in case it is being covered by many other words.

If a word is repeated, the first instance of the word is found and the font size is increased, the word moves closer to the center, and then the color of the word is blended weighted by how many times each user has spoken the word. By doing this all the important words in the conversation grow to be very big and they also are clustered around the center of the screen. When a user comes back to the image later, all one has to do is look at what words appear larger and around the center. This resembles a tag cloud, where word sizes are increased by frequency of use and the ordering, for the most part, is based on the time the word was said. Words that are repeated “transcend” time in a way because they move away from the order in which they were said and move out of place towards the center. For archival, the themes of conversations can easily be looked up by looking at an image. If multiple images exist, one can look to see if there are themes that span the conversations.

The images also can be used as artifacts. The program allows the users the take screenshots at any point during the conversation. So a user can make as many screenshots as they want of the conversation. This will provide the user with the ability to make a gallery of their conversations and use them just as they would a normal photograph album.