JARVIS: Real-Time Speech Recognition Engine for Personal Assistant on FPGA

CSEE 4840 SPRING 2016
PROJECT PROPOSAL

Rishikanth Chandrasekaran(rc3022)
Gikku Stephen Geephilip (gg2624)
Richa Glenn Netto(rn2388)
Ashwin Karthik Tamilselvan (at3103)

I.   Objective

Speech recognition is an increasingly popular choice of input for interacting with devices around us intuitively. This allows users to naturally interact with technology by simply talking to it. Though the concept is an intuitive one, the implementation involved for such a system is complex. Major challenges involved are the processing complexities and its associated latency in terms of response time for recognition. We propose an interactive portable device that works as a digital personal assistant.

The general intuition behind our system is to implement a hardware accelerated/enhanced algorithm that runs a speech recognition algorithm on voice inputs obtained via a microphone. The computational burden of the algorithm is managed by the FPGA and the application back-end for implementing the digital assistant features are implemented in software on the ARM processor.

II.   Functionality of Personal Assistant:
   - Setting an alarm
   - Scheduling appointments on Google Calendar
   - LCD indicator for weather abnormalities
   - Task reminder on LCD display

III.   Algorithms (tentative):

- Feature Extraction:
    i.   Windowing
    ii.   Fast Fourier Transform (FFT)
    iii. Mel Binning
    iv.   Discrete Cosine Transform (DCT)

- Training:
    i.   Hidden Markov Model (HMM) – Viterbi Algorithm
    ii.   Gaussian Mixture Model (GMM)

- Classification:
    i.   Forward Backward Algorithm
    ii.   Other deep learning neural network algorithms