

# Memory and CPU requirements for the ASR1600v2



Author: Michael Vanslembrouck

Revision History:

- June 18, 1998 : Version 1.0

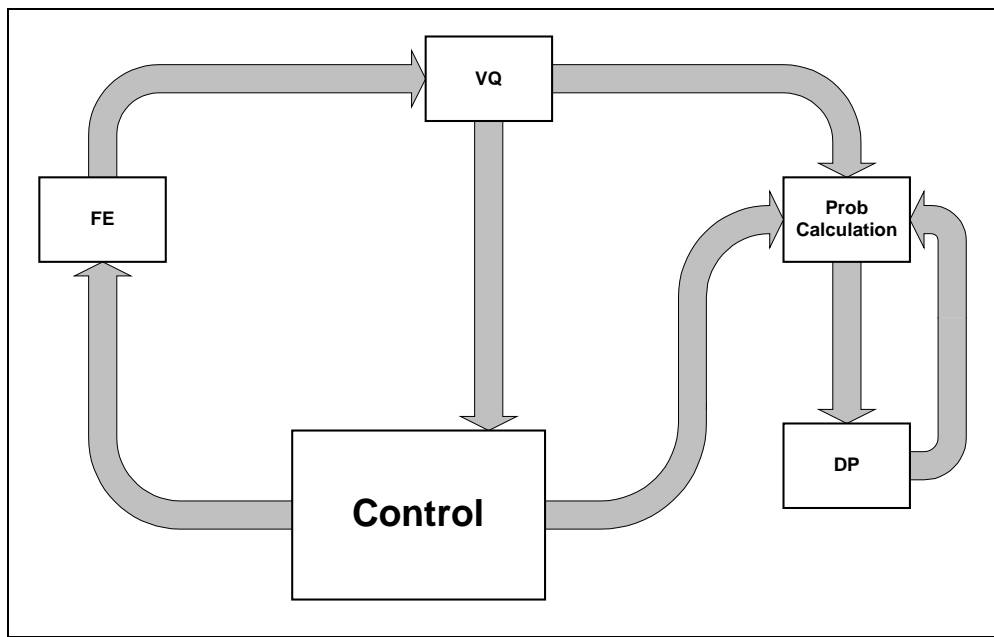
## 1 Table of Contents

Memory and CPU requirements for the ASR1600.....1	4	Language Buffer Size .....2
1 Table of Contents.....1	5	Dynamic Memory Allocations.....2
2 Introduction .....1	6	CPU Requirements.....4
3 Code and constant data size.....2		

## 2 Introduction

This document describes the memory and CPU requirements for the ASR1600 recognition engine. The memory and CPU requirements are given for the different blocks used in the algorithm. The algorithm consists of the following blocks (See Figure 1):

- Feature Extraction (FE)  
The feature extraction is FFT based. For each 10 ms of speech a FFT of 512 points has to be calculated. Other components in this block are: hamming window, IFFT of size 21, ... . In the current version all calculations are done in floating point. A fixed point version of this block is available.
- Vector Quantisation (VQ)  
Mainly distance calculations between feature vectors and cluster vectors. In the current version floating point calculations are used. A fixed point version of this block is available.
- Probability Calculation  
Mainly fixed point additions.
- Dynamic Programming (DP)  
Search algorithm. A lot of branching, comparisons and additions. All calculations are in fixed point.
- Control  
Contains the control diagram and the decision logic for training and verification.



**Figure 1 : Block diagram of ASR1600**

The CPU measurements are done on a PC with a Pentium Pro 200 with 256 kByte of Cache and 64 MB of memory in Windows NT 4.0. The compiler used is Visual C++ 5.0 optimized towards speed. There is currently no assembly code used. The code is written in ANSI C. No operating system functions are called directly in the library. All needed operating system functionality (memory allocation, thread synchronization) is given by the application by a set of callback functions

Version used: **ASR1600v2 beta**

### **3 Code and constant data size**

Code Size	83232
Constant Data	223009
<b>Total</b>	<b>306241</b>

### **4 Language Buffer Size**

German

8 Bit Models	1404452
4 Bit Models	NA

### **5 Dynamic Memory Allocations**

Memory sizes are given in bytes. The maximal memory allocation is given for each block separately. All recognitions are done with the default parameter settings, except for start detection and trailing silence detection, which are disabled. The context is a continuous digits context. The memory of the queues (FIFO) used in the datapath (See Figure 1) are included in the Control block.

Input File	3.1 s	9.8s
Control	59868	69725
Feature Ext.	38132	38132
VQ	384	384
DistCalc	12180	12180
Recog	48358	67942
<b>Max Total</b>	<b>156104</b>	<b>176196</b>

## 6 CPU Requirements

All time measurements are given in milliseconds. The times given are the average times required for processing 10 ms of input speech (note that this does include a second pass for the Prob calculation and DP ).

Utterance 3.1 s :

Number of frames processed : 310

Total time spent in blocks : 358.772 ms

Average time spent on a frame : 1.157 ms

Block	Average	Relative
Windowing	0.024	2.08
FFT + Modulus	0.111	9.55
Vad	0.021	1.80
LookAhead	0.011	0.93
Mel (*)	0.022	1.90
SpecSub (*)	0.040	3.44
Normalisation (*)	0.051	4.25
DCT+Lifter (*)	0.023	1.95
Mean	0.013	1.11
Vad	0.019	1.65
Stationarity	0.012	1.06
Trigger	0.009	0.81
Gender	0.041	3.51
Combiner	0.015	1.31
VQ	0.089	7.66
DistCalc	0.068	5.88
DP	0.418	36.09
<b>Total</b>	<b>1.157</b>	

(\*) numbers for a single gender.