

SpeechAnalyser

Kurzbeschreibung:

Die Schaltung soll ausgesprochene Wörter erfassen, auswerten und erkennen können. Eine erfolgreiche Erkennung bzw. die dazu notwendige Zwischenschritte werden über ein angeschlossenes 128x64 Punkte Graphic Display angezeigt. (Bild 1)

Zum „Lernen“ der Wörter wird ein zusätzlich erstelltes Visualisierungsprogramm auf einem PC eingesetzt (Bild 3). Dieser ist über die serielle Schnittstelle mit der Erkennungshardware verbunden wird.

Die Hardware überträgt die erfaßten Sprachdaten an den PC. Die PC-Software beinhaltet den identischen Erkennungsalgorithmus wie die Software des R8C, so daß dieser Vorgang bereits beim „Lernen“ beobachtet werden kann. Jedes Sprachmuster wird mittels der PC-Software mit einem editierbaren Namen versehen. Sind alle gewünschten Sprachmuster zusammengestellt, werden sie über die serielle Schnittstelle in das Flash-Rom der Hardware übertragen, die daraufhin autonom weiterarbeiten kann.

Jeder nun erfaßte Laut wird mit dem im Flash gespeicherten Mustern verglichen und bei Übereinstimmung durch den vereinbarten Namen angezeigt.

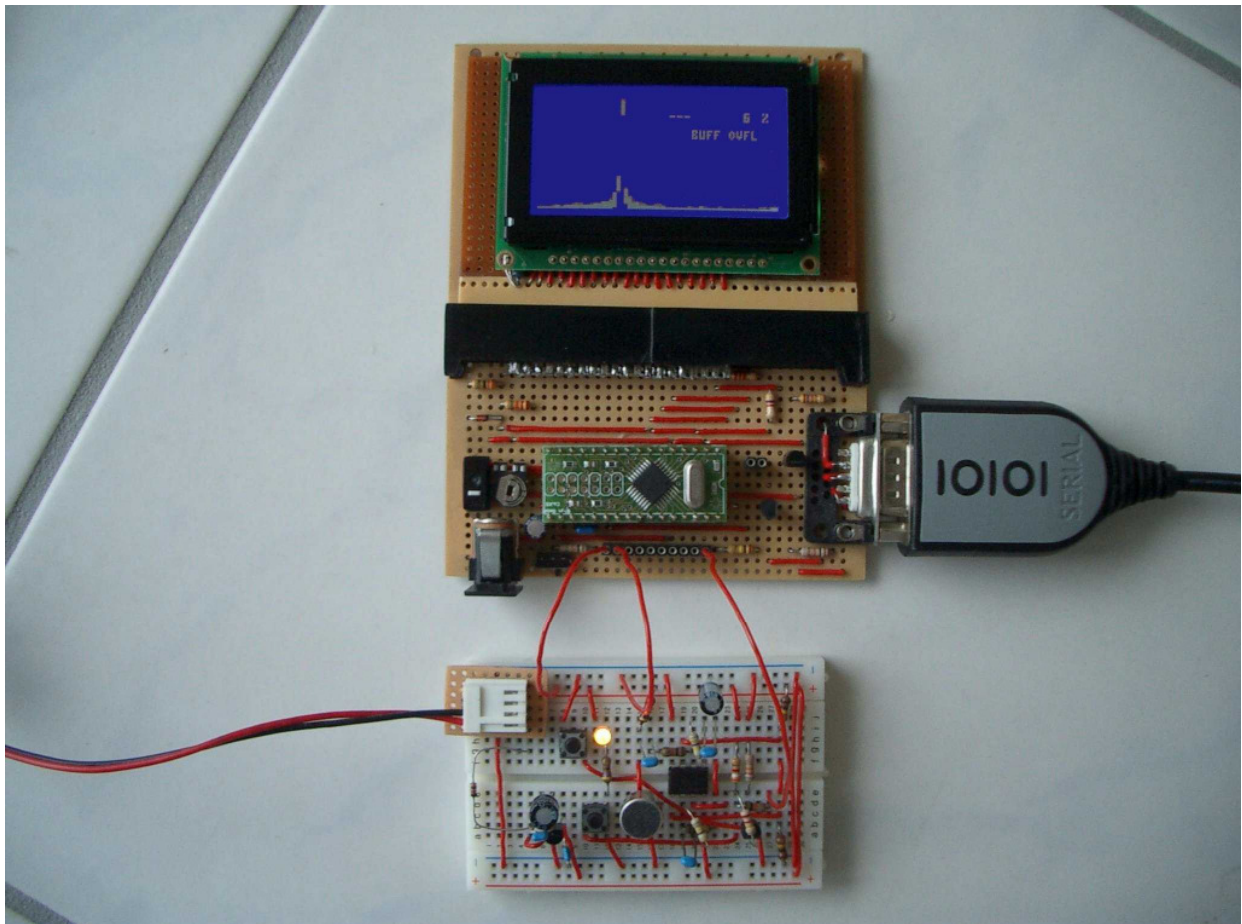


Bild1

Das Erfassen der Sprache geschieht über ein Elektretmikrophon mit nachgeschaltetem 2fach OpAmp Verstärker und Transistor Endstufe zur Spannungsanpassung an den A/D Eingang des R8C. Der Audioteil kommt ohne Abstimpmpotentiometer aus. Die korrekte Einstellung kann mit Hilfe der verschiedenen Display Anzeigen leicht kontrolliert werden. Der Digitalteil besteht im wesentlichen aus der im Dezemberheft der Elektor vorgestellten Minimallösung, ergänzt durch ein handelsübliches LCD Display mit KS0108B Controller:



Das Programm selektiert mit Hilfe des Lautstärkeverlaufes, Laute mit maximal einer Sekunde Länge und vergleicht die durch die FFT errechneten Frequenzspektren mit denen der abgespeicherten Sprachmuster. Die Ähnlichkeit zwischen einem gelernten Muster und einem Prüfmuster wird in Prozent angegeben. Bei einem Lernvorgang ergeben sich stets 100%, da das gerade aufgenommene Muster exakt mit dem Prüfmuster übereinstimmt. Jedes gelernte Muster mit mehr als 50% tiger Übereinstimmung wird als Kandidat herangezogen. Das Muster mit der besten Übereinstimmung gewinnt.

Über Taster S3 kann eine Einstellung gewählt werden, in der die Frequenzdaten im Bereich 100-1000Hz über die serielle Schnittstelle an den angeschlossenen PC versendet werden.

Der Lernvorgang wird mit Hilfe der PC-Software durchgeführt, die es ermöglicht, die übertragenen Sprachmuster in Form eines drehbaren 3D-Kennfeldes anzusehen und zu speichern. Es können 16 Wörter aufgenommen werden. Vor dem Versenden der gesammelten Sprachmuster an den Mikrocontroller, wird dieser über Taster S4 in den Empfangsmodus versetzt. Eigentlich könnte der R8C die Speicherung der Sprachmuster auch selbst durchführen. Aber es scheint, daß vor jeder Modifikation des FlashRoms eine initiale Löschung von mind. 2KByte Größe nötig ist. Dies macht die interne Verwendung des FlashRoms etwas unflexibel.

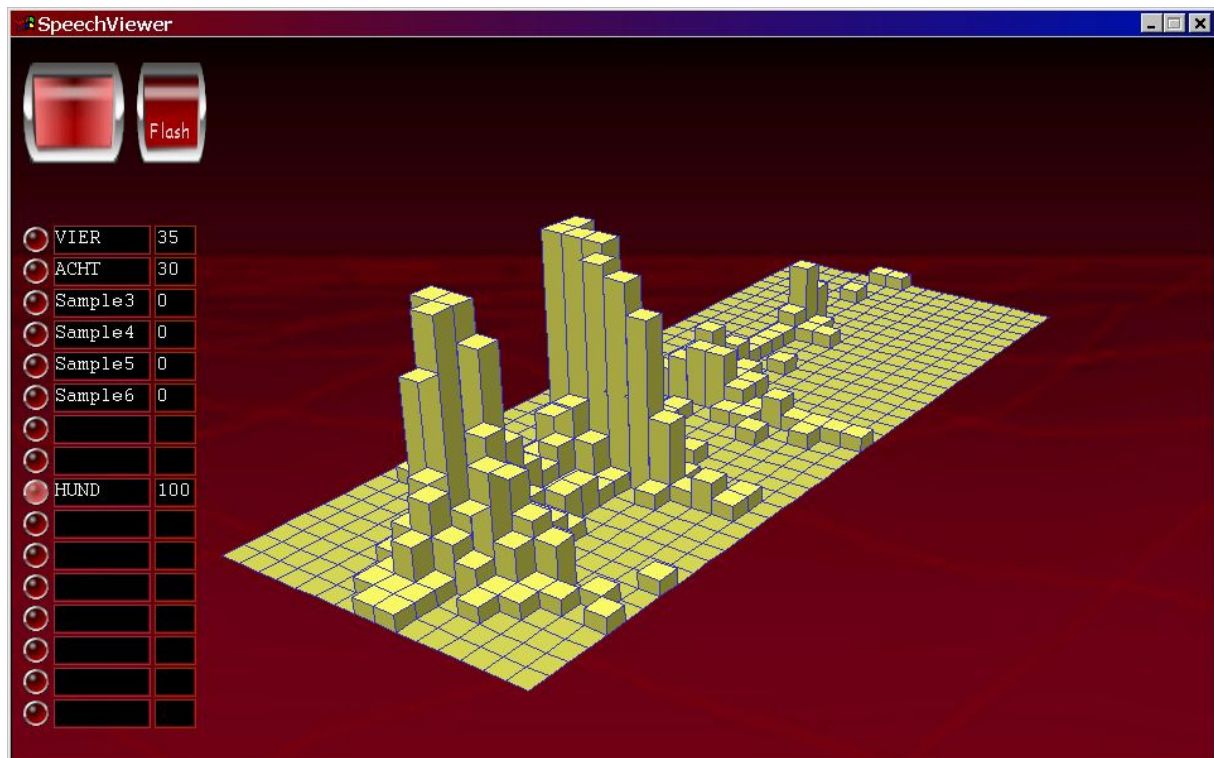


Bild3

Die Visualisierungssoftware SpeechViewer ist in C++ mit Hilfe des frei verfügbaren cygwin Toolkit's geschrieben (www.cygwin.com) und sollte auf jeder Windows Version zwischen Windows 95 und Windows XP laufen. Außer dem SpeechViewer Executable liegt auch der zugehörige Source Code bei. Er ist prinzipiell auch mit Windows spezifischen Compilern übersetzbar. Kleinere Anpassungen sind dann aber nötig.

Einschränkungen:

Zum aktuellen Zeitpunkt erkennt das System zwar problemlos gleich ausgesprochene Wörter, allerdings wie man mit Hilfe des SpeechViewers leicht erkennen kann, ist der Begriff "gleich" im wahrsten Sinne des Wortes wörtlich zu nehmen. Der gleiche Begriff führt bereits durch leicht geänderte Betonung zu deutlich anderen Frequenzverläufen. Störgeräusche und Hall sollten in jedem Fall vermieden werden.

Zumindest bei meinem Notebook stellte ich fest, daß der Ladevorgang des Akkus über das Netzteil starke Störungen über den USB/Seriell Adapter verbreitet. Entfernen des Akkus oder noch besser reiner Akkubetrieb ohne Netz beseitigen diesen Störfaktor.