

Robotron Elektronik Dresden
Stammbetrieb im Kombinat Robotron

Dr.-Ing.Lothar Seveke

Dresden, den 12.11.1985

B E S C H R E I B U N G

eines

S i g n a l e r k e n n u n g s - Z u s a t z e s

für

U 880 - Mikrorechner

Inhalt:

1. Systemberblick
2. Funktionsprinzip
 - 2.1. Signalanalyse
 - 2.2. Erkennung
 - 2.3. Lernvorgang
3. Gerätetechnik des Zusatzes
4. Programme für den Zusatz
 - 4.1. Programmstruktur
 - 4.2. Programmschnittstellen
 - 4.3. Einbindung in den Wirtsrechner
5. Bedienungsanleitung

1 . Systemüberblick

Der Signalerkennungszusatz dient der Erkennung von Signalklassen, von denen ihm in einem Lernvorgang Repräsentanten mit dem jeweils zugeordneten Klassennamen (Symbol) übermittelt wurden. Die Erkennung erfolgt durch den Vergleich einer numerischen Beschreibung des Eingangssignals (Meßwertmuster) mit den Mustern aus einem Referenzdatensatz, der während des Lernvorganges generiert wurde.

Das Symbol, das dem Muster zugeordnet ist, welches dem Eingangsmuster in einem bestimmten Sinne am ähnlichsten ist, wird dem Wirtsrechner übergeben.

Der Frequenzbereich der Signale am Eingang des Erkenners sollte innerhalb der Grenzen von 100 Hz bis 6 kHz liegen; die optimale Aussteuerung des Analysators nach (1) ist $U_{eff}=300$ mV.

Einmalige Vorgänge dürfen eine Dauer von 1,8s nicht überschreiten. Aus Dauersignalen können Abschnitte durch ein zusätzliches Steuersignal ausgeblendet werden (z.B. Tastenbetätigung am Anfang des Signalabschnittes, Ende durch Überlauf des Eingabepuffers 1,8s danach). die zeitliche Dynamik des Eingangssignals sollte so sein, daß die Meßwertvektoren (siehe 2.1) Signalabschnitte von 25 ms ausreichend genau beschreiben (abh. von notwendiger Trennschärfe).

Im Falle von menschlicher Lautsprache kann die sprachliche Einheit (Wort oder Wortfolge) eine Dauer von max. 1,8s haben. Die einzelnen sprachlichen Einheiten müssen durch Sprechpausen von mindestens 200 ms getrennt werden. Die Wiedererkennung erfolgt innerhalb dieser Zeit.

Der Modul ist einer Speicherkapazität von 4K Bytes RAM in der Lage, bis zu 200 verschiedene Muster aus einem determinierten Prozeß sicher wiederzuerkennen.

Bei einem stochastischen Erzeugungsprozeß, wie er der Lautsprache zugrunde liegt, muß die Symbolzahl wesentlich geringer als diese Musteranzahl sein, um Varianten einer Klasse speichern zu können. Maximal 50 Symbole sind für freizügig wählbare Wortschätze und für ausnahmearmes Sprecherpersonal als nutzbar anzusehen. Um natürlich bedingte Streuungen zu verringern, müssen die Sprecher im Lern- und Erkennungsvorgang deutlich und ungezwungen gleichmäßig sprechen und einen konstanten Mikrofonabstand einhalten. Dies kann eine längere Übungsphase mit dem Gerät erfordern.

Bei Verwendung des Kopfbügelmikrofons SP 75 (Hersteller VEB Funkwerk Köllda) sollte das Umgebungsgeräusch das in einem Büro übliche Maß (55 dB) nicht übersteigen. Höhere Störpegel erfordern den Einsatz spezieller, mundnah getragener Nahbesprechungsmikrofone mit Umgebungsschallkompensation (z.B. HME1019 Sennheiser, BRD) oder Schallisationsmaßnahmen für den Arbeitsplatz.

Die erzielbare Erkennungssicherheit hängt beim gegenwärtigen technischen Stand wesentlich vom kooperativen Verhalten des Sprechers, dem phonetischen Abstand der zu erkennenden Worte, den Umgebungsgeräuschen und der exakten Durchführung des Lernvorgangs ab.

Das Anlernen des Erkenners auf jede Signalquelle erfolgt in einer speziellen Form des Erkennungsvorganges. Die Repräsentanten der zu erkennenden Signalklassen werden solange in den Erkennenner eingegeben, bis er allen das richtige Symbol zuordnet (bei Lautsprache etwa 10 mal je Wortklasse). Der Bediener bestätigt oder korrigiert dazu die vom Erkennenner ausgegebenen Symbole. Der Lernvorgang ist durch das Einlesen des Referenzdatensatzes von einem externen Speichermedium ersetzbar.

Der Signalerkennungszusatz kann so für die Klassifizierung von Signalen eingesetzt werden, die sich durch ihre Spektren bzw. ihre Folgen von Kurzzeitspektren im o.g. Frequenzbereich unterscheiden. Die Analyseschärfe entspricht etwa der des menschlichen Ohres (ohne neuronale Nachverarbeitung).

Damit sind akustische Signale (z.B. Maschinen- oder Schweißgeräusche) aber auch entsprechend gewandelte Signale nicht-akustischen Ursprungs (z.B. Lichtintensitätsschwankungen) verarbeitbar.

Im Falle der Spracheingabe dient der Modul der Wiedererkennung von einzeln gesprochenen Worten von einem systembekannten Sprecher. Er kann als sogenannte "phonetische Tastatur" überall dort zur Anwendung kommen, wo herkömmliche Tastaturen aufgrund der Einsatzbedingungen nicht geeignet oder wenig effektiv sind. Solche Einsatzfälle sind beispielsweise Arbeitsplätze, an denen die Hände und der Gesichtssinn des Bedieners durch andere Aufgaben gebunden sind (z.B. grafische und Mikroskopiearbeitsplätze) oder Beleuchtungs- oder Witterungsverhältnisse oder Verschmutzung den Betrieb einer Tastatur erschweren. Beim Einsatz der Spracheingabe sollte jedoch auch beachtet werden, daß sie nicht als "Krücke" statt eines optimal geeigneten Mediums fungiert, daß z.B. der Sprecher nicht nur als opto akustischer Transformator an einem Problem arbeitet, das durch eine optische Zeichenerkennung oder Bildverarbeitung vollautomatisch zu lösen wäre.

Beim Betrieb als Spracheingabezusatz werden folgende Leistungsparameter erreicht:

- Wortschatzgröße : max. 50 Wortklassen
- Erkennungsquote bei kooperat. Sprecher verh.: 93%
- Rückweisungsquote 3%
- max. Dauer einer sprachl. Einheit : 1,8s
- min. notw. Pause zw. zwei Worten : 0,2s
- Reaktionszeit nach Wortende : 0,2s

Im Gegensatz zu Einplatinen-Signalerkennern, die als Peripherieprozessoren arbeiten (z.B. ESE K7824 von Robotron), nimmt der Zusatzmodul Kapazitäten des Wirtsrechners in Anspruch (Speicherraum, ZE-Zeit). Er kann nur eingesetzt werden, wenn während der Eingabe eines Signals keine anderen Forderungen an den Wirtsrechner bestehen. Für den Betrieb des Zusatzes werden 6K Bytes aus dem Speicherraum des Wirtsrechners (2K promfähiges Programm, 4K RAM) und drei aufeinanderfolgende Adressen aus dem EA-Adreßraum benötigt. Zur Speicherung des Referenzwissens für jeden Sprecher (4K Bytes) muß ein peripheres Speichermedium betrieben werden können. Zu einem autarken Signalerkennern kann der Zusatz durch Ergänzung mit handelsüblichen U880-Mikrorechnerbaugruppen (ZVE, Speicher, Interface) ausgebaut werden.

2. Funktionsprinzip

Die Signalerkennung beruht auf einer Struktur, die einen Signalanalysator, einen Speicher für das Referenzmuster und einen Vergleicher enthält. Für den notwendigen Lernvorgang existiert ein Referenzdatenverwalter, der die vom Signalanalysator gelieferten Meßwertmuster in geeigneter Weise in Referenzmuster umwandelt und im Referenzspeicher ablegt.

2.1. Signalanalyse

Das vom Signalwandler in elektrische Spannungsschwankungen gewandelte Signal wird vom Analysator des Zusatzmoduls durch eine Folge von Nulldurchgangshistogrammen numerisch beschrieben. Dazu wird das Eingangssignal in äquidistante Zeitabschnitte von 25 ms Dauer segmentiert. Die Häufigkeiten der Abstände zwischen benachbarten Nulldurchgängen der Signale in zwei Frequenzbereichen (unterhalb bzw. oberhalb 1 kHz), die zu einer von acht Abstandsklassen gehören, bilden das Histogramm für jedes Segment. Es entspricht so in grober Näherung einem 3-Kanal-Kurzzeit-Filterspektrum (Grenzen der Abstandsklassen = Bandpaßgrenzen).

Die anschließende Transformation, eine lineare Zeitnormalisierung, wandelt die Folge der Nulldurchgangshistogramme, deren Länge durch die Dauer des eingegebenen Signalabschnittes bestimmt ist, in eine Folge mit konstanter Länge um (16 Histogramme). Da jede Komponente der Histogramme durch ein Bit dargestellt wird, erhält man als Ergebnis der Signalanalyse eine numerische Beschreibung von 16 Bytes Speicherbedarf für jeden eingegebenen Signalabschnitt.

Dieser Vorgang wird begleitet von der Erkennung des Signalanfangs und -endes. Als Parameter dafür werden die Signal- und die Pausendauer herangezogen, die sich aus der Zahl von Segmenten ergeben, in denen die Komponentensummen der Histogramme vorgegebene Schwellwerte durchgängig über- bzw. unterschreiten.

2.2 Erkennung

Die Wiedererkennung eines eingegebenen Signalabschnittes beruht auf dem bitweisen Vergleich seiner Histogrammfolge mit allen (gleichartig organisierten) Folgen aus dem Referenzspeicher. Das Symbol, für dessen zugeordnete Folge das Abstandsmaß den geringsten Wert aufweist, wird als Erkennungsergebnis ausgegeben. Unterschreiten die Abstandswerte aller Muster im Referenzspeicher einen bestimmten Schwellwert nicht, wird die Entscheidung abgelehnt (Rückweisung).

2.3. Lernvorgang

Der Erkenner arbeitet am sichersten nur für die Signalquelle, auf die er angelernt wurde.

Für Lautsprache als Signal bedeutet dies, daß er sprecherabhängig ist. Jeder Sprecher paßt ihn durch einen Lernvorgang an den Wortschatz, der benötigt wird, die Sprache (beliebige Fremdsprachen außer Tonhöhen Sprachen), seine persönliche Sprechweise und die akustischen Umgebungsbedingungen an.

In diesem Lernvorgang wird das für die Erkennung notwendige Referenzwissen vollständig erworben. Die Änderung des Klassenvorrates (Wortschatz) erfordert also nur einen erneuten Lernvorgang.

Das Lernen erfolgt unter Rückkopplung mit dem Erkenner und besteht in der Abspeicherung der Histogrammfolgen der Lernprobe gemeinsam mit dem jeweils vom Bediener zugeordneten Klassensymbol. Nachdem ein Signalabschnitt der Lernprobe eingegeben wurde, gibt der Erkenner auf der Grundlage des bisher erworbenen Referenzwissens ein Symbol als Erkennungsergebnis aus. Mit Hilfe der Tastatur verifiziert der Bediener das Ergebnis oder korrigiert es durch Eingabe des richtigen Symbols. Die Aufnahme der Histogrammfolge des gerade eingegebenen Signalabschnittes in den Referenzspeicher erfolgt nur, wenn sie zu den bisher schon für diese Klasse abgespeicherten Folgen in einem bestimmten Masse unähnlich ist. Referenzmuster, die bei der Erkennung nur selten eine hohe Ähnlichkeit mit dem Muster eines eingegebenen Signalabschnittes der zugehörigen Klasse erreichen, werden aus dem Referenzspeicher wieder gestrichen. Dieses Streichen kann nur in dem Teil des Referenzsatzes erfolgen, der mit der gerade aktuellen Probenkennung, die am Beginn jedes Lernvorgangs angegeben werden muß, gekennzeichnet ist. Damit bleiben Muster, die von früheren Lernvorgängen stammen, erhalten.

Der Referenzspeicher enthält so nach dem Lernvorgang für jede Klasse mehrere Muster (Folgen von Nulldurchgangshistogrammen), die im Falle von Lautsprache je nach Geschick des Sprechers seine Aussprachevarianten für ein Wort repräsentieren.

3. Gerätetechnik für den Zusatz

Die Zusatzschaltung hat folgende Aufgaben zu erfüllen:

1. Anpassung an den Signalwandler
2. Selektierung des Signals in zwei Frequenzkanäle
3. absolute Begrenzung der Signale in beiden Kanälen
4. Bestimmung der Mulldurchgangsabstände in beiden Kanälen (programmgestützt)

5. Anpassung an den Rechnerbus
6. optische Bereitschaftsanzeige (optional)

Für den Betrieb als Worterkenner mit dem o.g. Kopfbügelmikrofon SP 75 liegt mit (1) ein entsprechender Schaltungsvorschlag vor. Er beinhaltet einen Mikrofonvorverstärker (1x B761), zwei aktive, RC-Filter (1/2 B4761), zwei Amplitudenbegrenzer (1/2 B4761, 1/3 DL014) und etwa 15 Kondensatoren und 30 Widerstände im Analogteil. Der Digitalteil benötigt einen programmierbaren Zähler/Zeitgeber (CTC U857), einen Adreßdeko­der (1x U8205) und geringfügige Randelektronik (U8286, DL...). Die Schaltung ist problemlos auf einer Leiterplatte der Kategorie IV mit den Massen 95 mm x170 mm aufzubauen.

4. Programme für den Zusatzmodul

4.1 Programmstruktur

Das Programmpaket für den Betrieb des. Signalerkenners umfaßt ein Hauptprogramm für das Anlernen mit je einem Eintrag für das Neulernen (leerer Referenzspeicher) und das Weiterlernen (aufbauend auf schon vorhandenem Referenzwissen) und ein Unterprogramm für die Erkennung eines Signalabschnittes. Das Programmpaket ist als Assembler-Quellprogramm verfügbar. Das zugehörige Maschinenprogramm mit einem Speicherbedarf von 2 KBytes ist promfähig und kann auf beliebige 1K-Grenzen im Speicherraum des Wirtsrechners gelinkt werden. Es wird ein Arbeitsspeicher von 4K Bytes benötigt, der an beliebigen 1K-Grenzen beginnen kann.

Das Programmpaket gliedert sich in drei wesentliche funktionelle Segmente:

1. Anpassung an Wirtsrechner
2. Steuerroutinen für Lernen und Erkennen mit Hilfsprogrammen
3. Unterprogramme für Signalanalyse, Erkennung und Lernen

Die Segmente 1 und 2 können durch den Nutzer modifiziert werden. In das Segment 3 sollte nicht eingegriffen werden. Es realisiert die folgenden Funktionen (verwendete Namen siehe (2)):

Jeder Signalabschnitt wird durch die Unterprogramme UP2, UP7, UP9...UP13 in ein Meßwertmuster der Länge VANZ abgebildet, das in XFELD abgespeichert wird. Um die Vergleichbarkeit von Mustern unterschiedlicher Länge zu gewährleisten, bildet das Unterprogramm UP5 XFELD auf YFELD ab. YFELD enthält dann immer Muster einheitlicher Länge.

Im Lernvorgang werden durch das Programm UP1 jedem verwendeten Klassensymbol ein oder mehrere Referenzmuster in RFELD zugeordnet. Zugleich erhält jedes Referenzmuster eine zusätzliche

Bewertung in BFELD, die über dessen Wichtigkeit innerhalb einer Klasse Auskunft gibt. Während des Lernens werden diese Bewertungen verstellt. Unterschreitet eine Bewertung eine vorgegebene Schwelle, führt das zum Streichen des betreffenden Referenzmusters. Die Zahl der Referenzmuster wird in der Variablen KLAZ abgespeichert.

Das Unterprogramm UP4 sucht aus den in RFELD abgespeicherten Referenzmustern das zu YFELD ähnlichste aus. Die entsprechende Ähnlichkeit wird auf die Variable MINHD geschrieben. Beim Lernen entscheidet die Größe von MINHD über die Eröffnung eines neuen Referenzmusters, beim Erkennen über das Erkennungsergebnis.

4.2 Programmschnittstellen

Die Einträge des Hauptprogramms LERNEN werden über folgende Adressen aufgerufen:

- NEULERN : Adr. PROM (Beginn des Programmbereiches)
(Lernen mit anfangs leerem Referenzspeicher)
- WEILERN : Adr. PROM+0CH (Lernen, aufbauend auf vorhandenem Referenzwissen)

Nach Abschluß des Lernprogramms (siehe. 5.) erfolgt ein Warmstart des Betriebssystems. In den 4K Bytes des Arbeitsspeichers liegt ein Referenzdatensatz vor, der auf ein Peripheres Medium ausgelagert werden kann bzw. als Basis für die Abarbeitung des Unterprogramms RECOG (Erkennung eines Signalabschnittes) dienen kann. Der Aufruf dieses Unterprogrammes durch den Nutzer erfolgt über die Adresse

PROM+19H

Eine Parameterübergabe beim Aufruf ist nicht notwendig. Der Referenzdatensatz muß vorhanden sein!). Mit dem Aufruf des UP wird die Eingabebereitschaft des Signalanalysators hergestellt. Das UP wird verlassen, nach Abschluß der Erkennung spätestens 200 ms nach Überlauf des Eingabepuffers für 1,8s Signaldauer oder nach Erkennung der Pause am Ende eines Signalabschnittes), wenn nicht auf Rückweisung (Code 00H) erkannt wird. Eine Rückweisung löst über das UP BEEP akustisches Signal aus (wenn entspr. programmiert) und führt zur erneuten Eingabebereitschaft. Nach der Rückkehr aus dem UP liegt im A-Register und in den Speicherzellen ERKER und XSYM das Symbol der erkannten Signalklasse vor.

Bit 7 dieses Bytes ist gesetzt, wenn die Signaleingabe durch eine Tastatureingabe abgebrochen wurde (wenn UP INKEY als Tastaturabfrage wirksam). Die Bits 0...6 des Bytes sind in diesem Fall nicht relevant.

4.3. Einbindung in den Wirtsrechner

Die Erläuterung der programmtechnischen Rechneranpassung erfolgt anhand der Markensymbole in (2), die EQU-Anweisungen, Unterprogramme bzw. Stringdefinitionen bezeichnen.

Folgende Speicheradressen sind entsprechend der Speicherbelegung im Wirtsrechner zu modifizieren:

PROM - Beginn des 2K-Bytes-Programmbereiches
(auf 1K-Grenze!)

RAMA - Beginn des 4K-Bytes-Arbeitsspeichers
(auf 1K-Grenze!)

OPSYS - Warmstart-Adresse des Betriebssystems

ERKER - Adresse der Speicherzelle für die Übergabe des Erkennungsergebnisses (1 Byte)

CTKO - Beginn des Freibereiches in der Interrupt-Tabelle des Betriebssystems (6: Bytes für CTC-Kanäle 0...2)

Drei aufeinanderfolgende Adressen im E/A-Adressraum sind zu suchen:

CTC0 - erste der drei E/A-Adressen. (Bits 0 und 1 = 0 !), die auch am Adressdecoder des Moduls einzustellen ist

Für die Kommunikation sind E/A-Symbole zu vereinbaren:

EE - Codes des Zeichens "Eingabeende" von der Tastatur

NL - Code des Zeichens, das die Funktion "Neue Zeile" auf der Anzeigeeinrichtung erzeugt

NULL - Code des Zeichens "0"

NEUN - Code des Zeichens "9"

(In den UP's HLLAD und ZZEIL in (2) wird für alle E/A-Symbole der ASCII-Zeichenvorrat vorausgesetzt. Sie sind also bei anderen Codes zu modifizieren.)

Folgende Kommunikations - Unterprogramme sind mit Hilfe des Betriebssystems des Wirtsrechners zu realisieren:

TEIBD - Eingabe eines Zeichens von der Tastatur

TAUBD - Ausgabe eines Zeichens zur Anzeigeeinrichtung

BEEP - Erzeugung eines akustischen Signals

BEREI - Einschalten einer Bereitschaftsanzeige z.B. über Kanal 3 des CCC und Blockierung aller anderen Interruptquellen im Wirtsrechner

BERAU - Ausschalten der Bereitschaftsanzeige und Freigabe der mit BEREI gesperrten Interruptquellen INKEY - Abfrage Tastaturstatus für den Abbruch einer laufenden Signaleingabe (nur über polling möglich!)

Durch den Nutzer können die Texte für die Kommunikation gewählt werden:

EING - Aufforderung zur Eingabe über die Tastatur

MUANZ - Prätext für die Anzeige der Musteranzahl
PRONR - Aufforderung zur Eingabe der Probennummer
SPREC - Aufforderung zum Sprechen
(Als Endekennzeichen für die Texte wird in (2) das gesetzte Bit 7 eines Textzeichens verwendet. Für andere Endekennzeichen ist das UP ANZM1 zu modifizieren.)

Sollte durch die Veränderung der Kommunikations-UPs und -texte beim Assemblieren die Marke UP1 (PROM+2E0H) überschritten werden, sind diese nach dem Ende der Tabelle KTAB anzuordnen. Die Programmlänge von: 2K Bytes wird dabei überschritten!

Neben diesen programmtechnischen Anpassungen können zwei interne Programmkonstanten durch den Nutzer experimentell variiert werden, um eine optimale Anpassung des Erkenners an die Einsatzbedingungen zu erreichen:

RUESW - Rückweisungsschwelle
Eine Vergrößerung des Standardwertes von 30 führt zu einer Verringerung der Fehlentscheidungen, verbunden mit einem Ansteigen der Rückweisungsquote.
(RUESW = 0 --> keine Rückweisungen)

SWCLU - Schwellwert für das Streichen von Mustern
Eine Vergrößerung des Standardwertes von 8 verringert die Musteranzahl, mit der eine Klasse im Referenzspeicher vertreten ist. Der notwendige Speicherbedarf für die Speicherung eines Klassenvorrates wird so gemindert, gleichzeitig wird jedoch die mögliche Variationsbreite innerhalb einer Signalklasse eingeschränkt.

Vor dem Start eines der Programme aus dem Erkennerpaket ist durch das Betriebssystem zu realisieren, daß:

- der Interruptmodus 2 eingestellt ist,
- der Stackbereich definiert ist und
- das I-Register den Wert HIGH(CTCKO) hat.

5. Bedienungsanleitung

Eine Bedienung ist nur für das Lernprogramm erforderlich: die Erkennung läuft ohne Kommunikation mit dem Nutzer ab.

Ist das Programm nicht auf ROM gespeichert, muß es vom peripheren Medium geladen werden. Anschließend ist entweder das Programm NEULERN oder das Programm WEILERN in geeigneter Weise (abh. vom Wirtsrechner) zu starten. Für die Abarbeitung von WEILERN ist ein Referenzdatensatz im Arbeitsspeicher erforderlich. Er muß ggf. vor dem Start des Programms vom Speichermedium geladen werden.

NEULERN beginnt das Lernen mit leerem Referenzspeicher und löscht einen eventuell vorhandenen Referenzdatensatz.
WEILERN baut einen bereits vorhandenen Datensatz weiter aus.

Beide Programme zeigen nach dem Start die Anzahl der im Arbeitsspeicher abgelegten Referenzmuster in Verbindung mit dem gewählten Kommunikationstext an (max. 200). Dem Nutzer wird damit eine Information über weitere Möglichkeiten zum Nachlernen gegeben. Durch einen Abbruch des Lernvorgangs und einen anschließenden Aufruf von WEILERN ist eine ständige Kontrolle des Füllstandes des Referenzspeichers möglich.

Anschließend wird der Nutzer zur Eingabe der Probenkennung aufgefordert. Die Probenkennungen 0...4 sind möglich. Ein Referenzsatz kann also aus max. 5 unabhängigen Lernproben aufgebaut werden. Die Kennzeichnung der Probe dient der notwendigen Separierung der von verschiedenen Nachlernvorgängen stammenden Referenzmuster, um Verdeckungseffekte zu vermeiden.

Die Programmvariante in (2) sieht für alle Eingaben bei der Kommunikation nur Ziffernfolgen mit Endekennzeichen vor. Alle Ersteingaben, die keine Ziffer bzw. Endekennzeichen darstellen, führen zum Abbruch des Lernvorgangs und zum Warmstart des Betriebssystems (einzige Möglichkeit zum Beenden des Anlernens). Alle Zeichen/Ziffern in einer Ziffernfolge werden ignoriert (Ausgabe "?"). Die Korrektur einer fehlerhaften Eingabe vor dem Abschluß mit dem Endekennzeichen kann durch die Wiederholung der zweistelligen Eingabe erfolgen, da nur die beiden letzten Ziffern signifikant sind.

Nach Eingabe der Probenkennung wird die Eingabebereitschaft des Erkenners angezeigt, und eine Realisierung einer Klasse ist einzugeben.

Nach dem Ende Signalabschnittes zeigt das Programm das Erkennungsergebnis in Form des Klassensymbols an (Durch Änderung von (2) hat der Nutzer die Möglichkeit, statt der Klassennummer hier einen Text als Klassennamen auszugeben.).

Die Klassennummer "0" hat die Bedeutung einer Rückweisung. Der Bediener korrigiert durch eine Zifferneingabe das Erkennungsergebnis, wenn rückgewiesen oder falsch erkannt wurde, oder bestätigt es durch Eingabe des Endekennzeichens. Die Eingabe einer "0" führt zur Nichtbeachtung des gerade eingegebenen Signalabschnittes (Korrekturmöglichkeit für fehlerhafte Eingaben). Es sind nur die Klassennummern 1...50 erlaubt.

Durch die immer sicherer werdende Erkennung mit zunehmender Zahl der Lernproben ist der Erfolg des Lernvorgangs zu beobachten. Ebenso wird eventuell deutlich, welche Klassen sich durch den Erkenner nicht unterscheiden lassen.

Im Falle der Lautsprache sind das phonetisch ähnliche Worte. Dann sollte eines der Worte durch ein Synonym ersetzt werden (z.B. ständige Verwechslung "zwei" mit "drei" --> "zwo" statt "zwei").

Fehleingaben von Klassensymbolen sind möglichst zu vermeiden, da sie zu falsch gekennzeichneten Mustern im Referenzspeicher führen, die nur durch viele Realisierungen der Klasse, die dem falsch eingegebenen Symbol entspricht, wieder aus dem Referenzdatensatz "verdrängt" werden können.

Der Lernvorgang kann abgebrochen werden, wenn alle Realisierungen der Lernstichprobe stabil richtig wiedererkannt werden. Ein objektives Maß für die Beendbarkeit des Lernens ist jedoch nur das Wiederabnehmen der Musteranzahl bei weiterer Eingabe vor Lernproben ("Erkennung weiß genug über die Signalquelle."), das über die Anzeige der Musteranzahl beim Start von WEILERN kontrollierbar ist. Konvergiert die Musteranzahl nicht, erzeugt die Quelle Signale mit zu großer Streuung (Sprecher artikuliert zu ungleichmäßig), oder die Signale sind durch Fremdeinflüsse (z.B. Umgebungsgeräusche) zu stark gestört. Der Lernvorgang muß dann bei gefülltem Referenzspeicher abgebrochen werden.

Im Anschluß an einen erfolgreichen Lernvorgang sollte das Referenzwissen (4K Bytes RAM) auf ein peripheres Speichermedium ausgelagert werden.

Vor der Nutzung eines solchen Referenzsatzes für eine Erkennung zu einer späteren Zeit empfiehlt es sich, über WEILERN kurz nachzulernen, um den Datensatz an die aktuelle Situation der Quelle zu adaptieren.

Von wesentlicher Bedeutung für die Erkennungssicherheit bei der Verarbeitung von Lautsprache ist die Auswahl eines geeigneten Mikrofons und die Arbeit mit diesem. Es ist unbedingt ein mundnah zu tragendes, nahbesprechbares Mikro (ähnlich dem o.g. SP 75) zu verwenden. Das Mikrofon darf seine Lage (etwa 1 cm seitlich vor dem Mund) durch Bewegung des Kopfes nicht verändern und muß diese Lage auch nach dem Ab- und Aufsetzen wieder einnehmen. Laryngophone (Kehlkopfmikrophone) sind aufgrund der stark beschränkten Höhenwiedergabe nicht einsetzbar.

In der Kommunikation auftretende Fehler rufen eine Wiederholung der Eingabeaufforderung hervor. Der Überlauf des Referenzspeichers führt zum Abbruch des Lernvorgangs und anschließendem Warmstart des Betriebssystems.

- (1) Schaltungsvorschlag für einen Spracheingabe-Zusatzmodul für den Kleincomputer Z9001 vom VEB Meßelektronik Otto Schön Dresden
Dr.-Ing. Ulrich Kordon, TU Dresden, Sektion 9, Bereich 4
Tel.: Dresden 463 2240
- (2) Assembler-Quellprogramm für den KC Z9001 zum Betrieb eines Spracheingabezusatzmoduls
Dr.-Ing. Lothar Seveke, Robotron Elektronik Dresden
Tel.: Dresden 487 2509