

Sonderdruck  
aus  
ARCHIV DER ELEKTRISCHEN ÜBERTRAGUNG

Phonetische Addiermaschine  
von HANS-GEORG MUSMANN und KARL-HEINZ STEINER

## Phonetische Addiermaschine

VON HANS-GEORG MUSMANN und KARL-HEINZ STEINER

Mitteilung aus dem Institut für Fernmelde- und Hochfrequenztechnik  
der Technischen Hochschule Braunschweig

(A. E. Ü. 19 [1965]. 502–510; eingegangen am 20. Mai 1965)

DK 621.14.-523.8:681.89

Bei bisher verwendeten Rechnern werden Ziffern und Befehle fast ausschließlich auf mechanischem Wege über eine Tastatur, eine Stecktafel oder über Lochstreifen eingegeben. Die automatische Erkennung gesprochener Worte ermöglicht eine akustische Eingabe von Ziffern und Befehlen. Dieser Rechner enthält daher im wesentlichen zwei Teile, das Rechenggerät und das Erkennungsgerät. Das angewandte Verfahren der automatischen Spracherkennung wird beschrieben. Die auftretenden Probleme, die erzielten Ergebnisse sowie Ideen zur Weiterentwicklung werden diskutiert. Das Rechenggerät wurde zur Studie gebaut und in Verbindung mit dem Erkennungsgerät nur zur Demonstration verwendet, es wird deshalb nur in großen Zügen beschrieben.

Into computers used so far digital information and commands have been introduced nearly invariably in a mechanical way via keyboard, patchboard, or perforated tape. Automatic recognition of the spoken word allows an acoustical introduction of digits and commands. This computer contains therefore essentially two parts, the computer and the recognition device. The adopted method of automatic speech recognition is described. The problems encountered, the results obtained as well as ideas for advanced development are discussed. The computer was constructed as a study and it is used in conjunction with the recognition device for demonstrations only; it is therefore described merely along broad lines.

### 1. Einführung

Die Funktionsweise eines digitalen Tischrechners wird beschrieben, der sowohl mechanisch durch Tastendruck als auch phonetisch durch gesprochene Befehle betätigt werden kann. Der Rechner enthält im wesentlichen zwei Teile, das Erkennungsgerät und das Rechenggerät, siehe Bild 1. Im Erkennungs-

gerät werden die akustischen Befehle über die automatische Erkennung in digitale Befehle gewandelt. Es werden die Ziffern 0 bis 9 erkannt, sowie die Befehle „plus“, „minus“, „start“, „löschen“. Im Rechenggerät werden die Befehle digital gespeichert und verarbeitet. Addition und Subtraktion werden in einem digitalen Addierwerk ausgeführt, die Ergebnisse dezimal angezeigt.

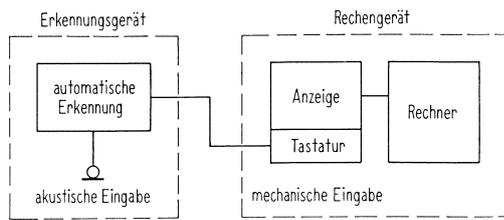


Bild 1. Blockschaltbild des Rechners.

In Arbeiten jüngster Zeit [1] – [5] werden Vorschläge zur automatischen Spracherkennung und untersuchte Verfahren beschrieben. Sie weisen auf die Schwierigkeiten automatischer Erkennung hin, die im wesentlichen in den linguistischen Eigenheiten der Sprecher liegen und im Unwissen über die Funktionen des menschlichen Gehörs in Zusammenarbeit mit dem Verstand. Wertvolle Hilfe bieten schon die bisher aus Untersuchungen der Physiologie des Hörens gewonnenen Erkenntnisse [6], [7], [8]. Die bisher entwickelten Geräte einschließlich des hier beschriebenen können nur als Versuche zur automatischen Spracherkennung bezeichnet werden, ihre Ergebnisse sind für eine praktische Anwendung noch meist unbefriedigend.

Das hier beschriebene Verfahren zur automatischen Spracherkennung basiert auf einer Analyse der Spannungsamplituden über Zeit und Frequenz des gewandelten Signals. Daran schließt sich die Umformung dieser Signale in binäre Informationen und das Aufsuchen charakteristischer Merkmale des gesprochenen Wortes, die von einer binären Logik erkannt werden. Die Beschreibung des Gerätes zur automatischen Spracherkennung und die Diskussion der erzielten Ergebnisse umfaßt den größten Teil der vorliegenden Arbeit.

## 2. Das Erkennungsgerät

### 2.1. Prinzip der Erkennung

Dem hier verwendeten Prinzip der Erkennung liegt der Gedanke zugrunde, nach Wandlung des akustischen Signals in elektrische Schwingungen durch das Mikrophon, die charakteristischen Merkmale des gesprochenen Wortes aus der jedem Wort eigenen Verteilung der Spannung über Frequenz und Zeit aufzusuchen. Diese Verteilung kann man sich als

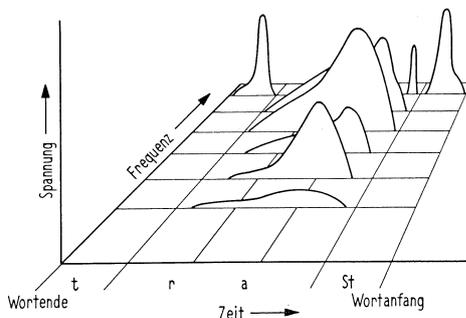


Bild 2. Relief der Spannungsverteilung des Wortes „start“ nach Vorverzerrung.

dreidimensionales Gebilde vorstellen, wie es zur Veranschaulichung in Bild 2 für das Wort „start“ skizziert ist. Um binäre Kriterien zu erhalten, wird dieses noch analoge Informationsvolumen digitalisiert [1]. Dafür wird die Frequenzachse durch zehn Filter in zehn Stufen, die Zeitachse durch einen Taktgeber in drei Stufen und die Spannungsamplitude durch einen Schwellwert in zwei Stufen geteilt. Der Schwellwertfühler, ein Schmitt-Trigger, liefert im Ruhezustand die Aussage „0“ und bei Erreichen des Schwellwertes die Aussage „L“, der Informationsgehalt entspricht einem Bit. Ordnet man jedem Filter einen Schmitt-Trigger zu und fragt diesen im Verlauf eines Wortes dreimal ab, so erhält man ein binäres Informationsvolumen von  $10 \times 3 = 30$  bit. Das Bild 3 zeigt dieses Informationsvolumen mit einer dem Wort „start“ charakteristischen Verteilung. Die gesamte binäre Information wird in einem Speicher von 30 bit gespeichert und von einer Bedeutungsmatrix entschlüsselt, damit ist das gesprochene Wort erkannt.

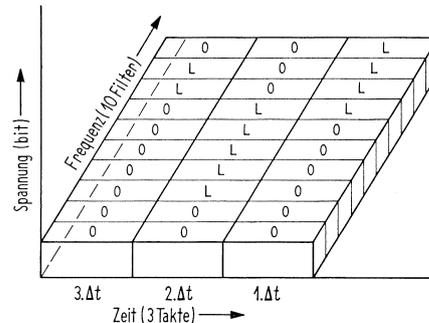


Bild 3. Charakteristisches binäres Informationsvolumen des Wortes „start“.

### 2.2. Prinzip der technischen Ausführung

Das elektrische Schwingungsgemisch des Mikrophons wird über einen Verstärker auf einen Satz von zehn parallelen Filtern geleitet, vgl. Bild 4. Die

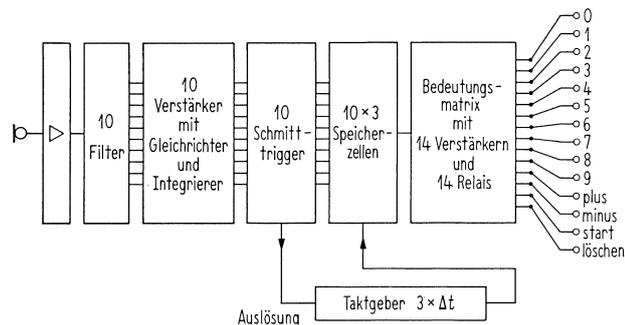


Bild 4. Blockschaltbild des Erkennungsgerätes.

Filterausgangsspannung eines jeden Filters wird gleichgerichtet, integriert und durch je einen Schmitt-Trigger in die Binärziffern „0“ oder „L“ gewandelt.

Mit Wortbeginn läuft ein Taktgeber an, der das gesprochene Wort in drei definierte Zeitintervalle  $\Delta t$  teilt. Der Taktgeber steuert den Informationsfluß

Schmitt-Trigger in den Speicher. Jedem Schmitt-Trigger sind drei Speicherzellen angeschlossen, von denen jede Zelle einem Zeitintervall zugeordnet ist und nur die „0“- oder „1“-Information dieses Zeitintervalls  $\Delta t$  aufnimmt. Der gesamte Speicher umfaßt entsprechend zehn Schmitt-Triggern  $10 \times 3 = 30$  Zellen bzw. bit. Nach Ablauf von drei Zeitintervallen ist der Speicher gefüllt, er enthält das Informationsvolumen in binärer Form.

Der Speicherinhalt wird im letzten Schritt der Erkennung von einer Bedeutungsmatrix, bestehend aus Und- und Oder-Verknüpfungen, entschlüsselt. Die Matrix hat für jedes erkennbare Wort einen definierten Ausgang mit Verstärker und angeschlossenem Relais, für 10 Ziffern und 4 Operationsbefehle insgesamt 14 Relais. Bei Erkennung zieht nur jeweils ein Relais an, es schaltet die erkannte Ziffer oder den Befehl in das Rechenwerk. Die Relaiskontakte liegen den Tasten der mechanischen Eingabe parallel.

2.3. Die Bausteine des Erkennungsgerätes und ihre Funktion

2.3.1. Mikrophon und Verstärker

In Voruntersuchungen wurde der Frequenzgehalt der Mikrophonspannung für die zu erkennenden Worte analysiert. Am Beispiel des Wortes „start“ zeigte sich, daß das Spektrum im wesentlichen nur die Grundfrequenz der Lautgruppe „ar“ und den ersten Formanten des Vokals „a“ enthielt. Die weiteren Formanten sowie die höheren Frequenzen der Laute „st“ und „t“ gehen praktisch unter. Bei einer binären Amplitudenstufung durch nur einen Schwellwert könnten diese Laute nicht erfaßt werden. Da das menschliche Ohr jedoch auch diese Laute höherer Frequenzen wahrnimmt, sie sogar als gleichlaut mit den tieffrequenten Lauten empfindet, liegt die Vermutung nahe, daß unsere Sprache sich dem Gehör angepaßt hat. Möchte man für eine automatische Spracherkennung gleiche Voraussetzungen, wie für die physiologische Erkennung schaffen, so muß das frequenzabhängige Hörempfinden des menschlichen Ohres bei der künstlichen Wahrnehmung durch eine Vorverzerrung des Signals berücksichtigt werden. Die größere Empfindlichkeit des Gehörs bei höheren Frequenzen muß durch eine Anhebung dieser Frequenzen, z. B. mittels eines RC-Gliedes, im Mikrophonverstärker ausgeglichen werden. In Bild 5 ist der experimentell gefundene Frequenzgang des

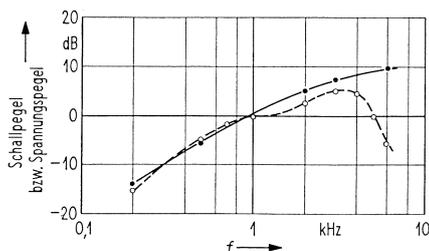


Bild 5. Frequenzgang des Mikrophonverstärkers (—) und an der Abszisse gespiegelte Kurve der Hörempfindung für 30 phon (---).

Verstärkers aufgetragen, der erforderlich war, um die Laute „st“ und „t“ bis zur gleichen Spannungsamplitude der Lautgruppe „ar“ anzuheben. Ein Vergleich des Frequenzganges mit der an der Abszisse gespiegelten Kurve der Hörempfindung [9] zeigt, daß die experimentell gefundene Lösung gleichzeitig eine über weite Bereiche gültige Kompensation der Hörempfindung bedeutet. Die eingangs erläuterten Gedanken werden dadurch bestätigt.

Die Empfindlichkeit des Mikrophonverstärkers kann bei unterschiedlichen Lautstärken der Sprecher geregelt werden.

2.3.2. Filter

Das am Ausgang des Mikrophonverstärkers gemessene Spektrum der Testwörter liegt zwischen 200 Hz und 6,6 kHz, dieser Frequenzbereich wird durch zehn Filter unterteilt. Die Wahl der Filterbandbreiten wurde ebenfalls den Verhältnissen des menschlichen Ohres angepaßt. R. FELDTKELLER und E. ZWICKER [7] haben in Untersuchungen einen Frequenzgruppeneffekt des Ohres festgestellt und empfehlen demzufolge für Analysen eine Filterstaffelung nach der durch diese Experimente gefundenen sogenannten melodischen Tonkala (Bild 6). In der Empfindung der Lautstärke faßt das Ohr jeweils benachbarte Frequenzkomponenten eines Schalles zusammen und bildet damit sogenannte Frequenzgruppen. Der gesamte Hörbereich unterteilt sich in 24 solcher Gruppen, die Breite jeder einzelnen beträgt bei allen Frequenzen konstant etwa 100 mel. In analoger Weise wurden für die digitale Zerlegung des verwendeten Frequenzbandes Filter konstanter Bandbreite in mel gewählt. Der Frequenzbereich von 200 Hz bis 6,6 kHz wurde dafür in zehn gleiche mel-Stufen von  $10 \times 180$  mel geteilt, die zugehöri-

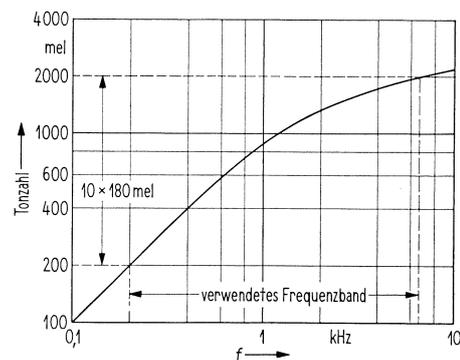


Bild 6. Melodische Tonkala nach FELDTKELLER und ZWICKER.

Filterabstufung

Filter	mel	Hz
1	200 — 380	200 — 390
2	380 — 560	390 — 590
3	560 — 740	590 — 810
4	740 — 920	810 — 1100
5	920 — 1100	1100 — 1480
6	1100 — 1280	1480 — 1980
7	1280 — 1460	1980 — 2600
8	1460 — 1640	2600 — 3400
9	1640 — 1820	3400 — 4600
10	1820 — 2000	4600 — 6600

gen Bandbreiten in Hz können dem Bild 6 entnommen werden.

Durch diese Staffelung der Filterbandbreiten werden die mit der Frequenz anwachsenden Frequenzstreuungen beim Sprechen im beschränkten Maße ausgeglichen.

### 2.3.3. Verstärkung, Gleichrichtung, Integration

Für die Abtastung durch Schmitt-Trigger müssen die Filterausgangsspannungen verstärkt und gleichgerichtet werden, siehe Bild 7. Eine exakte Spitzengleichrichtung erwies sich als ungünstig, da kurzzeitige Geräusche bzw. zufällige Spannungsspitzen

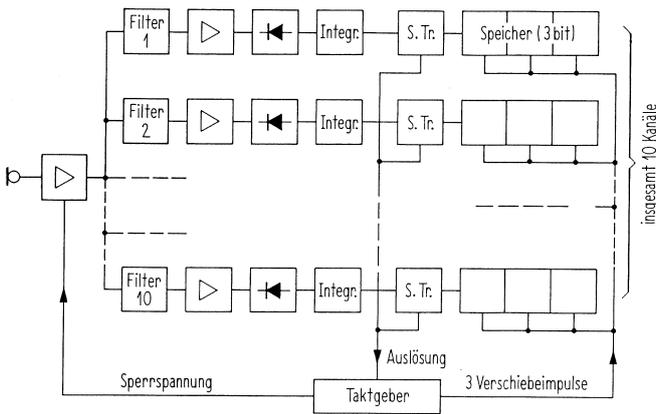


Bild 7. Digitale Wandlung und Speicherung.

die Schmitt-Trigger auslösen und damit unsichere Informationen liefern. Diese impulsartigen Geräusche werden vom menschlichen Gehör nicht bewußt wahrgenommen und tragen nicht zur Erkennung bei. Sie liegen unterhalb der Tonkennzeit  $t_T$  (vgl. Bild 8), der Mindestzeit, die das Ohr braucht, um einen Ton

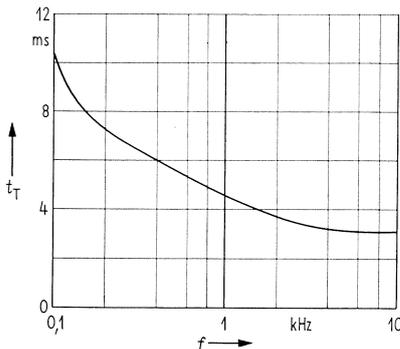


Bild 8. Minimal erforderliche Tonkennzeiten in Abhängigkeit von der Frequenz nach TRENDLEBURG [6].

bestimmter Frequenz als solchen zu erkennen [6]. Zur Unterdrückung dieser Störerscheinungen wurden deshalb Ladewiderstand  $R_a$  und Kapazität  $C_a$  des Gleichrichters Bild 9 als Integrierglied dimensioniert, die Zeitkonstante  $\tau_a = R_a C_a = 0,5 t_T$  gewählt. Da die Tonkennzeit  $t_T$  frequenzabhängig ist, mußten für die Filter unterschiedlich bemessene Integrierglieder verwendet werden. Die Entladezeitkonstante wurde  $\tau_e = R_e C_a = 2,3 t_T$  gewählt.

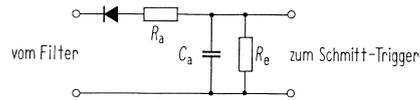


Bild 9. Gleichrichter mit Integrierglied.

Auf diese Weise wird eine Amplitudenglättung der Signale erzielt, ohne daß für die Erkennung wichtige Informationen unterdrückt werden.

### 2.3.4. Schmitt-Trigger und Speicher

Alle Schmitt-Trigger sind auf den gleichen festen Schwellwert eingestellt. Der erste mit Wortbeginn kippende Schmitt-Trigger löst den Taktgeber aus.

Beim Kippen schreibt der Ausgangsimpuls des Schmitt-Triggers in den als Schieberegister geschalteten dreizehligen Speicher ein „L“ ein. Nach Ablauf eines Zeitintervalls wird der Speicherinhalt um eine Stelle nach rechts verschoben, jedem Schmitt-Trigger steht damit für jedes neue Zeitintervall eine leere Speicherzelle zur Einspeicherung eines Bits zur Verfügung. Ein einmal angeklungener Laut erzeugt beim Übergang von einem in das nächste Zeitintervall keine neue L-Information, da die Schmitt-Trigger nur zu Beginn eines Lautes einen Ausgangsimpuls liefern. Mit dem dritten Zeitintervall ist die digitale Wandlung und Einspeicherung beendet. Der Mikrophonverstärker wird vom Taktgeber für die Zeitdauer der sich anschließenden Entschlüsselung gesperrt.

### 2.3.5. Taktgeber

Der Wahl der drei Taktzeiten lag der Gedanke zugrunde, sie so zu bemessen, daß Wörter wie „start“, „sechs“, „zehn“ mit ihren Anfangskonsonanten in das erste, mit dem Vokal in das zweite und mit dem Endkonsonanten in das dritte Intervall fallen. In Bild 10 ist für das Testwort „start“ die Häufigkeits-

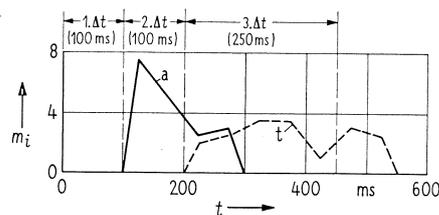


Bild 10. Häufigkeitsverteilung für den zeitlichen Beginn des Vokals „a“ und den Endkonsonanten „t“ beim Wort „start“ ( $\sum m_i = 18$  Sprecher).

verteilung für den zeitlichen Beginn des Vokals „a“ und des Endkonsonanten „t“ aufgetragen; dazu die gewählten Taktzeiten. Infolge der Überschneidung der Kurven für „a“ und „t“ läßt sich die gewünschte Aufteilung nicht für alle Sprecher verwirklichen. Man kann deutlich eine Gruppe langsam und schnell sprechender Testpersonen unterscheiden. Die erheblichen zeitlichen Streuungen dieser beiden Gruppen können nur durch gleitende Taktzeiten ausgeglichen werden, auf die später noch einmal eingegangen werden soll.



Speicherzelle 30 oder 20. Mittels einer disjunktiven Verknüpfung erhält man für die Erkennung von „plus“ eine sichere Aussage. Die Gleichung für „plus“ lautet dann:

$$1 \ \& \ \overline{10} \ \& \ \overline{22} \ \& \ (30 \vee 20) = L \quad \text{bedeutet: plus.}$$

Als Matrixelement wurden Gatter mit maximal vier Eingängen verwendet. Bild 11 zeigt einen Teil der Bedeutungsmatrix des Sprechers A mit den Verknüpfungen für „start“ und „plus“ in symbolischer Darstellung.

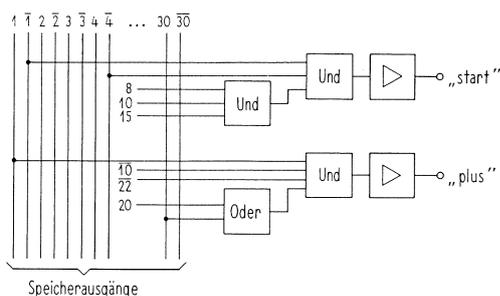


Bild 11. Matrixelemente für „start“ und „plus“.

Aus statistischen Untersuchungen mit 10 Versuchspersonen wurden die optimalen Matrizen sowohl für die Einzelpersonen als auch für die Gesamtheit der Versuchspersonen ermittelt. Infolge der im angewandten Verfahren noch nicht gänzlich eliminierten linguistischen Eigenheiten der Sprecher unterscheiden sich diese Matrizen in einzelnen Verknüpfungen. Die Erkennungssicherheit nimmt mit der Größe der Sprechergruppe ab.

### 2.4. Ergebnisse

Zur Erprobung des Erkennungsgerätes wurde bei der Aufnahme der Versuchsreihe eines Sprechers die Verstärkung des Mikrophonverstärkers fest eingestellt und auch bei Lautstärkeschwankungen des Sprechers nicht nachgestellt. Während der Untersuchungen betrug der Abstand vom Sprecher zum Mikrofon etwa 40 cm. Die 14 Testwörter einer Versuchsreihe wurden in willkürlicher Reihenfolge gesprochen und an der Anzeige des Rechengerätes kontrolliert. Die Tabelle III enthält die Er-

Tabelle III. Testergebnisse.

		Versuchsreihen je Person	Mittelwert der richtigen Entscheidungen
1	Sprecher A	21	97%
2	10 Sprecher	3	87%

gebnisse zweier Untersuchungen. Im Test 1 wurde mit Sprecher A die für ihn optimale Matrix gemäß Tabelle II erprobt. Ähnliche Ergebnisse können mit anderen Sprechern erzielt werden, wenn stets eine für den Sprechertyp optimale Matrix verwendet wird. Die Tendenz der Zuordnung von Erkennungsmatrizen zu Sprechergruppen mit verwandten Eigenschaften zeichnet sich ab und damit vielleicht

die Möglichkeit mit einer beschränkten Zahl von Matrizen einen größeren Sprecherkreis zu erfassen. Im Versuch 2 wurden zehn Sprecher nach ihren Eigenarten in drei Gruppen aufgeteilt und jede Gruppe mit der für sie ermittelten Erkennungsmatrix untersucht.

Die stärkeren linguistischen Streuungen bei einer Gruppe von Sprechern, insbesondere Schwankungen der Lautstärke und Schnelligkeit des Sprechers, sind die Ursachen für die Verringerung der Erkennungssicherheit von 97% auf 87%. Die Fehlentscheidungen enthalten Verwechslungen und Ausfälle. Mit einer sich automatisch regelnden Anpassung des Gerätes an die Eigenheit des Sprechers kann die Fehlerquote der Erkennung vermindert werden.

### 2.5. Gedanken für eine Weiterentwicklung

Die in Untersuchungen an diesem Gerät gewonnenen Erfahrungen erbrachten einige Vorschläge zur Weiterentwicklung der automatischen Spracherkennung.

Für die Kompensation der Lautstärkeschwankungen beim Sprechen muß eine automatische Verstärkungsregelung des Mikrophonverstärkers vorgesehen werden.

Der Zeittakt muß der Sprechschnelligkeit des jeweiligen Sprechers angepaßt werden. Es besteht die Möglichkeit, den Takt automatisch durch den Wechsel von stimmlosen zu stimmhaften Lauten oder umgekehrt steuern zu lassen. Die Kriterien für eine Unterscheidung solcher Laute sind bekannt [2]. Die technische Realisierung ist relativ einfach.

Ausfälle der Erkennung können reduziert werden, wenn die Entscheidung nicht mehr von der starren Erfüllung der Erkennungskriterien abhängig gemacht wird, sondern durch Vergleich der eingesprochenen Informationen mit den Erkennungskriterien der Matrix nach größter Ähnlichkeit gefällt wird. In einem weiteren Schritt können auch unsichere Informationen zur Erkennung mit verwendet werden, wenn man ihnen nur entsprechend ihrer Unsicherheit eine geringe Bedeutung beimißt. Die Verknüpfungselemente der Matrix müßten dafür eine unterschiedliche Bewertung erhalten. Die Möglichkeiten der Realisierung solcher Logik werden zur Zeit in einer theoretischen Studie untersucht.

## 3. Das Rechengerät

Das Rechengerät war als Demonstrationsmodell zum Erkennungsgerät gedacht, es sollte damit aber auch eine Studie über einen kleinen elektronischen Tischrechner verbunden sein. Im folgenden wird dieses Gerät in großen Zügen beschrieben. Für die Demonstrationszwecke genügte es, den Stellenumfang auf maximal vier Stellen zu beschränken. Das Gerät kann auch nur die Operationen Addition und Subtraktion ausführen. Da aber die beiden fehlenden Grundoperationen, die Multiplikation und Division, im allgemeinen auf fortlaufende Additionen und Subtraktionen zurückgeführt werden, hätte eine zusätzliche Berücksichtigung auch dieser

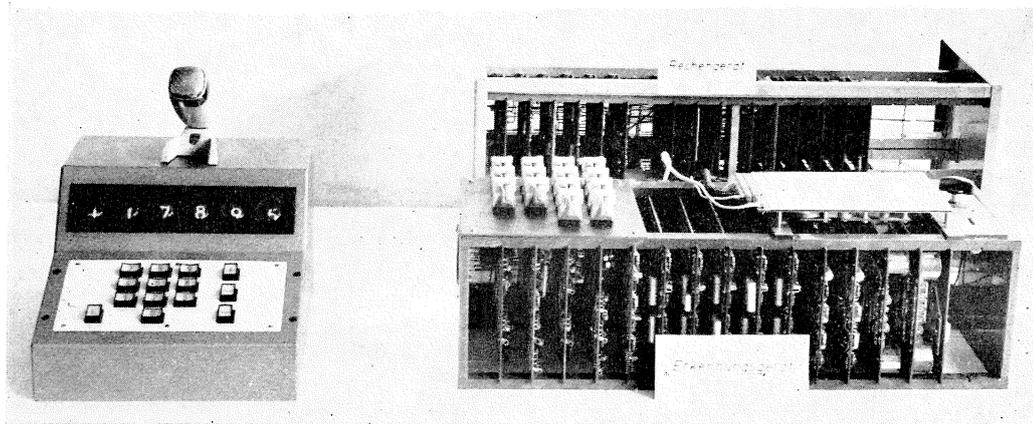


Bild 12. Ansicht der phonetischen Addiermaschine.

Operationen lediglich Aufwand aber kaum prinzipielle Erkenntnis gebracht. Das Gerät ist im wesentlichen mit Transistoren ausgerüstet, nur in den langsamen Teilen der Eingabe und Ausgabe wurden noch Relais vorgesehen. Die Eingabe der Ziffern erfolgt durch Handtasten oder durch das Spracherkennungsgerät, die Ergebnisse werden durch Ziffernanzeigeröhren angezeigt. Bild 12 vermittelt eine Ansicht der äußeren Ausführung des

Rechners. In einem Pult sind die Handtastatur und die Anzeige untergebracht; Erkennungsgerät und Rechenggerät befinden sich in getrennten Gestellen.

3.1. Innere Organisation

Die einzelnen Ziffern werden im dual verschlüsselten Dezimalcode verarbeitet. Jede Ziffer wird also durch eine Kombination von 4 bit (Tetrade) dargestellt. Die Subtraktion wird zurückgeführt auf

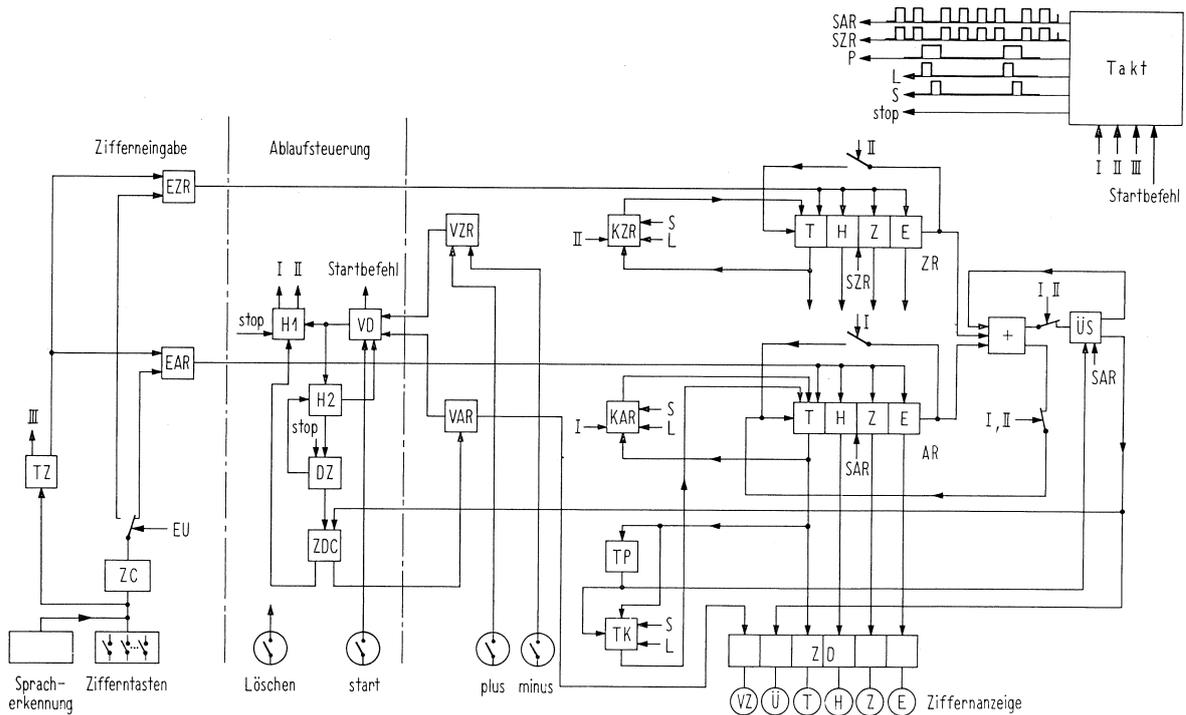


Bild 13. Prinzipschaltbild des Rechengertes;

- |     |                        |                                 |                        |     |                    |
|-----|------------------------|---------------------------------|------------------------|-----|--------------------|
| AR  | Akkumulatorregister,   | TZ                              | Tastendruckzähler,     | VZ  | Vorzeichen,        |
| ZR  | Zweitregister,         | EAR                             | Einschreibgatter AR,   | Ü   | Übertrag,          |
| KAR | Komplementierer AR,    | EZR                             | Einschreibgatter ZR,   | T   | Tausender,         |
| KZR | Komplementierer ZR,    | EU                              | Einschreibumsteuerung, | H   | Hunderter,         |
| VAR | Vorzeichenspeicher AR, | VD                              | Vorzeichendecodierer,  | Z   | Zehner,            |
| VZR | Vorzeichenspeicher ZR, | H <sub>1</sub> , H <sub>2</sub> | Hilfsspeicher,         | E   | Einer,             |
| TP  | Tetradenprüfer,        | DZ                              | Durchlaufzähler,       | SAR | Schiebeimpulse AR  |
| TK  | Tetradenkorrektur,     | ZDC                             | Zählerdecodierer,      | SZR | Schiebeimpulse ZR, |
| ÜS  | Übertragspeicher,      |                                 |                        | P   | Prüfimpulse,       |
| ZC  | Zifferncodierer,       |                                 |                        | L   | Löschimpulse,      |
| ZD  | Zifferndecodierer,     |                                 |                        | S   | Schreibimpulse.    |

eine Addition des komplementierten Minuenden zum Subtrahenden. Als Komplement wurde das Neuner-Komplement gewählt. Bild 13 zeigt einen Stromlaufplan mit den wesentlichsten Elementen des Rechengerätes. Als Speicher und Transporteur der Zahlen dienen zwei Schieberegister aus jeweils 16 bistabilen Stufen ( $4 \times 4$  bit). Das Akkumulatorregister (AR) nimmt die erste der beiden zu verknüpfenden Zahlen auf, das Zweitregister (ZR) die jeweilige zweite. Die Verknüpfung geschieht in einem Serienaddierer; Schiebeimpulse von einem Taktgeber schieben die Informationen beider Register durch den Addierer, das Ergebnis erscheint nach einem Kreislauf im AR. Die bei der Verknüpfung der einzelnen Ziffern möglichen auftretenden Pseudotetraden (Ziffern 10 bis 19) werden von einer Tetradenprüfschaltung am AR erkannt und korrigiert. Die Prüfung und Korrektur erfolgt beim Durchlaufen der Ergebnisse durch das AR in der höchst vorkommenden Stelle, der Tausender-Stelle. Die Schiebeimpulse für beide Register sind so beschaffen, daß nach jeweils vier Impulsen — entsprechend einer Ziffernverschiebung — eine Pause eintritt (siehe Impulsplan des Taktgebers). In dieser Pause wird vom Taktgeber ein Prüfpuls (P) abgegeben; eine Tetradenprüfschaltung (TP) entscheidet, ob im Register eine Pseudotetrade ansteht oder nicht. Erkennt die Schaltung eine solche, wird sie während der Taktpause durch eine +6-Addition von einer Tetradenkorrekturschaltung (TK) korrigiert und in einem Unterzyklus wieder ins Register eingeschrieben. Dem Einschreiben der korrigierten Tetrade geht eine Löschung der Pseudotetrade im Register voraus. Dazu dient ein vom Taktgeber in der Schiebeimpulspause gelieferter Löschimpuls (L), das Einschreiben geschieht durch einen zeitlich dem Löschimpuls folgenden Schreibimpuls (S). — Nach einem ähnlichen Prinzip wie die Tetradenkorrektur arbeitet die Tetradenkomplementierung für die Subtraktion. Jedem der beiden Register ist wieder in der höchst vorkommenden Stelle eine Komplementierschaltung zugeordnet (KAR), (KZR). Welche von beiden Schaltungen bei einer Subtraktion in Funktion tritt, hängt vom Vorzeichen der Zahlen im AR und ZR ab. Den Entscheid hierüber liefert die Ablaufsteuerung. Zur Komplementierung wird das jeweilige Register im Ring geschaltet, die Ausgänge hinter dem Addierer werden aufgetrennt. Nach einem vollen Durchlauf erscheint die Zahl als Komplement wieder im Register. Die Komplementierung der einzelnen Ziffern geschieht dabei in den Taktpausen der Schiebeimpulse in einem Unterzyklus mit Löschen und Wiedereinschreiben. — Das zusätzliche Komplementieren bei der Subtraktion bedingt einen dreifachen Zeitaufwand für die Operation Subtraktion gegenüber der Operation Addition. In einem ersten Durchlauf erfolgt die Komplementierung einer der beiden Zahlen, in einem zweiten die Verknüpfung und in einem dritten die Rückkomplementierung oder eine Übertragsverarbeitung je nach Vorzeichen des erreichten Ergebnisses. — Die Steuerung der einzelnen Abläufe ge-

schieht über eine Ablaufsteuerung, die bei der Behandlung eines Rechenablaufes noch verdeutlicht wird.

### 3.2. Eingabe und Ausgabe

Die von Hand oder durch das Spracherkennungsgerät eingegebenen Ziffern werden von einem Zifferncodierer (ZC) in den dualen Dezimalcode gewandelt und die Dualstellen einer Ziffer jeweils parallel in die Register eingeschrieben. Die erste Zahl gelangt in das AR, die zweite in das ZR. Der Eingabe der zweiten Zahl geht stets die Wahl der Verknüpfungsart (plus oder minus) voraus, die gleichzeitig die Eingabe-Umsteuerung (EU) auf das ZR auslöst. Die Eingabe geschieht in der gewohnten Reihenfolge, beginnend mit der höchstvorkommenden Dezimalstelle. Ein Tastendruckzähler (TZ) sorgt zusammen mit den Einschreibgattern (EAR), (EZR) für die stellengerechte Einschreibung in die Register. Die erste Ziffer wird jeweils in die höchste Registerzelle eingeschrieben. Bei Zahlen mit weniger als vier Stellen muß (zur Vermeidung von vorausgehenden Füllnullen) eine Stellenverschiebung nach Eingabebeendigung vorgenommen werden. Die Kriterien dazu werden ebenfalls vom Tastendruckzähler abgeleitet. Die stellengerechte Verschiebung der Zahl im AR geschieht bei Betätigung der Operationstasten, die der Zahl im ZR bei Auslösung der Starttaste, bzw. bei den jeweils äquivalenten akustischen Befehlen.

Die Ziffernanzeigeröhren liegen über einem Zifferndecodierer (ZD) an den Parallelausgängen des AR. Außer den vier Anzeigeröhren für die zu verarbeitende Stellenzahl ist noch eine weitere vorgesehen, die einen Überlauf in die 5. Dezimale registriert. Das Vorzeichen des Ergebnisses wird aus dem Vorzeichenspeicher des AR (VAR) gewonnen. Während der Betätigung der Starttaste liegt das Anzeigefeld an den Parallelausgängen des ZR. Damit ist eine kurze Kontrolle der korrekten Eingabe der zweiten Zahl möglich.

### 3.3. Ablauf

Eine kurze Beschreibung eines Rechenablaufes verdeutlicht die Funktion des Rechners.

1. Eingabe der ersten Zahl.  
Über Zifferncodierer und Einschreibgatter (EAR) gelangt die erste Zahl in das AR.
2. Wahl der Operationsart (plus oder minus).  
Das gewählte Vorzeichen wird im Vorzeichenregister des ZR (VZR) gespeichert. Gleichzeitig werden Schiebeimpulse nur für das AR ausgelöst zur stellengerechten Verschiebung der Zahl im AR, ferner wird die Eingabeumsteuerung (EU) zur Aufnahme der zweiten Zahl eingeleitet.
3. Eingabe der zweiten Zahl.
4. Startbefehl.  
Die Eingabeumsteuerung wird zurückgestellt, Schiebeimpulse nur für das Zweitregister werden zur stellengerechten Verschiebung der zweiten Zahl ausgelöst.

Beide Vorzeichenspeicher (VAR), (VZR) werden auf den Vorzeichendecoder (VD) geschaltet. Da bei Kettenrechnungen ein positives Vorzeichen (VZ) oder ein negatives im AR stehen kann, muß im VD der Entscheid gefällt werden, welche Rechenart vom Rechner zu leisten ist.

- a) Beide VZ positiv oder negativ: Addition.  
Das Ende des Startbefehls löst ein volles Impulsprogramm aus, Rechenablauf Addition. Das Vorzeichen des Ergebnisses steht bereits im VAR.
- b) Vorzeichen gemischt.  
Beispiel: VZ VAR positiv, VZ VZR negativ.  
Im Hilfsspeicher  $H_2$  wird generell der Subtraktionsbefehl eingespeichert. Hilfsspeicher  $H_1$ , Ausgang I veranlaßt die Ringschaltung des AR in sich und die Abtrennung des Addierers. Der Taktgeber wird vorbereitet zur Aussendung von Schiebeimpulsen nur für das AR. Bei vertauschten Vorzeichen übernimmt Ausgang II diese Funktionen für das ZR. Das Ende des Startbefehls löst den Ablauf der Komplementierung der Zahl im AR aus (Komplementierung des Minuenden). Ein Stopimpuls nach vollem Durchlauf stellt den Hilfsspeicher  $H_1$  zurück und beendet seine Funktionen, gleichzeitig wird der Durchlaufzähler (DZ) weiterschaltet.  $H_2$  speichert noch die Information „Subtraktion“, daher erneute Auslösung des gesamten Impulsprogrammes und Verknüpfung beider Zahlen.  
Nach dem zweiten Durchlauf ist vom Zählerstand des DZ eine Kontrolle der höchsten AR-Tetrade vorbereitet.
- α) Keine Pseudotetrade, d. h.: Rückkomplementierung des Ergebnisses, Vorzeichen positiv.  
Die Information wird über ZDC in  $H_1$  eingespeichert, Ausgang I bereitet die erneute Ringschaltung des AR und die Rückkomplementierung im dritten Durchlauf vor. Ein positives VZ wird im VAR eingeschrieben.
- β) Pseudotetrade, d. h.: Übertragsverarbeitung in einem dritten Durchlauf, Vorzeichen negativ.  
Anstelle der Rückkomplementierung wird in einem dritten Durchlauf der Übertrag zum AR-Inhalt addiert; ein negatives VZ wird im VAR eingeschrieben.

Nach dem dritten Durchlauf stellt der Stopimpuls den Durchlaufzähler in Ruhestellung, der Subtraktionsbefehl in  $H_2$  wird gelöscht, es erfolgt kein weiterer Ablauf. Das Ergebnis steht nach Beendigung des Rechenablaufes im AR und wird von den Anzeigelampen angezeigt.

Das beschriebene Rechenggerät arbeitet mit einer Taktfrequenz von 100 kHz. Zur Addition zweier Dezimalziffern sind damit 40  $\mu$ s erforderlich, mit Berücksichtigung der Tastlücke 60  $\mu$ s. Da für eine Addition zweier Zahlen unabhängig von ihrer Stellenzahl aber stets ein voller Durchlauf erforderlich ist, so beträgt die Additionszeit 240  $\mu$ s, die Subtraktionszeit wegen drei Durchläufen 720  $\mu$ s.

Die Konzeption des beschriebenen Rechenggerätes stellt zweifellos nur eine von vielen möglichen Lösungen dar. Es sollte eine grundsätzliche Studie sein, ohne daß dabei nennenswerte Optimierungen oder wirtschaftliche Lösungsgesichtspunkte eine entscheidende Rolle spielten.

Dem Herrn Abteilungsvorsteher Prof. Dr.-Ing. R. ELSNER danken wir für anregende Diskussionen. Bei den Untersuchungen hat eine Reihe von Diplomprüfungskandidaten der Technischen Hochschule Braunschweig mitgewirkt [10].

#### Schrifttum

- [1] MEYER-EPPLER, W., Grundlagen und Anwendungen der Informationstheorie, Springer-Verlag, Berlin 1959.
- [2] STEINBUCH, K., Taschenbuch der Nachrichtenverarbeitung, Springer-Verlag, Berlin 1962.
- [3] STEINBUCH, K., Automatische Spracherkennung, Nachrichtentech. Z. **11** [1958], 446–454.
- [4] SAKAI, T. und DOSHITA, S., The phonetic typewriter, Information Processing 62, Proceedings of the IFIP Congress 1962, North Holland Publishing Co., Amsterdam 1963.
- [5] KUSCH, H., Automatische Spracherkennung gesprochener Zahlen (Ziffern), Nachrichtentech. Z. **18** [1965], 57–62.
- [6] TRENDELENBURG, F., Einführung in die Akustik, Springer-Verlag, Berlin 1950.
- [7] FELDTKELLER, R. und ZWICKER, E., Das Ohr als Nachrichtenempfänger, S. Hirzel Verlag, Stuttgart 1956.
- [8] Aufnahme und Verarbeitung von Nachrichten durch Organismen (Vorträge aus dem Gebiet der Kybernetik, gehalten auf der NTG-Fachtagung in Karlsruhe am 11. und 12. 4. 1961), S. Hirzel Verlag, Stuttgart 1961.
- [9] HÜTTE, Theoretische Grundlagen; Physiologische und musikalische Akustik, S. 366, Verlag W. Ernst & Sohn, Berlin 1955.
- [10] CURTHS, P., Untersuchungen des Zeit-Amplituden-Spektrums von Sprachsignalen.  
GERDES, D., Bau eines Gerätes zur Erkennung gesprochener Befehle.  
HEISS, D. und OTTENROTH, K., Bau des Additionsteiles eines elektronischen Ziffernrechners.  
Diplomarbeiten am Institut für Fernmelde- und Hochfrequenztechnik der Technischen Hochschule Braunschweig, 1964.