MASTER

Patroonherkenning : overzicht van de huidige stand van de wetenschap inzake tekenherkenning

van Binsbergen, P.R.M.

Award date:
1969

AFDELING DER ELECTROTECHNIEK
TECHNISCHE HOGESCHOOL
EINDHOVEN

Groep Telecommunicatie B

PATROONHERKENNING
Overzicht van de huidige stand van de wetenschap inzake tekenherkenning.

P. R. M. van Binsbergen

ECB

Afstudeerverslag onder leiding van Prof. ir. A. Heetman.
Mentor: ir. C. P. J. Schnabel.

Juni 1969

Afdeling der Electrotechniek
Technische Hogeschool.
Eindhoven.

Groep Telecommunicatie B

Patroonherkenning

Overzicht van de huidige stand van de wetenschap inzake tekenherkenning.

P. R. M. van Binsbergen

Afstudeerverslag onder leiding van
Prof. ir. A. Heetman.
Mentor: ir. C. P. J. Schnabel.

Juni 1969

Inhoudsopgave

blz.

1. Inleiding 3

2 Overzicht van de stand der techniek voor tekenherkenning 7
   2.1. Algemene beschouwing 7
   2.2. Klassifikatie methoden. 10
      2.2.1. Lineaire Klassifikatie 13
      2.2.2. Correlatie 17
      2.2.3. Maximale waarschijnlijkheid. 23
      2.2.4. De Anderson - Bahadur methode. 25
      2.2.5. Discriminant Analyse 28
      2.2.6. Lerende Klassifikatie 32
         2.2.6.1 Perceptron 33
         2.2.6.2 Sequentiële beslissingen. 35
      2.2.7. Potentiaal Funkties. 36
   2.3. Zelf corrigerend systeem. 38
   2.4. Groepen analyse. 44
   2.5. Destillatie van Kenmerken. 47

3 Adaline 51
   3.1 Grenzen bij toepassing van drempelwaarde-element. 58

4 Tekenherkenning met behulp van coherente optische systemen. 65
   4.1. Toepassing van matched filters in tekenherkenning 70
   4.2. Voor- en nadelen. 76
   4.3. Toepassing van optische herkenningssystemen. 78

5. Lezen en herkennen van handgeschreven tekens. 82

5.1 Moeilijkheden bij de herkenning van handgeschreven letters. 83

5.2 Herkenningssysteem voor geschreven symbolen. 87

5.3 Machine gedrukte tekens. 89

5.4 Herkenning van handgeschreven tekst 92

5.5 Enige mogelijkheden voor de toekomst. 98

6 Herkenningsproces 102

6.1 Een toepassing van het verwerken van cijfergecodeerde informatie. 102

6.2 Procesopbouw. 105

6.3 Reductie, normalisatie en transformatie 108

6.4 Preprocessing 113

6.5 Lijnreductie 120

6.6 Slotopmerkingen 121

7 Conclusies 125

7.0 Slotwoord. 127

8 Appendix 128

9 Literatuur 137

1 Inleiding

De mens onderscheidt zich van de rest van de schepselen door het verstand. Het denken is een faktor die de mens in zeer veel hoedanigheden verheven heeft boven zijn omgeving. Dit proces heeft er toe geleid, dat allerlei middelen worden aangewend om het leven te vereenvoudigen. Het aantal toepassingen als funktie van de tijd verloopt als een e-macht. Beginnend bij het vuur is de lengte van de rij der toepassingen a.h.w. als een kettingreactie toegenomen.

Het physisch gedrag was vroeger toonaangevend in het levenspatroon van de mens. De evolutie had een toename van het gebruik van de ratio tot gevolg. Tot op heden zette deze tendens zich voort, met als gevolg een verschuiving van de waardering van geestelijke en lichamelijke arbeid en een hiermee samenhangende evenwichtsverstoring. Vroeger gold voor lichamelijke inspanning: arbeid; nu is hetzelfde van toepassing op geestelijke inspanning, terwijl physieke inspanning daarentegen vaak ontspanning is geworden.
De mens heeft door de tijden heen hulpmiddelen ontworpen, die zijn lichamelijke arbeid geheel of gedeeltelijk kunnen vervangen. Tegenwoordig verschaft de mens zich ook hulpmiddelen om zijn geestelijke arbeid te vereenvoudigen. De aard van de hulpmiddelen zijn sterk afhankelijk van de verschuiving in de bepalende faktoren van het gedragspatroon. De meest recente en opmerkelijke hulpmiddelen zijn de rekenmachines. Zij vereenvoudigen in sommig opzicht het denkwerk van de mens. Vaststaande denkpatronen en algorithmen kunnen, eenmaal in de machine vastgelegd, door het apparaat moeiteloos herhaald worden. Deze gedachte heeft er toe geleid, de rekenmachine in te schakelen om vermoeiende, langdurige vaste handelingen en eenduidig bepaalde denk- en handelingsschema's uit te voeren.

Een van de toepassingsgebieden is de patroonherkenning. Het herkenningsproces speelt een voorname rol in het menselijk

leven. Het wordt overal en op ieder tijdstip toegepast. Het menselijk gedrag is te beschouwen als een regelproces waarbij de patroonherkenning een zeer belangrijke rol speelt. Het herkennen van kleuren, figuren, letters, cijfers, gebruiksvoorwerpen, toestanden etc. is iets dat de mens bijna volkomen beheerst.
De vraag is nu: „In hoeverre is het mogelijk en nuttig een deel van dit herkenningsproces door de machine te laten overnemen?"
Het is onmogelijk op deze vraag een juist antwoord te geven. Er valt echter wel zeer veel over het herkenningsproces en zijn toepassing te zeggen; het theoretisch aspect, de praktische moeilijkheden in de toepassing op speciale gebieden worden door mij in dit bestek behandeld.
Om de uitgebreidheid en algemeenheid van het herkenningsgebied tot uitdrukking te brengen, is het noodzakelijk de zeer verschillende problemen van het proces onder de loep te nemen. Iedere techniek die hierbij gebruikt wordt, kan gezien worden als een zelfstandige groep van het proces. Enige van deze groepen zijn:

schakeltechniek	algorithmen in de trant van het perceptronprinciepe.
automatentheorie	
set-theorie	kenmerken destillatie
contrôle-theorie	preprocessing
informatietheorie	taalwetenschap
correlatie methoden	cybernetica
discriminant analyse	studie van het zenuwstelsel
groepentheorie	psychologie
statistische beslissingstheorie	physiologie
wiskundig programmeren.	psychiatrie.

Deze lijst is beslist niet komplect en geeft dan ook alleen een idee van de complexheid van het herkenningsproces.

Ga ik de praktische verwezenlijking van het patroonherkenningsproces bekijken, dan kom ik tot de konklusie, dat om louter economische reden het alleen verantwoord is een gedeelte van de <u>tekenher</u>kenning te verwezenlijken. Dit vindt zijn oorzaak o.a. in:

a/ het relatief uitgebreide vooronderzoek

b/ de complexheid van de toe te passen procedures
c/ de vereiste verwerkingssnelheid.
d/ de nauwkeurigheid die aangehouden moet worden.
e/ het leerproces dat in veel gevallen langdurig kan zijn.
f/ het klein toepassingsgebied, dat in geen verhouding staat met de zeer uitgebreide en langdurige opzet van het systeem.

Alleen als de verschillende punten op elkaar zijn afgestemd is een bruikbare en rendabele toepassing mogelijk.
Ik denk bijvoorbeeld aan het urgent probleem van letter en cijferherkenning, dat zijn toepassingsgebied o.a. heeft bij:

a/ Bankwezen en giro.
Automatisch lezen en verwerken van gedrukte of handgeschreven informatie. De nauwkeurigheid moet zeer groot zijn, omdat het om geldzaken gaat.

b/ Post.
Automatisch lezen en sorteren van adressen.

c/ Bibliotheek
Automatisch katalogiseren en administreren van informatie en informatiebronnen.

Minder urgent is de toepassing op medisch en biologisch gebied. De aanpak van de problemen richt zich dan ook vooral op de letter-en cijferherkenning, ook al om reden dat op dit gebied de meeste economische winst te behalen is.

Ik ga in dit artikel eerst een algemeen overzicht geven betreffende de mathematische theorieën en methoden. Het zou te ver voeren, diep op deze stof in te gaan. Het is alleen de opzet een overzicht te geven wat betreft de theoretische achtergronden en hun toepasbaarheid in de praktijk. Voor een eventueel diepgaande studie in een bepaald facet, moet ik naar de literatuur verwijzen. Aan de hand van het overzicht, worden enige speciale modellen met hun mogelijkheden uitgewerkt. De theoretisch opgezette achtergronden van de „Adaline" zijn bedoeld om, aan de hand van een paralel-model, een juist inzicht te geven in de mogelijkheden. De toepassing van deze paralelsituatie ligt ook bij soortgelijke technieken, zoals leermatrix en perceptron. Na deze modellen komen praktisch te verwezenlijken herkenningsprocessen aan de orde, die gekozen zijn uit een groot aantal technieken, die o.a. in de literatuur aangeduid of beschreven staan. Het zijn juist deze systemen, omdat ik voor de methoden een grote en vruchtbare toekomst zie wat betreft efficiënte en veelvuldige toepassing in de praktijk.

2. Overzicht van de stand der techniek voor tekenherkenning

2.1. Algemene beschouwing

Bij de analyse van een tekenherkenningsproces, is de hoofdlijn te verdelen in twee stukken:

a) leerproces

b) herkenningsproces.

In enkele gevallen kan een kombinatie van beide optreden:

c) lerend herkenningsproces.

De techniek en methoden die toegepast worden, moeten een gunstig en snel resultaat hebben in zowel het leer- als het herkennings-proces. Uiteraard wordt eerst gekeken naar de opzet van het leerproces, maar daarbij wordt beslist de toepasbaarheid voor herkenning niet vergeten. Zo zal de opname van de gepresenteerde symbolen op een voor a en b identieke manier geschieden. De analysering van de symbolen, door opname van de, voor het verwerkingsproces belangrijke gegevens (attributen of parameters), geschiedt in principe ook op dezelfde manier. In het leerproces echter, zal de analysering in de tijd verbeterd worden; wanneer een vooraf vastgestelde analyse-nauwkeurigheid is bereikt, zal het leerproces worden afgesloten. Het herkenningsproces kan beginnen met een analysering, die een nagenoeg konstante herkenningsnauwkeurigheid heeft. De symbolen, die in de leerfase gebruikt worden, moeten representatief zijn voor de in de herkenningsfase voorkomende symbolen. Hoe lager de correlatie tussen beide groepen symbolen, des te groter is het aantal foutieve herkenningen.

Het lerend herkenningsproces heeft een zich asymptotisch verbeterende analyse.

Niet ieder systeem heeft een duidelijk herkenbaar leerproces. Het komt wel eens voor dat het leerproces geheel of

gedeeltelijk bij de mens ligt. In geval van het ontbreken van een leerproces in het systeem, moet de methode van waarneming, het opnemen en verwerken van de attributen door de mens worden vastgelegd. Dit kan natuurlijk alleen voor vrij eenvoudige processen, die door de mens gemakkelijk te overzien zijn.
Als voorbeeld geef ik de fotocel in een lift, die moet detecteren of „herkennen", dat er een object nog niet geheel binnen of buiten de lift is, zodat de deuren nog niet moeten sluiten. Het systeem heeft nu het onderscheid moeten leren tussen wel of geen object in de deuropening. Dit proces heeft de mens vertaald in : „Wel of geen onderbroken lichtstraal." In hoeverre deze vertaling opgaat, is een maat voor de nauwkeurigheid van het „leerproces." Er bestaat een mogelijkheid dat er zich een object in de deuropening bevindt, terwijl de lichtstraal niet onderbroken wordt. In dit geval gaat de vertaling niet op; er is een onnauwkeurigheid in het leerproces. De mens is zich hiervan bewust en heeft dan ook voor dit geval een extra beveiliging aangebracht in de deur.

Eén van de moeilijkste problemen in de tekenherkenning is het bepalen van de juiste attributen. Het totaal van meetgrootheden of attributen wordt meetvector genoemd, waarbij de meetvector bepalend is voor de klassifikatie van het te herkennen symbool.
Drie belangrijke faktoren in het proces zijn:

1) procestijd : t
2) nauwkeurigheid : n
3) aantal attributen : p

Deze faktoren zijn niet onafhankelijk van elkaar. Er bestaat een optimum in de minimale procestijd van een herkenning, als funktie van het aantal lineair onafhankelijke attributen en de nauwkeurigheid :

$t = t(p, n)$

t groot voor p groot en n groot

t klein voor p klein en n klein

De nauwkeurigheid n is ook een funktie van het aantal attributen, in die zin, dat p de nauwkeurigheid begrenst aan de bovenzijde. De onderlinge afhankelijkheid is geen triviale zaak. Het is dan ook van urgent belang dat voor het herkenningsproces, met een vereiste nauwkeurigheid in de beslissing, de juiste en het juiste aantal attributen worden genomen, die tesamen kenmerkend zijn voor een klasse. Het is makkelijk veel attributen te nemen; het is een ingewikkelde zaak een minimum aantal efficiënte attributen te vinden. Dit laatste is de basis waarop de hele tekenherkenning gebaseerd is. Uitgaande van de gegeven attributen, is de techniek zo ver, dat een aangepast herkenningssysteem ontwikkeld kan worden.
De rol van de rekenmachine is in deze zeer belangrijk:

1) Bepalen en selecteren van attributen met behulp van computer gegenereerde statistieken van attributenoptimalisatie.
2) testen en selecteren van min of meer succesvolle criteria.
3) testen van handgeschreven tekst.
4) uitvoeren van bepaalde herkenningshandelingen.
5) uitvoeren van algorithmen.
6) procesbesturing.

Aan de hand hiervan is de vraag belangrijk, of een proces moet worden uitgevoerd met behulp van een centrale computer of een „special purpose" computer.

2.2 Klassifikatiemethoden.

Het primaire doel van het tekenherkenningsproces is, dat het systeem werkt voor een vooraf vastgestelde verzameling van tekens, waarbij er natuurlijk naar gestreefd wordt, deze verzameling zo algemeen mogelijk te houden. De leerfase van het systeem wordt doorlopen met een voldoend groot aantal trainingssymbolen, die allen bruikbaar en representatief zijn voor de verzameling van testsymbolen. Het is nu in dit geval voldoende, het systeem te beschrijven aan de hand van de trainingssymbolen. Hierbij moet echter bedacht worden, dat in het algemeen het oefenschema voor tekenherkenning te klein is, zodat via extrapolatie van de bruikbare trainings-symbolen, op stochastische wijze, nieuwe symbolen moeten worden geschapen. Het is, zo blijkt in de praktijk, mogelijk om aan de hand van de waarschijnlijkheidsverdeling van de voorgaande informatie, betreffende een eindige oefen-set, het gedrag te voorspellen van de test-set. Echter óók het waarschijnlijkheidsmodel moet voortkomen uit een eindige trainingsset. Deze raakt dusdanig opgezettend, komt er een hele brok stochastiek op papier, die uitgewerkt, een zeer gecompliceerde koppeling kan verzorgen tussen oefen- en test-symboolverzameling. Het is echter een groot goed het systeem zo eenvoudig mogelijk te houden, zodat een duidelijk overzicht gegarandeerd wordt. Een praktische maar onnauwkeurige regel is: „Hoe gecompliceerder de methode, des te meer tijd is nodig om het systeem op te bouwen, en des te vlugger worden er fouten gemaakt."

Ik ga het systeem beschrijven aan de hand van de leer- of oefenfase, en sluit dan direkt op het principe van de klassifikatie. Om een kompleet beeld te geven moet de niet-binaire of conventionele klassifikatie genoemd worden. Ik wil een klein voorbeeld geven om het klassebegrip iets te verduidelijken.
In het algemeen is een herkenningsprobleem een multiklasprobleem. Het is mogelijk dit probleem te ontleden in een aantal twee-klas-

problemen. Dit is theoretisch mogelijk, maar op economische gronden meestal niet toegestaan.
Stel een trainingsset van 2N tekens, verdeeld in 2 klassen. (zie fig. 1):

klas K_1 : $(x_1^1, y_1^1), (x_2^1, y_2^1), \ldots, (x_N^1, y_N^1)$
klas K_2 : $(x_1^2, y_1^2), (x_2^2, y_2^2), \ldots, (x_N^2, y_N^2)$

gemiddelde vector van K_1 : $\frac{1}{N}\left\{\sum_{i=1}^{N} x_i^1, \sum_{i=1}^{N} y_i^1\right\} = \bar{\mu}_1$
gemiddelde vector van K_2 : $\frac{1}{N}\left\{\sum_{i=1}^{N} x_i^2, \sum_{i=1}^{N} y_i^2\right\} = \bar{\mu}_2$

variantie $x_i^1 = \sigma^2(x_i^1) = E({x_i^1}^2) - E^2(x_i^1)$
covariantie $(x_i^1, y_i^1) = 0$, want x_i^1 en y_i^1 worden onafhankelijk verondersteld.
Geldt de onafhankelijkheid niet, dan wordt de cov.-matrix complexer.
Bijvoorbeeld:
De cov. matrix van K_1 heeft de vorm:

$$A_1 = \begin{pmatrix} \sigma^2(x^1) & cov.(x^1, y^1) \\ cov.(y^1, x^1) & \sigma^2(y^1) \end{pmatrix}$$

Iedere klasse heeft nu zijn eigen specifieke cov.-matrix, zodat deze matrix een kenmerk is voor zijn klasse.
Dr. David Allais heeft aangetoond:
Om een cov.-matrix betrouwbaar te kunnen gebruiken, moet het aantal trainingssymbolen tien maal zo groot zijn als de dimensie van de symboolruimte. Wordt daar niet aan voldaan, dan kan het mogelijk zijn dat twee verschillende clusters eenzelfde cov.-matrix geven.
N.B.: de dimensie van de symboolruimte wordt bepaald door het aantal onafhankelijke attributen.
Deze stelling laat duidelijk zien, dat deze cov.-matrix-methode aan zware eisen moet voldoen. Er komen andere methoden aan de orde, die minder zware voorwaarden stellen, zodat dit niet als erg geschikte methode gekwalificeerd kan worden. Het is echter geenszins de bedoeling dit principe te verwerpen, daar er beslist toepassing mogelijk is in kleinere sub-problemen.

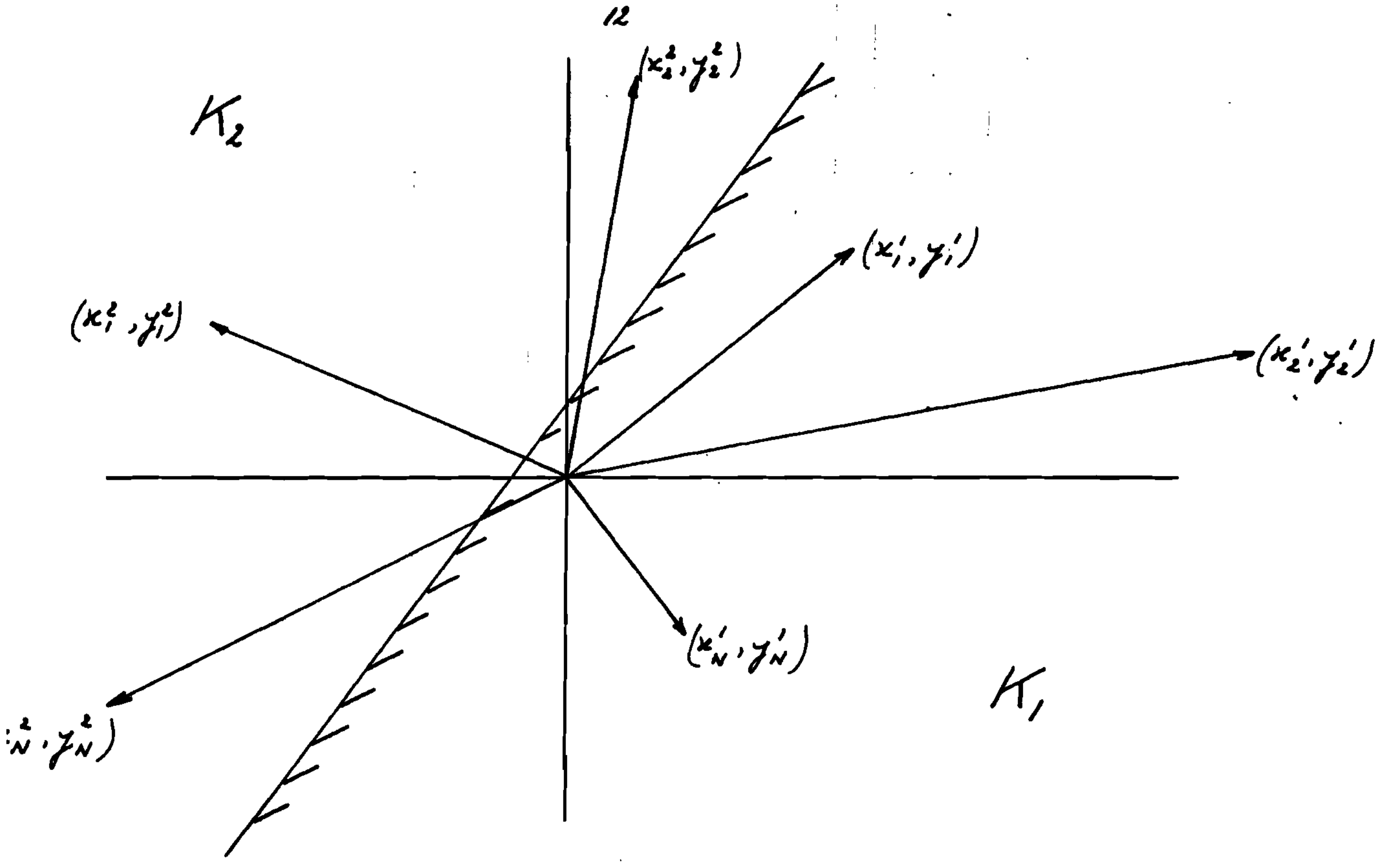

12
K_2
(x_2^2, y_2^2)
(x_1^1, y_1^1)
(x_1^2, y_1^2)
(x_2^1, y_2^1)
(x_N^1, y_N^1)
K_1
(x_N^2, y_N^2)

FIG.-1

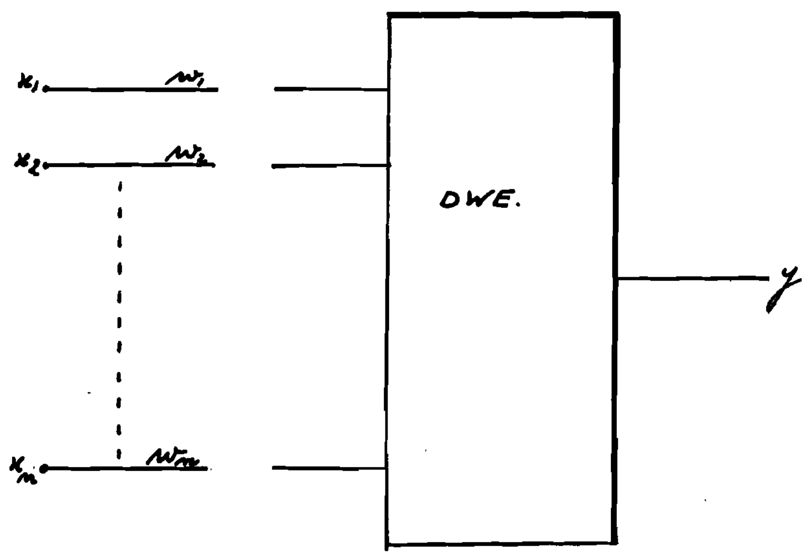

x_1
w_1
x_2
w_2
x_m
w_m
DWE.
y

FIG.-2

2.2.1. Lineaire Klassifikatie.

Stellen $x_1, x_2, \ldots, x_n$ het aantal attributen voor, en $\bar{x} = (x_1, x_2, \ldots, x_n)$ de meetvector, dan wordt de lineaire klassifikatie gedefiniëerd door de bewerking die:

$\bar{x}$ toewijst aan K_1 als $\bar{x}.\bar{w} > \theta$
$\bar{x}$ toewijst aan K_2 als $\bar{x}.\bar{w} < \theta$
en $\bar{x}$ onbepaald laat als $\bar{x}.\bar{w} = \theta$

De vector $\bar{w}$ heeft dezelfde dimensie als $\bar{x}$:

$$\bar{w} = (w_1, w_2, \ldots, w_n)$$

en wordt wel de gewichtsvector genoemd, omdat daarmee de relatieve invloed van ieder attribuut bij de bepaling van de klas wordt aangegeven. De scalar θ wordt drempelwaarde genoemd.
Een analoog systeem is een drempelwaarde-element (DWE.), dat o.a. gebruikt wordt bij de nog te behandelen „Adaline". Zie ook figuur 2. De uitgang y wordt als volgt gedefinieerd:

$$y = \begin{cases} 1 & \text{als } \bar{x}.\bar{w} \geq \theta \\ 0 & \text{als } \bar{x}.\bar{w} < \theta \end{cases}$$

De w_i kan in het schema bijvoorbeeld een potentiometer zijn.
De paralelsituatie is, dat K_1 ongeveer gerepresenteerd wordt door de waarde „1" en K_2 exact door de waarde „0".
Het hypervlak $\bar{x}.\bar{w} = \theta$ scheidt dus de twee klassen. In de twee-dimensionale ruimte, als $n=2$, is het hypervlak een rechte. (zie figuur 3), die loodrecht staat op de vector $\bar{w}$, en een afstand θ heeft tot de oorsprong.
Dit twee klassen systeem wordt genormaliseerd door:

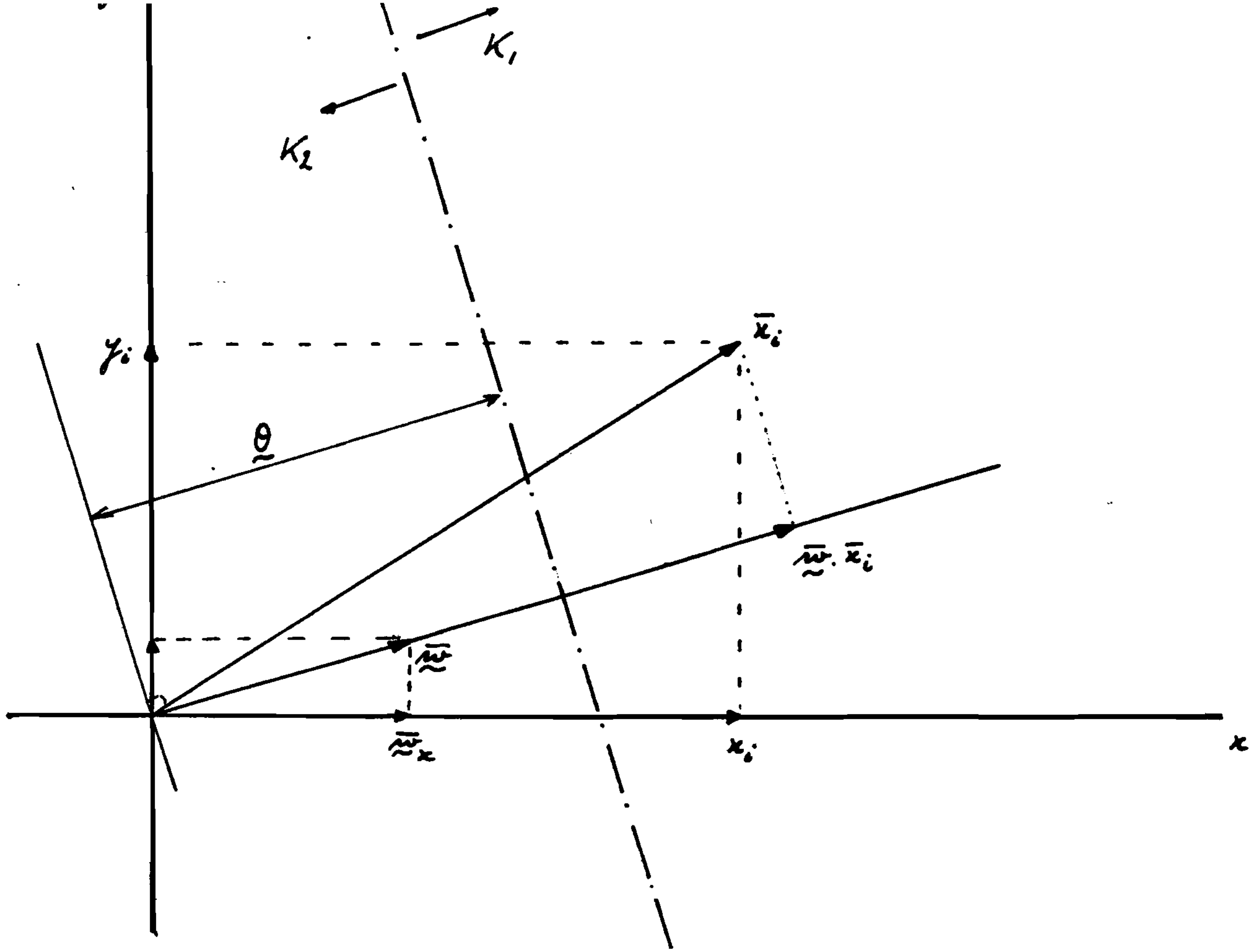

K_1
K_2
y_i
$\underset{\sim}{\theta}$
$\bar{x}_i$
$\bar{\underset{\sim}{w}} \cdot \bar{x}_i$
$\bar{\underset{\sim}{w}}$
$\bar{\underset{\sim}{w}}_x$
x_i
x

FIG.-3

$$\bar{w} := \frac{\bar{w}}{|\bar{w}|} = \underset{\sim}{\bar{w}}$$

te nemen. De drempelwaarde gaat dan over in

$$\theta := \frac{\theta}{|\bar{w}|} = \underset{\sim}{\theta}$$

Het is duidelijk dat het systeem bepaald wordt door gewicht en DW., die beide met hetzelfde getal vermenigvuldigd mogen worden, zonder dat er iets aan het systeem wijzigt.
In figuur 3 ligt de vector $\bar{x}_i$ in klasse K_1.
Stel dat het hier beschreven systeem een oneindig aantal normaal-verdeelde vectoren bevat. De projectie van de vectoren op de gewichtsvector $\bar{w}$ is dan ook normaal verdeeld met:

gemiddelde : $\bar{\mu}_i = \bar{w} . \bar{\mu}_i$
variantie : $\bar{w}^T . A_i . \bar{w}$
met : A_i = cov. matrix van K_i
$\bar{w}^T$ = getransponeerde van $\bar{w}$

Bedenk : dichtheidsfunktie van een normaalverdeling is :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Stel :

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{p^2}{2}} dp$$

$$E(z) = \bar{w}^T . \bar{\mu}_i$$
$$\sigma^2(z) = \bar{w}^T . A_i . \bar{w}$$

dan geldt:

$$\Phi\left\{\frac{\theta - E(z)}{\sigma(z)}\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \frac{e^{-\frac{\{\theta - E(z)\}^2}{2\sigma^2(z)}}}{\sigma(z)} . d\theta$$

$$= \frac{1}{\sigma(z).\sqrt{2\pi}} \cdot \int_{-\infty}^{y} e^{-\frac{\{\theta - E(z)\}^2}{2\sigma^2(z)}} d\theta$$

$$= \int_{-\infty}^{y} f(z) dz = \text{kans dat } (z = y)$$

$$= P\left[\bar{x}_{\text{geprojecteerd op } \bar{w}} \text{ valt voor } x = \theta\right]$$

De kans op een foutieve klassifikatie is dus in het geval van de twee klassen:

$$P[\text{fout}] = \tfrac{1}{2}\cdot\left\{P[\text{geprojecteerde } \bar{x} \text{ ligt in } K_1] + 1 - P[\text{geprojecteerde } \bar{x} \text{ ligt in } K_2]\right\}$$

$$= \tfrac{1}{2}\left[\Phi\left\{\frac{\theta - \bar{w}^T\cdot\bar{\mu}_1}{(\bar{w}^T\cdot A_1\cdot\bar{w})^{1/2}}\right\} + 1 - \Phi\left\{\frac{\theta - \bar{w}^T\cdot\bar{\mu}_2}{(\bar{w}^T\cdot A_2\cdot\bar{w})^{1/2}}\right\}\right]$$

De gewichtsvector wordt bepaald uit de trainingsset, en omdat $\bar{w}$ en θ de verdeling van de ruimte in klassen (subruimten) bepalen, is de klassenverdeling afhankelijk van de keuze van de trainingsset. Aan de hand van een aantal trainingssymbolen kan volgens bovenstaande formule de kans op een fout berekend worden. Al naar gelang de toelaatbare fout, kan dus een juiste trainingsset samengesteld worden, die resulteert in een juiste keuze van de gewichtsvectoren en de drempelwaarde.
De theory van de lineaire beslissing en verdeling is bekend, zodat de klassifikatie geen probleem is. Het hedendaagse knelpunt ligt bij de trainingsmethoden die verbeterd moeten worden.

Omdat hier steeds van lineaire klassifikatie sprake is, wil ik nu enige opmerkingen maken over niet lineaire klassifikaties. Uitgezonderd enige speciale artikelen is er over dit gebied weinig bekend gemaakt. Cooper, Chow, Cover en Munson hebben zich in beperkte mate met dit onderwerp bezig gehouden. De laatste heeft iets gedaan aan „nearest neighbor decision". Hierbij wordt in de leerfase een patroon aan die klasse toegevoegd, waarmee hij de meeste overeenkomst vertoont. (bv. op metrische basis.) Dat dit niet de ideale oplossing is, kan een voorbeeld het beste toelichten:

Een slecht geschreven „e" kan metrisch even veel afwijken van de standaard „i" als van de „e", zodat de toewijzing niet eenduidig te bepalen is.

De toepassing van deze methode is alleen mogelijk in kombinatie met andere systemen.

2.2.2. Correlatie.

De overeenstemming, in metrische zin, tussen twee symbolen kan ook weergegeven worden door een correlatie. Dit soort correlatie kan alleen dan pas worden toegepast, als de meetvectoren genormaliseerd zijn. Het slagen van een herkenningsproces, op grond van correlatie, hangt voor een groot deel af van de juiste normalisatie. Het is vaak noodzakelijk ingewikkelde bewerkingen, moeilijke omwegen en slimme trucjes toe te passen om op een goede en efficiënte manier te kunnen normaliseren. Het begrip is dan ook een vakgebied op zichzelf. Ieder herkenningsproces heeft zijn eigen specifieke normalisatie. Het is hier niet de juiste plaats om er dieper op in te gaan, maar in de loop van het betoog komt het vaak en uitgebreider terug.

Een andere toepassing is de correlatie - bepaling tussen een onbekend symbool en de gemiddelde representanten van alle klassen, aan de hand waarvan een klassifikatie kan plaatsvinden. Het principe is gelijk aan dat van de lineaire klassifikatie, alleen wordt na ieder nieuw symbool $\bar{x}$ de waarden $\bar{w}$ en θ opnieuw bepaald. (een zich corrigerend proces.).
Stel representant van K_1 is $\bar{\mu}_1$ en van K_2 is $\bar{\mu}_2$.
De gewichtsvector krijgt de waarde

$$\bar{w} = \bar{\mu}_1 - \bar{\mu}_2$$

en de drempelwaarde wordt

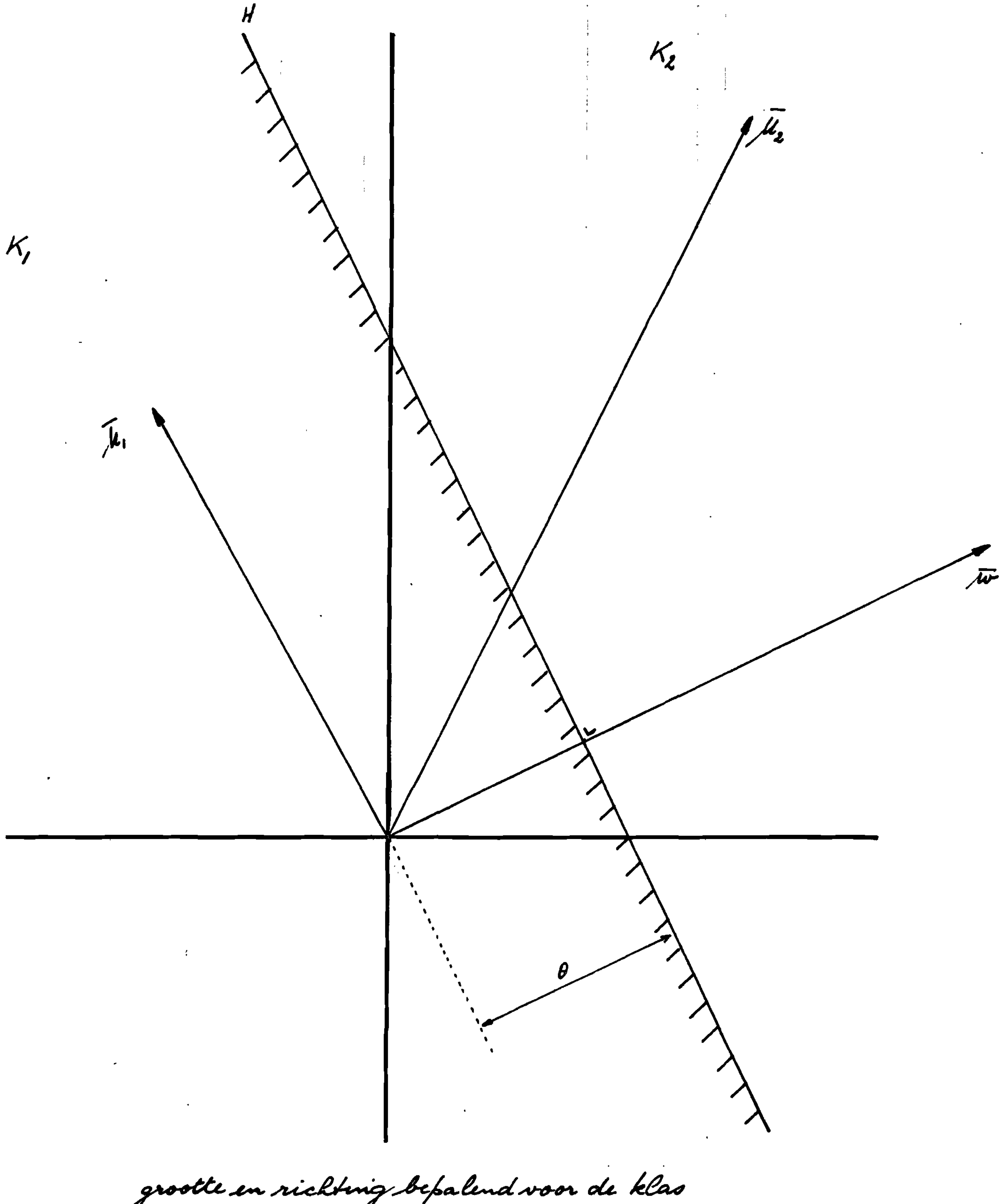
$$\theta = \tfrac{1}{2}\cdot\left\{(\bar{\mu}_1+\bar{\mu}_2)\cdot(\bar{\mu}_1-\bar{\mu}_2)\right\}$$

zie figuur 4.

Is het systeem van symbolen zo opgebouwd, dat de hoek tussen de verschillende vectoren van belang is, dan wordt als gewichtsvector en drempelwaarde genomen:

$$\bar{w} = \frac{\bar{\mu}_1}{|\bar{\mu}_1|} - \frac{\bar{\mu}_2}{|\bar{\mu}_2|} = \underset{\sim}{\bar{\mu}_1} - \underset{\sim}{\bar{\mu}_2}$$

$$\theta = 0$$



grootte en richting bepalend voor de klas

$\bar{\mu}_1 = \frac{1}{N_1} \sum_{x_i \in K_1} \bar{x}_i$ $\quad$ $\bar{\mu}_2 = \frac{1}{N_2} \sum_{x_i \in K_2} \bar{x}_i$ $\quad$ met N_i : aantal vectoren in K_i

FIG.-4

Natuurlijk gaat het scheidingsvlak in dit geval door de oorsprong. (zie figuur 5).
Het afstudeerverslag van H. Ponsen, T.H.D. juni 1966, beschrijft een methode voor het automatisch klassificeren van de meet-vector $\bar{x}$ aan de hand van de hoek die $\bar{x}$ maakt met de gemiddelde klassevectoren. Het systeem is als volgt (zie figuur 6):

$$\alpha = \angle(\bar{\mu}_1, \bar{x}) \qquad \cos\alpha = \frac{(\bar{\mu}_1 \cdot \bar{x})}{|\bar{\mu}_1| \cdot |\bar{x}|}$$

$$\beta = \angle(\bar{\mu}_2, \bar{x}) \qquad \cos\beta = \frac{(\bar{\mu}_2 \cdot \bar{x})}{|\bar{\mu}_2| \cdot |\bar{x}|}$$

Beslissingskriterium:

$\bar{x}$ behoort tot K_1 als: $\alpha < \theta_1$ en $\alpha < \beta$

$\bar{x}$ behoort tot K_2 als: $\beta < \theta_2$ en $\beta < \alpha$

Stel:

$$\underset{\sim}{\bar{\mu}}_1 = \frac{\bar{\mu}_1}{|\bar{\mu}_1|}$$

$$\underset{\sim}{\bar{\mu}}_2 = \frac{\bar{\mu}_2}{|\bar{\mu}_2|}$$

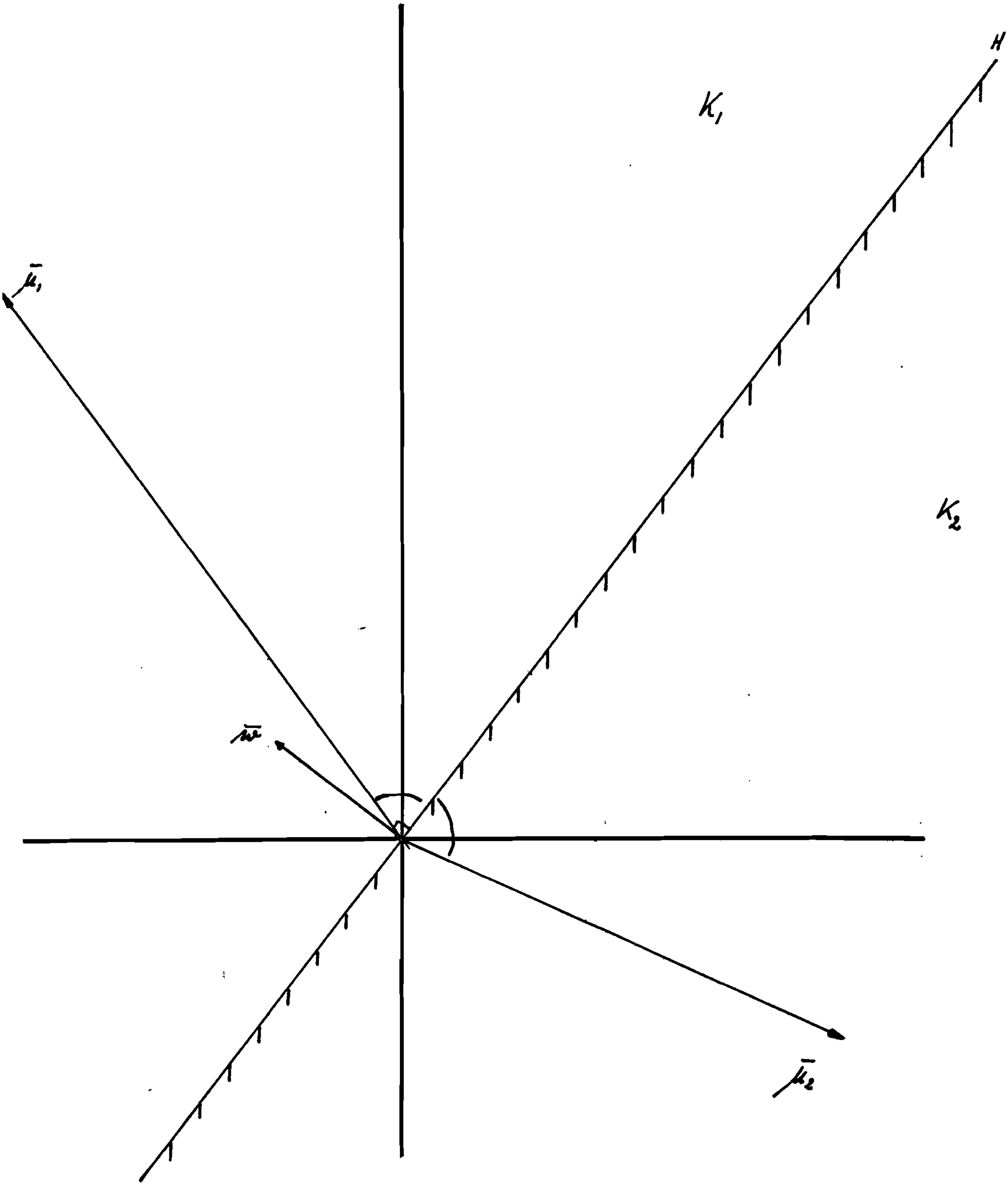
$$\underset{\sim}{\bar{x}} = \frac{\bar{x}}{|\bar{x}|}$$

Het beslissingskriterium voor bijvoorbeeld K_1 gaat dan over in:

$\bar{x}$ behoort tot K_1 als: $(\underset{\sim}{\bar{\mu}}_1 \cdot \underset{\sim}{\bar{x}}) < \cos\theta_1$

en $(\underset{\sim}{\bar{\mu}}_2 \cdot \underset{\sim}{\bar{x}}) - (\underset{\sim}{\bar{\mu}}_1 \cdot \underset{\sim}{\bar{x}}) > 0$

Een heel andere techniek van correleren wordt toegepast bij het vergelijken van een in digitale vorm getransformeerd symbool met een aantal standaardsymbolen. Deze correlatie kan uitgevoerd worden met behulp van „software" of met „hardware." In het algemeen wordt, om de efficiency, gekeken naar de overeenkomstige digitale bits die niet dezelfde informatie-inhoud hebben. Deze gevallen worden gesommeerd, en als de som een vooraf vastgestelde drempelwaarde overschrijdt, wordt het correlatieproces afgebroken, en wordt een nieuwe vergelijking opgezet met het volgend standaardsymbool. Het te herkennen symbool wordt tot die klasse benoemd, waarmee hij met de gemiddelde klassevector een



richting bepalend voor de klas.

FIG.-5

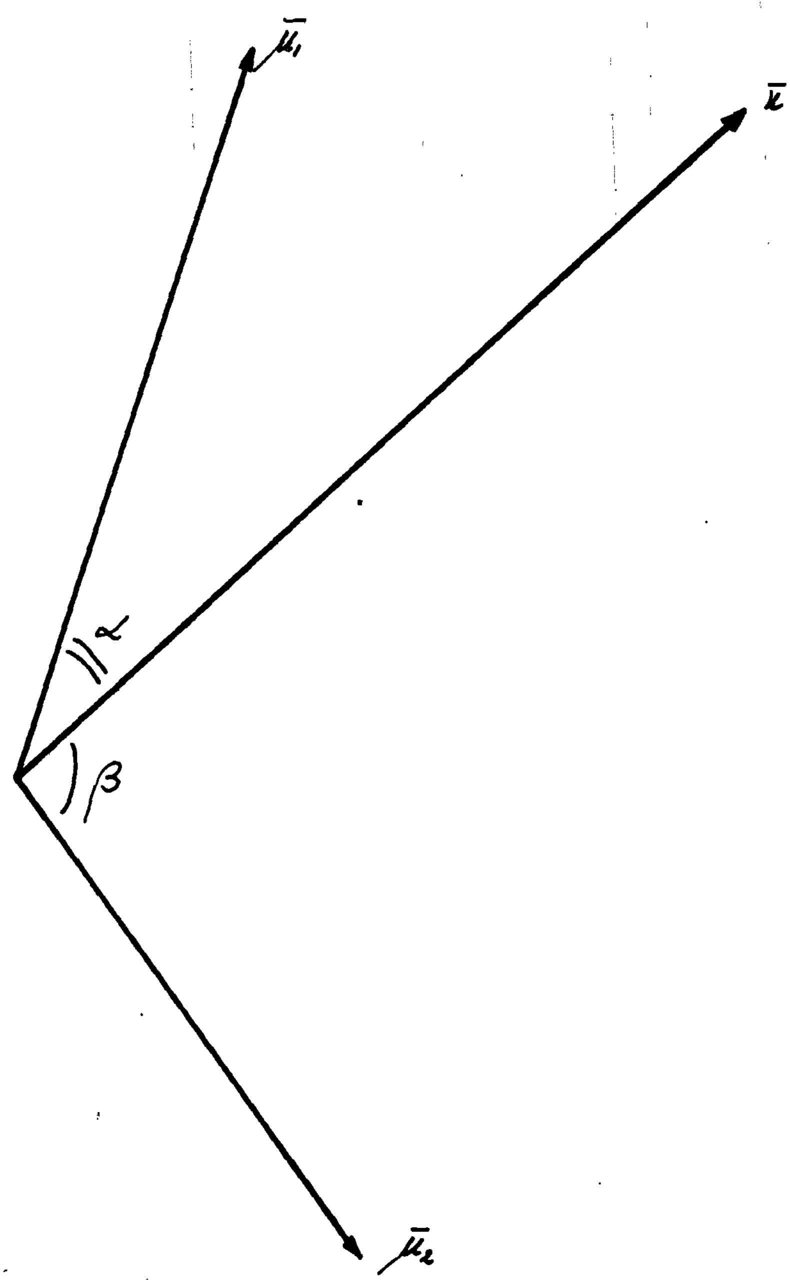


FIG.-6

minimum aantal bitfouten maakt. Het is triviaal, dat voor een groot aantal standaardsymbolen (bijvoorbeeld een alfabet van 50 symbolen) deze methode te veel tijd in beslag zal nemen.

Een continu variabele correlatie met de klasse-gemiddelden is door Steinbuch uitgevoerd met behulp van de „Lernmatrix".

Een zeer duidelijke toepassing van de correlatie wordt gevonden in de optische herkennings- en filter-methoden. Als speciaal onderwerp verwijs ik naar de optische filter-toepassing.

In het algemeen is op grond van theoretische beschouwingen de correlatiemethode alleen efficiënt toepasbaar, als de verdeling van de symbolen voldoet aan speciale symmetrische voorwaarden. Is hier aan voldaan, dan is de methode makkelijk toe te passen.

2.2.3. Maximale waarschijnlijkheid.

De methode van de scheidingsvlakken en die van de maximale waarschijnlijkheid zijn sterk aan elkaar verwant. In het afstudeerverslag van A.C.J. de Leeuw, T.H.E. januari 1968, is aangetoond, dat voor een kwadratisch kriterium de beide methoden identiek zijn. Het is volgens mij zelfs aan te tonen, dat de identiciteit ook geldt voor hogere graads criteria. De maximale waarschijnlijkheidsmethode oftewel optimale- of Bayes-methode wordt gekenmerkt door het probleem:

Als $\bar{x}_i$ de meetvector is, behoort $\bar{x}_i$ dan tot K_1 of tot K_2.
Anders gezegd:

Wat is groter:

$$P[K_1/\bar{x}_i] \quad \text{of} \quad P[K_2/\bar{x}_i]$$

Dit moet bij meerdere klassen uitgebreid worden tot:

$$P[K_j/\bar{x}_i]$$

Als nu geldt

$$P[K_1] = P[K_2]$$

dan is volgens de formule van Bayes:

$$P[K_1/\bar{x}_i] \lesseqgtr P[K_2/\bar{x}_i]$$

$$P[K_1 . \bar{x}_i] \lesseqgtr P[K_2 . \bar{x}_i]$$

$$\frac{P[K_1 . \bar{x}_i]}{P[K_1]} \lesseqgtr \frac{P[K_2 . \bar{x}_i]}{P[K_2]}$$

$$P[\bar{x}_i/K_1] \lesseqgtr P[\bar{x}_i/K_2]$$

Zijn de attributen x_{ij} statistisch onafhankelijk:

$$P[\bar{x}_i/K_1] = \prod_j P[x_{ij}/K_1] \qquad \ldots (1)$$

In de praktijk bevat x_{ij} altijd wel informatie die ook in x_{ik} $(j \neq k)$

is opgeslagen, zodat alleen bij benadering aan de statistische onafhankelijkheid is voldaan. Een benadering van dit ideale geval is een vorm van een Markovproces, waarbij het j^e attribuut alleen afhankelijk is van het $(j-1)^e$ attribuut.
Als x_{ij} binair is, dan kan worden bewezen:

$$P[K_1/\bar{x}_i] \geq P[K_2/\bar{x}_i] \iff \bar{x}_i^T . \bar{w} \geq \theta$$

met $\bar{w} = (w_1, w_2)$

en gewichtovector in richting j:

$$w_j = \ln \frac{P[x_{ij}=1/K_1].P[x_{ij}=0/K_2]}{P[x_{ij}=1/K_2].P[x_{ij}=0/K_1]}$$

$$\theta = \sum_j \ln \frac{P[x_{ij}=0/K_2]}{P[x_{ij}=0/K_1]}$$

zodat we via de Bayes-methode weer komen tot het scheidingsvlakken principe.

Het grote nadeel van de Bayes-methode is, dat alle kansen

$$P[K_n/\bar{x}_i]$$

uit het leerproces bepaald moeten worden, alvorens een deugdelijke beslissingsregel op te zetten. Deze kansen moeten in het geheugen van een rekenmachine opgeborgen worden. Een ander bezwaar is de fout die geïntroduceerd wordt door de attributen statistisch onafhankelijk te beschouwen.

2.2.4. De Anderson - Bahadur methode.

Deze methode wordt meerdere malen toegepast in de praktijk, en kan in het overzicht beslist niet weggelaten worden, ook al zijn de achtergronden sterk wiskundig.

Lineaire procedures worden bekeken voor twee multivariabele normaalverdelingen, waarbij de gemiddelde vector en de covariantie-matrix verschillend zijn. Aangenomen wordt dat alle parameters bekend zijn.
De twee verdelingen van de vectorvariabele x met p komponenten wordt gegeven door:

$$N(\mu_1, \Sigma_1) \qquad N(\mu_2, \Sigma_2)$$

μ_i = gemiddelde vector van K_i
Σ_i = cov.-matrix van K_i

De dichtheidsfunktie van de i^e klasverdeling is:

$$n(x/\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{p/2} \cdot |\Sigma_i|^{1/2}} \exp.\left[-\tfrac{1}{2}(x-\mu_i)^T \cdot \Sigma_i^{-1} \cdot (x-\mu_i)\right]$$

De theoretisch beste klassifikatie is gebaseerd op de waarschijnlijkheidsfunktie:

$$b = \frac{n(x/\mu_2, \Sigma_2)}{n(x/\mu_1, \Sigma_1)}$$

x behoort tot K_1 als $b < \theta$
x behoort tot K_2 als $b > \theta$

Als nu $\Sigma_1 = \Sigma_2$, dan is b een lineaire funktie van x, nl. de discriminantfunktie.
Als $\Sigma_1 \neq \Sigma_2$, dan is b een kwadratische funktie van x.
Zijn de normaalverdelingen univariabel dan is de log. van b:

$$\log\frac{\sigma_1}{\sigma_2} + \frac{1}{2}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)x^2 - \left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2}\right)x + \frac{1}{2}\left(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2}\right)$$

met $\Sigma_1 = \sigma_1^2$
$\Sigma_2 = \sigma_2^2$
$\sigma_2^2 > \sigma_1^2$

Het klassifikatieprobleem moet toegepast worden in een situatie waarbij de symbolen verdeeld zijn volgens een (benaderde) normaalverdeling rond hun respectievelijke gemiddelde klassevector. Het scheidingsvlak wordt gerepresenteerd door een (hyper-) vlak, zodat de hele klassifikatie lineair opgezet kan worden.

Stel vector $\bar{w} = (w_1, \dots, w_p) \neq 0$
scalar θ
$\bar{x} \in K_1$, als $\bar{w}^T \cdot \bar{x} \leq \theta$
$\bar{x} \in K_2$, als $\bar{w}^T \cdot \bar{x} > \theta$

Het belangrijkste verschil tussen K_1 en K_2 komt tot uitdrukking door $\bar{\mu}_1 - \bar{\mu}_2$ $(\neq 0)$.
Verder wordt gesteld $\Sigma_1 \neq \Sigma_2$. Dus $\bar{w}^T \cdot \bar{x}$ heeft een normaalverdeling met:

$$\text{gemiddelde} : \varepsilon_i \bar{w}^T \cdot \bar{x} = \bar{w}^T \cdot \bar{\mu}_i$$

$$\begin{aligned} \text{variantie} &: \varepsilon_i \left(\bar{w}^T \cdot \bar{x} - \bar{w}^T \cdot \bar{\mu}_i\right)^2 = \\ &= \varepsilon_i \bar{w}^T (\bar{x} - \bar{\mu}_i)(\bar{x} - \bar{\mu}_i)^T \cdot \bar{w} \\ &= \bar{w}^T \cdot \Sigma_i \cdot \bar{w} \end{aligned}$$

Kans op foutieve klassifikatie, als symbool van eerste verdeling komt:

$$P_1\left[\bar{w}^T \cdot \bar{x} > \theta\right] = P_1\left[\frac{\bar{w}^T \cdot \bar{x} - \bar{w}^T \cdot \bar{\mu}_1}{(\bar{w}^T \Sigma_1 \bar{w})^{1/2}} > \frac{\theta - \bar{w}^T \cdot \bar{\mu}_1}{(\bar{w}^T \cdot \Sigma_1 \bar{w})^{1/2}}\right] = 1 - \Phi\left\{\frac{\theta - \bar{w}^T \cdot \bar{\mu}_1}{(\bar{w}^T \cdot \Sigma_1 \cdot \bar{w})^{1/2}}\right\}$$

$$\text{met } \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}t^2} dt$$

Kans op foutieve klassifikatie, als symbool van tweede verdeling komt:

$$P_2\left[\bar{w}^T \cdot \bar{x} \leq \theta\right] = 1 - \Phi\left\{\frac{\bar{w}^T \cdot \bar{\mu}_2 - \theta}{(\bar{w}^T \cdot \Sigma_2 \cdot \bar{w})^{1/2}}\right\}$$

Het doel is nu, deze kansen minimaal te maken.
Hoe dit o.a. gedaan kan worden is beschreven in:

„ Annual of Math. Statistics " vol. 33. pp. 420-431 1962
„ Classification into two multivariate normal distributions with different covariance matrices. "

De Anderson - Bahadur methode is de statistische paralel van de lineaire beslissingsfunkties van Highleyman. Welke van de twee het beste kan worden toegepast in een konkreet probleem moet bepaald worden uit een experimentele vergelijking.
De Bayes' beslissingsmethode geeft betere resultaten dan de Anderson- Bahadur methode, ofschoon de laatste meer gewichten en bewerkingstijd nodig heeft. De eerste methode gebruikt de waarschijnlijkheden voor iedere meting in iedere klas, hetgeen met minder testsymbolen bereikt kan worden dan met de cov.-matrix die de Anderson - Bahadur methode gebruikt.

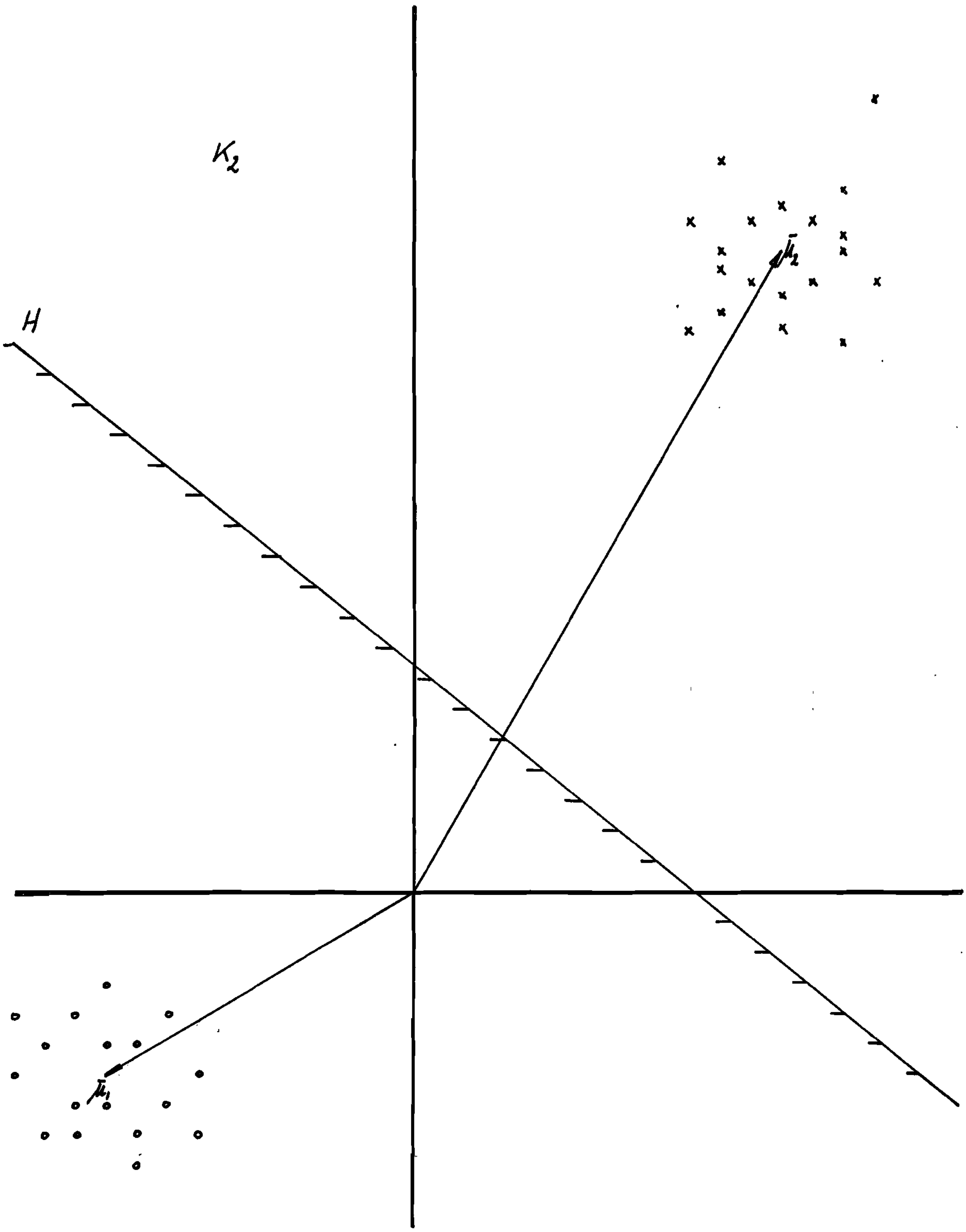
2.2.5. Discriminant Analyse.

Wanneer de waarschijnlijkheidsmatrix, die de verdeling van de symbolen aangeeft, niet bekend is, dan is het onmogelijk het hypervlak dat de klassen scheidt, via analytische weg te berekenen. Het is echter in zo'n geval toch nuttig te weten in welke richting de klassevectoren staan, en hoe groot ze zijn. Er moet getracht worden:

1/ de klassevectoren van verschillende klassen zo ver mogelijk uit elkaar te laten vallen.

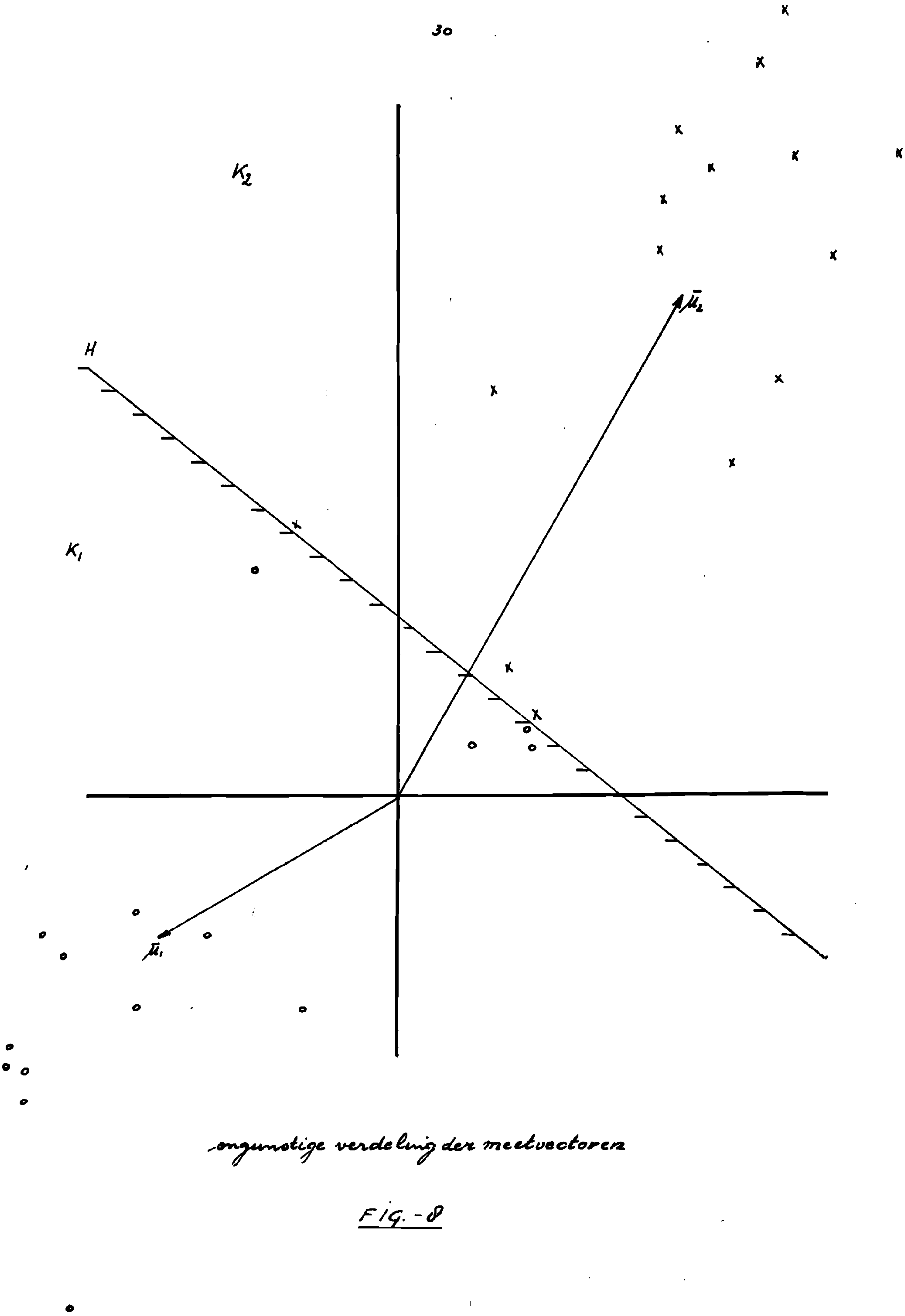
en 2/ de afstand tussen klassevectoren van dezelfde klas moet minimaal zijn.

Met deze probleemstelling houdt de Discriminant Analyse zich bezig.
In figuur 7 is een „ideale" verdeling geschetst. De vectoren van K_1 en K_2 liggen op grote afstand van elkaar, terwijl de onderlinge afstand van de vectoren, behorend tot één klas, klein is. In een dergelijke situatie kan niet gauw een verkeerde toewijzing plaatsvinden. Anders is het in de situatie van figuur 8. De afstand tussen μ_1 en μ_2 is weliswaar even groot als de overeenkomstige afstand in fig. 7, maar de onderlinge afstanden der klassevectoren is erg groot. Vanwege de grote spreiding in een klas, is de μ_i wel nog een gemiddelde vector, maar geen echte representant voor K_i. In dit geval is de kans op een foutieve toewijzing zeer groot.
In de clustertheorie geldt hetzelfde probleem. Goede toewijzing is pas mogelijk als aan 1/ en 2/ voldaan is; dus een verdeling in de trant van figuur 7. Komt de verdeling der attributen zo ongunstig uit als in figuur 8 geschetst, dan moet getracht worden het probleem op te lossen met andere meet-attributen. Het is natuurlijk mogelijk dat ook dit geen resultaat oplevert; in dat geval moet de oplossing gezocht worden met behulp van een heel ander systeem, bijvoorbeeld met filtertheorie of met niet lineaire transformatie van de symboolruimte.
Wordt de zaak van de formele kant bekeken, dan moet



wenselijke verdeling der meetvectoren

FIG.-7



ongunstige verdeling der meetvectoren

FIG. - 8

$$\sum_{\substack{\bar{x}_i \in K_1 \\ \bar{x}_j \in K_2}} |\bar{w}^T \bar{x}_i - \bar{w}^T . \bar{x}_j|$$

maximaal zijn, oftewel:

$$\sum_{\substack{\bar{x}_i \in K_1 \\ \bar{x}_j \in K_2}} (\bar{w}^T . \bar{x}_i - \bar{w}^T . \bar{x}_j)^T . (\bar{w}^T . \bar{x}_i - \bar{w}^T . \bar{x}_j) =$$

$$= \sum_{\substack{\bar{x}_i \in K_1 \\ \bar{x}_j \in K_1}} (\bar{w}^T . \bar{x}_i - \bar{w}^T . \bar{x}_j)^T . (\bar{w}^T . \bar{x}_i - \bar{w}^T . \bar{x}_j) + \sum_{\substack{\bar{x}_i \in K_2 \\ \bar{x}_j \in K_2}} (\bar{w}^T . \bar{x}_i - \bar{w}^T . \bar{x}_j)^T . (\bar{w}^T . \bar{x}_i - \bar{w}^T . \bar{x}_j)$$

moet maximaal zijn.
Met behulp van Lagrange vermenigvuldigers is $\bar{w}$ op te lossen.
Verder moet nog minimaal zijn

$$\sum_{\bar{x}_i \in K_1} |\bar{x}_i - \bar{\mu}_1| = \sum_{\bar{x}_i \in K_1} (\bar{x}_i - \bar{\mu}_1)^T . (\bar{x}_i - \bar{\mu}_1)$$

en $$\sum_{\bar{x}_j \in K_2} (\bar{x}_j - \bar{\mu}_2)^T . (\bar{x}_j - \bar{\mu}_2)$$

Het kan worden aangetoond dat $\bar{w}$ de eigenvector, met de grootste eigenwaarde, is van de vergelijking

$$(M_1 M_2^{-1} - \lambda I)\, \bar{w} = 0$$

met:

I = identiteitsmatrix

$$A = \sum_{\bar{x}_i \in K_1} (\bar{x}_i - \bar{\mu}_1)^T . (\bar{x}_i - \bar{\mu}_1) + \sum_{\bar{x}_i \in K_2} (\bar{x}_i - \bar{\mu}_2)^T . (\bar{x}_i - \bar{\mu}_2)$$

$$B = \sum_i \left(\bar{x}_i - \frac{\bar{\mu}_1 + \bar{\mu}_2}{2}\right)^T . \left(\bar{x}_i - \frac{\bar{\mu}_1 + \bar{\mu}_2}{2}\right) - A$$

Het is dus via deze berekeningen mogelijk om met behulp van de testdata, in niet al te ongunstige gevallen, een klassenindeling te maken. Deze methode geeft dan, evenals het principe van Anderson en Bahadur, een laag percentage foutieve beslissingen.

2.2.6. Lerende klassifikatie.

De tot nu toe behandelde systemen gaan uit van een data-set, waarbij de klassifikatie wordt bepaald door een simultane beschouwing van alle meetvectoren.
Het principe van de lerende klassifikatie gaat er van uit dat de meetvectoren een voor een beschouwd worden, waarbij na iedere vector, indien noodzakelijk, de gewichtsvector $\bar{w}$ wordt verbeterd. Indien een vector de juiste klassifikatie aantreft, blijft $\bar{w}$ ongewijzigd; is dit niet het geval dan wordt $\bar{w}$ aangepast aan alle voorafgaande vectoren $\bar{x}$ + de nu beschouwde vector. Zijn alle testvectoren doorlopen, dan kan het leerproces gecontinueerd worden door weer van voren bij de testvectoren te beginnen.
Stel er zijn N meetvectoren.
Bij de j^e stap beschouw ik de i^e vector, waarbij geldt: $i = j \bmod N$.
De aanpassing van de gewichtsvector geschiedt als volgt:

$$\bar{w}_{j+1} := \bar{w}_j + \bar{x}_j \;\;/\;\; \bar{w}_j \bar{x}_j \leq \theta \quad \wedge \quad \bar{x}_j \in K_1$$

$$\bar{w}_{j+1} := \bar{w}_j - \bar{x}_j \;\;/\;\; \bar{w}_j \bar{x}_j \geq \theta \quad \wedge \quad \bar{x}_j \in K_2$$

in alle andere gevallen: $\bar{w}_{j+1} := \bar{w}_j$

In tegenstelling tot de statistische methodes, zijn er voor deze methode geen formele hypothesen, maar men kan aannemen dat het eindresultaat ongeveer gelijk zal zijn aan dat van vorige methoden. Echter, dit principe stelt hogere eisen aan de „hardware"
De grote vraag is natuurlijk of dit systeem zal convergeren naar een eindwaarde, en hoe snel is die convergentie?

1) De ligging van de testsymbolen is in deze erg belangrijk. Een grote spreiding zal een gemiddeld langzaam verlopende convergentie tonen.
2) Een andere volgorde van de te doorlopen testsymbolen kan ook een ander verloop in de convergentiesnelheid

tot gevolg hebben.
Er zijn verschillende algorithmen die de convergentie kunnen aantonen. Een van de eerste is het „error correcting algorithm" van F. Rosenblatt. Naast dit algorithme zijn er nog zes of zeven andere bewijzen voor convergentie.
In het algemeen kan in een eindig aantal stapjes een oplossing gevonden worden, als het klassifikatieprobleem tenminste een lineaire oplossing heeft. De convergentietheorema's kunnen in dat geval dan ook toegepast worden. Het is te bewijzen dat er een oplossing bestaat dan en slechts dan als de test-symbolen lineair onafhankelijk zijn. Het is geen eenvoudige zaak de lineaire onafhankelijkheid aan te tonen. Een mogelijkheid is het simuleren van de lineaire klassifikatie op een rekenmachine, waarbij getracht wordt de gewenste klassifikatie aan te leren. Lukt dit na enkele pogingen niet, dan kan aangenomen worden dat niet voldaan is aan de lineaire onafhankelijkheid. In dat geval moet ik er op wijzen, dat het niet noodzakelijk is, dat het systeem convergeert naar een klassifikatie die een minimum aantal fouten zal maken.

2.2.6.1. <u>Perceptron.</u>

Het is de bedoeling enkele kenmerken van het Perceptron te vermelden; voor een uitgebreide beschrijving moet ik naar de literatuur verwijzen:

„Perceptron Simulation Experiments" Rosenblatt.
Proceedings of the IRE blz. 301 e.v. 1960
„The Perceptron: A Model for Brain Functioning" H. D. Block.
Reviews of Modern Physics vol. 34 No 1 jan. 1962.

Het Perceptron is een verwezenlijking van een lerend klassifikatie

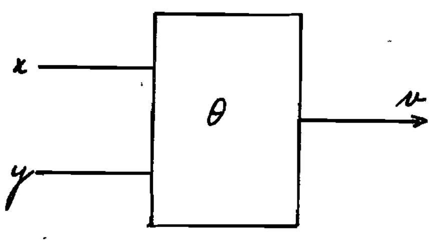
systeem.
De beschrijving van een simulatieprogramma voor bestudering van patroonherkenning moet voldoen aan:

1/ De simulatie moet opgezet worden aan de hand van een theoretische analyse.
2/ Geschikte metingen van het gedrag moeten worden opgesteld. Dus de uitgangsgrootheden moeten worden getransformeerd in goed te interpreteren signalen of grootheden.
3/ De opzet van de proef moet dusdanig zijn, dat er controlemogelijkheden aanwezig zijn voor triviale of tegenstrijdige resultaten.

Het perceptron is gedefinieerd door:

a/ Topologische opbouw van de elementen en hun verbindingen.
b/ De dynamische gedragingen van het systeem (elementen). Bijvoorbeeld: drempelwaarde, signaalsterkte, geheugenfunkties.

Een eenheid van een perceptron kan ik schematisch voorstellen door:



θ = drempelwaarde

$$v = \begin{cases} 1 & \text{als } x + y \geq \theta \\ 0 & \text{als } x + y < \theta \end{cases}$$

x = „prikkelpuls" = positieve puls
y = „verbodspuls" = negatieve puls
v is een stochastische grootheid met uitkomstenruimte $(0,1)$
In principe kan ik nu een signaal v krijgen met behulp van
$\Sigma x + \Sigma y = X + Y$

2.2.6.2. Sequentiële beslissingen.

Ook deze techniek gaat niet uit van een simultane beschouwing der meetvectoren, maar werkt met een sequentieel proces. Als stelregel geldt:

> De opvolgende meting wordt altijd daar gekozen, waar de meeste informatie wordt gegeven over de klas met de hoogste waarschijnlijkheid op een fout.

Volgens deze stelregel worden dus steeds betere meetvectoren in het beslissingsproces betrokken, zodat de convergentie-snelheid in de loop van het proces zal toenemen. Het proces zet zich voort totdat:

a) Er geen gebied meer is waar de waarschijnlijkheid op een fout groter is dan een van te voren vastgestelde waarde θ.

b) Alle metingen in een bepaald gebied zijn uitgevoerd.

Deze methode lijkt veel op dynamisch programmeren, en is dus zeer geschikt voor uitvoering met behulp van een rekenmachine.

2.2.7. Potentiaalfunkties.

Deze methode heeft een zeer mathematische grondslag. Vooral de definitie doet denken aan systemen die vaak in de exacte logica worden gebruikt. Er wordt dan een een vervangingsfunktie gecreëerd, die het voorstellingsvermogen meer aanspreekt.

Als de potentiaalfunktie $\psi(\bar{x})$ gedefinieerd is in een symboolruimte X, waarin geen overlappingen zijn van de verschillende verdelingen, dan geldt per definitie:

$$\psi(\bar{x}) > 0 \quad \text{voor } \bar{x} \in K_1$$
$$\psi(\bar{x}) < 0 \quad \text{voor } \bar{x} \in K_2$$
$$\psi(\bar{x}) = 0 \quad \text{voor } (\bar{x} \notin K_1 \wedge \bar{x} \notin K_2)$$

Er wordt nu aangenomen dat de potentiaalfunktie vervangen kan worden door een sommatie van een eindig aantal termen van funkties in een orthonormaal systeem. Stel dit aantal is m.

$$\psi(\bar{x}) = \sum_{i=1}^{m} w_i \, \phi_i(\bar{x})$$

Ga ik een transformatie toepassen van de n-dimensionale symboolruimte X naar de m-dimensionale linearisatieruimte Z, dan gaan de coordinaten van $\bar{x}$ over in:

$$(z_1, z_2, \ldots, z_m)$$
$$\text{met } z_i = \phi_i(\bar{x})$$

Verder gaat het scheidingsvlak $\psi(\bar{x}) = 0$ over in:

$$\sum_{i=1}^{m} w_i z_i = 0$$

Dit principe geeft een methode aan om de moeilijk te interpreteren en te bewerken potentiaalfunktie te vervangen door een beeld in een m-dimensionale ruimte, waardoor ná het opnemen van een eindig aantal trainingssymbolen, de

coëfficiënten w_i bepaald kunnen worden. Deze methode is gedeeltelijk ook toegepast door A.C.J. de Leeuw. (T.H.E. januari '68, afstudeerverslag: „Beschouwingen over Methoden der patroonherkenning".)
In het algemeen wordt deze methode toegepast als:

1/ de testsymbolen niet onderverdeeld kunnen worden in klassen.

2/ de klassen elkaar overlappen.

Het klassieke orthonormale systeem voldoet niet goed in het voorzien van optimale onafhankelijke kenmerken. De funktie $\sum_{i=1}^{m} w_i \phi_i$ moet immers zoveel mogelijk overeenstemmen met $\psi(\bar{x})$, zodat de vrijheidsgraad in de keuze van de onafhankelijke kenmerken beperkt wordt. Dit is nogal bezwaarlijk omdat de optimaliteit een van de eerste vereisten is.
Deze methode werpt veel praktische moeilijkheden op, zodat een ruime toepassing in de symboolherkenning achterwege is gebleven.

2.3 Zelf-korrigerend systeem.

Is een trainingsgroep representatief voor de te verwachten symbolen, en is het aantal symbolen groot genoeg, dan is, zoals in het voorgaande is gebleken, een klassifikatiemethode op te zetten. Is aan deze voorwaarden niet voldaan of verandert het karakter van de symbolen in de loop der tijd, dan moet een ander systeem worden toegepast. In aanmerking komen:

1/ willekeurig lerende systemen.
2/ zelf-korrigerende systemen.
3/ adaptieve systemen.

In de praktijk komen de onder twee genoemde systemen het meest voor. Als voorbeeld van een zelfkorrigerend systeem noem ik het overseinen van gecodeerde informatie met behulp van een pulstrein. Het niet waarnemen van een puls of een pulsflank, kan er toe leiden dat de gehele verdere informatie verkeerd geïnterpreteerd wordt. Als korrigerende faktor zou er bijvoorbeeld een synchronisatiesignaal tussen de informatie geplaatst kunnen worden. De plaats kan tijd- of woord-afhankelijk genomen worden. Treedt er nu een foutieve informatie op, dan kan de fout, ten gevolge van deze verkeerde beslissing, hoogstens door blijven werken tot het volgende synchronisatiesignaal.

Bij een variatie van het karakter van de symbolen kan men zich afvragen of alle karakterbepalende eigenschappen veranderen, of slechts een gedeelte daarvan. In het eerste geval moet een zelfkorrigerend systeem toegepast worden; in het tweede geval hoeft dit niet noodzakelijk te zijn. Indien de stabiele karakter-eigenschappen genoeg attributen opleveren om een klassifikatie op te bouwen met een voldoend hoge zekerheid, dan hoeven we ons alleen maar te oriënteren op deze eigenschappen; in

hoeverre de variabele kenmerken veranderen is dan niet meer van direkt belang.
Een heel ander probleem is het convergentievraagstuk. Het wijzigen van het symboolkarakter heeft een negatieve invloed op de convergentie. Ook is de convergentie een funktie van de wijzigingssnelheid. Ligt deze snelheid boven een drempelwaarde, dan kan er geen convergentie meer optreden.
Op theoretische basis zijn er enige studies gedaan met behulp van verschillende algorithmen die toegepast werden op niet herkende symbolen. Het resultaat is, dat de variaties der attributen binnen beperkte grenzen moet blijven.

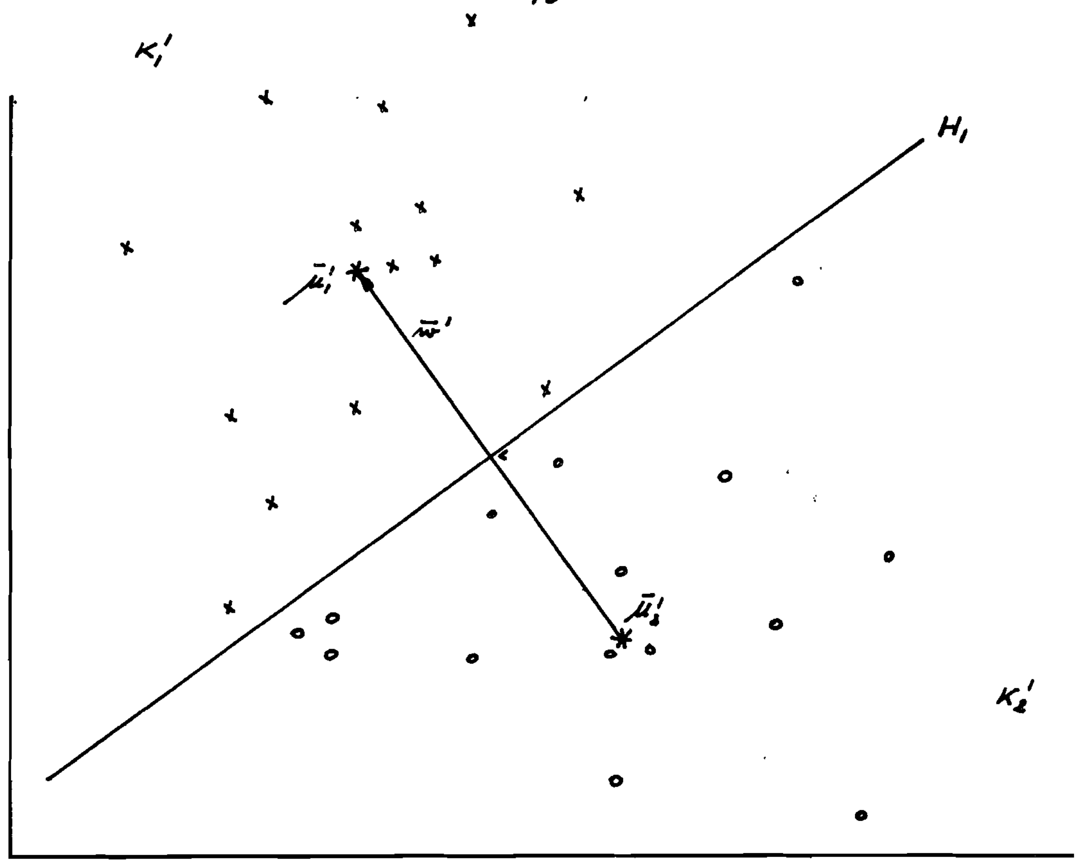
1/ translatie is toegestaan.
2/ rotatie over een hoek van maximaal ± 16°.
3/ de attributen moeten blijven voldoen aan een bepaalde groep van kenmerken.

Het komt bij handgeschreven tekst ter sprake, dat in geval van een sterk variërend karakter van de symbolen, een beroep kan worden gedaan op de context. In het algemeen kan dit alleen als de symbolen een onderlinge afhankelijkheid vertonen. In de praktijk blijkt echter, dat het zeer moeilijk is deze afhankelijkheid vast te leggen en in een herkenningproces te gebruiken.
In het algemeen moet getracht worden, situaties die aanleiding geven tot variaties in het karakter van de symbolen te vermijden. Het is dus in dit geval vooral belangrijk de informatie producerende bron aan te passen aan het herkenningssysteem. Zo komt men ook tot de konklusie:

„ Bij herkenning van handgeschreven tekst, moet de verbetering vooral liggen bij de bron: de mens."

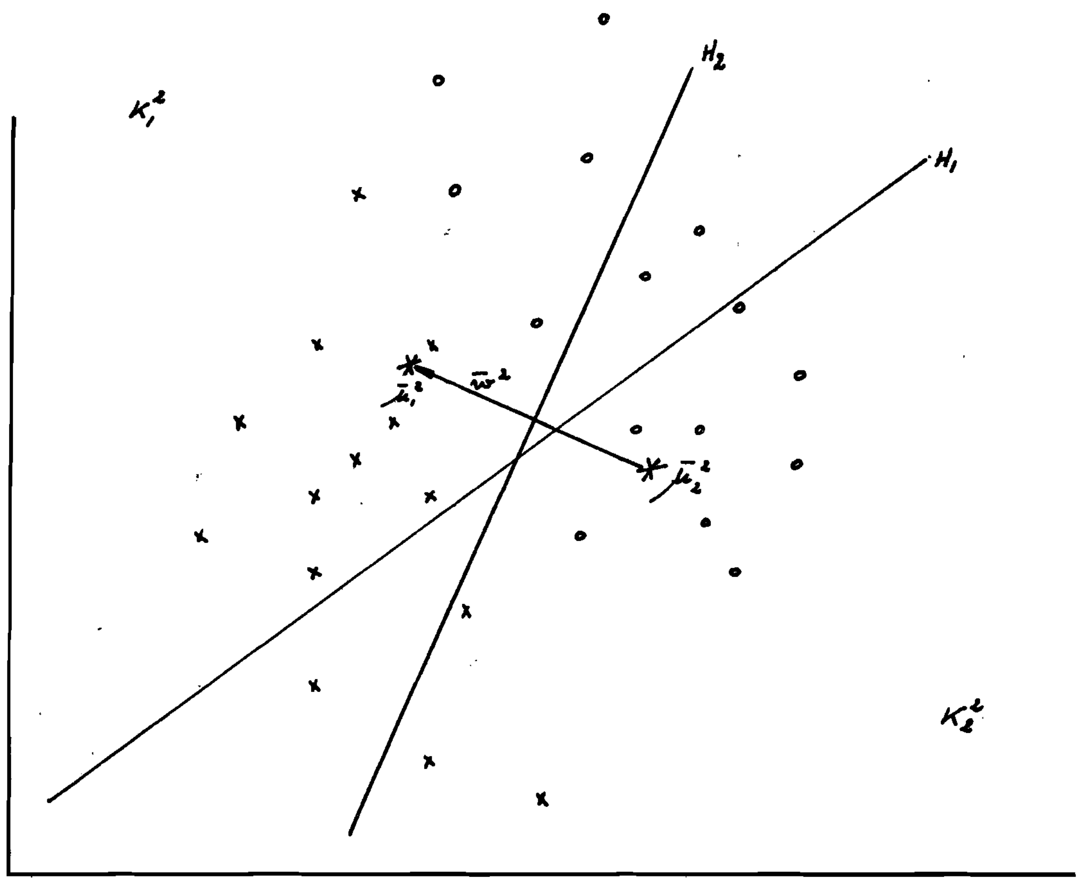
Een universeel schrift met een kleine spreiding zou heel wat moeilijkheden omzeilen.

Als voorbeeld van een zelf korrigerend systeem, zal een adaptief algorithme voor groepen gegevens besproken worden, waarbij de attributen verdeeld zijn in groepen met onderling verschillende kenmerken. Deze methode is afgebeeld in fig. 9.
Aan de hand van de eerste verdeling der twee groepen



K_1^1
H_1
$\bar{\mu}_1^1$
$\bar{w}^1$
$\bar{\mu}_2^1$
K_2^1

FIG - 9 a



K_1^2
H_2
H_1
$\bar{\mu}_1^2$
$\bar{w}^2$
$\bar{\mu}_2^2$
K_2^2

FIG - 9 b

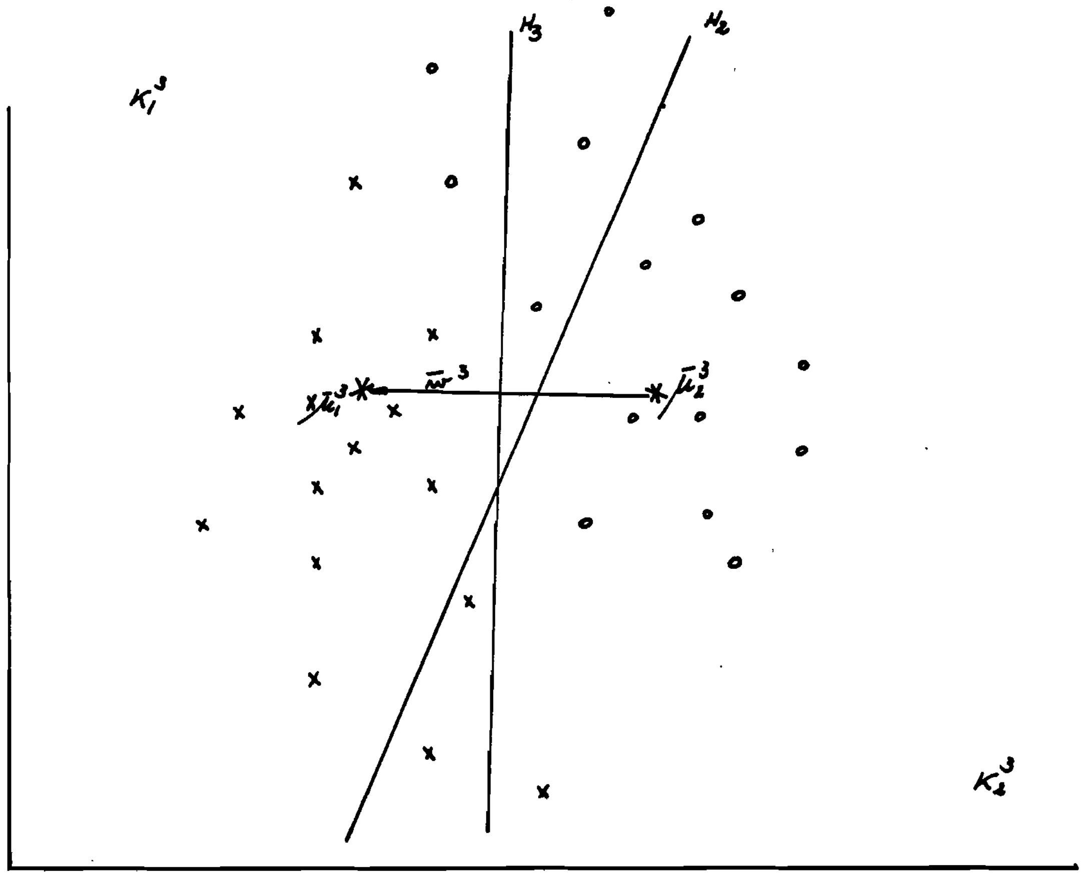

H3
H2
K_1^3
$\bar{w}^3$
$\bar{\mu}_1^3$
$\bar{\mu}_2^3$
K_2^3

FIG-9c

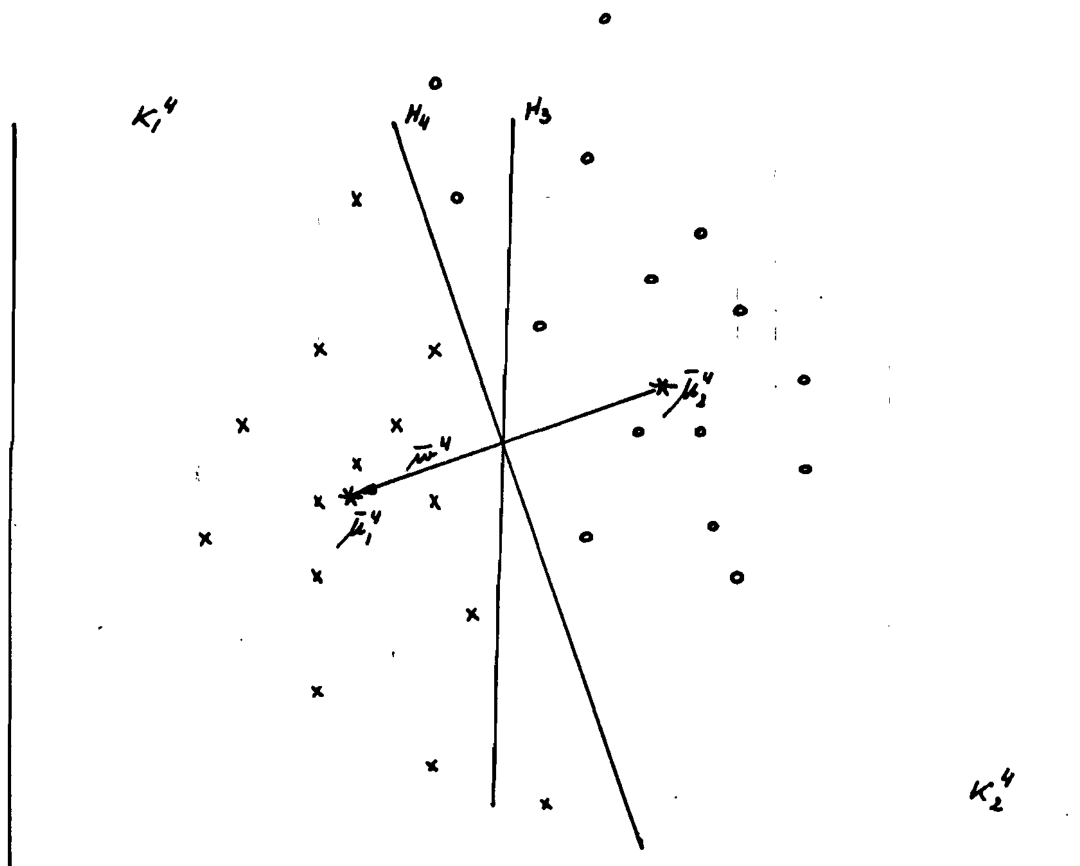

H4
H3
K_1^4
$\bar{w}^4$
$\bar{\mu}_1^4$
$\bar{\mu}_2^4$
K_2^4

FIG-9d

elementen wordt met behulp van

$$\bar{w} = \bar{\mu}_1 - \bar{\mu}_2$$
$$\theta = \tfrac{1}{2}\,|\bar{w}|$$

het scheidingsvlak H, bepaald.
Na wijziging van de verdeling der attributen gaat deze verdeling niet meer op. Er wordt een iteratieproces opgezet, totdat er weer een korrekte klassifikatie is:

$$\bar{\mu}_1^{j} = \frac{1}{N_1^{j-1}} \sum_{i \in K_1^{j-1}} \bar{x}_i$$ ≡ gemiddelde van de attributen liggend in K_1^{j-1} bij de j^e meting.

$$\bar{\mu}_2^{j} = \frac{1}{N_2^{j-1}} \sum_{i \in K_2^{j-1}} \bar{x}_i$$

$$\bar{w}^{j} = \bar{\mu}_1^{j} - \bar{\mu}_2^{j}$$

$$\theta^{j} = \tfrac{1}{2}\,|\bar{w}^{j}|$$

waarbij: N_1^{j-1} = aantal attributen, op moment j, liggend in de klas K_1^{j-1}

$N_1^{j-1} + N_2^{j-1} = N$ = totaal aantal attributen.

In de figuur neemt het aantal foutieve klassifikaties af:

figuur 9_b : 8
9_c : 3
9_d : 0

Zoals te zien is, is het verdelingspatroon in figuur b, c en d identiek. Dit wil niet zeggen dat er geen verandering der attributen meer is; het iteratieproces heeft een dusdanige snelheid, dat de attributen tijdens een itteratieproces niet veranderen. Het totale proces is een sommatie van dergelijke iteratieprocessen, waarbij er slechts een wijziging van de verdeling der attributen is per itteratie. De wijziging kan natuurlijk ook frequenter plaatsvinden, maar dit maakt in principe niets uit.
Wat betreft de convergentie kan het volgende gezegd worden. Stel het gedeelte der symbolen van klas K_i, dat

wordt toegewezen aan K_j is P_{ij}. Nu zal het systeem convergeren als voldaan is aan:

$$P_{11} \cdot P_{22} > P_{12} \cdot P_{21}$$

Enkele variaties van het hier beschreven algorithme zijn door Nagy toegepast bij tekenherkenning.

2.4. Groepen Analyse.

Tot nu toe waren de symbolen bepaald door een klasse-indeling. Bij groepen weten we niets van deze indeling; zelfs het aantal groepen en de aard daarvan is meestal onbekend. De groepenindeling is afhankelijk van een complex aantal attributen. Voor het bepalen van de juiste meetvectoren met behulp van foutendetectie is nog geen objectief uitvoeringsschema opgezet, zodat de verwezenlijking van een standaardmethode voor groepsherkenning nog niet mogelijk is.

Een belangrijke metrische parameter is de afstand. De overeenkomst tussen elementen en groepen kunnen we weergeven met behulp van afstandsfunkties. Het is nl. mogelijk een groep in een metrische ruimte te plaatsen, zodat de onderlinge afstanden der elementen een maat zijn voor de groepseigenschappen. De afstand tussen twee elementen kan dan ook beschouwd worden als een maat voor de correlatie.

Zijn de afstanden te klein zodat er geen significant onderscheid waar te nemen is, dan is het aan te bevelen de symboolruimte dusdanig te transformeren. dat de onderlinge afstanden in grootte-orde gelijk zijn aan een vooraf gestelde maat. De afstandsmeting bij binaire symbolen kan het beste verwangen worden door de funktie z („Hamming-afstand"):

$$y = (\bar{x}_i + \bar{x}_j) \text{ mod. } 2$$

z = aantal „enen" in y = gewicht van y

b.v.: $\bar{x}_i = 0110 \quad \bar{x}_2 = 01101$

$$y = 00011 \longrightarrow z = 2$$

waarbij z minimaal is bij maximale correlatie en minimale afstand.

Is de informatie, verkregen door analyse van de attributen, onvoldoende, dan kan nog altijd een factor- en/of componenten-analyse toegepast worden, die als hoofddoel heeft, het verkrijgen van meer informatie omtrent de opbouw en struktuur van een verzameling gegevens. Een voorbeeld van een dergelijke analyse is de „Ketting methode." In principe is de werking als volgt:

Ga uit van een exemplaar e_1,
wijs een groep toe: $e_1 \in$ groep $1 = G_1$,
bepaal afstand van e_1 tot e_2: x_{12}
als $x_{12} > \theta$ dan $e_2 \in G_2$
anders $e_2 \in G_1$

bepaal afstand van e_i tot e_j: x_{ij} met $i = 1, 2, \ldots, j-1$
en $(e_1 \cup e_2 \cup \ldots \cup e_{j-1}) \in (G_1 \cup \ldots \cup G_k)$
als $x_{ij} > \theta$ voor iedere $i < j$ dan $e_j \in G_{k+1}$
anders $e_j \in G_p$ met $p \in (1, 2, \ldots, j-1)$

Aan de hand hiervan kan een symmetrische vergelijkings-matrix opgezet worden:

$$A[a_{11}, \ldots, a_{nn}]$$

met:

$$a_{ij} = \begin{cases} 1 \text{ als } x_{ij} < \theta \\ 0 \text{ als } x_{ij} \geq \theta \end{cases}$$

Is het aantal te verwachten groepen bekend, dan worden de groepen zo ingedeeld, dat de gemiddelde spreiding in een groep minimaal is. Voor deze indeling zijn algorithmen ontwikkeld door o.a.: Casey en Braverman.
In de praktijk zal rekening moeten worden gehouden met faktoren zoals ruis en drukfouten. Als bijvoorbeeld ruis-informatie ver van een groep verwijderd is, kan er als het ware een „brug" geslagen worden tussen groep en ruis, zodat een niet gewenst informatiegedeelte bij de groep wordt gevoegd. Dit kan de herkenning bemoeilijken en zelfs onmogelijk maken. Het is dus zaak om deze stoorfaktoren niet in het beslissingsproces op te nemen. Er moeten dus in het algorithme

voorwaarden worden opgenomen, die er voor zorgen, dat deze faktoren geen rol spelen in het proces.
Een toepassing is het herkennen van drukletters, die ondergebracht zijn in een n x m raster.
De ruisinformatie kan hierbij storend werken doordat bijvoorbeeld de „P" als „R" geïnterpreteerd wordt. De vraag is nu, of de restrictieregels, die aan het herkenningsalgorithme zijn toegevoegd, hier een oplossing kunnen brengen. Dit blijkt meestal niet het geval. Er moet dan in een dergelijke situatie een ander systeem worden toegepast, bijvoorbeeld ruisfilters. Ik kom daar later nog op terug.

2.5. Destillatie van Kenmerken.

Het is van urgent belang om uit een gegeven trainingspopulatie juist die kenmerken te destilleren, die aanleiding geven tot een optimaal herkenningsproces. De meetvectoren bepalen in hoge mate de herkenningsnauwkeurigheid en de herkenningstijd. In de praktijk wordt een nauwkeurigheid en een procestijd aangegeven, aan de hand waarvan de meetattributen bepaald moeten worden.
Bij symbolen die in zeër dicht bij elkaar gelegen klassen liggen, is het moeilijk om een onderscheid te maken. In zo'n geval moet een sterk niet-lineair systeem worden toegepast om voldoende onderscheid te kunnen maken tussen de verschillende symbolen. Een essentieel andere aanpak is:

het systeem zoveel mogelijk vereenvoudigen, zodat een onderscheid kan worden gemaakt op eenvoudige basis.

Het vereenvoudigen wordt o.a. gedaan door:

a/ transformeren van symboolruimte, zò dat de symbolen minder variabelen hebben, terwijl de variabelen die overgebleven zijn onderling een groot verschil vertonen.
b/ reduceren van symboolruimte, met het risiko dat de beslissingsfunkties ingewikkeld worden.

Het vereenvoudigen van het systeem is een vorm van preprocessing. Dit is een heel belangrijk proces, omdat hiermee de herkenning vaak sterk vereenvoudigd kan worden. Als voorbeelden van o.a. in preprocessing toepasbare technieken wil ik noemen:

1/ Ruis filteren.
Eenvoudige filters verwijderen de niet belangrijke informatie.

Alleen het essentiële blijft over.

2/ Normalisatie van afmetingen

Het lineair transformeren van symbolen is erg belangrijk. Er wordt getracht de symbolen genormaliseerd te produceren, zodat de transformatie gedeeltelijk kan wegvallen. Als voorbeeld noem ik het schrijven van cijfers in aangegeven vakjes.

Vooral voor correlatiemethoden is een voorafgaande normalisering zeer belangrijk. De normalisatie kan uitgevoerd worden met behulp van „hardware" of „software."

Er zijn algorithmen die de lengte, breedte en de richting normaliseren. Het bezwaar van de „software"-methode is de noodzaak van het gebruik van een al of niet „special purpose" computer.

3/ dimensie-reduktie

Dit geeft een mogelijkheid de verwerkingstijd te verkleinen. Het is echter een moeilijke en niet triviale zaak, welke dimensies weggelaten kunnen worden. Een juiste reduktie kan de nauwkeurigheid van het systeem verhogen, omdat de stoorfaktoren geen aandeel meer hebben in het beslissingsproces. Wordt verkeerd gereduceerd, dan kan de onnauwkeurigheid zo toenemen, dat herkenning onmogelijk wordt.

4/ Selectief algorithme.

Dit algorithme is opgesteld voor het vinden van een optimaal bruikbare set meetattributen.

5/ Intuïtieve en Random metingen.

Het is een bekend feit dat er vele malen intuïtief een aantal meetgrootheden is opgezet voor een herkenningsproces. In hoeverre de keuze juist was, moest blijken uit de herkenningsresultaten. Dit is uiteraard niet de meest elegante oplossing, maar helaas is vaak geen duidelijk algorithme op te stellen voor het bepalen van de juiste kenmerken. In zo'n geval is het noodzakelijk om met verschillende attributen vergelijkende

herkenningstesten op te zetten, om dan met behulp van de rekenmachine de meest optimale keuze te bepalen. Het hoeft niet, dat de gevonden oplossing de absoluut optimale oplossing is. Bij deze processen kan uiteraard gebruik worden gemaakt van de uit andere proefnemingen naar voren gekomen resultaten.
In het algemeen kan gezegd worden, dat het destilleren van kenmerken nog te veel gebeurt op basis van intuïtiviteit, omdat andere mogelijkheden niet toepasbaar zijn. Het is dan ook van urgent belang algorithmen op te stellen die dit probleem zullen oplossen. Dit zal echter geen eenvoudige zaak zijn vanwege de genuanceerdheid van de optredende problemen.

6/ Geometrische en topologische kenmerken.
Geometrische en topologische kenmerken worden gebruikt in het herkenningssysteem van de mens. Het is dus de moeite waard eigenschappen van patronen vast te leggen, die betrekking hebben op dit gebied. Het vooronderzoek, dat noodzakelijk is, om dergelijke kenmerken te analyseren, moet een zo breed mogelijk scalar van symbolen bestrijken. Duidelijk is dan, dat dit, ondanks gebruik van een rekenmachine, een tijdrovende bezigheid is. Dat dit echter noodzakelijk is, zal o.a. blijken uit de beschrijving van een methode voor cijferherkenning. In deze methode wordt ook de zogenaamde „strook-informatie" toegepast, waarbij het symbool in horizontale of vertikale stroken verdeeld wordt, terwijl de informatie per strook wordt bekeken. Dit laatste is een vorm van dimensie-uitbreiding.

7/ Scannen.
Dit kan o.a. gebeuren met:

a een computer gestuurde „flying spot scanner"

b een door hardware of software gestuurde „lijn volger"

c „on-line scanning"

d op informatie terugkaatsende lichtbundel die via een lenzenstelsel op fotocellen valt.

De mate van uitvoerbaarheid van een bepaalde toepassing is inherent aan de mate van moeilijkheid. Het herkennen van een symbool dat in de praktijk op zeer verschillende manieren gepresenteerd wordt, terwijl al deze symbolen herkend moeten worden als één symbool, is een van de moeilijkste punten bij tekenherkenning. Het vinden van herkenningspunten die bij het merendeel der symbolen voorkomen is de oorsprong van alle moeilijkheden.
Tot slot dient nog te worden opgemerkt, dat niet alles bij tekenherkennig op theoretische basis opgezet kan worden. Vele punten zullen experimenteel bepaald moeten worden. Om deze experimenten goed op te zetten is echter een theoretische kennis noodzakelijk.

3. Adaline

Het gehele proces van tekenherkenning is te verdelen in leer- en herkenningsfase. Het leerproces is te verwezenlijken volgens verschillende technieken; een daarvan is het adaptief leren.
Als duidelijk voorbeeld van een leer- en herkenningsfase kan de Adaline dienen. (Adaline is een afkorting voor: adaptive linear threshold element.). Ook de verschillende problemen, zoals convergentie, mogelijkleden en toepassing zijn aan de hand van dit model duidelijk te bespreken. De drempelwaarde-elementen en drempel-logica (threshold logic) spelen een grote rol.

Beschrijving van de Adaline.

Zie figuur 10.
De Adaline heeft $n+1$ ingangen x waarvoor geldt

$$x_i = \pm 1 \qquad \text{voor } i \neq 0$$
$$x_0 = +1$$

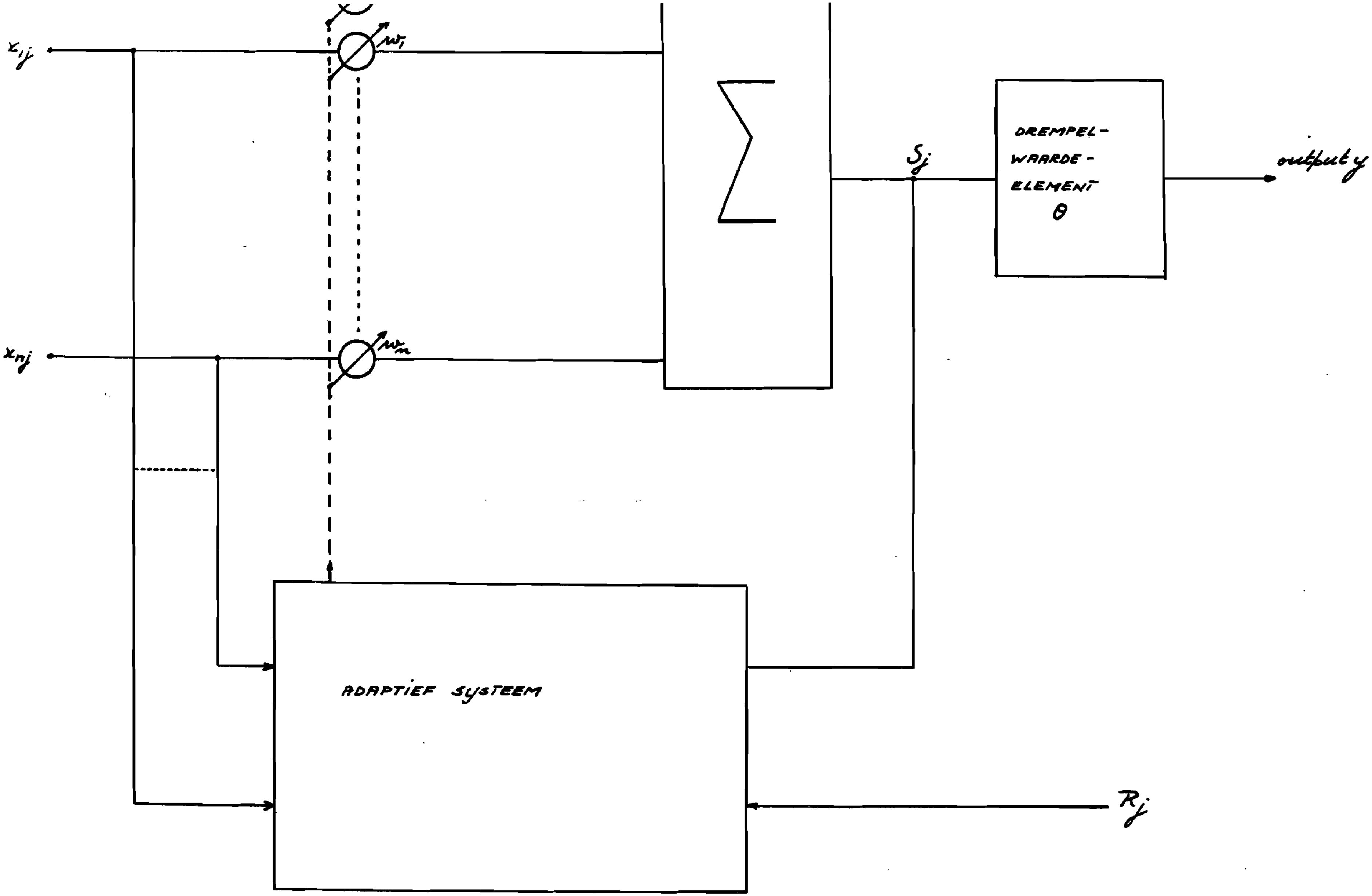
De ingang x_i gaat via een potmeter w_i naar een sommator, die als uitgang: S heeft. De output kan, afhankelijk van S en van een DWE (drempelwaarde-element) een uitgangsspanningsniveau geven. In de leerfase heeft het systeem een terugkoppeling via een adaptief systeem, werkend op de potentiometers w_0 t/m w_n. De gewichten w_i geven de relatieve importantie van x_i t.o.v. S aan.
Het te leren of te herkennen patroon wordt geconverteerd in een vector

$$\bar{x} \ (x_1, \ldots, x_n)$$

waarbij x_i aan bovenstaande voorwaarde moet voldoen.
Het k^e aangeboden patroon is

$$\bar{x}_k = (x_{1k}, \ldots, x_{nk})$$

Het toevoegen van x_0 is te zien als een dimensievergroting.



ADAPTIEF LINEAR THRESHOLD ELEMENT : ADALINE

FIG.-10

De belangrijkheid daarvan komt in de appendix ter sprake. (Door x_0 toe te voegen, voldoen de hypervlakken aan de algemene opstelling.). De x_0 kan ook gezien worden als een relatieve verschuiving van de drempelwaarde.

De vectorcomponent x_i wordt vermenigvuldigd met de gewichtsfactor w_i. De som S wordt dus:

$$S_k = \sum_{i=0}^{n} a_i\, x_{ik} = \bar{a}^T . \bar{x}_k$$

De drempelwaarde is θ_k. De output y van het OWE wordt gedefinieerd door:

$$y_k = \begin{cases} 1 \text{ als } \sum_{i=0}^{n} a_i x_{ik} \geq \theta_k \\ 0 \text{ als } \sum_{i=0}^{n} a_i x_{ik} < \theta_k \end{cases}$$

Wil ik bij het j^e patroon een sommatie-uitgang $S = R_j$ krijgen, dan moet het adaptief systeem een korrektiefaktor C toepassen, met zoals zal blijken:

$$C = \frac{R_j - S}{n+1}$$

De werking van het adaptief systeem is als volgt:

Bepaling van $R_j - S$ en $C = \frac{R_j - S}{n+1}$

Daarna verandering van de potmeters volgens:

$$\left.\begin{array}{l} w_i := w_i + C \quad \text{als } x_i = 1 \\ w_i := w_i - C \quad \text{als } x_i = -1 \end{array}\right\} \longrightarrow w_i := w_i + C . x_i$$

Het gevolg is : $S := R_j$, want:

$$S_i := \sum_{i=0}^{n} (w_i + C . x_i) . x_i$$

$$= \sum_{i=0}^{n} w_i\, x_i + \sum_{i=0}^{n} C$$

$$= S + (n+1)\, C = S + R_j - S = R_j$$

Het is dus mogelijk met behulp van het adaptief systeem de potmeters zo in te stellen, dat de gewenste uitgang bij S verschijnt. Bij gegeven inputpatroon $\bar{x}_j$ en gewenste waarde R_j stellen de potmeters zich in op de waarde $\bar{w}_j = (w_{0j}, w_{1j}, \dots, w_{nj})$.
Na dit leerproces geef ik als inputpatroon $\bar{x}_{j+1}$. Het is dan duidelijk dat bij gewenste R_{j+1} de potmeters niet goed staan. Het leerproces moet opnieuw gestart worden, met als resultaat $\bar{w}_{j+1}$ $(\neq \bar{w}_j)$.
Hierna wordt $\bar{x}_j$ weer toegevoegd. etc.
Het blijkt nu, dat de $\bar{w}$ naar een vaste waarde convergeert. Voor het bewijs hiervan moet ik naar de literatuur verwijzen, omdat dit een vrij omslachtige mathematische bewijsvoering is. Voor de eenvoud wordt de convergentie hier dus zonder meer aangenomen.
De convergentie zal in de loop van het proces steeds minder snel gaan, en de eindwaarde hoeft niet exact de gewenste waarde aan te nemen. Deze twee feiten zijn niet zozeer van belang, omdat het DWE er voor zorgt, dat als $|S - R_j| < \varepsilon$, er een output $y = 1$ verschijnt. De waarde van ε kan zelf gekozen worden. Het DWE kan dus gezien worden als een stabilisatie van S, omdat bij voldoend nauwkeurige S altijd een vaste uitgang $y = 1$ komt.
Een vraag die wel behandeld zal worden is: Hoeveel verschillende patronen $\bar{x}$ kunnen aan de ingang gezet worden, opdat een goede herkenning nog verzekerd blijft?

Stel dat de herkenningsfunktie in een $(n+1) \times m$ matrix wordt gezet, met als kolomvectoren de patroonvectoren:

$$\bar{x}_j = (x_{0j}, x_{1j}, \dots, x_{nj})$$

De funktie wordt dus bepaald door een aantal patroonvectoren met $(n+1)$ elementen. Voor het j^e patroon geldt:

gewenste uitgangssignaal: R_j
optredende fout $f_j = R_j - \sum_{i=0}^{n} w_{ij} x_{ij}$

Vanwege het teken wordt ook wel eens bekeken:

$$f_j^2 = \left(R_j - \sum_{i=0}^{n} w_{ij} x_{ij}\right)^2$$

Totale fout over de m te herkennen patronen

$$F = \sum_{j=1}^{m} f_j = \sum_{j=1}^{m} \left\{ R_j - \sum_{i=0}^{n} w_{ij} x_{ij} \right\}$$

Bekijk ter vereenvoudiging het geval dat $n = 2$:

$$f_j = R_j - \{ w_{0j} + w_{1j} x_{1j} + w_{2j} x_{2j} \}$$

$$f_j^2 = R_j^2 + w_{0j}^2 + w_{1j}^2 x_{1j}^2 + w_{2j}^2 x_{2j}^2 - 2R_j (w_{0j} + w_{1j} x_{1j} + w_{2j} x_{2j}) + \\ + 2w_{0j}(w_{1j} x_{1j} + w_{2j} x_{2j}) + 2 w_{1j} x_{1j} w_{2j} x_{2j}$$

$$\frac{\partial f_j^2}{\partial w_{0j}} = 2w_{0j} - 2R_j + 2w_{1j} x_{1j} + 2w_{2j} x_{2j} = -2f_j$$

$$\frac{\partial f_j^2}{\partial w_{1j}} = 2\{ w_{1j} x_{1j}^2 - R_j x_{1j} + w_{0j} x_{1j} + 2x_{1j} w_{2j} x_{2j} \} = \\ = -2x_{1j} f_j$$

$$\frac{\partial f_j^2}{\partial w_{2j}} = -2x_{2j} f_j$$

Algemeen geldt:

$$\frac{\partial f_j^2}{\partial w_{ij}} = -2x_{ij} \cdot f_j$$

Dus alle partiële afgeleiden van de fout op moment $t+1$ hebben dezelfde 1e graads afhankelijkheid van de fout op moment t. Het teken wordt bepaald door het teken van:

1/ de fout
2/ de input.

Widrow en Hoff gebruikte „Bootstrap learning", hetgeen langzamer gaat, maar omdat het adaptief leren van ieder individueel patroon niet altijd op de juiste en kortste manier wordt gedaan, heeft het desondanks een groter toepassingsgebied. Er wordt gebruik gemaakt van de minimale gemiddelde kwadratische fout-methode, hetgeen neerkomt op het minimaliseren van:

$$\sum_{j=1}^{m} \varepsilon_j^2$$

In de t^e adaptiecyclus is de verandering van de gewichtsvector van het j^e patroon:

$$\Delta \bar{w}^{\,t} = \frac{\lambda}{n+1} \cdot \varepsilon(j) \cdot \bar{x}(j)$$

$$\text{met } \lambda = \frac{\text{aantal keren dat } \bar{x}_j \text{ is gecorrigeerd}}{\text{totaal aantal keren.}}$$

Het leerproces is dus afhankelijk van het symbool en de gewenste uitgang op ieder moment t.
Als de binaire inputsignalen totaal niet gecorreleerd zijn met elkaar en met de gewenste uitgang R, dan gaan de gewichtswaarden als een exponentiële funktie naar een eindwaarde; de tijdkonstante voor L.M.S. (least-mean-square error) is:

$$T = \frac{n+1}{\lambda} \cdot \text{aantal adaptie cycli.}$$

De leerkromme van MSE (mean-square-error), uitgezet tegen het aantal adaptie-cycli, is in geval van ongecorreleerde ingangsignalen een exponentiële funktie. De gekwadrateerde fout f^2 heeft dus een tijdkonstante:

$$T = \frac{n+1}{2\lambda} \cdot \text{aantal adaptie cycli.}$$

Hoe minder het aantal patroonkomponenten, des te sneller zal het leerproces verlopen.
In dit geval is uitgegaan van het gemiddelde kwadratische fout-kriterium. Het is belangrijk, ook voor de leersnelheid, van wat voor kriterium wordt uitgegaan. Het zou mogelijk zijn de invloeden van de verschillende kriteria op te tijdkonstante prakties en theoretisch te bepalen. Dit is hier niet gedaan.

Na het leerproces kan het adaptief systeem verwijderd worden en kan tot herkenning worden overgegaan. De adaline werkt dan als een klassifikatie-element. De uitbreiding van de toepassing ligt in het paralel of serie schakelen van adalines. Het systeem wordt dan Madaline genoemd. In de praktijk wordt meestal een Madaline gebruikt vanwege de vele mogelijkheden. Wat die mogelijkheden zijn, zal in het volgend gedeelte worden besproken.

3.1 Grenzen bij toepassing van drempelwaarde-element(en)

De beschouwing van een Adaline heeft een parallel beschouwing in een n-dimensionale Euclidische ruimte E^n. Om nu de verschillende mogelijkheden van de Adaline te bepalen is het noodzakelijk een goed beeld te hebben van hypervlakken en van mathematische manipulaties in E^n. Daar de wiskundige afleidingen niet iedereen zullen interesseren, en omdat het moeilijk voor te stellen zaken zijn, heb ik de berekeningen in appendix A gezet. De grote lijn en de resultaten zullen echter wel hier vermeld worden.

Stelling 0

Als m hypervlakken $F_1, F_2, \ldots, F_m$ in E^n voldoen aan de algemene opstelling, gedefiniëerd in de appendix, dan wordt het aantal n-dimensionale gebieden (subruimten), dat de hypervlakken in E^n opdelen, gegeven door:

$$B_n^m = 2 \sum_{i=0}^{n-1} \binom{m-1}{i}$$

De Adaline heeft (n+1) ingangen. Iedere ingang kan in principe twee waarden aannemen. Bekijk ik het (n+1) tuple

$$w_0 \ , w_1 x_1 \ , w_2 x_2 \ , \ \ldots \ , w_n x_n$$

waarbij dus $x_0 = 1$, dan zijn er 2^n verschillende tuples mogelijk. Het maximaal aantal verschillende ingangspatronen, per Adaline, is dus:

$$m = 2^n$$

Het absoluut maximum aantal subruimten, waarin 2^n hypervlakken E^{n+1} verdelen is:

$$B_{n+1}^{2^n} = 2 \sum_{i=0}^{n} \binom{2^n - 1}{i}$$

Stelling 1

Als er met behulp van 1 DWE, m ingangscombinaties worden beschouwd, dan kunnen er door één DWE maximaal

$$B^{m}_{n+1} < \frac{2m^{n}}{n!}$$

essentieel verschillende funkties gerealiseerd worden.

Met behulp van:

$$V = \frac{\text{totaal aantal drempelfunkties}}{\text{totaal aantal mogelijke funkties}}$$

$$= \frac{\frac{2m^{n}}{n!}}{2^{m}} = \frac{2m^{n}}{n!\,2^{m}}$$

wordt afgeleid:

<u>Stelling 2</u>
Als $m > 3n$, dan kan geen selectie gedaan worden met behulp van één DWE. Voor grote n zullen de meeste funkties, die door meer dan $3n$ inputcombinaties gedefinieerd zijn, niet te herkennen zijn.

<u>Stelling 3</u>
Is voldaan aan:

a/ $n \to \infty$

b/ Een set van m inputcombinaties wordt random gekozen uit 2^{n} mogelijkheden.

c/ en een funktie wordt met deze inputcombinaties random uit de 2^{m} mogelijkheden gedefinieerd.

dan nadert de waarschijnlijkheid, dat de resulterende funkties door één DWE kunnen worden gerealiseerd naar:

1 voor $m < 2n$

½ voor $m = 2n$

0 voor $m > 2n$.

Het resultaat van stelling 3 kan als volgt geïnterpreteerd worden:
Als de waarschijnlijkheid van herkenning met behulp van één DWE uit wordt gezet tegen het genormaliseerd aantal patronen ($m/n+1$, $\approx \alpha$), met als parameter $n+1$, dan wordt de

grafiek, afgebeeld in figuur 11, gevonden.
p = waarschijnlijkheid dat m of meer random gekozen patronen lineair te onderscheiden zijn.
De kromme is symmetrisch t.o.v. $\left(\frac{m}{n+1}, p\right) = (2, \frac{1}{2})$.
In het limietgeval geldt de asymptoot. Algemeen is:

$p = 1$ voor $0 \leq m \leq n$

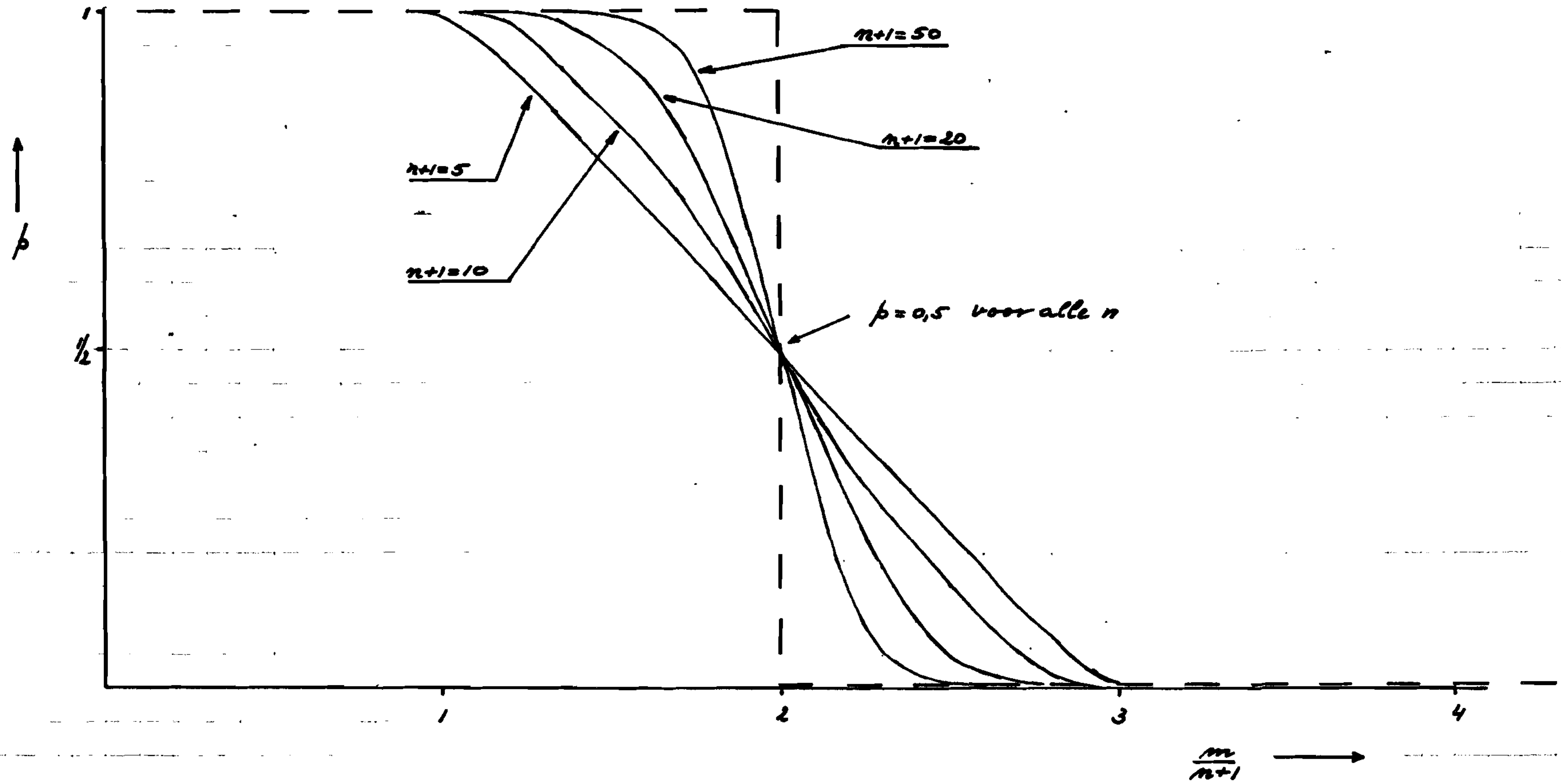
$p = 0$ voor $m \geq 3n$

p neemt af van 1 naar 0 voor $n \leq m \leq 3n$.

Uit het plaatje van figuur 11 kunnen we nu een DWE-„capaciteit" definiëren, als zijnde het getal $2n+2$, òf voor grote n: $2n+2 \approx 2n$.

Één patroon met minder dan $2n$ inputfunkties kan normaal gerealiseerd worden met een enkel DWE; zijn er meer dan $2n$ inputfunkties, dan is realisatie niet meer mogelijk.
De vraag, of deze afleiding, met verschillende benaderingen, bruikbaar is, kan bevestigend beantwoord worden. De benaderingen zijn allen in ongunstige zin uitgevoerd, wat betreft de uitkomst. Bij exacte afleiding zullen de krommen hoogstens iets naar rechts verschuiven; dit houdt in, dat de capaciteit van $2n$ zeker zal gelden. De praktijk-proeven hebben de faktor $2n$ reeds lang als waarschijnlijke capaciteit naar voren gebracht. Deze bewijsvoering is een bevestiging van dit vermoeden.

Aan de hand van de gevonden resultaten is het mogelijk de juiste toepassing van de Adaline of een kombinatie bij gegeven aantal herkenningsfunkties te bepalen. Hiermee is bereikt, dat vooraf bekend is de kombinatie het herkenningsproces tot een goed einde zal brengen. Het bespaart de teleurstelling die op kan treden, als een systeem met een te lage capaciteit is opgebouwd, zodat herkenning onmogelijk voldoende goed kan worden uitgevoerd. Ook kunnen op economische gronden de verschillende kombinatiemogelijkheden tegen elkaar afgewogen worden. Hier kan een optimalisatie-methode van nut zijn, vooral als het om grotere systemen gaat.
Tot slot geef ik twee voorbeelden van kombinaties.



n+1 : aantal ingangen van 1 DWE.
m : aantal random gekozen ingangspatronen.
p : waarschijnlijkheid dat m of meer random gekozen patronen lineair te onderscheiden zijn.

FIG.-11

19

Voorbeeld 1

Stel k DWE'n van $(n+1)$ ingangen. De respektievelijke uitgangen gaan allen naar de ingang van het $(k+1)^e$ DWE. Zie figuur 12. Ieder DWE kan $\frac{2m^n}{n!}$ funktie's realiseren, waarbij m het aantal te gebruiken ingangskombinaties is.
Het $(k+1)^e$ DWE kan $\frac{2m^k}{k!}$ funkties realiseren.
Het totaal aantal te realiseren funkties is:

$$\left(\frac{2m^n}{n!}\right)\left(\frac{2m^k}{k!}\right) = \frac{2^{k+1} \cdot m^{(n+1)k}}{(n!)^k \cdot k!}$$

Voorbeeld 2

Stel k DWE'n die zonder terugkoppeling zijn geschakeld. Het j^e element heeft als ingangen (zie figuur 13):

a/ o t/m. n basisingangen.

b/ $y_1, y_2, \cdots, y_{j-1}$

Totaal aantal ingangen van het j^e DWE:

$$(n+1) + (j-1)$$

Aantal te realiseren funkties door het j^e DWE:

$$\frac{2 \cdot m^{(n+j-1)}}{(n+j-1)!}$$

Het totaal aantal realiseerbare funkties door de gehele schakeling is:

$$\prod_{j=1}^{k-1} \frac{2 \cdot m^{(n+j-1)}}{(n+j-1)!} = \frac{2^k \cdot m^{\{nk + \frac{k(k-1)}{2}\}}}{(n!)^k}$$

Als laatste moet ik verwijzen naar:
„A Critical Comparison of Two Kinds of Adaptive Classification Networks." K. Steinbuch en B. Widrow.
IEEE Trans. on Elec. Computer okt. '65 pp. 737-740.

waar een vergelijkende studie is gemaakt tussen leermatrix en Madaline. De leermatrix is eenvoudiger toe te passen, terwijl de capaciteit heel duidelijk vastligt. In het algemeen kan echter gezegd worden, dat het leerproces een optimaal proces is, zodat de toepassingsgebieden van leermatrix en madaline verschillend zijn: beide hebben hun eigen gebied waarin ze optimaal toe te passen zijn.

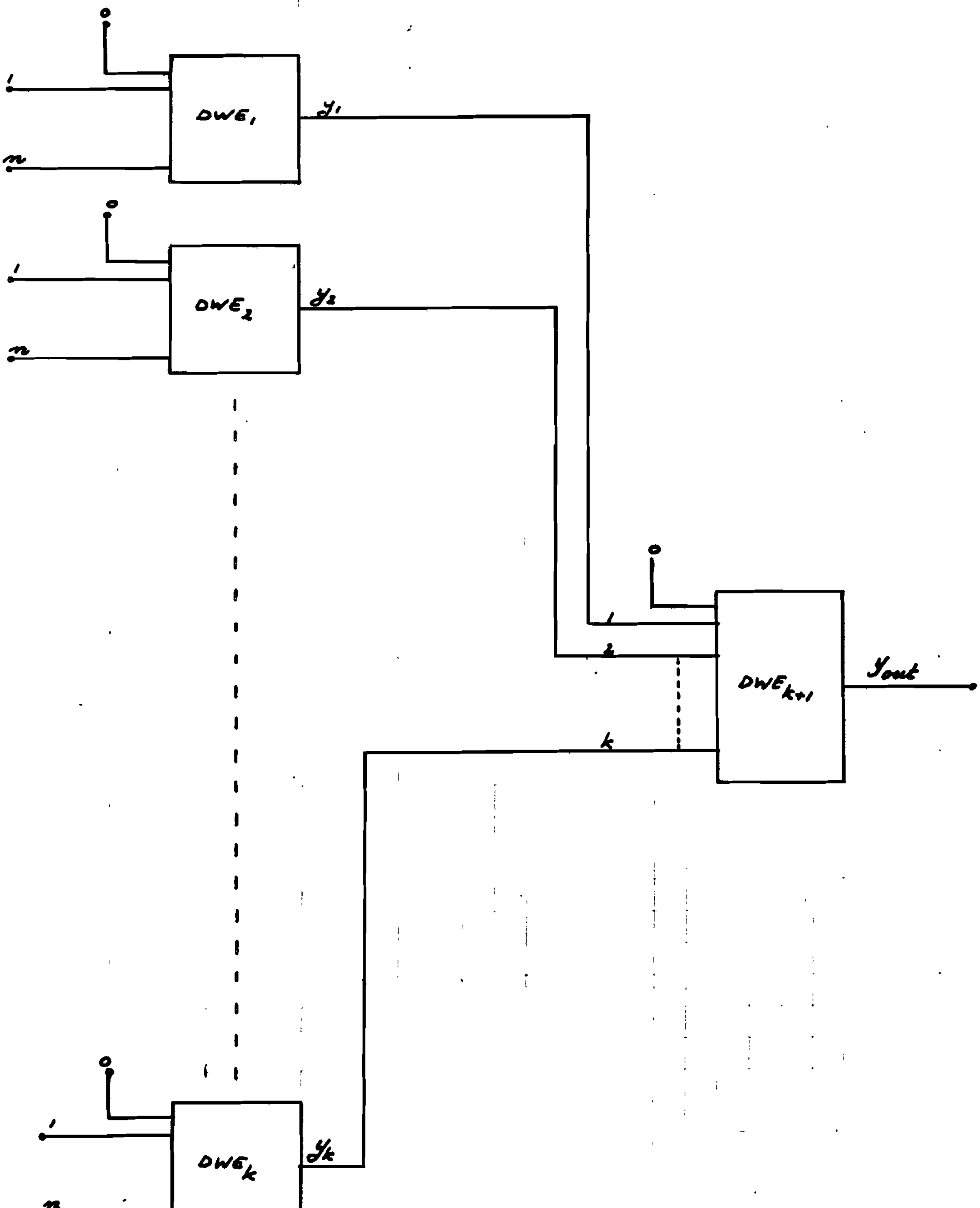


FIG.-12

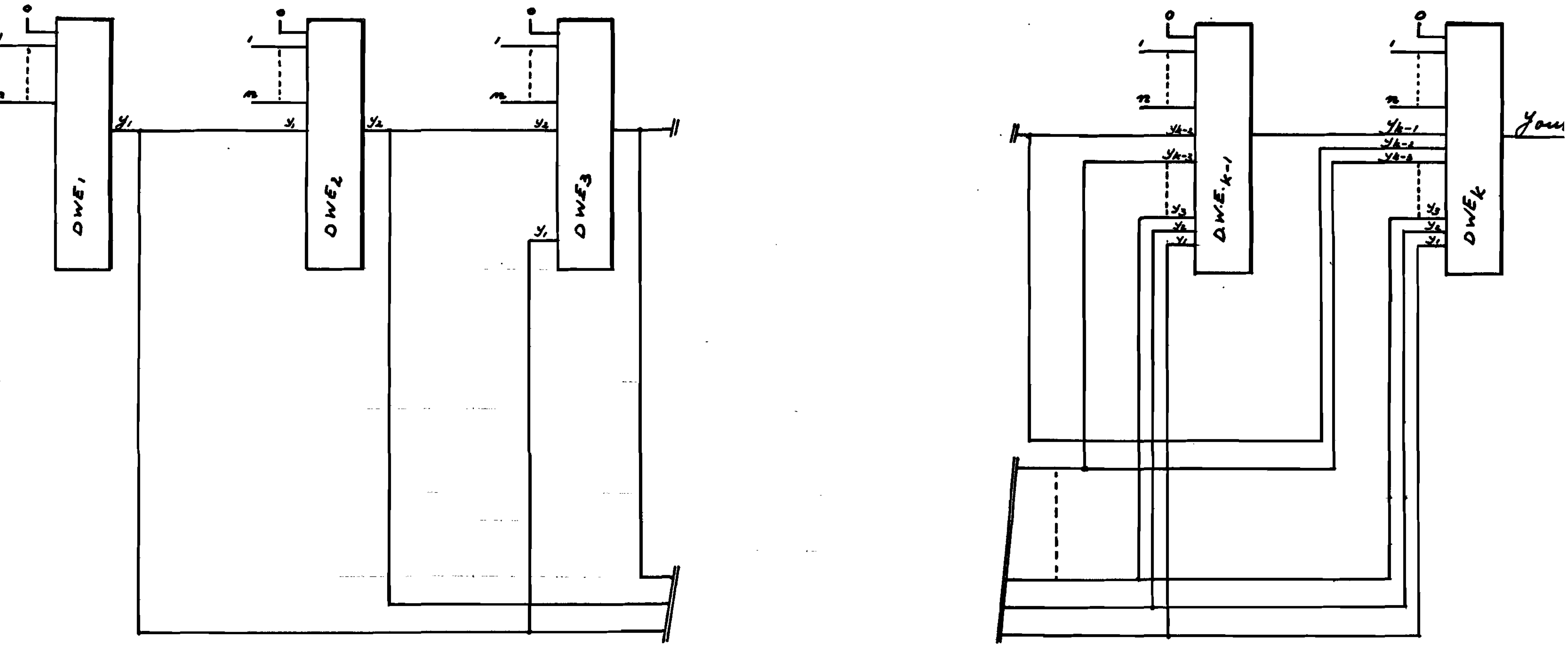


FIG. -13

4. Tekenherkenning met behulp van coherente optische systemen.

Eén van de belangrijkste punten bij de herkenning is het scheiden van het symbool en de ruis. Onder ruis wordt dan verstaan: alle overbodige informatie. Het scheiden kan bijvoorbeeld gebeuren met behulp van „matched"-filters in coherente optische systemen. Voor deze techniek is het noodzakelijk dat complexe filters op eenvoudige wijze te maken zijn. De sterke analogie tussen ruimtelijk filteren in coherente optische systemen en optimale lineaire filters in de communicatietheorie opent een nieuwe techniek om deze frequentiedomein-filters te realiseren. De meest voor de hand liggende toepassing is de fotografische film, die echter geen negatieve funkties kan opnemen, zodat eventueel gebruik moet worden gemaakt van fase-filters.

Veel moeilijkheden bij tekenherkenning, die bij veel andere systemen moeilijk op te lossen zijn, treden bij deze methode nauwelijks of niet op, of zijn op de een of andere manier te omzeilen:

- Een fout in de positie van het symbool is niet rampzalig. De Fouriertransformatie van een getransleerd patroon blijft, op een fasefaktor na, gelijk.
- Een behoorlijke spreiding kan worden toegestaan in afmeting en richting.
- De graad van selectiviteit van twee symbolen kan worden nagekeken door de frequentie-inhoud van het filter te veranderen.
- De output van het systeem is een twee dimensionale verdeling van het licht, afhankelijk van de input.

Het gebruik van coherente systemen brengt ook moeilijkheden met zich mee:

Het geschikt maken voor presentatie aan het optisch

systeem is een ingewikkelde zaak.
Andere problemen liggen bij de output. Hoe geef ik de informatie op de beste manier door naar bijvoorbeeld een computer of een tape.

Voordat een praktisch voorbeeld ter sprake komt, wordt eerst de mathematische achtergrond bekeken.

Filteroverdrachtsysteem.

Zie figuur 14.
Stel het signaal: $x(t)$ en de ruis: $n(t)$
De ingangsfunktie is een sommatie van het signaal en de ruis.
De output van het systeem is $y(t)$. Het verschil tussen signaal en output is:

$$\varepsilon(t) = x(t) - y(t)$$

Het kiezen van de filterfunktie $h(t)$ kan volgens verschillende kriteria gebeuren:

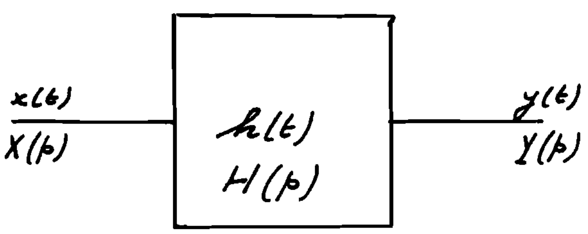
1. Bij het zogenaamde „matched" filter is de verhouding der vermogens van het signaal en de ruis zo gunstig mogelijk.
2. De optimale filter is verwezenlijkt als

$$\int_0^\infty \varepsilon^2 \, dt$$

minimaal is.
Andere integraal-fout-kriteria zijn o.a.:

$$\int_0^\infty \varepsilon^2 t \, dt$$

$$\int_0^\infty |\varepsilon| t \, dt \quad (\text{ITAE})$$



De fouriergetransformeerde van $x(t)$ is $X(p)$
Idem voor:

$$\mathcal{F}\{y(t)\} = Y(p)$$
$$\mathcal{F}\{h(t)\} = H(p)$$

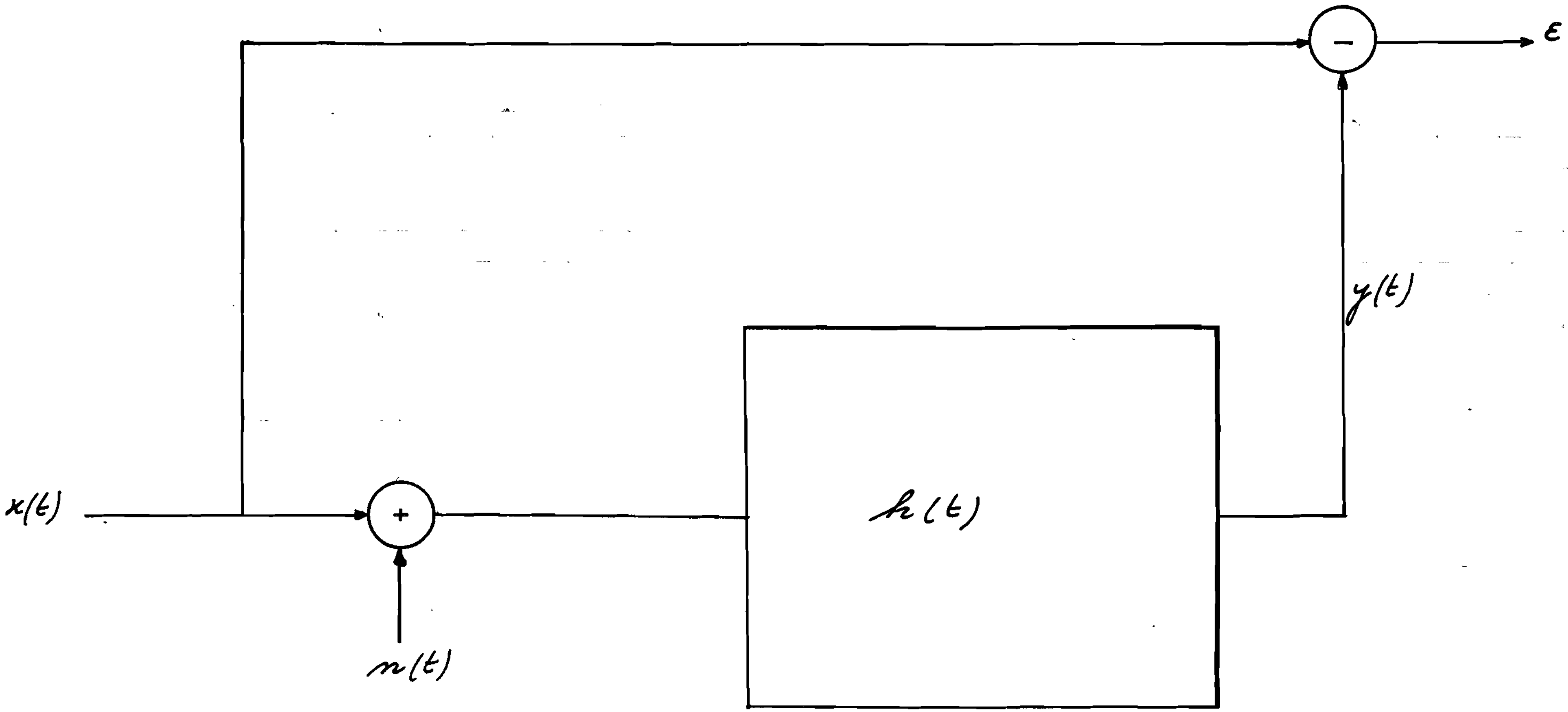


FIG. - 14.

De fouriertransformatie geef ik in het vervolg aan met „ ⟶ ",
dus: $x(t) \longrightarrow X(p)$
Voor het op de vorige bladzijde geschetste systeem geldt:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)\,.\,h(t-\tau)\,d\tau = x(t) * h(t) = \text{convolutie van } x(t) \text{ en } h(t).$$

$$\therefore\ Y(p) = X(p)\,.\,H(p)$$

Kruiscorrelatiefunktie:

$$\psi_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{2T}\int_{-\infty}^{\infty} x_T(t)\, y_T(t+\tau)\, dt$$

als:

$$x_T(t) = \begin{cases} x(t) \text{ als } |t| \leq T \\ 0 \quad \text{als } |t| > T \end{cases}$$

Kruisvermogenspectrum:

$$\Phi_{xy}(p) = \lim_{T\to\infty}\frac{1}{2T}\left\{\int_{-\infty}^{\infty} x_T(t)\, e^{jpt}\, dt \int_{-\infty}^{\infty} y_T(\theta)\, e^{-jp\theta}\, d\theta\right\}$$

Wiener-Khintchin relaties:

$$\psi_{xy}(\tau) \longleftrightarrow \Phi_{xy}(p)$$

Nu geldt:

$$\psi_{xx}(\tau) = x(t) * x'(t)$$

met $x'(t) = x(-t)$

$x'(t) \longleftrightarrow X^*(p)$

$$\Phi_{xx}(p) = X(p)\,.\,X^*(p)$$

met $X^*(p)$ is toegevoegd complex van $X(p)$

Bekijk nu het schema van figuur 15. Hierin is mathematisch de kruiscorrelatiefunktie voor f en g bepaald via enige manipulaties. Alleen de term $\psi_{fg}(\tau-\alpha)$ is in dit geval belangrijk. Deze term is gescheiden van de rest van de output van de berekening, en een maat van overeenstemming tussen het referentiesignaal $f(t)$ en de inputfunktie $g(t)$. Het mathematisch schema kan m.b.v. de optica verwezenlijkt worden.

referentiesignaal : $g(t)$
inputfunktie : $f(t)$

$$g(t) \longrightarrow G(p)$$

toevoeging van een vlakke golf : $+\; e^{jbp}$

$$G(p) + e^{jbp}$$

↓ maat voor vermogensinhoud van deze vorm.

$$\left(G(p)+e^{jbp}\right)\left(G(p)+e^{jbp}\right)^* \quad \ldots (I)$$

(I) is in physische zin de lichtverdeling in vlak P_5 (figuur 17)

$$f(t) \longrightarrow F(p) \quad \ldots (II)$$

(I) en (II) worden vermenigvuldigd:

$$\left\{1+G(p)G^*(p)\right\} F(p) + G(p)\cdot F(p)\cdot e^{-jbp} + G^*(p)\, F(p)\, e^{jbp}$$

↓ ↓ ↓ fouriertransformatie door L_2

$\ldots$ $\ldots$ $\psi_{fg}(\tau - b)$

output : $r(x,y)$

FIG. - 15

4.1. Toepassing van matched filters in tekenherkenning

Zoals reeds gezegd is een van de moeilijkheden bij herkennen het scheiden van teken en ruis. Stel nu dat teken en ruis lineair onafhankelijk zijn, en dat de ruis homogeen is.
Stel: tekensignaal $s(x,y)$
ruis $n(x,y)$
$s(x,y)$ is bekend; de Fouriergetransformeerde $S(p,q)$ is:

$$S(p,q) = \iint_{-\infty}^{\infty} s(x,y)\, e^{j(px+qy)}\, dx\, dy$$

Ruis is een random funktie, gekarakteriseerd door zijn autocorrelatiefunktie $R_n(x,y)$, met

$$R_n(x,y) = \iint_{-\infty}^{\infty} n(u,v)\, n^*(u+x, v+y)\, du\, dv.$$

en zijn spectrale dichtheidsfunktie $N(p,q)$, met

$$N(p,q) = \iint_{-\infty}^{\infty} R_n(x,y)\, e^{j(px+qy)}\, dx\, dy$$

Dit is de Fouriergetransformeerde van de autocorrelatiefunktie.

Het is af te leiden, dat het optimale filterproces wordt voorgesteld door:

$$H(p,q) = \frac{S^*(p,q)}{N(p,q)} \qquad \ldots (2)$$

Bekijk nu het volgend proces:

$f(x,y)$ is de inputfunktie, met als Fouriergetransformeerde $F(p,q)$.
Vermenigvuldig $F(p,q)$ met $H(p,q)$ en neem dan de Fouriergetransformeerde.

Dit proces beeldt de input als funktie van de filtering $H(p,q)$ punt voor punt uit in de output. Een coherent optisch systeem, zoals afgebeeld in figuur 16, voert dit proces uit.
De output wordt gegeven door:

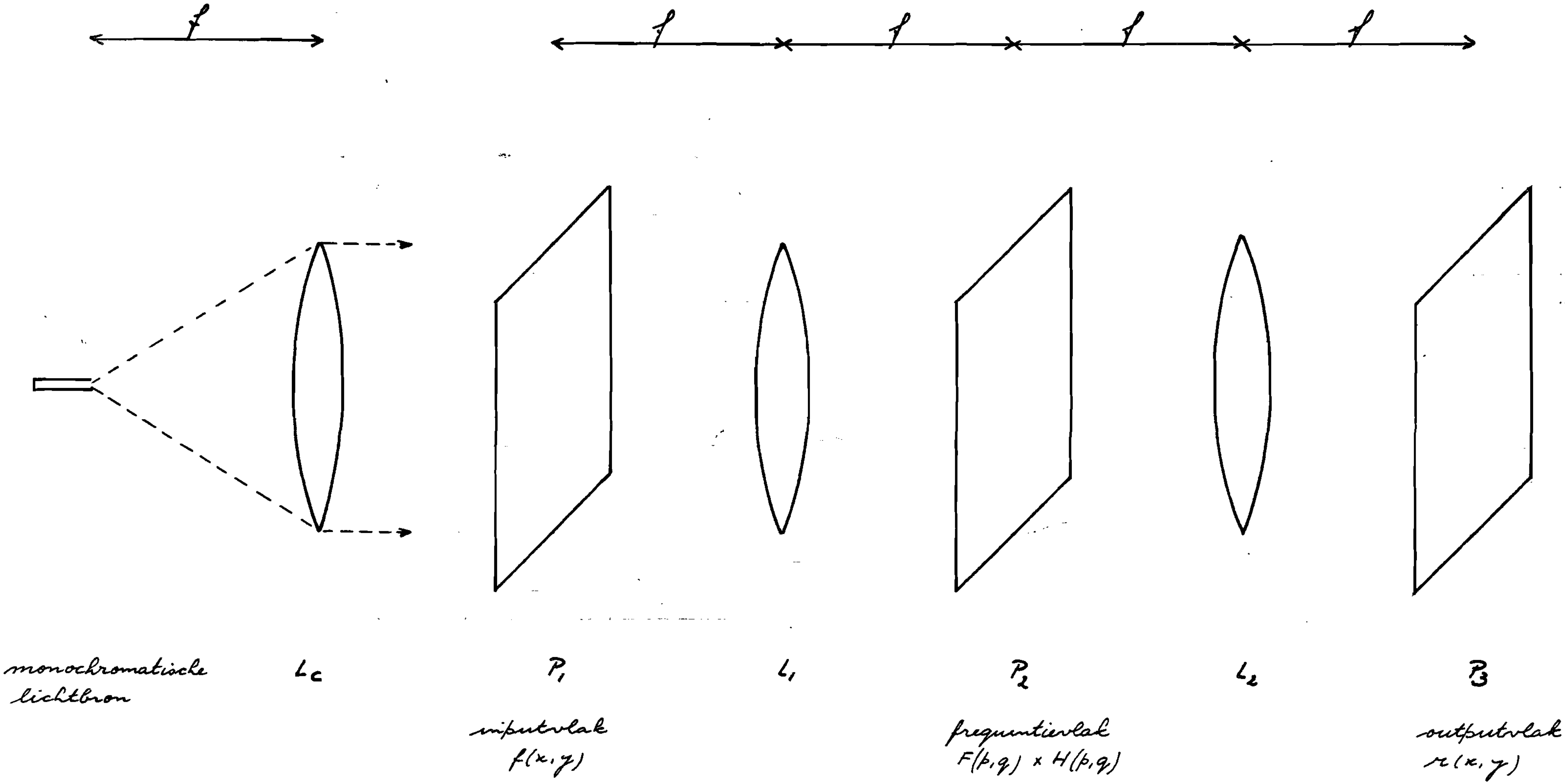

f
f
f
f
f
monochromatische lichtbron
L_c
P_1
inputvlak
$f(x,y)$
L_1
P_2
frequentievlak
$F(p,q) \times H(p,q)$
L_2
P_3
outputvlak
$r(x,y)$

FIG. -16

$$r(x,y) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} F(p,q)\, H(p,q)\, e^{j(px+qy)}\, dp\, dq.$$

Het systeem is dus te verwezenlijken als het filter te maken is. De filter kan mathematisch beschreven worden met behulp van de kennis omtrent het signaal, maar moet ook prakties te verwezenlijken zijn. Dit kan op verschillende manieren. Ik neem de methode waarbij gebruik wordt gemaakt van een fotografisch negatief.
Het fotografisch negatief is gevoelig voor energie, maar is als zodanig niet gevoelig voor fase-informatie. Dit wordt opgelost door de fase-informatie om te zetten in intensiteit-informatie, zodat dit wel kan worden opgenomen. De uitvoering hiervan geschiedt door een complexe funktie op te vatten als een reële funktie, gemoduleerd door een draaggolf.
De filterfunktie van vergelijking (2) herschrijf ik als:

$$H(p,q) = \frac{k_1}{N(p,q)} \cdot k_2 \cdot S^*(p,q) \qquad \ldots (3)$$

k_1 en k_2 moeten zo gekozen worden, dat het filter passief is.
In vergelijking (3) is $N(p,q)$ de Fouriergetransformeerde van een autocorrelatiefunktie, dus een niet negatieve funktie, zodat $k_1/N(p,q)$ eenvoudig te realiseren is.
Het verwezenlijken van $H(p,q)$ resulteert dus nu in het realiseren van de term : $k_2\, S^*(p,q)$.
De realisatie kan op verschillende manieren gebeuren, waarvan er een is afgebeeld in figuur 17.

Het signaal $s(x,y)$ wordt in vlak P_4 geplaatst, zodat de helft van de straal het signaal belicht. Lens L_r centreert de andere helft van de lichtstraal precies in vlak P_4, zodat de afstand tussen L_r en P_1 precies f_r moet zijn. De afstand tussen het centreerpunt in P_1 en het midden van het signaal is b. Deze afstand werkt door tot in de kruiscorrelatiefunktie van referentiesignaal en input-signaal in de output van het systeem van figuur 15 en 16.
De lichtverdelingsfunktie in vlak P_4 is:

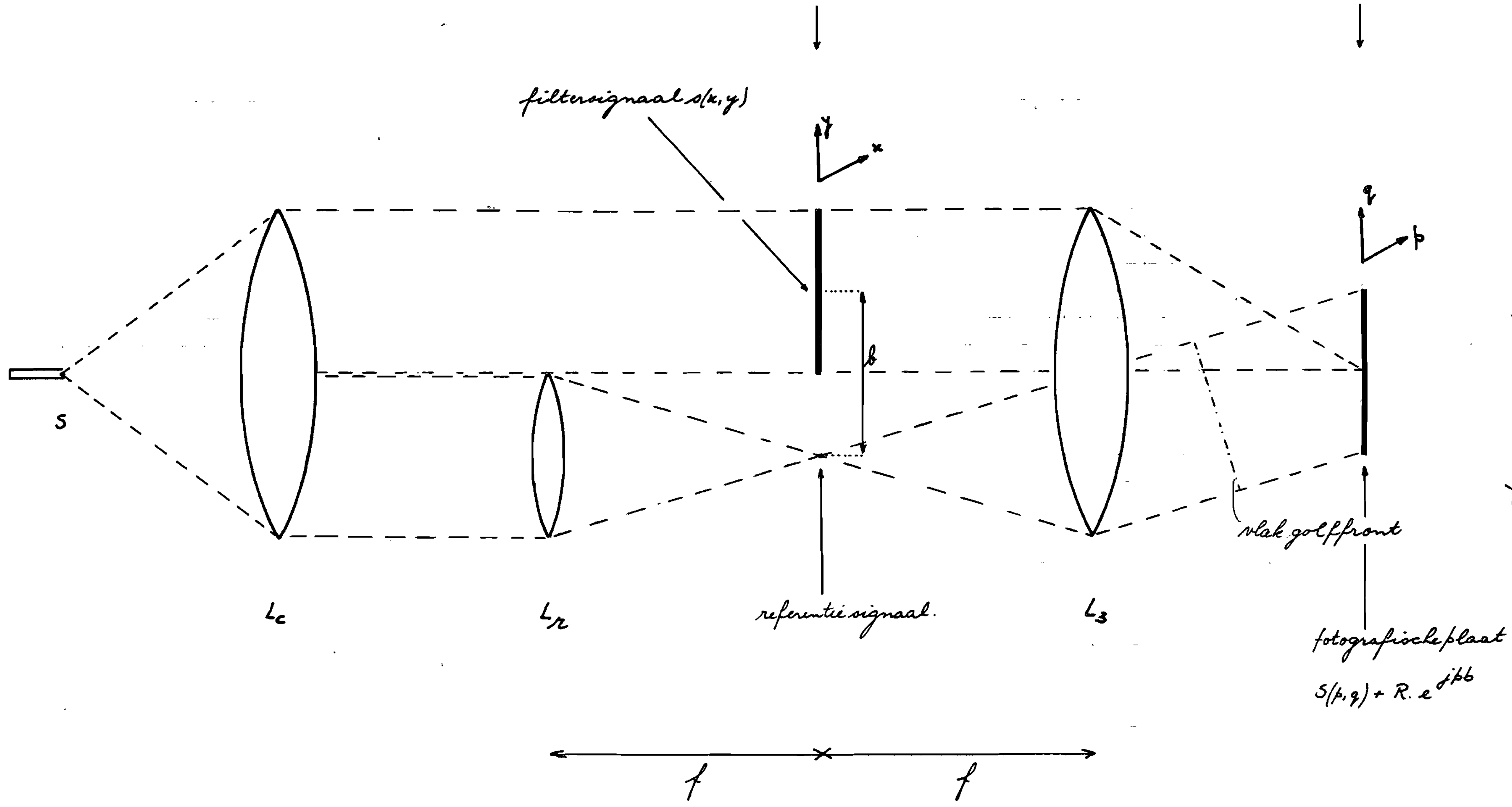

filtersignaal s(x,y)
y
x
q
p
b
S
L_c
L_r
referentie signaal.
L_3
vlak golffront
fotografische plaat
$S(p,q) + R.e^{jpb}$
f
f

FIG.-17

73

$$g(x,y) = s(x,y) + R\,\delta(x-b,\,y)$$

met $R\,\delta(x-b,y)$: verdeling t.g.v. L_R

De afbeelding in P_5 wordt dan:

$$G(p,q) = S(p,q) + R\,e^{jbp}$$

Plaats ik in P_5 een fotografische plaat, dan wordt $|G|^2$ opgenomen:

$$|G(p,q)|^2 = |S(p,q) + R\,e^{jbp}|^2$$

$$= |S(p,q)|^2 + R^2 + R\,e^{-jbp}\,S(p,q) + R\,e^{jbp} \cdot S^*(p,q) \qquad \ldots (4)$$

De laatste term heeft de vorm van de gevraagde funktie, vermenigvuldigd met e^{jbp}, zodat bij terugtransformatie, door middel van L_2, de lineaire faseterm er voor zorgt, dat de laatste term gescheiden wordt afgebeeld in P_3.
Het is duidelijk dat de term $k_2\,S^*(p,q)$ verwezenlijkt is in de term $R.e^{jbp} \cdot S^*(p,q)$. Voeg nu aan deze term $k_1/N(p,q)$ toe, door een fotografisch negatief met dichtheidsfunktie $k_1/N(p,q)$ er mee in serie te zetten, dan is de filterfunktie, gegeven in (3), gerealiseerd.
Wordt $k_1/N(p,q)$ in serie gezet met $|G(p,q)|^2$, dan is het resultaat, wat betreft de gewenste funktie, hetzelfde; alleen de plaats van afbeelding is nu van belang, omdat er ook funkties worden afgebeeld, die niet ter zake doende zijn (zie ook uitdrukking (4)).

Stel nu dus de nieuwe filterfunktie

$$J(p,q) = \frac{k_1}{N(p,q)} \cdot |G(p,q)|^2 \qquad \ldots (5)$$

In het outputvlak P_3 wordt afgebeeld:

$$r(x,y) = \mathcal{F}\{F(p,q) \times J(p,q)\} =$$

$$= \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} F(p,q) \cdot \frac{k_1}{N(p,q)} \cdot |G(p,q)|^2 \cdot e^{j(px+qy)}\, dp\, dq =$$

neem $R = k_2$

$$= \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \frac{k_1 . F(p,q)}{N(p,q)} . \left\{ |S(p,q)|^2 + k_2^2 \right\} . e^{j(px+qy)} dp . dq +$$

$$+ \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} F(p,q) . H^*(p,q) \, e^{j\{p(x-b)+qy\}} dp \, dq +$$

$$+ \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} F(p,q) H(p,q) \, e^{j\{p(x+b)+qy\}} dp \, dq. \qquad \ldots (6)$$

Term 1 wordt afgebeeld in het optisch centrum.
Term 2 wordt afgebeeld rond $x=-b$ en $y=0$. Hij stelt de kruiscorrelatiefunktie voor van de inputfunktie en de geconjugeerde impulsresponsie van het filter.
Term 3 wordt afgebeeld rond $x=b$ en $y=0$. Het stelt de convolutie voor van de inputfunktie met de impulsresponsie van het filter. Dit is dus de gewenste output, die over een afstand $x=b$ is verschoven.
Wordt b voldoende groot gekozen, dan zullen de signalen van (6) elkaar niet overlappen.
Wil ik dit proces weer onderverdelen in leer- en herkenningsgedeelte, dan kan ik de operatie gegeven door figuur 17 als leerproces, en de operatie gegeven door figuur 16 als herkenningsproces beschouwen.

4.2. Voor- en nadelen.

De afmetingen van de te verwerken patronen hebben een boven- en een onder-grens. De bovengrens wordt onder andere beperkt door de afmetingen van de lenzen. De maximale hoogte of breedte van het patroon mag niet groter zijn dan ongeveer de helft van de diameter van de lenzen. De ondergrens wordt onder andere bepaald door het scheidend vermogen van de lens en het fotografisch negatief. Het gebruik van een laserstraal werkt gunstig, vanwege de grote concentratie van het licht in een punt; hierdoor is het contrast van een patroon veel beter.
In het algemeen moet er gezocht worden naar een maxi-male aanpassing van : informatie, optisch systeem en opnameproces.
Voorbeelden van opname-elementen zijn :

1/ amplitude modulator :
   a/ fotografische film
   b/ oliefilm.
2/ fase/polarisatie modulator
   a/ thermoplastische film.
   b/ electro-optische film.
   c/ magneto-optische film.
   d/ foto-elastisch materiaal.

Wordt het herkenningssysteem bekeken, dan is de toepas-sing van deze techniek zeer eenvoudig, als de afmetingen, positie wat betreft rotatie, en de kwaliteit van het patroon niet al te sterk afwijken van elkaar. Zijn de afwijkingen te groot, dan is deze methode totaal onbruikbaar, omdat dan het correlatieproces niet meer opgaat.
Een eventuele rotatie van het patroon hoeft geen bezwaar te zijn. Als ik het filter namelijk laat draaien in het vlak, loodrecht op de optische as, dan wordt het patroon toch gedetecteerd.

In hoeverre de verschillende intensiteiten te onderscheiden zijn is nog een grote vraag. Het scheidend vermogen moet in dit geval bij de fotocellen en drempelwaarde-elementen liggen. Het is natuurlijk noodzakelijk dat de scheiding der intensiteiten zo goed mogelijk moet zijn.
Bij patroonherkenning kan ieder basispatroon apart op een filter gezet worden. Het bezwaar is dan dat de filters mechanisch verwisseld moeten worden, hetgeen te veel tijd in beslag neemt. Één vaak gebruikte techniek is het plaatsen van alle basispatronen op één enkel filter. Bij deze toepassing kan er een grotere tolerantie toegestaan worden in de positie van het inputsymbool.
Als nu het serie-gewijs presenteren van de te herkennen patronen eenvoudiger te verwezenlijken is, dan het verwisselen van de filters in de eerste methode, dan heeft de tweede techniek sterk de voorkeur.

4.3 Toepassing van optische herkenningssystemen.

In het algemeen kan gezegd worden, dat toepassing van deze systemen mogelijk is, -daar waar de te herkennen symbolen een kleine spreiding hebben in de representatie-vorm. De toepassing van de systemen is dan relatief eenvoudig.
Een mooi voorbeeld is het herkennen van chromosomen. Omdat de verschijningsvorm van chromosomen slechts een kleine tolerantie heeft, kan bij het herkennen gebruik worden gemaakt van een optisch herkenningssysteem. Ook sorteermachines kunnen een optisch systeem hebben. Bij bank- en post-wezen zal deze toepassing in de toekomst beslist van nut kunnen zijn. Bij dergelijke toepassingen zou een inleessysteem kunnen worden opgezet, zoals afgebeeld in figuur 18. Dit systeem is „real time".
De werking is als volgt:

P_0 wordt via L_0 afgebeeld op P_1.
Vlak P_1 bevat een systeem, dat door licht van een bepaalde frequentie geactiveerd wordt, zodat de informatie van P_0 in P_1 komt te staan. P_1 moet deze informatie even vasthouden, opdat dit vlak als inputvlak kan werken in het systeem: L_c, P_1, L_1, P_2, L_2, P_3.
Dit laatste systeem is eerder reeds besproken.

Alles draait dus om de werking van P_1. Er is een glassoort ontwikkeld, dat door bepaalde frequenties wordt geactiveerd, zodat het ondoorlaatbaar wordt voor bijvoorbeeld een laserstraal met een bepaalde frequentie. Bij kortstondige belichting van het glas met behulp van ultraviolet licht, wordt de aangeslagen toestand van het glas opgeheven: het glas wordt a.h.w. schoongeveegd.
De capaciteit van het systeem moet voldoende zijn. Stel ik heb 50 tekens: 26 letters, 10 cijfers en 14 andere tekens. Een

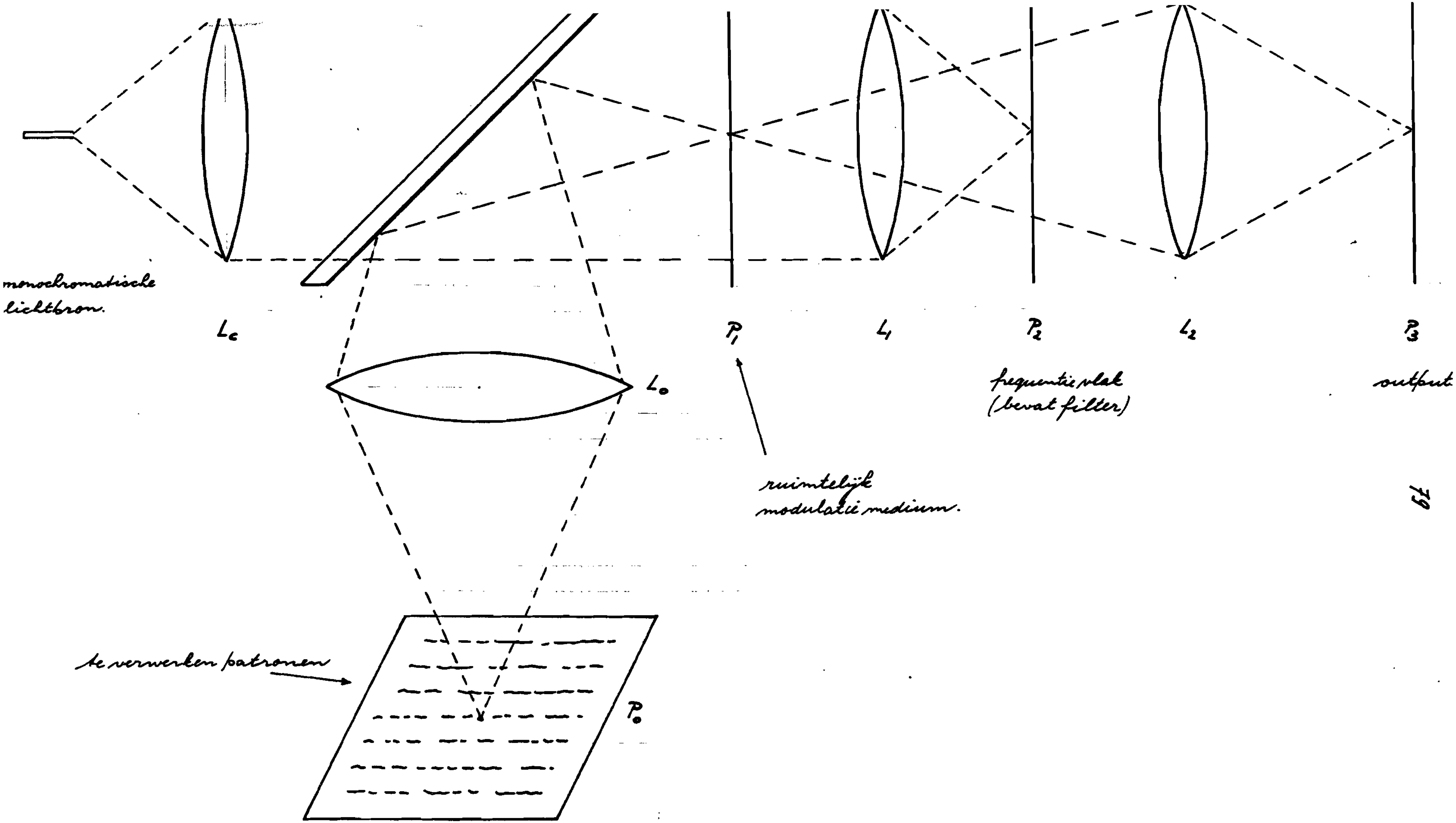

monochromatische
lichtbron.
L c
P 1
L 1
P 2
L 2
P 3
frequentie vlak
(bevat filter)
output
L o
ruimtelijk
modulatie medium.
te verwerken patronen
P o

FIG. -18

regel bevat 100 tekens, en een bladzijde bevat 35 regels. De output moet dan 175.000 punten kunnen detecteren. Voor het detecteren moet een drempelwaardesysteem worden opgezet, zodat de verschillende lichtpuntintensiteiten onderscheiden kunnen worden. De output kan natuurlijk aan een rekenmachine gehangen worden, om de resultaten op te slaan en/of te verwerken.
Andere toepassingen zijn:

1) Kwaliteitscontrôle. Bijvoorbeeld aan de hand van afmetingen.

2) Verwezenlijking van optische procesregeling en besturing. Bijvoorbeeld met optische terugkoppeling:
Zie figuur 19.
Het filter bevat het patroon van het produkt. Ligt het produkt op een lopende band, dan bevat het filter het patroon van (produkt + lopende band.).
In het outputvlak bevinden zich fotocellen. Per produktcyclus moeten deze cellen een bepaalde (maximale) hoeveelheid licht ontvangen. Gebeurt dit niet, dan moet de terugkoppelfunktie er voor zorgen, dat het produktieproces bijgeregeld wordt.
De terugkoppelfunktie zal een integrerende werking moeten bevatten, die zich uitstrekt over een produktperiode, waarbij onder produktperiode de tijd tussen het klaar komen van twee opeenvolgende produkten wordt verstaan. De integrerende werking moet „gereset" worden na iedere periode.

3) Luchtverkeerscontrôle

4) Uitzoeken van bijvoorbeeld: vingerafdrukken
chromosomen.

5) Katalogiseren van artikelen en boeken.

6) Sorteren van post bij gecodeerde adressering.

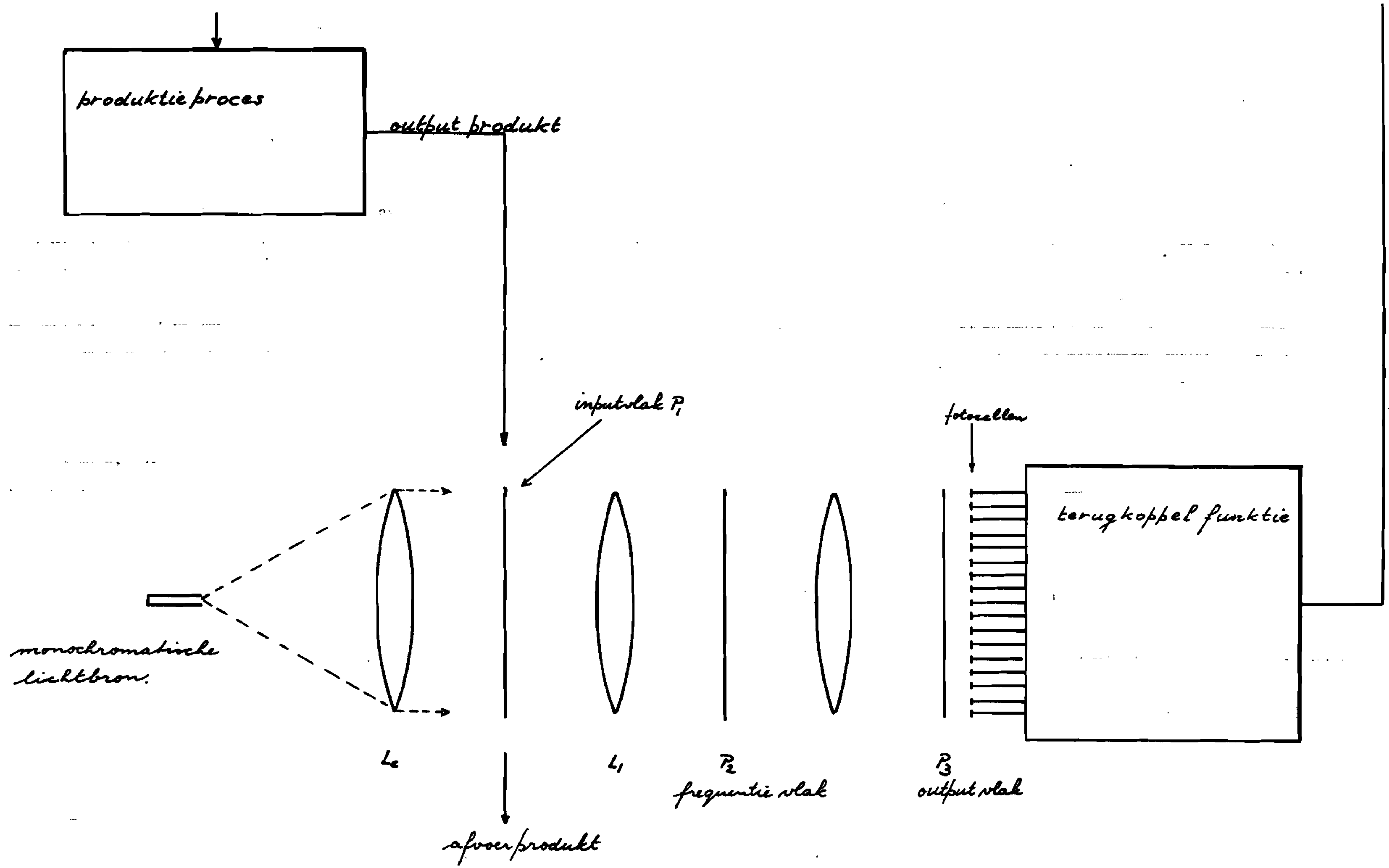

produktie proces
output produkt
inputvlak P_1
fotocellen
terugkoppel funktie
monochromatische lichtbron.
L_c
L_1
P_2
frequentie vlak
P_3
output vlak
afvoer produkt

FIG.-19

5. Lezen en herkennen van handgeschreven tekens.

In de praktijk is er een grote vraag naar systemen, die gedrukte en handgeschreven tekens, bijvoorbeeld letters en cijfers, kunnen herkennen. Tot op heden is een systeem voor herkenning van handgeschreven tekens nog niet voldoende goed ontwikkeld. Vele onderzoekingen en proefnemingen zijn in deze richting reeds gedaan, maar de resultaten zijn nog niet zo, dat een praktisch systeem kan worden opgezet. Het gedrukte, min of meer standaardschrift, kan reeds met een zekerheid van 95 tot 99 % worden herkend, terwijl de herkenningstijd binnen redelijke grenzen blijft.
Het handgeschreven schrift heeft een sterke spreiding in presentatievorm. Bedenk dat het schrift van tien verschillende personen evenzoveel verschillende schrijfwijzen naar voren brengt. Het is eenvoudig in te zien dat met behulp van de huidig ontwikkelde methoden, de herkenning van een dergelijk sterk variërend scalar van symbolen onmogelijk effectief uit te voeren is.
De enige oplossing bestaat hierin, dat er in die sterk variërende symbolen eigenschappen gevonden moeten worden die voor een bepaald symbool altijd aanwezig zijn. Deze eigenschappen worden dan de attributen waarop het hele herkenningsproces gebouwd kan worden.
Het probleem schuilt dus in het vinden van die min of meer representatieve kenmerken, die bij handgeschreven symbolen, moeilijk te destilleren zijn.

5.1. Moeilijkheden bij de herkenning van handgeschreven letters.

Alvorens me te wagen aan het veel complexere beeld van handgeschreven tekst, wil ik de optredende moeilijkheden bij handgeschreven letter-herkenning onder de loep brengen.

De grote struikelblok bij het leer- en herkenningsproces is wel de zeer grote spreiding van de typen stijlen. Iedere persoon heeft zijn eigen, karakteristieke schrijfstijl, die in de loop der jaren niet identiek blijft. Het is eenvoudig bij ons zelf na te gaan hoe de schrijfstijl in de tijd veranderd is; het schrift is een levend object, afhankelijk van de persoon en zijn omgeving. Het is zelfs zo levendig, dat het situatie afhankelijk is. De gemoedstoestand van een persoon kan zich o.a. uiten in het schrift. Een karakteristiek voorbeeld is het handschrift van Beethoven. Het schrift bevat informatie over de schrijver, terwijl de hoofdfunktie een informatieoverdrachts- of een geheugen-funktie is van feiten van de schrijver en niet over de schrijver. Het is nu zaak om de informatie over de schrijver te elimineren, zodat alleen de gegevens van de auteur overblijven.

In figuur 20 is te zien, dat de sterke variatie in de schrijfwijze van een bepaalde letter, afgezien van de normeringsmoeilijkheden, een grote belemmering is om te komen tot uniform geldende herkenningsregels voor een bepaalde letter:

De grote en kleine gesloten krul zijn niet overal aanwezig.
Het aantal snijpunten van krommen, de hoeken die ze met elkaar maken en de verschillende hellingen kunnen sterk variëren.
De verhouding zwart/wit in een vakje heeft een te grote spreiding. enz.

Al deze punten kunnen niet leiden tot een eenduidige herkenning. De vraag rijst dan natuurlijk: „Zijn er herkenningspunten (parameters of attributen) die wèl een eenduidigheid bewerkstelligen?" Helaas moet ik dan zeggen, dat dergelijke punten nog niet bekend zijn. Het probleem wordt nog zwaarder, als de letters niet afgepast in een uniform hokje geschreven worden, maar eenvoudigweg op een lijn geplaatst moeten worden. De interpretatie wordt dan veel gecompliceerder, omdat het noodzakelijk is, de letter eerst in een afgepast hokje te projecteren (normaliseren), waarna we zijn teruggekeerd tot het probleem dat in figuur 20 wordt aangeduid. Zoals figuur 21 duidelijk aantoont, moet de lengte en de hoogte van de letter reeds vóór de normalisatie juist aangenomen worden, wil er geen foutieve interpretatie volgen. De hoogte kan zo gekozen worden, dat een „e" als een „l" (en omgekeerd) geïnterpreteerd wordt.
De lengte van bijvoorbeeld de „i", n" en „m" zijn allen verschillend en moeten bij de normalisatie tot een standaardlengte teruggebracht worden. De moeilijkheid is echter dat in eerste instantie de letter, en dus de lengte, niet bekend is, zodat bij aaneengeschreven letters, het onmogelijk is, het schrift onder te verdelen in afzonderlijke letters. Zo heeft het woord „in" ongeveer dezelfde lengte als de letter „m". De normalisatie brengt dus ook onoverkomenlijke moeilijkheden met zich mee. Een vereenvoudiging van het proces geeft de eis, dat iedere letter in een aangegeven vakje geplaatst moet worden. In hoeverre deze eis in praktijk te verwezenlijken valt, blijft een open vraag, die door een gericht onderzoek beantwoord kan worden.
De invoering van het schrijfvakje geeft geen oplossing in het onderscheiden van letters die veel op elkaar lijken, zoals: „i", „e" en „r". De ruimten die deze letters in het mathematisch model representeren overlappen elkaar gedeeltelijk. Een scheiding van deze symbolen, ondanks dit laatste bezwaar, wordt door de mens ver-

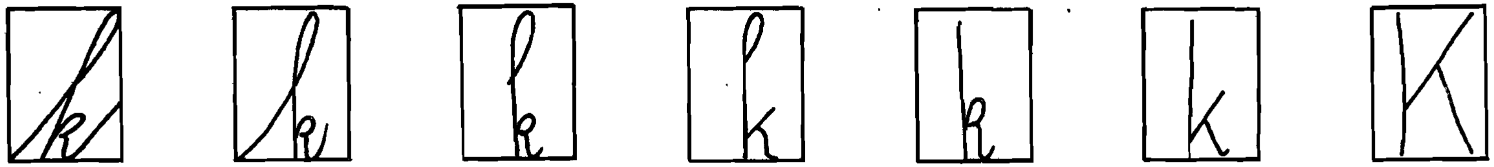


FIG.-20

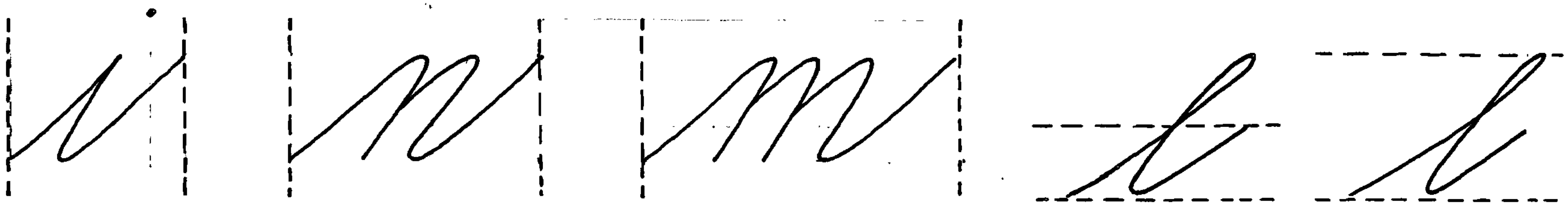


FIG.-21

verwezenlijkt, door de context in het herkenningsproces te betrekken. Hier kom ik later nog op terug.

Stelling:

Zonder de omgeving in het herkenningsproces mee te nemen, is het praktisch onmogelijk om alleenstaande handgeschreven letters (eventueel ook letterkombinaties) voldoende nauwkeurig te identificeren.

Ik laat hierbij open, dat de methode die bij cijferherkenning beschreven wordt een mogelijkheid biedt, een min of meer acceptabel resultaat in de herkenning te behalen. In hoeverre het resultaat in de praktijk toepasbaar is, zullen onderzoekingen in de toekomst moeten uitwijzen.

Mocht bovenstaande beschouwing op het handgeschreven schrift in zijn meest uitgebreide verschijningsvorm, het herkennen van één bepaald type schrijfstijl ondervindt aanzienlijk minder moeilijkheden omdat de spreiding van de verschillende symbolen veel kleiner is. Het is in dit verband beslist mogelijk om via een leerproces het handschrift van een bepaalde persoon te herkennen. Dit is dus een leren en herkennen van letters, geschreven in één stijl.

Voor blokletters en drukletters geldt dezelfde opmerking, omdat de spreiding der letters nog kleiner is.

5.2. Herkenningssysteem voor geschreven symbolen.

Het is de kunst om, ondanks de vele reeds genoemde moeilijkheden, een systeem op te zetten, dat voldoet aan een redelijk nauwkeurige herkenningsfunktie. Een methode tot het verwezenlijken van een proces kan gebaseerd worden op de vormherkenning. In eerste instantie moet dan de vorm van het symbool omgezet worden in digitale informatie, opdat de verwerking met behulp van de computer kan geschieden. De omzetting wordt verwezenlijkt met behulp van een 10 x 15 fotocel-matrix. Als output worden 75 paren random gekozen cellen genomen, die ieder vier toestanden kunnen aannemen. Bij het leer- en herkennings-proces blijkt dat bij een bepaald symbool verschillende kombinaties van paren niet kunnen voorkomen. Dit feit is de kern van de methode die o.a. door W. W. Bledsoe en J. Browning beschreven is in:

„Proceedings of The Eastern Joint Computer Conference 1959, pp. 225-232."

Met deze methode is het reeds mogelijk om symbolen te herkennen die:

- a niet precies aan de standaardvorm voldoen.
- b niet de juiste positie hebben. (translatie.)
- c een kleine hoekverdraaïng hebben ondergaan.

De eigenlijke herkenning wordt gedaan door de meetresultaten (fotocel-paren) te vergelijken met de gegevens van het leer-proces. Deze vergelijking kan met behulp van een combinatorische schakeling, een correlatie of een programma verwezenlijkt worden. Ik zal hier niet verder op ingaan, omdat deze zaken elders uitgebreid aan de orde komen.
Uiteraard is deze methode slechts een kleine stap in de goede richting; er zijn methoden bekend die tot veel betere resultaten leiden. De essentie van dit systeem ligt niet zozeer in het percentage herkende symbolen,

dan wel in de gegevensreductie, die het mogelijk maakt een vrij eenvoudig proces met relatief goede responsie op te zetten.

De resultaten geven te zien, dat het percentage herkenning steeds een stuk beneden de 80 % ligt, hetgeen vooral te wijten is aan de iets te simplistisch opgezette informatie-overdracht met behulp van paren fotocellen. Er wordt dus bij dit systeem een groot deel van de informatie niet gebruikt. Het systeem is dan ook niet toepasbaar voor het herkennen van meerdere handschriften.

5.3. Machine gedrukte tekens.

De door een machine (flexowriter, typemachien, telex) op papier gezette symbolen hebben een kleine variatie in verschijningsvorm. Beperk ik het herkenningsgebied tot slechts een type letter, dan is de spreiding in de afmetingen zeer gering. Het lees- en herkenningsproces wordt dus relatief eenvoudig, omdat o. a. :

1. De normalisatiemethode is eenvoudig - of kan in een gunstig geval zelfs geheel of gedeeltelijk weggelaten worden.
2. Er kan herkend worden met behulp van correlatie methoden.
3. De symbolen worden in standaardvorm gepresenteerd, zodat alleen de ruis, ten gevolge van misdruk, breuk, vuil, variatie in papierkwaliteit e.d., er uit gefilterd moet worden.
4. Natuurlijk moet bij de presentatie van het symbool er op gelet worden, dat de juiste positie t. o. v. het lees-systeem wordt aangenomen. Het vinden van de juiste positie is bij gedrukte tekens relatief eenvoudig.

Vanwege de kleine spreiding in de presentatievorm van de symbolen, kunnen de herkennings-eisen hoger gesteld worden. Met name de nauwkeurigheid van herkenning kan hoger liggen. Dit is belangrijk voor toepassingen in het bank- en post-wezen: manipulaties met geldbedragen. Het is tot op heden alleen mogelijk gebleken met behulp van gedrukte tekens een dusdanige herkenningsnauwkeurigheid te krijgen, dat toepassing in financiële zaken gerechtvaardigd is.
Het is duidelijk d[illegible] bij een hoge nauwkeurigheid verschillende stoorfaktoren gaan meetellen, die bij een minder exacte herkenning te verwaarlozen waren.
Als voorbeeld noem ik :

print- en papierkwaliteit.

drukfout en letterkwaliteit.
externe stoorfactoren zoals bevuiling
kleine storingen in opname of verwerking
systeem.

Ook de processnelheid kan bij gedrukte tekens verhoogd worden. Alleen al de vereenvoudigde normalisatie levert een behoorlijke tijdwinst op. Een belangrijke vraag, die met behulp van proefnemingen beantwoord kan worden is: „Wat levert de kortste procestijd op, een systeem op basis van
„software"
„hardware"
of een kombinatie van beide?".
Verder is het totaal aantal meetattributen belangrijk voor de snelheid; in het algemeen geldt dat minder parameters een lagere processnelheid en een lagere nauwkeurigheid tot gevolg hebben. Het aantal parameters is dus een kwestie van de te stellen eisen.

Op het gebied van gedrukte teken-herkenning zijn reeds veel experimenten gedaan, mede omdat er op dit ogenblik in dit genre een grote behoefte is en wordt gekweekt. De belangrijkste experimenten zijn gepubliceerd in:

1 IEEE Trans. on system science and cybernetics.
vol. 4 no. 2. juli 1968
„An Experimental Study of Machine Recognition of Hand-Printed Numerals.
2 IEEE Trans. on Electronic Computer
vol. EC-15 no. 6. december 1966.
„An Experimental Investigation of a Mixed-Font Print Recognition System."

Het eerste artikel gaat weliswaar over handgeschreven symbolen, maar het bevat desondanks zeer nuttige informatie, die bij herkenning van gedrukte symbolen

reker gebruikt kan worden.
Het tweede artikel behandelt de herkenning van verschillende drukstijlen. Enkele methoden worden toegepast en de resultaten besproken. Wanneer de herkenning zich tot slechts een type drukletter beperkt, is het percentage goed herkende symbolen uiteraard hoger dan in het geval van meerdere typen drukletters. Als een van de best te herkennen stijl komt „Hermes-Techno Elite" uit de bus.

Een reeds in de handel gebracht systeem is de IBM 1275 optische lezer en sorteerder. De machine herkent het standaardschrift 150-A of 150-B met een nauwkeurigheid van 99%. De gedrukte cijfers worden met een snelheid van 6,7 m/s langs de leeskop gevoerd, die 72 fotocellen bevat. Per symbool wordt vijf keer gescand, zodat een 5x72 matrix ontstaat. Na normalisatie wordt de matrix vergeleken met de 14 aanwezige patroonmaskers of sjablonen. Per seconde worden door 14 maskers tegelijkertijd 1.900.000 vergelijkingen gemaakt. Als de totale vergelijking achter de rug is, wordt bepaald met welk symbool dit teken de grootste totaal overeenkomst heeft. Voldoet de overeenkomst niet aan bepaalde eisen, dan wordt het teken als niet herkend beschouwd. De IBM 1275 herkent 2700 tekens per seconde. Nauwkeurigheid is onontbeerlijk voor een leesmachine voor financiële instellingen. Ondanks onvolkomenheden in de presentatie der symbolen, moet de herkenning toch goed zijn. Tijdens de herkenning in de leerfase worden de sjablonen met behulp van een computer getoetst aan een groot aantal slecht gedrukte symbolen. De geweigerde beelden worden door de mens geanaliseerd en naar aanleiding daarvan worden de sjablonen verbeterd. Hier komt weer heel duidelijk de wisselwerking tussen mens en machine naar voren.

5.4. Herkenning van handgeschreven tekst.

Voor het opstellen van een lees- en herkenningssysteem voor handgeschreven tekst is een zeer goede kennis noodzakelijk van de psychische en fysiologische processen die zich bij het lezen bij de mens afspelen.

De huidige systemen zijn allen te verdelen in leerproces en herkenningsproces. Deze twee fasen doorloopt de mens ook; immers het is onmogelijk om een symbool te herkennen, als het in een soortgelijke vorm niet voorheen aan de mens gepresenteerd en aangeleerd is. Het woord spreekt trouwens voor zichzelf: „her-kennen." Geldt deze paralel voor de grote lijn in het gedraagspatroon van leer- en herkenningsproces, ook wat betreft de details is dit door te voeren. Bedenken we daarbij dat een groot deel van de hedendaagse techniek een zo zuiver mogelijke kopie is van de natuur, dan is het alleen al op grond van dit feit, dat de paralelsituatie aanwezig moet zijn.

De vraag rijst nu: „Hoe werkt het herkenningssysteem in zijn totaliteit bij de mens?" Fysiologisch gezien komt dit neer op een beschrijving van:

1 receptoren ; bijvoorbeeld: het oog
2 transmissiekanaal ; de zenuw
3 verwerkingseenheid : de hersenen.

De grote onbekende is punt drie. De bekendheid omtrent de werking van de hersenen zal er beslist toe bijdragen, dat de herkenningssystemen geperfectioneerd zullen kunnen worden, immers: het menselijk systeem werkt nagenoeg perfect. Een ander punt is het algorithme dat de mens bij herkenning toepast. Het is een kwestie van tijd en niet van onmogelijkheid om dit te onderzoeken. Het is de taak van de psycholoog om het gedragspatroon en het algorithme van reductie, deductie en

klassificatie vast te leggen.
Enige vragen die opgelost moeten worden:

a Hoe onderscheidt men symbolen die veel op elkaar lijken, zoals:

e en i
h en k
a en o
u en v

b Het is mogelijk dat één letter, geschreven door 50 verschillende personen, allen herkend worden als die bewuste letter, terwijl er sterk verschil in schrijfstijl kan zijn. Wat zijn dan de herkenningspunten en vormen.

c In hoeverre speelt de „context" een rol? Hoe wordt het woord in zijn verband geplaatst.

d. Het is aan te tonen, dat een punt van grote invloed kan zijn. Zie figuur 22, waar van boven naar beneden door 80% van de vijftig proefpersonen werd gelezen: „lus, lees, lies, leis".
Dit woord in een context geplaatst, gaf een responsie van 100% voor het woord „lees".

<u>Stelling</u>:
De context van het te herkennen symbool heeft in geval van handgeschreven schrift een grote invloed op het herkenningsproces van de mens, en kan in geval van onzekerheid de doorslag geven.

Opmerking:
Ik durf zelfs te stellen, dat de context zo'n sterke invloed kan uitoefenen, dat het symbool verkeerd geïnterpreteerd kan worden.

Uit voorgaande stelling kan gekonkludeerd worden

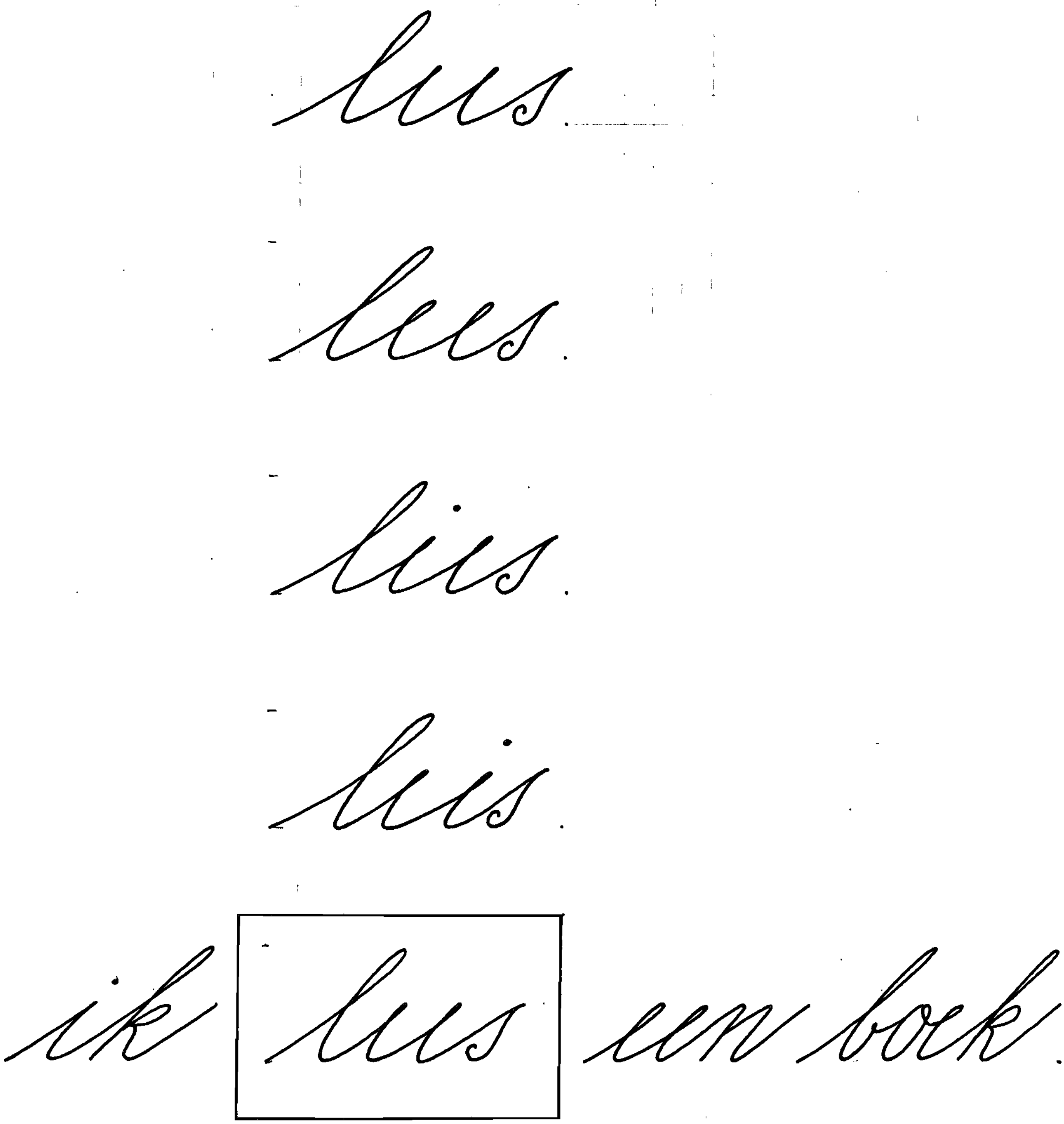


FIG. - 22.

dat bij goede herkenning van handgeschreven tekst het noodzakelijk is, dat de omgeving van het te herkennen symbool (letter) mede bepalend is.

Bekijk dit probleem eens van de stochastische kant.
Stel symboolruimte E^n
Deelruimte van E^n welke een symbool representeert : D.
Iedere deelruimte heeft een gemiddelde vector, die het standaardsymbool representeert : toewijzingsvector $\vec{d}$.
Meetvector : $\vec{x}$

$$\left(\vec{x} \text{ wordt gerepresenteerd door } \vec{d}_i \;/\; \vec{x} \in D_i \right)$$

of in andere notatie :

$$\left(\vec{x} := \vec{d}_i \;/\; \vec{x} \in D_i \right)$$

Bij stochastische onafhankelijkheid der metingen geldt voor de kansverdeling :

$$P\left\{ \vec{x}[t] \,/\, \vec{x}[0], \ldots, \vec{x}[t-1] \right\} = P[\vec{x}[t]]$$

met $\vec{x}[t] \equiv t^e$ meetvector.

$$P\left\{ \vec{x}[t] := d_i \,/\, \vec{x}[0] := d_j, \ldots, \vec{x}[t-1] := d_k \right\} = P\left\{ \vec{x}[t] := d_i \right\}$$

waarbij $j, \ldots,$ en k aan elkaar gelijk mogen zijn.

Het is duidelijk dat de toewijzingskans van het herkenningproces, dat gebruik maakt van de context, stochastisch afhankelijk is. In principe wordt de kansverdeling bij het herkennings-proces van de mens als volgt weergegeven :

$$P\left\{ \vec{x}[t] := d_i \,/\, \vec{x}[t-j] := d_{k_1}, \vec{x}[t-j+1] := d_{k_2}, \ldots, \vec{x}[t-1] := d_{k_j}, \vec{x}[t+1] := d_{k_{j+1}}, \vec{x}[t+2] := d_{k_{j+2}}, \ldots, \vec{x}[t+l] := d_{k_{j+l}} \right\}$$

Vereenvoudiging door alleen de voorafgaande context te beschouwen geeft :

$$P\left\{ \vec{x}[t] := d_i \,/\, \vec{x}[t-j] := d_{k_1}, \vec{x}[t-j+1] := d_{k_2}, \ldots, \vec{x}[t-1] := d_{k_j} \right\}$$

Het herkenningproces wordt een Markov-proces als $j = 1$

wordt genomen:

$$P\{\vec{x}[t]:=d_i / \vec{x}[t-1]:=d_{k_1}\}$$

Dit is de eenvoudigste vorm van de kansverdeling van een herkenningsproces, dat rekening houdt met de context, namelijk het vooraf bepaalde symbool. Bedenk echter dat deze kansverdeling nog steeds een gekompliceerde zaak is. Om de kansen in een concreet geval te bepalen, is een langdurig leerproces noodzakelijk.
Het is mogelijk deze kansen uit te zetten in een matrix.
Stel het aantal toewijzingsvectoren is $2b$:
$d_1, d_2, \ldots, d_{2b}$, waarbij ze de symbolen van het alfabet representeren.
In de i^e rij, j^e kolom komt dan te staan:

$$P\{\vec{x}[t]:=d_i / \vec{x}[t-1]:=d_j\}$$

Zoals bekend, moet in de matrix gelden:

$$\sum_{i=1}^{2b} \left[P\{\vec{x}[t]:=d_i / \vec{x}[t-1]:=d_j\}\right] = 1$$

Dus de som van de termen in een kolom is altijd „een".

Uit dit alles is zonder meer te concluderen, dat tekenherkenning niet te sterk gebaseerd moet worden op de context, omdat de techniek, om het op een dergelijke manier te verwerkelijken, niet aanwezig is. Het ligt dus voor de hand dat het goed herkennen van handgeschreven tekst voorlopig onmogelijk is, omdat de herkenning te veel moet steunen op de context.
Het probleem van de context is niet van toepassing bij standaard gedrukt schrift, omdat de herkenningscriteria wegens de eenduidigheid geen gebruik hoeven te maken van de omgevingsinformatie.

Wil ik toch handgeschreven informatie verwerken, dan moet de presentatie in cijfers geschieden, omdat:

1. De onderlinge correlatie der verschillende cijfers is klein genoeg om de verschillende symbolen te onderscheiden.
2. De correlatie van hetzelfde symbool, in verschillende stijlen geschreven, is gemiddeld groter bij cijfers dan bij letters.
3. De context speelt totaal geen rol:

$$P\{\vec{x}[t] := d_i \,/\, \vec{x}[t-j] := d_k\} = P\{\vec{x}[t] := d_i\}$$

als $j \neq 0$

Opmerking:

Het zou een nuttige studie zijn, om de correlatie te bepalen tussen:

a. dezelfde cijfers in verschillende stijl geschreven.
b. verschillende cijfers in dezelfde stijl geschreven.
c. dezelfde letters in verschillende stijl geschreven.
d. verschillende letters in dezelfde stijl geschreven.

Dit onderzoek kan gedaan worden met behulp van „spatial filtering." Zie daarvoor o.a.:

„Optical and Electro-Optical Information Processing" bladzijde 136, 137.

5.5 Enige mogelijkheden voor de toekomst.

Het systeem voor herkenning van handgeschreven schrift vindt een onoverkomenlijke moeilijkheid in de grote spreiding van de lettertekens. Er kunnen in principe twee dingen gedaan worden:

a Het systeem dusdanig inrichten, dat deze sterke spreiding geen essentieel bezwaar meer is. Het herkenningssysteem zal dan een beslissing opzetten aan de hand van o.a. topologische kenmerken.

b De spreiding in het schrift moet drastisch verkleind worden. In hoeverre dit praktisch te verwezenlijken is, laat ik even open, daar dit situatie-afhankelijk is.

Een kombinatie van beide is natuurlijk het meest aan te bevelen.

Als besprekingsvoorbeeld wil ik het lezen van adressen aanhalen. In dit geval zou het dus denkbaar zijn, dat alle adressen in blokletters geschreven worden, hetgeen een kleinere spreiding geeft in de lettertekens. Een stap verder is het toevoegen van een codenummer aan de straat- en stadnaam, bijvoorbeeld:

Leidsestraat 23   2501
Amsterdam   02000

Iets gecompliceerder wordt het voor de schrijver als codenummers en huisnummer geplaatst worden in een van de hoeken van de envelloppe, bijvoorbeeld rechter benedenhoek, waar ze geschreven worden in aangegeven vakjes. Vanwege de normalisatie is een eis, dat de cijfers binnen het vakje geplaatst worden, terwijl het vakje maximaal gevuld moet worden.

Bijvoorbeeld:

Leidsestraat 23
Amsterdam.

2000 2501 023

Voor het stadsnummer worden vier hokjes gereserveerd; bijvoorbeeld het telefoonnummer, zonder de eerste „0" en met een eventuele aanvulling tot vier cijfers.
Het straatnummer bevat vier cijfers; bijvoorbeeld een cijfer als wijknummer en drie cijfers als straatnummer in de betreffende wijk.
Het huisnummer beslaat drie nummers, decimaal genoteerd.
Totaal dus elf hokjes, die op een vaste afstand van de onderkant en rechter zijkant van de enveloppe zijn gedrukt. Als nu de afmetingen van de hokjes ook vastgelegd worden, dan ligt de plaats van ieder cijfer vast t.o.v. de rechter benedenhoek. Dit is natuurlijk van belang bij het automatisch lezen van de informatie.
Zie ook figuur 23.

Dit systeem eist een aanpassing wat betreft:
De gemeente; zij moet de straten en wijken coderen.
De enveloppenfabrikant; hij moet op de juiste plaats en in de juiste grootte de elf hokjes laten drukken.
De briefschrijver; hij moet zich aanpassen wat betreft het onthouden van de code en het plaatsen van de cijfers binnen de daarvoor bestemde vakjes.
De aanpassing wat betreft de briefschrijver zal doorslaggevend zijn. Blijkt dat deze eisen te zwaar zijn, dan is dit systeem niet praktisch te verwezenlijken.
Deze methode geeft geen moeilijkheden bij automatische adresseermachines en bij de adressering van de P.C.G.-dienst.
Het besprekingsvoorbeeld impliceert een herkenning van handgeschreven cijferschrift. Een methode wordt speciaal voor cijferherkenning uitgewerkt. De toepassing bij letters is natuurlijk in principe ook mogelijk, al

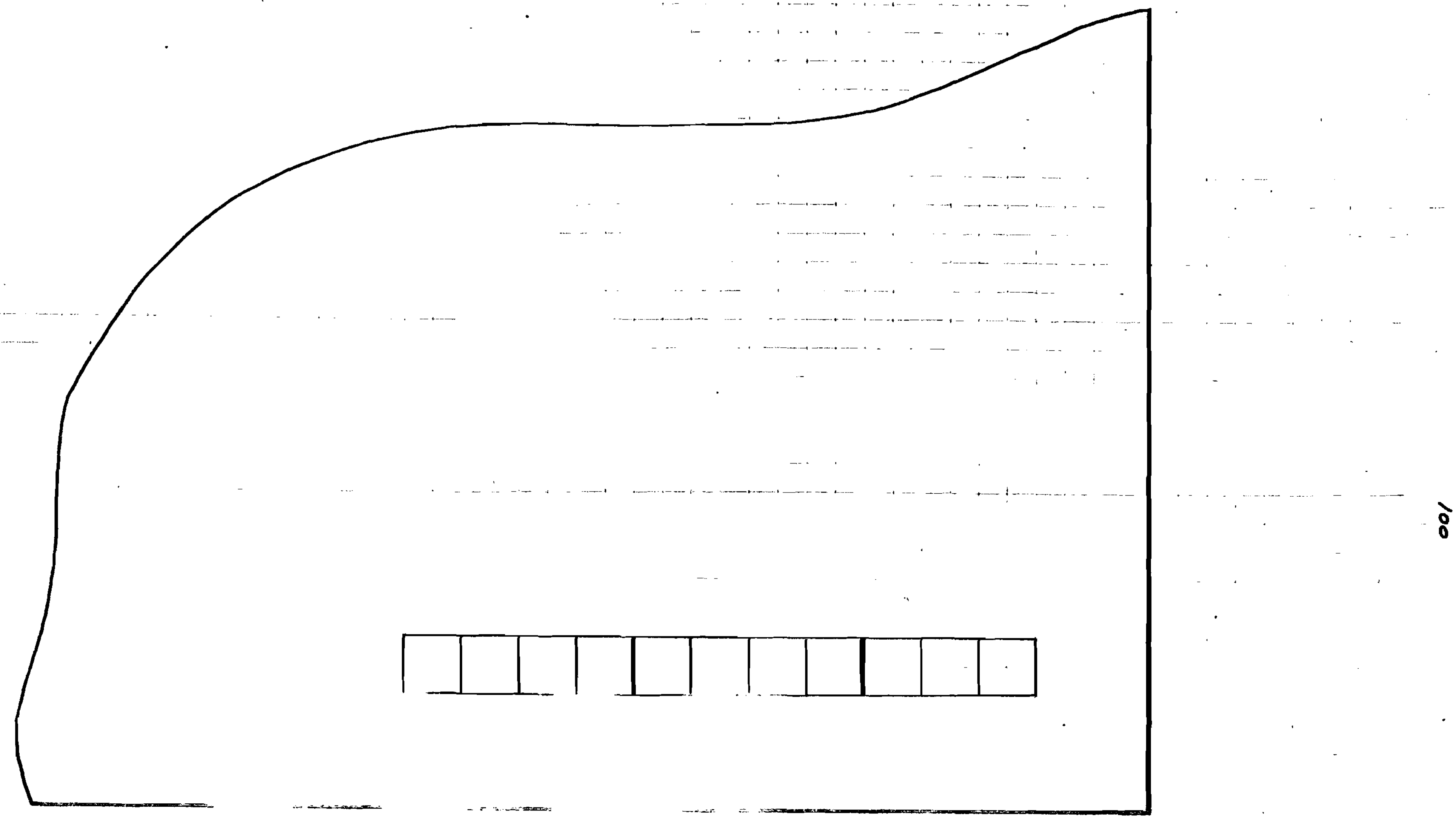


FIG.-23

zal het resultaat dan niet zo gunstig uitvallen. Daarbij komt nog, dat de eisen die gesteld worden bij het schrijven van de symbolen, voor letters veel moeilijker te verwezenlijken zijn dan voor cijfers. Immers: cijfers worden reeds los van elkaar geschreven, terwijl letters meestal aaneen geschreven worden zodat de plaatsing van letters in vakjes een grote aanpassing vereist van de schrijver. De grotere spreiding in de verschijningsvorm van dezelfde letters in verschillende stijlen, en de kleinere spreiding van verschillende letters van dezelfde stijl t.o.v. cijfers is natuurlijk ook van invloed; dit laatste beïnvloedt niet zozeer de presentatie dan wel de herkenningsprocedures.

6. Herkenningsproces

6.1. Een toepassing van het verwerken van cijfergecodeerde informatie.

De toename van de post is ongeveer 7% per jaar. Het automatiseren van de posterijen is dan ook van urgent belang. Het grote stagnatiepunt is het sorteren van brieven e.d. Het ontwikkelen van automatische sorteerders met behulp van cijfercodes mag dan ook gezien worden als de sleutel van de postautomatisering.
Het sorteren kan in twee delen gesplitst worden:

1e sorteren naar de plaats

2e sorteren naar de straat (of wijk.).

De eerste sortering vindt vaak plaats in de trein. Het is dan noodzakelijk dat het sorteerapparaat niet al te groot is. Bij de sortering kan dan ook geen gebruik worden gemaakt van een computer.
De tweede sortering geschiedt in het algemeen in de plaats van bestemming. De grootte van het apparaat is nu willekeurig te kiezen. Ook kan een computer bij het proces worden ingeschakeld.
Bij het ontwerp van een systeem moet rekening worden gehouden met de toelaatbare afmetingen van het apparaat.
De codering wordt al uitgebreid toegepast:

U.S.A.: vijf digit ZIP code

Verenigd Koninkrijk: 6 character alphanumerische code.

Duitsland: vier digit „Postleitzahlen."

Het op- nen van de code kan het beste gebeuren met behulp van optische systemen:
Laat ik licht op de symbolen vallen, dan kan het teruggekaatste licht via een lensensysteem naar een rij fotocellen gevoerd worden. De rij cellen bestaat

uit n delen, die een vertikale strip van het symbool opnemen. Beweegt het symbool in horizontale richting, dan kan, per tijdseenheid een opname gemaakt worden van dat deel van het symbool, dat juist onder het lenzensysteem is. Stel dat per symbool m keer een opname wordt gemaakt. Het op een fotocel vallend licht is een maat voor de helderheid van een bepaald stukje symbool. Met behulp van een drempelwaarde-element kan de helderheid omgezet worden in zwart of wit, respectievelijk „een" of „nul". Het uiteindelijke resultaat is een n x m matrix, die het symbool met behulp van „enen" en „nullen" representeert.

Dit is een van de vele optische opnamesystemen. De techniek is zo ver, dat ongeveer alle moeilijkheden die zich hierbij voordoen opgelost kunnen worden.

Iets anders ligt het bij de herkenning. De grootste moeilijkheden bij herkenning van handgeschreven cijfers bij een post-code zijn:

1. Het schrijfinstrument kan sterk varieren, zodat er een grote spreiding is in o.a. de „zwartheid" van het symbool en de dikte van de lijn.
2. De variatie (spreiding) in de vorm van de symbolen.

De dikte van de lijn kan genormaliseerd worden. Hiervoor is echter een extra procedure nodig. Hier kom ik later op terug.

Het sorteerproces moet een foutenpercentage hebben, dat beneden de 1 a 2 % ligt. Deze grens moet worden aangehouden omdat dit het percentage fouten bij handsortering is.

Een vr. onderzoek is noodzakelijk. De faktoren die daarbij bekeken moeten worden zijn o.a.:

1. Het gebruik van de verschillende soorten schrijfgereedschap.
2. De meetvectoren en kenmerken die in aanmerking

kunnen komen. Er moet een aantal kenmerken worden vastgesteld, die representatief zijn voor de populatie proefsymbolen. Het aantal testsymbolen moet zeker 4000 per type symbool zijn om een voldoend goede representatie te krijgen.

3 Het gedrag van de mens als hij aan bepaalde voorwaarden, wat betreft de presentatie van de symbolen, moet voldoen. Ik denk bijvoorbeeld aan de eis dat een cijfer binnen een vakje geschreven moet worden. „Is er een grote kans dat een hoog percentage van het aantal gepresenteerde symbolen aan die eis voldoet?"

Dit zijn slechts enkele punten. In het algemeen wordt aan de hand van een voorlopig opgezet systeem een vooronderzoek opgesteld. Aan de hand van de resultaten van dat vooronderzoek kunnen er korrekties worden toegepast op het systeem.

6.2. Procesopbouw

De opname van het symbool kan, zoals reeds aangehaald, zeer eenvoudig gebeuren met behulp van een optisch systeem. Uiteraard zijn er vele andere mogelijkheden. De keuze van het opnamesysteem wordt gedaan aan de hand van de gestelde eisen en de representatievorm van de symbolen.
In figuur 24 is, aan de hand van het herkennen van de postcode-cijfers, een schema opgezet, dat in grote trekken het proces weergeeft.
De belangrijkste stappen in het proces zijn:

opname
reductie
normalisatie
transformatie
herkenningsprocedure.

Het systeem berust op topologische kenmerken van de cijfers. Aan de hand van een trainingsset worden topologische basisfiguren geanaliseerd, waarmee alle cijfers te construeren zijn. Een stap verder is het analyseren en ontbinden van de topologische figuren in standaardfiguren. Er moet een minimaal aantal standaardfiguren gevonden worden waarmee de basisfiguren te construeren zijn.
De standaardfiguren worden aan de hand van een aantal 3×3 matrices (in het algemeen $p \times q$), die uit de inputmatrix zijn gedistilleerd, herkend door een combinatorisch netwerk. Deze gedachtegang, waaraan het gehele proces ten grondslag ligt, is weergegeven in figuur 25, waar de analyse van een acht uitgebeeld wordt. Het gepresenteerde symbool wordt dus in stukjes verdeeld zoals in d; daarna wordt alles samengesteld (via c en b) tot een totaal cijfer.
Deze methode is zo belangrijk, omdat er grote voordelen zijn:

1. In het algemeen mag er een grote spreiding zijn

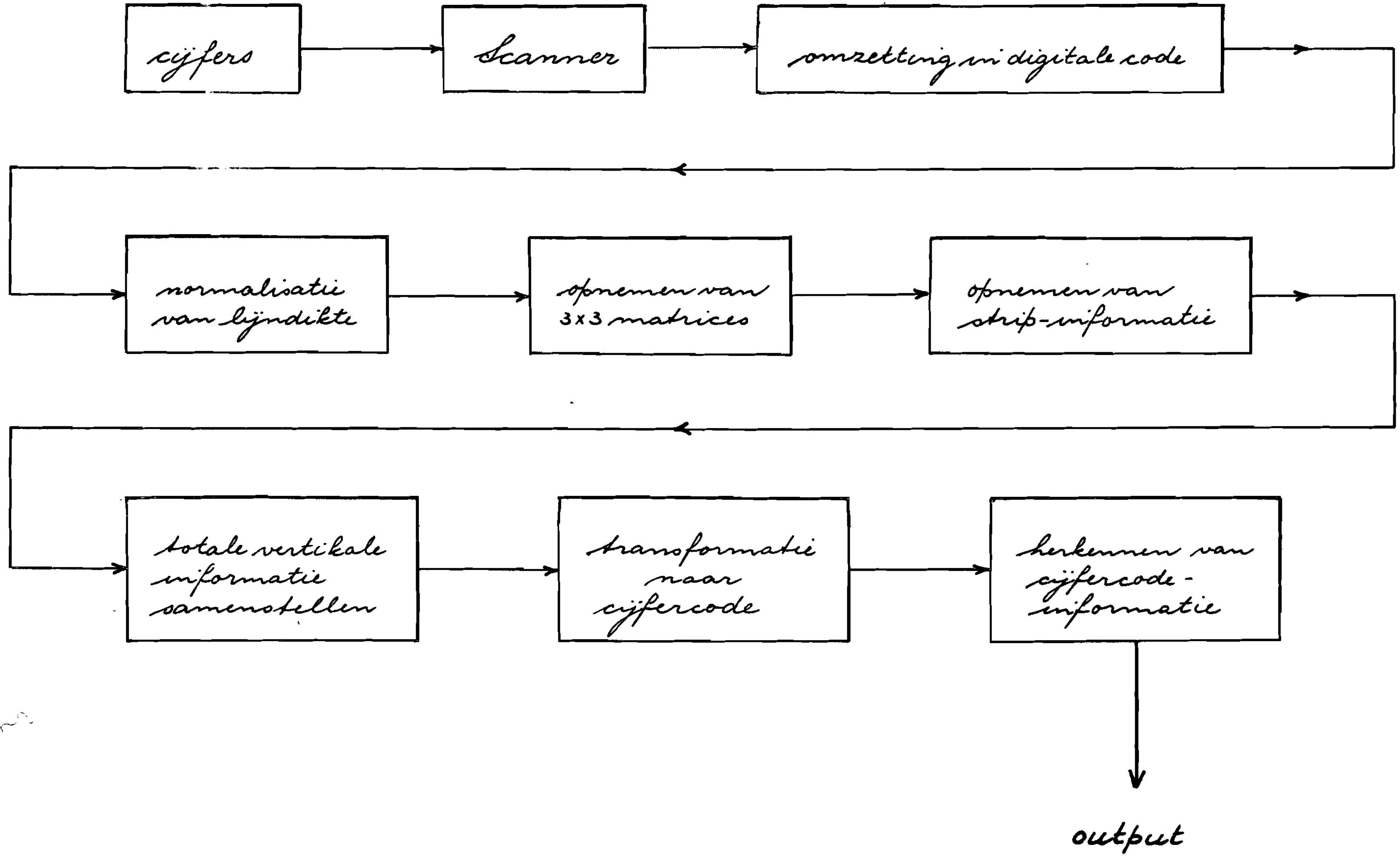

cijfers
Scanner
omzetting in digitale code
normalisatie van lijndikte
opnemen van 3x3 matrices
opnemen van strip-informatie
totale vertikale informatie samenstellen
transformatie naar cijfercode
herkennen van cijfercode-informatie
output

FIG. - 24

stripinformatie
(basisfiguren)

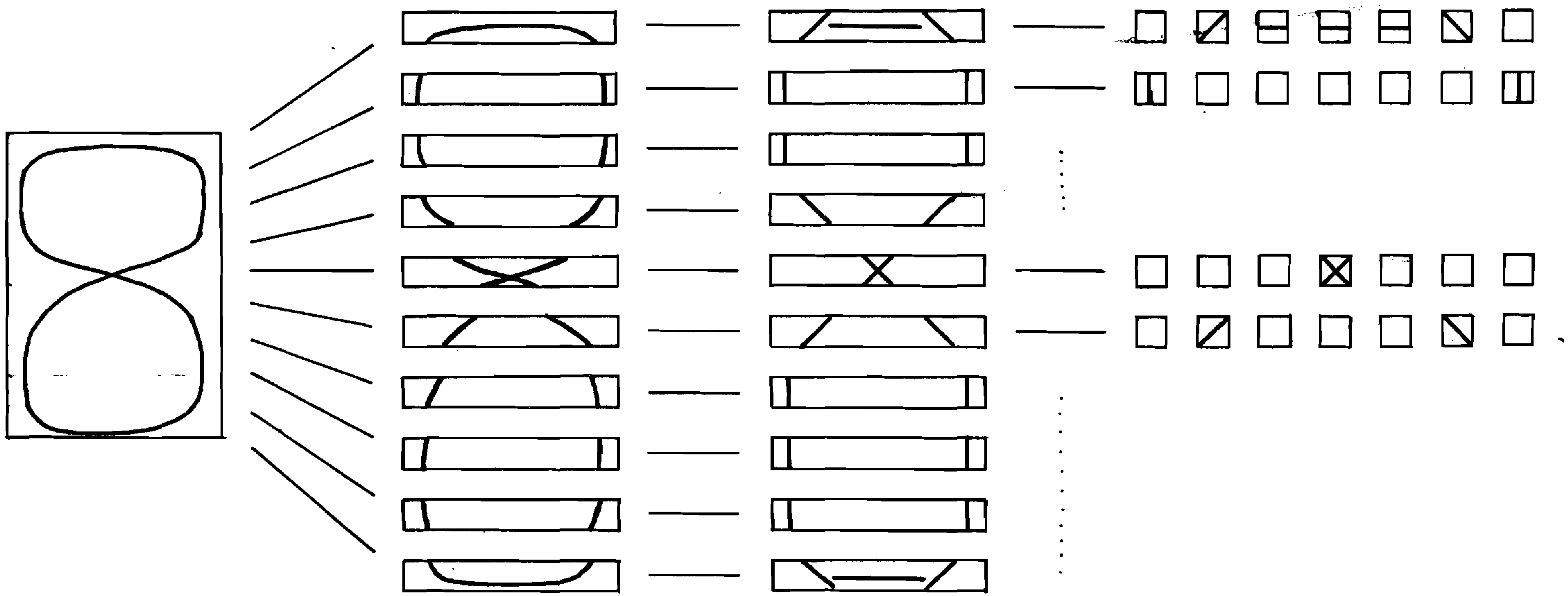


FIG.-25

in de presentatievorm.

2. Zeer speciaal is er een sterke richting en grootte onafhankelijkheid bij de symboolherkenning.
3. De normalisatie van de symbolen hoeft slechts tot het minimum beperkt te worden.

Het toepassingsgebied strekt zich verder uit dan cijfer of letter-herkenning. Ieder topologisch figuur kan in principe herkend worden.

6.3. Reductie, normalisatie en transformatie

Een uitgebreid onderzoek heeft aangetoond dat de informatiedichtheid bij cijfers in horizontale richting groter is dan in vertikale richting. Hiermee moet rekening worden gehouden met de informatie-reductie. De informatiehoeveelheid zal bij vertikale reductie minder afnemen dan bij horizontale reductie. Hiermee is in het proces rekening gehouden, door de $n \times m$ inputmatrix, na lijnnormalisatie, te verdelen in $(n-2)$ matrices in horizontale richting en $(\frac{m-1}{2})$ matrices in vertikale richting. In figuur 26 zijn drie reductieprincipes aangegeven. De (3×3) matrices bevatten een standaardfiguur, zodat de $n \times m$ matrix omgezet wordt in een $\{(n-2) \times (\frac{m-1}{2})\}$ standaardfiguren matrix. Deze laatste matrix wordt onderverdeeld in $(\frac{m-1}{2})$ horizontale strippen. Iedere strip bevat een basisfiguur die wordt samengesteld uit de horizontale standaardfiguren. Zie ook figuur 25.
Om te komen tot een juiste omzetting van

(3×3) matrix naar standaardfiguur.
standaardfiguur naar basisfiguur

REDUCTIE PRINCIPES

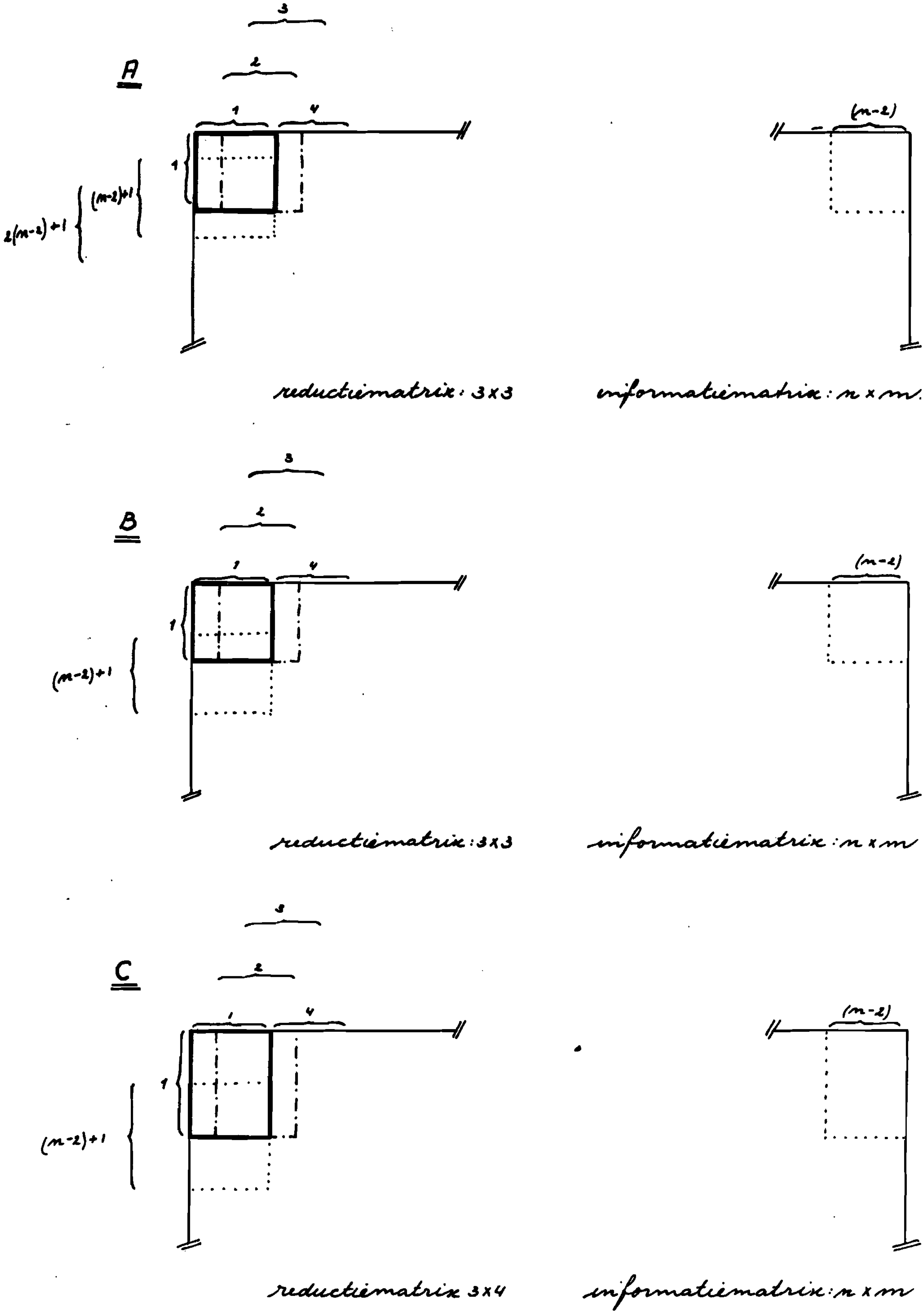


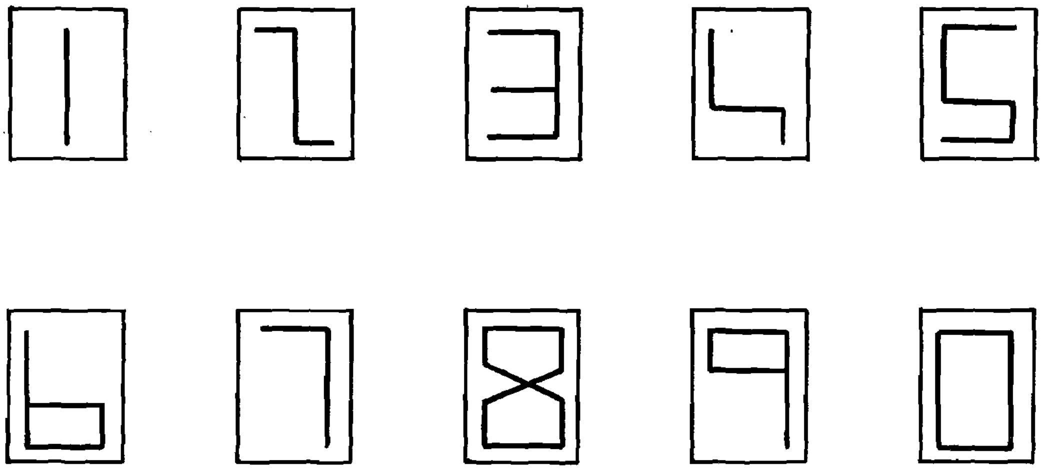
FIG.-26

is een uitgebreid onderzoek aan de hand van een trainingsset noodzakelijk.
Uit dit vooronderzoek wordt bepaald:

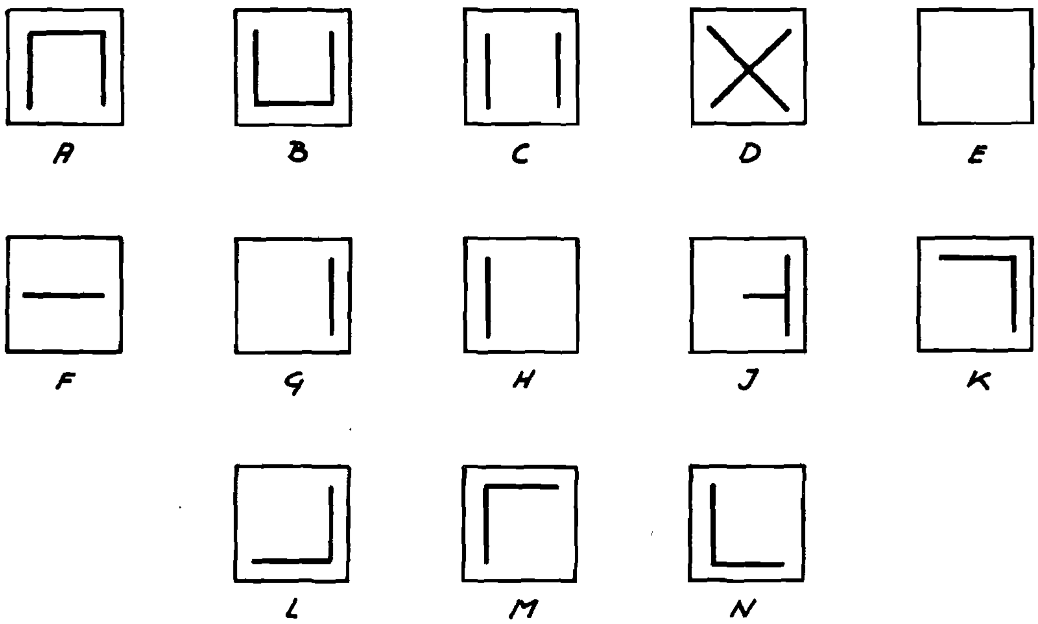
a welke en hoeveel standaardfiguren
b welke en hoeveel basisfiguren
c voorgenoemde transformaties

Nadat de stripinformatie is samengesteld, moeten de gegevens aan elkaar gekoppeld en getransformeerd worden om tot een herkenning te komen. Er zijn ongeveer $(\frac{n}{2}-2)$ basiskenmerken. Deze kunnen met behulp van „software" eventueel gereduceerd worden. Met behulp van een toestandsdiagram kan het geheel daarna omgezet worden in een cijfercombinatie, die in volgorde de doorlopen toestanden (basiskenmerken) aangeeft. De cijfercombinatie is nu representatief voor het te herkennen symbool, zodat aan de hand van de combinatie het symbool wordt bepaald.

In figuur 27 is een aantal genormaliseerde cijfers gegeven; stel dat de symbolen representatief zijn voor een trainingsverzameling. Aan de hand van deze cijfers worden er dertien basisfiguren gecreëerd, waarmee de cijfers geconstrueerd kunnen worden. Uit de dertien basisfiguren worden zeven standaardfiguren gedistileerd; met behulp van deze figuren (zie figuur 28) kunnen alle dertien basisfiguren en dus ook alle cijfers samengesteld worden.
Het toewijzen van de zeven standaardfiguren aan de 2^9 binaire (3×3) matrices kan o.a. geschieden met behulp van een Adaline-leerproces.

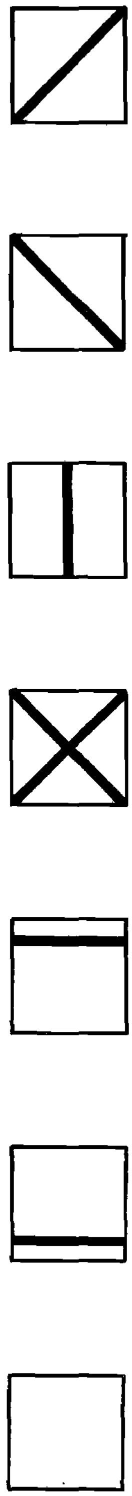


genormaliseerde cijfers



basisfiguren.

FIG.-27



standaardfiguren

FIG.-28

6.4. „Preprocessing"

a Onderzoek naar voorkomen van cijfervormen.
Als eerste moet er een groot aantal handgeschreven cijfers verzameld worden. Hiervoor moeten formulieren ontworpen worden, waarop de mensen de cijfers 0 t/m 9. in aangegeven vakjes, kunnen invullen. De vraag: „In welke mate voldoet men aan de eis, dat het gehele vakje wordt opgevuld door het cijfer." is aan de hand van de enquête eenvoudig te beantwoorden.

b Uit de resultaten van de enquête wordt onderzocht hoeveel en welke topologische basisfiguren opgesteld moeten worden. (figuur 27). Deze taak moet zeer grondig verricht worden, daar zij bepalend is voor de snelheid en nauwkeurigheid van het totale herkenningsproces. Het onderzoek kan met behulp van de rekenmachine uitgevoerd worden.

c De strip- en/of standaardfiguren kunnen beschouwd worden als bouwstenen van de cijfers. De standaardfiguren zijn eenvoudig te bepalen als de basisfiguren eenmaal vast liggen. Ook dit kan met de rekenmachine, en in eenvoudige gevallen zelfs met de hand geschieden. (zie ook figuur 28).

d De cijfers worden gepresenteerd in een binaire $(n \times m)$ matrix, die verdeeld wordt in een aantal redundante (3×3) of (3×4) matrices. Het aantal verschillende (3×3) matrices is 2^9. Het probleem is nu, aan iedere matrix één van de zeven standaardfiguren toe te wijzen. Een juiste toewijzing verhoogt de nauwkeurigheid van het herkenningsproces.

e Een toestandsdiagram zet de basisinformatie om in een cijfercombinatie.
Er zijn 2 mogelijkheden:
1 Voor ieder mogelijk cijfer (tien in totaal) wordt

een toestandstabel opgezet.
Zijnde bij het j^e stripfiguur van de vertikale rij basisfiguren, vindt de overgang van toestand k naar l plaats, als het $(j+1)^e$ stripfiguur een bepaalde stripfiguur is. Zo zijn er overgangen die beslist niet mogelijk zijn als het symbool goed gedetecteerd was; vindt zo'n overgang echter tóch plaats, dan is dat een indicatie voor een fout.
Is het aantal fouten hoger dan een bepaalde drempelwaarde (bijvoorbeeld 2) dan:

is het symbool fout gedetecteerd

of

ik heb een toestandstabel van een cijfer, dat niet het te herkennen cijfer is.

In dat geval wordt de volgende toestandstabel getest. De doorlopen toestanden worden achter elkaar genoteerd. Is het aantal fouten beneden twee, dan is het te herkennen cijfer bekend.
Geven alle toestandstabellen een aantal fouten boven de drempelwaarde, dan is het cijfer niet te herkennen.

2 Voor alle cijfers wordt één toestandstabel opgezet.
Ook hier weer een drempelwaarde-element. Echter, als het aantal fouten boven de drempelwaarde ligt, wordt het symbool als niet te herkennen beschouwd. Is dit niet het geval, dan is de cijfercombinatie een maat voor het symbool. In dat geval moeten deze cijfercombinaties nog geïnterpreteerd worden.

In de figuren 29, 30 en 31 zijn enkele toestandsdiagrammen opgebouwd aan de hand van de gestandaardiseerde cijfers van figuur 27.
In figuur 29 is een diagram gemaakt voor de herkenning van het cijfer acht. De basisfiguren worden in het herkenningsschema (zie figuur 25: basisfiguren) van boven naar beneden doorlopen; beginnend bij E en eindigend bij E.

Het diagram kan op verschillende manieren doorlopen worden:

111223445670009 = 123456709 = cijfer acht.

112445 79 = 124579 = cijfer acht.

In figuur 30 is een diagram voor de acht en de nul opgesteld. Figuur 31 geeft een diagram voor de acht en de negen.

Enige andere mogelijkheden van cijfercombinaties:

Bij figuur 30:

1 1 2 2 4 4 4 4 4 4 0 0 9 =

= 1 2 4 0 9 = cijfer nul.

1 3 4 4 4 4 7 9 = 1 3 4 7 9 = cijfer nul.

Bij figuur 31:

1 2 4 4 4 11 12 12 9 =

= 1 2 4 11 12 9 = cijfer negen.

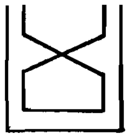
1 1 2 2 3 4 4 10 10 12 12 12 9 9 =

= 1 2 3 4 10 12 9 = cijfer negen.

Het zal nu wel duidelijk zijn, dat een diagram voor tien cijfers een complex geheel is.

Het interpretatieve algorithme moet in de trainingsfase met behulp van een rekenmachine worden vastgelegd.

Een anders opgezet toestandsdiagram is te zien in figuur 32. Nu worden de toestanden die doorlopen worden opgeschreven in letters. Raakt het systeem even in een stabiele toestand, dan wordt dat toch maar door een letter aangegeven, zodat hier automatisch een reductie plaats vindt. In hoeverre deze methode beter toepasbaar is dan de cijfermethode, is moeilijk te zeggen.

Diagram voor: 

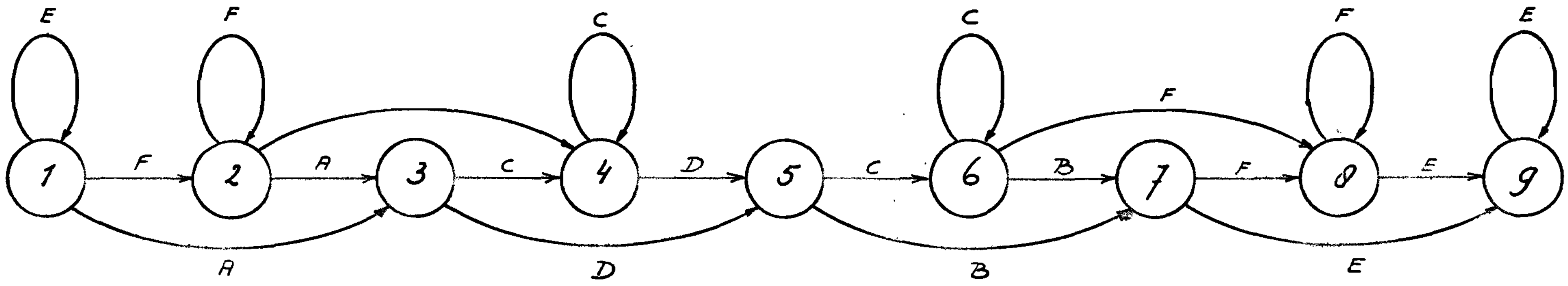
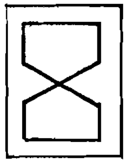
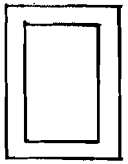


Fig.-29

Diagram voor:  en 

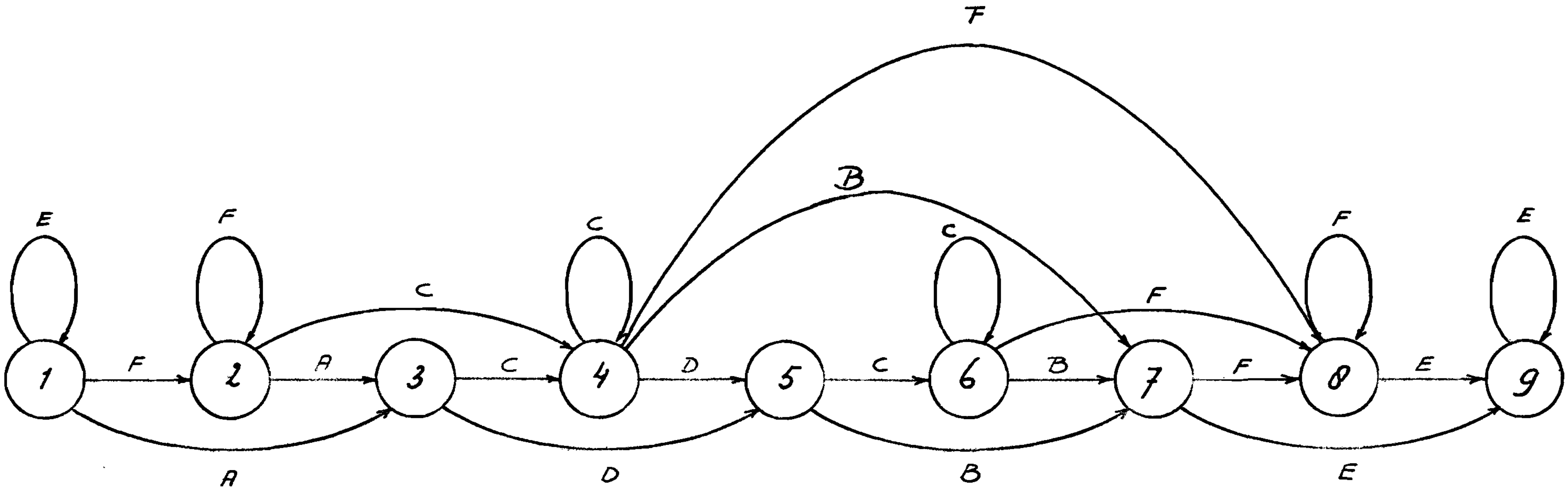
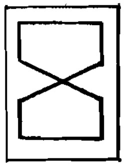
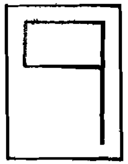


FIG.-30

117

Diagram voor:  en 

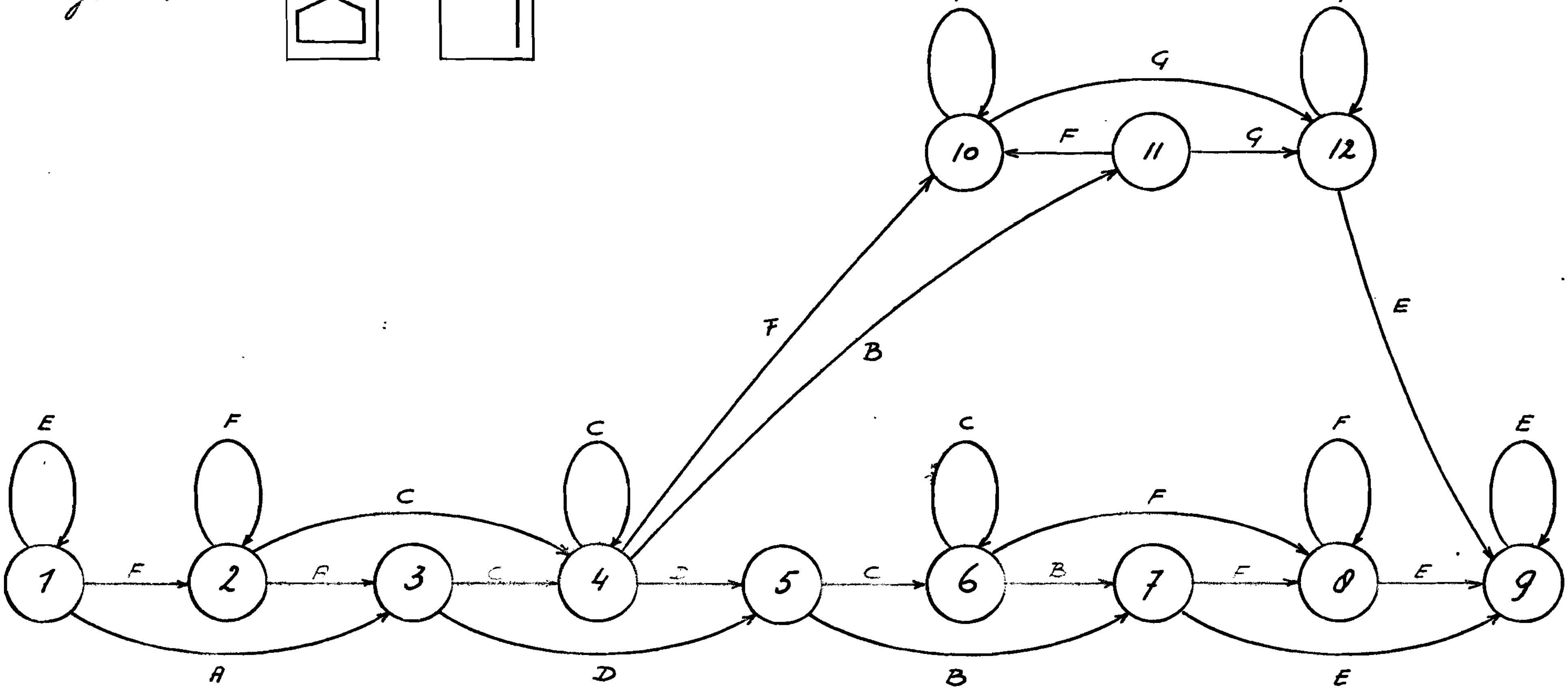


FIG.-31

Vervangingsdiagram voor figuur 31.

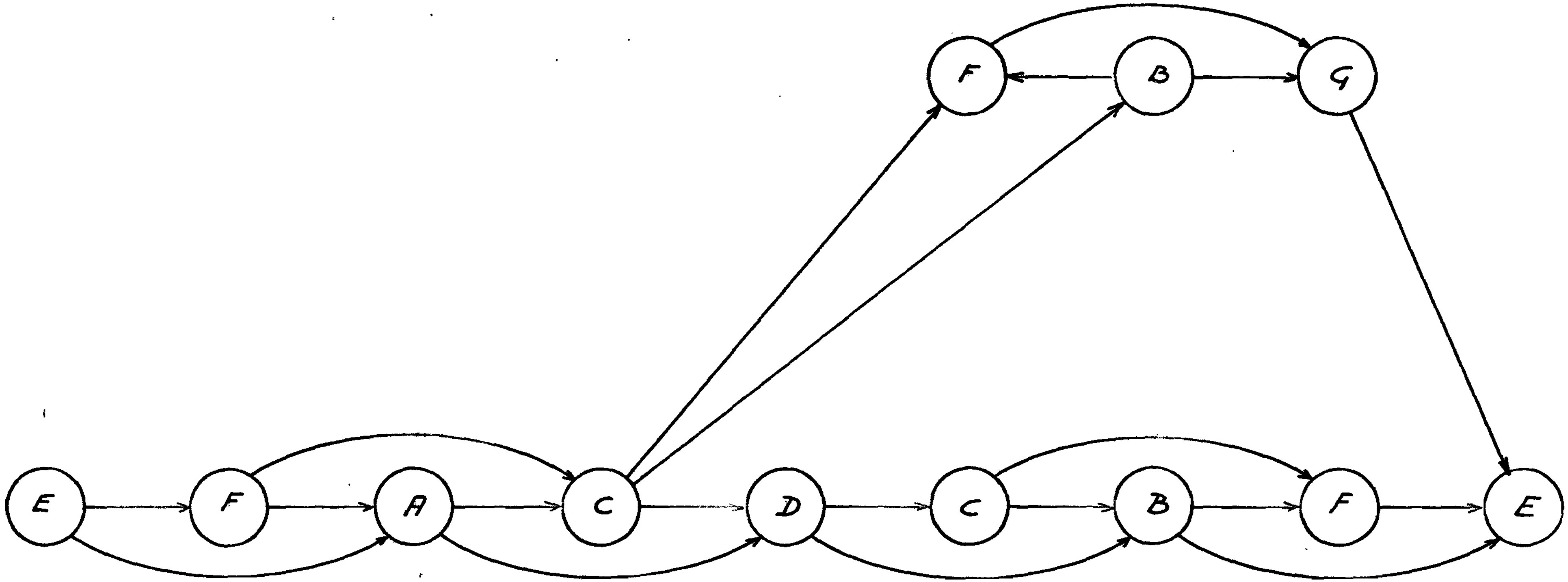


FIG.-32

6.5. Lijnreductie.

Een van de grote bezwaren bij de gepresenteerde cijfers is de spreiding in de lijndikte. Er kan een variatie in dikte optreden van ongeveer 0,2 tot 1,0 mm. Het patroon dat opgenomen en getransformeerd wordt naar de binaire $(n \times m)$ matrix kan een resultaat geven zoals in figuur 33 is te zien. Het is wenselijk dat de lijndikte gereduceerd wordt tot 1 a 2 bits.

In figuur 26 A, B en C zijn drie reductieprincipes aangegeven waarmee de lijndikte gereduceerd kan worden. Deze principes zijn aan de hand van de cijfermatrix van figuur 33 uitgewerkt.
In het algemeen wordt de $(n \times m)$ matrix volgens een bepaald reductieprincipe verdeeld in een aantal redundante $(p \times q)$ matrices.
Als de sommatie van de inhoud van de $(p \times q)$ reductiematrix groter is dan de drempelwaarde θ,
dan : wordt de matrix vervangen door 1
anders : wordt de matrix vervangen door 0
De parameters zijn p, q en θ.
In figuur 34 en 35 is genomen :

$p = q = 3 \qquad \theta = 6$

$p = q = 3 \qquad \theta = 6$

$p = 3 \quad q = 4 \qquad \theta = 7$

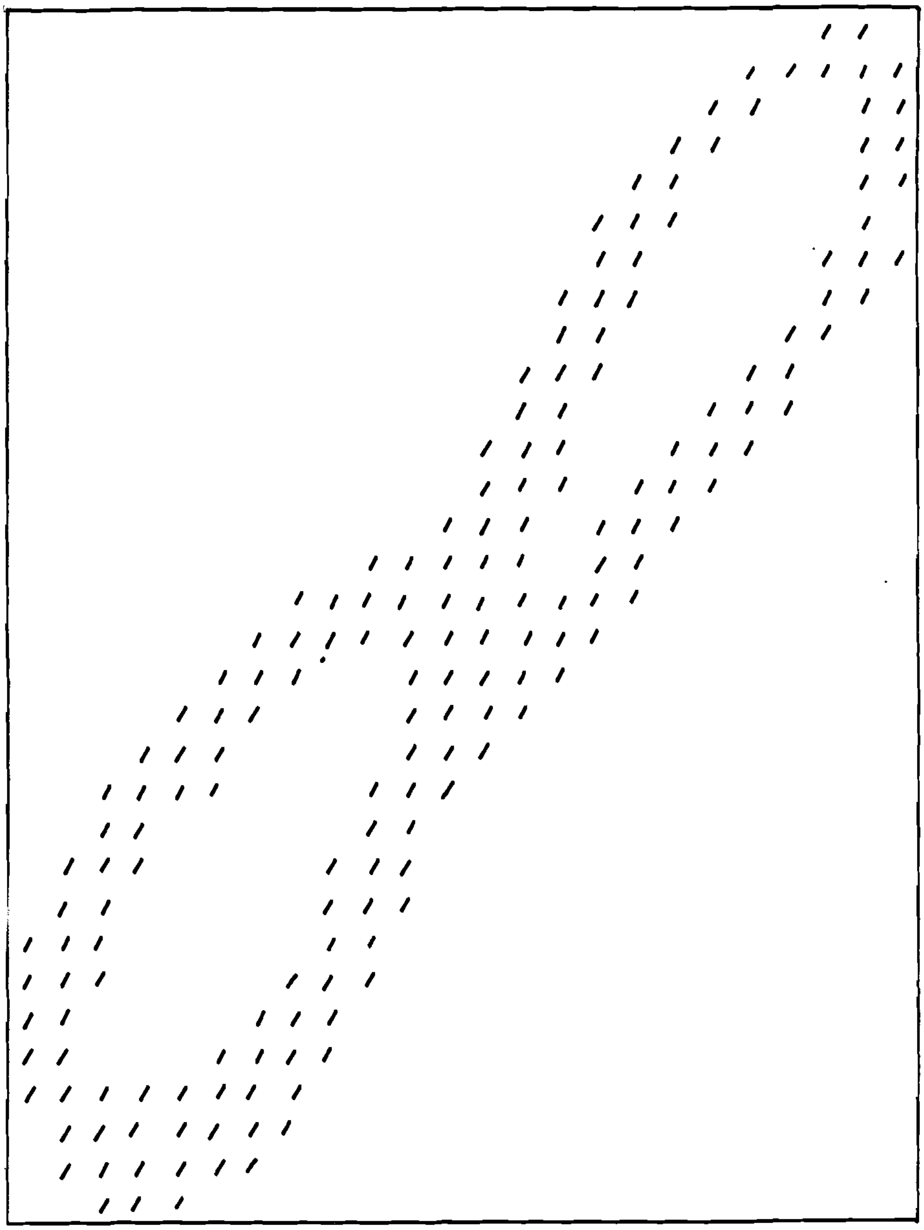
Door deze reductie toe te passen is de lijndikte afgenomen.
Een te sterke reductie doet lijnstukken wegvallen.
Een te kleine reductie zal een te brede lijn geven.
Tussen deze twee mogelijkheden zit een optimale reductie.
Er zijn twee bewerkingen mogelijk :

1 Neem $p = q = 3$ en $\theta = 6$. De reductie wordt een aantal malen achter elkaar uitgevoerd, totdat de lijndikte de gewenste waarde heeft.

2 De waarden van p, q en θ worden zó bepaald, dat in één reductiestap tot de gewenste lijndikte wordt gereduceerd.
De meest efficiënte methode kan met behulp van simulaties op de rekenmachine bepaald worden.

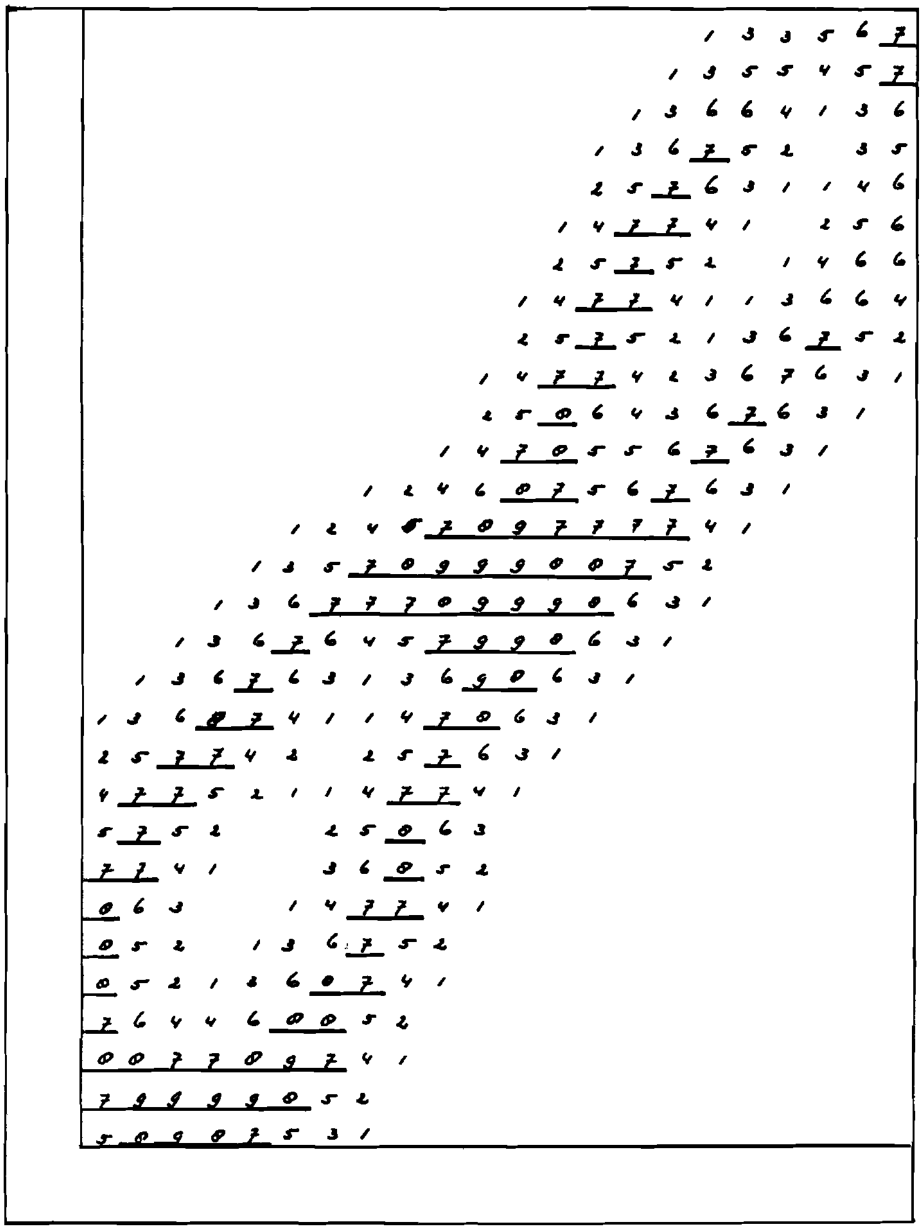
6.6 Slotopmerkingen

Het beschreven herkenningsproces is een complexe zaak. Bij het verwezenlijken van een dergelijk systeem zal een duidelijk diagram moeten worden opgesteld, waarin in chronologische volgorde de verschillende taken worden uitgezet. De cybernetica leert ons dat een juist opgesteld diagram een grote tijdsbesparing oplevert.
Het voorgaande laat duidelijk zien, dat het verwezenlijken van een goed herkenningsproces een moeilijke en omslachtige taak is, die de kennis vereist van een zeer uitgebreid gebied van de wetenschap.



Informatiematrix : n x m

FIG.-33



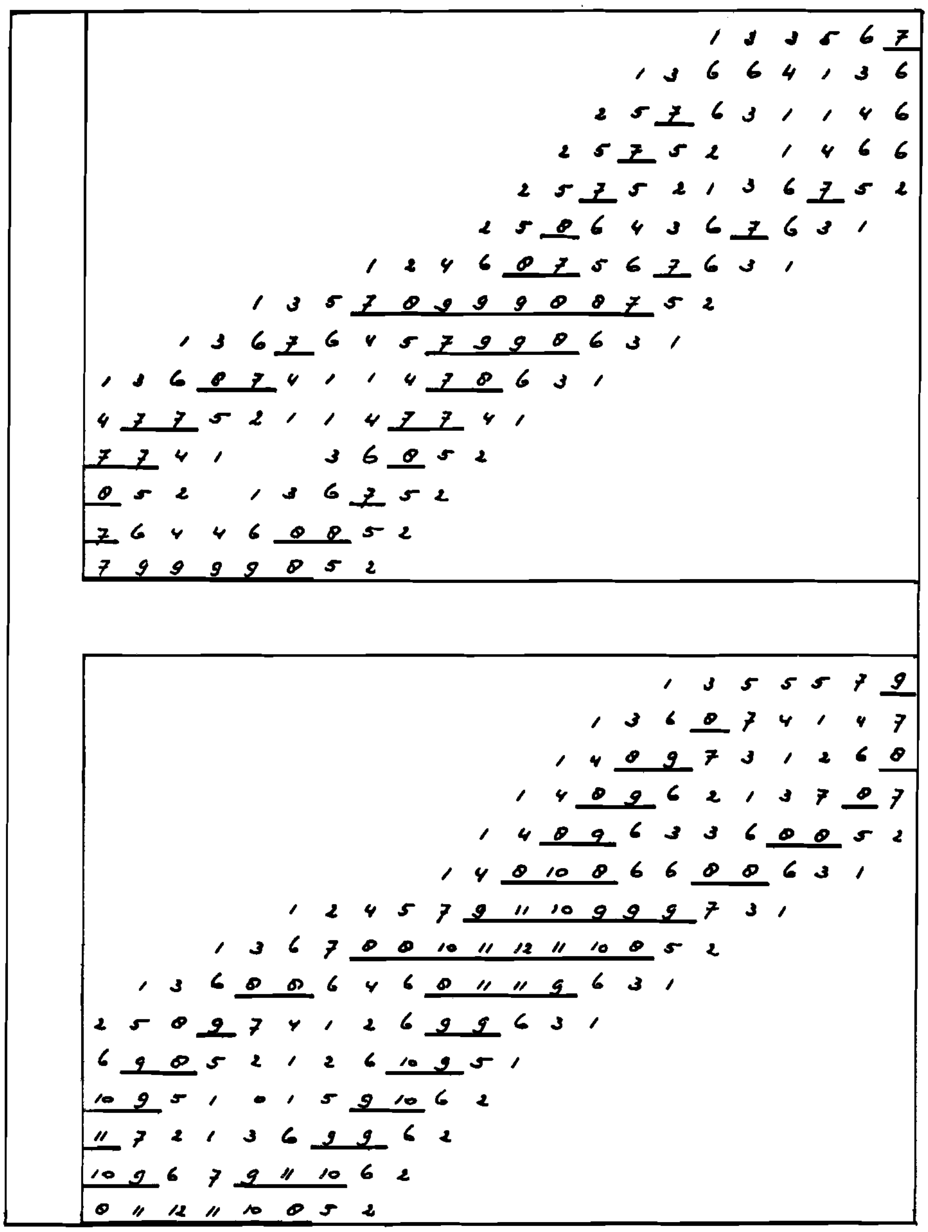
Reductiematrix : 3x3

Informatiematrix: n x m.

Drempelwaarde : 6

reductieprincipe : A

FIG.-34



boven: reductiematrix: 3x3 Informatiematrix: n x m

Drempelwaarde: 6 reductieprincipe: B

beneden: reductiematrix: 3x4 Informatiematrix: n x m

Drempelwaarde: 7 reductieprincipe: C

FIG.-35

7. Conclusies

De tekenherkenning zal in de toekomst een grote vlucht gaan nemen. Mede dankzij de zeer snelle ontwikkeling van de computersystemen is de mogelijkheid geschapen om een zo complex proces als tekenherkenning binnen redelijke grenzen te verwezenlijken. De complexheid van het proces vraagt een snelle uitvoering van de handelingen en algorithmen om tot een redelijke totaaltijd van de procesuitvoering te komen. Aan deze eis is voldaan door de ontwikkeling van zeer snelle systemen. De nauwkeurigheid hangt sterk af van:

1. de presentatievorm der symbolen.
2. de aard en hoeveelheid van bewerkingen.

Het vinden van een juist principe en herkenningssysteem is heden de grondslag van slagen of mislukken. De aparatuur speelt hierbij een ondergeschikte rol.

Het normeren van leer- en herkenningssymbolen kan nog steeds verbeterd worden. Een juiste aanpak van de normalisatie zal het proces sterk vereenvoudigen. Van even veel belang is het reduceren van gegevens (dimensiereductie.); de oorspronkelijke informatie moet behouden blijven, zover deze voor het herkenningsproces nodig is. De reductie moet dusdanig uitgevoerd worden, dat het aantal parameters wordt geminimaliseerd, terwijl de kans op herkenning optimaal gemaakt wordt. Dit is nog steeds een zeer moeilijke opgave. Simulatieprogramma's, uitgevoerd op de rekenmachine, kunnen er toe bijdragen dat de ideale reductie zoveel mogelijk benaderd wordt.

Het is dus een zaak van proefnemingen aan de hand van de verschillende theorema's. Dat het ontwerpen van een herkenningssysteem veel werk betekent, zal nu wel duidelijk zijn. Het opzetten van een herkenningssysteem

is van een dusdanige omvang, dat de verwezenlijking alleen kan geschieden met behulp van groepswerk. Dit komt vooral tot uiting in het laatste hoofdstuk.
Bij dit alles moet niet vergeten worden gebruik te maken van de resultaten van reeds uitgevoerde proeven.

De toepassing van de tekenherkenning ligt vooral bij het bank- en postwezen. Verwezenlijking van systemen op dit gebied kan voor de bedrijven financiëel heel aantrekkelijk zijn. De industrie concentreert zijn ontwikkeling dan ook vooral op dit gebied. Dit houdt in dat er veel aan cijfer- en letterherkenning wordt gedaan. Wat betreft de herkenning van spraak, patronen, wolkenformaties, interpretatie van E.E.G. enz., is op het ogenblik de activiteit iets teruggenomen.

Het herkennen van één type machineschrift is een procedure, die in een gunstig geval met een hoge zekerheid (90-99%) te verwezenlijken is. Het herkennen van meerdere typen machineschrift is ook mogelijk; de nauwkeurigheid zal met het toenemen van het aantal typen afnemen. Een goede herkenning is mogelijk met acht verschillende typen machineschrift.
De methode, in hoofdstuk zes besproken, heeft een grote toekomst, vooral op het gebied van cijfer- en machine-schrift-herkenning. De optische methode, met behulp van „spatial filtering", heeft een zeer beperkt toepassings-gebied: daar waar de symbolen in een standaardafmeting gepresenteerd worden. Als deze methode te gebruiken is, is de toepassing heel eenvoudig te verwezenlijken.

7.0 Slotwoord.

Nu met dit afstudeerwerk mijn studie aan de Technische Hogeschool is beëindigd, wil ik eenieder die mij tijdens de studie gesteund heeft vanaf deze plaats bedanken. De afstudeerperiode was voor mij zeer leerzaam, mede door een intensief kontakt met de gehele groep. Bijzondere dank geldt voor prof. ir. A Heetman die mij, ook buiten de zuivere wetenschap, veel geleerd heeft. Ook ben ik erkentelijk voor de vruchtbare gesprekken met mijn coach ir. C. P. J. Schnabel en ir F. H. Groen. De prettige verstandhouding met de leden van de groep en de afstudeerders is een belangrijke steun geweest bij de uitwerking van mijn opdracht.

juni 1969 P. R. M. van Binsbergen.

8 Appendix.

De capaciteit van een Adaline

Onder de capaciteit van een Adaline wordt het aantal funkties verstaan, dat door de Adaline te herkennen is. Er wordt een benadering opgezet, die, met behulp van een berekening in een $(n+1)$ dimensionale ruimte, de capaciteit bepaalt.
Een n-dimensionale Euclidische ruimte wordt voorgesteld door E^n.
Een hypervlak in E^n wordt voorgesteld door F.
Stel m hypervlakken:

$$F_1 \subset E^n,\ F_2 \subset E^n,\ \dots,\ F_m \subset E^n.$$

Defenitie van „algemene opstelling" in E^n:
Als de doorsnijding van k hypervlakken uit de verzameling $\{F_1, \dots, F_m\}$
voor $k = 1, 2, \dots, n$ de dimensie $(n-k)$ heeft
en voor $k > n$ de dimensie nul,
dan voldoen de k hypervlakken in E^n aan de algemene opstelling.

<u>Stelling 0</u>
Als m hypervlakken in E^n voldoen aan de algemene opstelling, dan wordt het aantal n-dimensionale gebieden, dat de hypervlakken in E^n opdelen gegeven door:

$$B_n^m = 2 \sum_{i=0}^{n-1} \binom{m-1}{i} \qquad \dots (1)$$

<u>Bewijs</u>
Met behulp van volledige inductie:
Als er een hypervlak F_1 is, dus $m=1$, dan moet gelden $B_n^1 = 2$

$$B_n^1 = 2 \sum_{i=0}^{n-1} \binom{0}{i} = 2.1 = 2.$$

Voor $m=1$ is aan (1) voldaan.

$$B_1^m = 2.1 = 2$$

Dit klopt, want alle hypervlakken moeten door de oorsprong gaan.
Stel dat (1) geldt voor alle $m < m'$, met $m' > 1$.
Aangenomen is, dat F_1 t/m F_m voldoen aan de algemene opstelling.
In dat geval voldoen ook F_2 t/m F_m aan deze voorwaarde; F_1 wordt even buiten beschouwing gelaten.
Het aantal opgedeelde gebieden is nu:

$$B_n^{m'-1}$$

Bekijk de subgebieden

$$G_i \equiv F_1 \cap F_i \quad \text{met } i = 2, 3, \ldots, m'$$

Uit de definitie blijkt dat G_i voldoet aan de algemene opstelling, dus F_1 is door de sub-gebiedenverzameling G_i verdeeld in $B_{n-1}^{m'-1}$ gebieden.
Ieder gebied hiervan deelt het gebied waar het toe hoort in twee nieuwe delen, omdat:

a) het een lineair systeem is

b) de grenzen worden gevormd door hypervlakken, zodat het systeem convex is.

Het optellen van de $B_{n-1}^{m'-1}$ delen bij de reeds bestaande delen geeft:

$$B_n^{m'} = B_n^{m'-1} + B_{n-1}^{m'-1} = 2\sum_{i=0}^{n-1}\binom{m'-2}{i} + 2\sum_{i=0}^{n-2}\binom{m'-2}{i} \qquad \ldots (2)$$

Nu moet nog bewezen worden:

$$\sum_{i=0}^{n-1}\binom{m'-2}{i} + \sum_{i=0}^{n-2}\binom{m'-2}{i} = \sum_{0}^{n-1}\binom{m'-1}{i} \qquad \ldots (3)$$

Met behulp van volledige inductie:

$$n=2: \quad \sum_{i=0}^{1}\binom{m'-1}{i} = 1 + (m'-1) = 1 + (m'-2) + 1 = \sum_{i=0}^{1}\binom{m'-2}{i} + \sum_{i=0}^{0}\binom{m'-2}{i}$$

dus vgl. (3) geldt voor $n = 2$
Stel (3) geldt voor $n < n'$, met $n' > 2$:

$$\sum_{i=0}^{n'-1}\binom{m'-2}{i} + \sum_{i=0}^{n'-2}\binom{m'-2}{i} = \sum_{0}^{n'-2}\binom{m'-2}{i} + \binom{m'-2}{n'-1} + \sum_{0}^{n'-3}\binom{m'-2}{i} + \binom{m'-2}{n'-2}$$

$$= \sum_{0}^{n'-2}\binom{m'-1}{i} + \left\{\binom{m'-2}{n'-1} + \binom{m'-2}{n'-2}\right\}$$

$$= \sum_{0}^{n'-2}\binom{m'-1}{i} + \left\{\frac{(m'-2)!\,(m'-n'+n'-1)}{(m'-n')!\,(n'-1)!}\right\}$$

$$= \sum_{0}^{n'-2} \binom{m'-1}{i} + \binom{m'-1}{n'-1}$$

$$= \sum_{i=0}^{n'-1} \binom{m'-1}{i}$$

Hiermee is (3) bewezen.

$$(2)(3) \longrightarrow \quad B_n^{m'} = 2\sum_{i=0}^{n-1} \binom{m'-2}{i} + 2\sum_{i=0}^{n-2} \binom{m'-2}{i} = 2\sum_{i=0}^{n-1} \binom{m'-1}{i}$$

Hiermee is (1) bewezen.

Bekijk het systeem van m hypervlakken $F_1, \dots, F_m$, in E^{n+1}, die allen door de oorsprong gaan. Het systeem voldoet aan de algemene opstelling. De ruimte E^{n+1} is nu maximaal verdeeld in:

$$B_{n+1}^{m} = 2\sum_{i=0}^{n} \binom{m-1}{i}$$

subruimten.

Een punt in de ruimte wordt voorgesteld door:

$$(\bar{w}^T, \bar{x}) = (w_0 x_0, w_1 x_1, \dots, w_n x_n)$$

met $x_i = \pm 1$ voor $i \neq 0$

De x_i kan dus voorgesteld worden als de i^e ingang van een drempel-waarde-element (DWE), waarvan de ingangen alleen de waarden ± 1 kunnen aannemen.

Een hypervlak wordt voorgesteld door de vergelijking:

$$\pm w_1 \pm w_2 \pm \dots \pm w_n = w_0$$

Het aantal verschillende hypervlakken is dus 2^n.

Twee punten representeren dezelfde funktie, dan en slechts dan als ze niet door een van deze hypervlakken gescheiden zijn.

De delen in E^{n+1} representeren dus ieder een DWE-funktie.

Een funktie wordt bepaald door een aantal ingangspatronen, die zijn genomen uit de verzameling van 2^n verschillende patronen. Het aantal DWE'n wordt voorgesteld door k.

Het aantal verschillende funkties is dan maximaal:

$$N(n,k) = \binom{2^n}{k} \cdot 2^k$$

k maximaal is: 2^n

Het absolute maximum van het aantal verschillende funkties bij $(n+1)^e$ graads DWE'n, die 2^n patronen hebben is:

$$N(n, 2^n)_{max} = \binom{2^n}{2^n} \cdot 2^{2^n} = 2^{2^n}$$

Het absolute maximum aantal subruimten, waarin 2^n hypervlakken E^{n+1} verdelen is:

$$B_{n+1}^{2^n} = 2\sum_{i=0}^{n}\binom{2^n-1}{i}$$

Bekijk het systeem van één DWE, dus $k=1$, en precies $m=2^n$ inputcombinaties. Er kunnen nu maximaal B_n^m essentieel verschillende funkties gerealiseerd worden.

$$B_{n+1}^{m} = 2\sum_{0}^{n}\binom{m-1}{i} = 2\binom{m-1}{n}\cdot\left\{\frac{n!}{(m-n)\ldots(m-1)} + \frac{n!}{(m-n)\cdot\ldots\cdot(m-2)} + \frac{n!/2!}{(m-n)\cdot\ldots\cdot(m-3)} + \ldots\right.$$

$$\left.\ldots + \frac{n(n-1)}{(m-n)(m-n+1)} + \frac{n}{m-n} + 1\right\}$$

Stel $m > 2n$

$$B_{n+1}^{m} < 2\binom{m-1}{n}\left\{1 + \frac{n}{m-n} + \frac{n^2}{(m-n)^2} + \ldots\right\} =$$

$$= 2\binom{m-1}{n}\frac{1}{1-\frac{n}{m-n}} = 2\binom{m-1}{n}\frac{m-n}{m-2n} \qquad \ldots (4)$$

aangezien $n < m-n \longrightarrow \frac{n}{m-n} < 1$, zodat de meetkundige reeks, tussen haken, convergeert naar $\frac{m-n}{m-2n}$.

Stel $m \geq 3n+2$ en $n \geq 2$

$$\therefore \{m-(n-1)\}\cdot\{m-n\}\cdot\frac{m-n}{m-2n} = \{m-(n-1)\}\cdot\left\{m + \frac{n^2}{m-2n}\right\}$$

$$\leq \{m-(n-1)\}\cdot\left\{m + \frac{n^2}{n+2}\right\} =$$

$$= m^2 - \frac{n^2+n-2-n^2}{n+2}\cdot m - \frac{n^2(n-1)}{n+2}$$

$$< m^2 \qquad \ldots (5)$$

$$(4)(5) \longrightarrow \quad B_{n+1}^{m} < 2\cdot\frac{(m-1)\cdot(m-2)\cdot\ldots\cdot(m-n+2)}{n!}\,m^2$$

$$< 2\frac{m^n}{n!} < m^n$$

<u>Stelling 1</u>

Als er m ingangscombinaties worden beschouwd, dan kunnen er door één DWE maximaal

$$B_{n+1}^{m} < \frac{2m^n}{n!} < m^n$$

essentieel verschillende funkties gerealiseerd worden.

Bekijk de verhouding van aantal drempelfunkties t.o.v. het totaal aantal mogelijke funkties:
Totaal aantal drempelfunkties, door m ingangscombinaties gedefinieerd, is kleiner dan $2m^n/n!$
Totaal aantal mogelijke funkties, gedefinieerd door m vaste punten: 2^m

$$V = \frac{\text{totaal aantal drempelfunkties}}{\text{totaal aantal mogelijke funkties}} = \frac{2m^n}{n!\,2^m}$$

De benadering van Stirling voor n! geeft voor grote n:

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot \left\{1 + \frac{1}{12n} + \dots\right\} \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

$$\therefore V \approx \frac{2m^n \cdot 2^{-m}}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} = \sqrt{\tfrac{2}{\pi}}\left(\frac{me}{n}\right)^n \cdot 2^{-m} \cdot n^{-1/2}$$

$${}^2\log V = {}^2\log\sqrt{\tfrac{2}{\pi}} - \tfrac{1}{2}\,{}^2\log n + n\,{}^2\log\left(\frac{me}{n}\right) - m$$

Substitueer $m = \alpha n$:

$$\therefore {}^2\log V = n\left\{-\alpha + {}^2\log \alpha e - \frac{1}{2}\,\frac{{}^2\log n}{n} + \frac{{}^2\log\sqrt{2/\pi}}{n}\right\}$$

Stel α = constant en $n \to \infty$:

$${}^2\log V \approx -\tfrac{1}{2}\,{}^2\log n + n\left\{{}^2\log \alpha e - \alpha\right\} \approx n\left\{{}^2\log \alpha e - \alpha\right\}$$

$$V = \left[2\,\text{exp.}\left\{-\alpha + {}^2\log(\alpha e)\right\}\right]^n$$

Als $\exp\left\{-\alpha + {}^2\log(\alpha e)\right\} < \tfrac{1}{2}$ d.w.z. $\alpha \geq 3{,}03$, dan nadert V naar nul.

$\alpha \geq 3{,}03 \rightarrow m \geq 3{,}03 \cdot n$ of : $m > 3n$

<u>Stelling 2</u>

Als $m > 3n$, dan kan geen selectie gedaan worden met behulp van één DWE. Voor grote n zullen de meeste funkties, die door meer dan 3n inputcombinaties gedefinieerd zijn, niet te herkennen zijn.

Deze uitspraak geldt in grote trekken, omdat er enige keren benaderd is.

<u>Stelling 3</u>

Als voldaan is aan:

a/ $n \to \infty$

b/ Een set van m inputcombinaties wordt random gekozen uit 2^n mogelijkheden,

c/ en een funktie wordt met deze inputcombinaties random uit de 2^m mogelijkheden gedefinieerd

dan nadert de waarschijnlijkheid dat de resulterende funkties door één enkel DWE kunnen worden gerealiseerd naar:

1 voor $m < 2n$

½ voor $m = 2n$

0 voor $m > 2n$

<u>Bewijs</u>

Aanname: Het aantal drempelfunkties, gedefinieerd met behulp van m inputcombinaties, wordt goed benaderd door de bovengrens B_n^m, als $m < 3n$ en n groot is.

Bewijs aanname:

Stel in E^{n+1} worden de hypervlakken F_1 t/m. F_{2^n} gegeven door

$$a_0 \pm a_1 \pm \dots \pm a_n = 0 \quad , \quad a_i = \pm 1 \text{ voor } i \neq 0$$

De hypervlakken voldoen niet noodzakelijkerwijs aan de voorwaarden voor algemene opstelling, zodat het aantal gebieden in E^{n+1} niet $B_{n+1}^{2^n}$ hoeft te zijn; het aantal kan kleiner zijn

$$R_n \leq B_{n+1}^{2^n} \leq 2 \sum_{i=0}^{n} \binom{2^n - 1}{i}$$

N.B.: $2\binom{2^n}{n}$ kan goed benaderd worden door $2 \sum_0^n \binom{2^n-1}{i}$, zodat een ruwe afschatting van de bovengrens is:

$$R_n < 2^{n^2} \quad (n > 1)$$

Er was reeds gevonden:

$$B^{\alpha n}_{n+1} = \binom{\alpha n-1}{n} 2^{-\alpha n+1} + \binom{\alpha n-1}{n-1} \cdot 2^{-\alpha n+1} + \dots$$

Bedenk: $\binom{\alpha n-1}{n} = \dfrac{(\alpha n-1)!}{n!\,(\alpha n-n-1)!} = \dfrac{\alpha n-n}{\alpha n} \cdot \dfrac{(\alpha n)!}{(\alpha n-n)!\,n!} = \dfrac{\alpha-1}{\alpha} \binom{\alpha n}{n}$

Bekijk het gedrag van de term

$$\binom{\alpha n}{n} \cdot 2^{-\alpha n} = \frac{\alpha n!\, 2^{-\alpha n}}{\{(\alpha-1)n\}!\, n!}$$

pas benadering van Stirling voor $n!$ toe:

$$\approx \frac{\sqrt{2\pi\alpha n}\, \left(\frac{\alpha n}{e}\right)^{\alpha n} 2^{-\alpha n}}{\sqrt{2\pi(\alpha-1)n}\, \left(\frac{(\alpha-1)n}{e}\right)\left(\frac{(\alpha-1)n}{e}\right)^{(\alpha-1)n} \cdot \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n}$$

$$= \sqrt{\frac{\alpha}{\alpha-1}}\, \frac{1}{\sqrt{2\pi n}} \cdot \left\{ \frac{(\alpha/2)^\alpha}{(\alpha-1)^{\alpha-1}} \right\}^n$$

Neem ik de limiet voor $n \to \infty$, dan zal de exponentiële vorm domineren.
Bekijk daarom:

$$f(\alpha) = \left(\tfrac{\alpha}{2}\right)^\alpha \cdot (\alpha-1)^{1-\alpha}$$

$$x^x = e^{x \ln x}$$

$$\frac{d}{dx} x^x = e^{x\ln x} \cdot \frac{d}{dx}(x \ln x) = x^x \cdot (\ln x + 1)$$

$$\frac{df}{d\alpha} = \frac{d}{d\alpha}\left[e^{\ln\{(\frac{\alpha}{2})^\alpha (\alpha-1)^{1-\alpha}\}} \right] = \frac{d}{d\alpha}\left[e^{\{\alpha \ln(\frac{\alpha/2}{\alpha-1}) + \ln(\alpha-1)\}} \right]$$

bedenk: $\alpha \geq 1$

$$= \frac{d}{d\alpha}\left[e^{\alpha \ln \frac{\alpha/2}{\alpha-1}} \right] \cdot e^{\ln(\alpha-1)} + \left(\frac{\alpha/2}{\alpha-1}\right)^\alpha \cdot \frac{d}{d\alpha}\left[e^{\ln(\alpha-1)} \right]$$

$$= (\alpha-1)\left[\left(\frac{\alpha/2}{\alpha-1}\right)^\alpha \cdot \frac{d}{d\alpha}\left\{ \alpha \ln \frac{\alpha/2}{\alpha-1} \right\} \right] + \left(\frac{\alpha/2}{\alpha-1}\right)^\alpha \cdot 1$$

$$= \left(\frac{\alpha/2}{\alpha-1}\right)^\alpha \cdot \left[1 + (\alpha-1) \cdot \left\{ \ln \frac{\alpha/2}{\alpha-1} + \alpha \cdot \frac{d}{d\alpha}\left(\ln \frac{\alpha/2}{\alpha-1}\right) \right\} \right]$$

$\dfrac{df}{d\alpha} = 0$ voor: $\quad 1 + (\alpha-1)\left\{ \ln \dfrac{\alpha/2}{\alpha-1} + \alpha \cdot \dfrac{-1}{\alpha(\alpha-1)} \right\} = 0$

$$\therefore \ln \frac{\alpha/2}{\alpha-1} = \frac{-1}{\alpha-1} + \frac{1}{\alpha-1} = 0 \longrightarrow \frac{\alpha/2}{\alpha-1} = 1 \longrightarrow \underline{\alpha = 2}$$

$f(\alpha)$ is maximaal voor $\alpha = 2$ (in het gebied $\alpha \geq 1$)
$f(2) = 1$
d.w.z.: voor $\alpha \geq 1$ en $\alpha \neq 2$ geldt: $f(\alpha) < 1$

Dus $\binom{\alpha n}{\alpha}.2^{-\alpha n}$ nadert exponentieel naar nul als $n \to \infty$.
In het geval dat $\alpha = 2$ gaat deze funktie als $\frac{1}{\sqrt{n}}$ naar nul.
Deze limieten gelden, als α constant is, ook voor

$$\binom{\alpha n-1}{\alpha}.2^{-\alpha n}$$

omdat

$$\binom{\alpha n-1}{\alpha}.2^{-\alpha n} = \frac{\alpha-1}{\alpha}\binom{\alpha n}{n}.2^{-\alpha n}$$

Terugkomend op de uitdrukking

$$\frac{B_{n+1}^{\alpha n}}{2^{\alpha n}} = 2^{1-\alpha n}.\left\{\binom{\alpha n-1}{0}+\binom{\alpha n-1}{1}+\dots+\binom{\alpha n-1}{n}\right\}$$

is voor $\alpha > 2$ de term $\binom{\alpha n-1}{n}.2^{1-\alpha n}$ de grootste term in deze uit-drukking, zodat

$$\frac{B_{n+1}^{\alpha n}}{2^{\alpha n}} < 2^{1-\alpha n}.n.\binom{\alpha n-1}{n}$$

De term $\binom{\alpha n-1}{n}.2^{1-\alpha n}$ gaat exponentieel naar nul, dus $\frac{B_{n+1}^{\alpha n}}{2^{\alpha n}}$ gaat exponentieel naar nul.

Stel $\alpha < 2$. Maak gebruik van $\binom{m}{p} = \binom{m}{m-p}$:

$$\frac{B_{n+1}^{\alpha n}}{2^{\alpha n}} = 2^{1-\alpha n}.\left\{\binom{\alpha n-1}{0}+\dots+\binom{\alpha n-1}{n}\right\}$$

$$= 2^{1-\alpha n}.\left\{\binom{\alpha n-1}{\alpha n-1}+\dots+\binom{\alpha n-1}{\alpha n-n-1}\right\}$$

bedenk: $2^x = (1+1)^x = \sum_{k=0}^{x}\binom{x}{k}1^k.1^{x-k}$

$$= 2^{1-\alpha n}\left[2^{\alpha n-1}-\left\{\binom{\alpha n-1}{0}+\dots+\binom{\alpha n-1}{\alpha n-n-2}\right\}\right]$$

$$= 1-2^{1-\alpha n}\left[\binom{\alpha n-1}{0}+\dots+\binom{\alpha n-1}{\alpha n-n-2}\right]$$

$$= 1-2^{-\alpha n}.B_{(\alpha-1)n}^{\alpha n}+\left[\binom{\alpha n-1}{\alpha n-n-1}+\binom{\alpha n-1}{\alpha n-n}\right].2^{-\alpha n}$$

In geval $n \to \infty$:

de laatste twee termen naderen naar nul
$\frac{\alpha}{\alpha-1} > 2$ $\Big\} \longrightarrow$

$\therefore$ 2e term nadert naar nul

$$\therefore \frac{B_n^{\alpha n}}{2^{\alpha n}} \longrightarrow 1$$

Stel $\alpha = 2$:
Bekijk V voor $n \to \infty$:

$$\frac{B^{\alpha n}_{n+1}}{2^{\alpha n}} = \frac{B^{2n}_{n+1}}{2^{2n}} = 2^{1-2n}\left[\binom{2n-1}{0}+\ldots+\binom{2n-1}{n-1}+\binom{2n-1}{n}\right]$$

$$= 2^{-2n}\left[\binom{2n-1}{0}+\ldots+\binom{2n-1}{n-1}+\ldots+\binom{2n-1}{2n-1}\right]+2^{1-2n}\cdot\binom{2n-1}{n}$$

$$= \frac{1}{2}+2^{1-2n}\binom{2n-1}{n}$$

Daar de tweede term naar nul nadert geldt:

$$\frac{B^{\alpha n}_{n+1}}{2^{\alpha n}} \longrightarrow 1/2 \qquad \text{q.e.d.}$$

Zie ook figuur 11.

9 Literatuur

Het eerste nummer geeft aan op welk hoofdstuk het boek of het artikel slaat.

2.1 „State of Art in Pattern Recognition"
G. Nagy
Proc. IEEE 56, № 5 pp. 836-862 (1968)

2.2 „Beschouwingen over Methoden der patroonherkenning."
A.C.J. de Leeuw.
Afstudeerverslag; groep Meten en Regelen; januari 1968; T.H.E.

2.3 „Het automatisch classificeren van handgeschreven cijfers."
Afstudeerverslag T.H.D.; laboratorium voor schakeltechniek en techniek der informatieverwerkende machines.
2 juni 1966

2.4. „Automat und Mensch"
K. Steinbuch.

2.5. IEEE Trans. vol. EC-12; december 1963; pp. 846-862, 822-835
Bekijk o.a. ook de index van dit tijdschrift.

2.6. „Legibility of Alphanumeric Characters and Other Symbols". A Reference Handbook.
Proeven om te komen tot een acceptabel alphanumerisch systeem van symbolen.

2.7. „An Autonomous Reading Machine"
Casey; Nagy.

2.8 Self Organizing Systems
M.C. Yovits, G.T. Jacobi, G.D. Goldstein. 1962.

2.9. Pattern Recognition
The Journal of the Pattern Recognition Society.
Pergamon Press.

3.1. „ A Critical Comparison of Two Kinds of Adaptive Classification Networks"
K. Steinbuch; B. Widrow.
IEEE Trans. on Elec. Computer. pp. 737-740. oktober 1965

3.2. „ An Estimate of the Complexity Requisite in a Universal Decision Network".
S.H. Cameron.
Bionics Symp. Dayton, Ohio; pp. 197-212. december 1960

3.3. „ Single stage threshold logic"
R.O. Winder.
Switching Circuit Theory and Logical Design.
AIEE Special Publications. S-134; september 1961.

3.4. „ Threshold Logic"
R.O. Winder.
Ph. D. dissertation, Princeton University; may 1962.

3.5. „ Logical elements on majority decision principle and complexity of their circuits."
S. Muroga.
Proc. International on Information Processing; Paris, june 1959; blz: 400-407.

3.6. „ Reliable, trainable networks for computing and control."
B. Widrow and J.B. Angell.
Aerospace Eng., vol. 21, pp. 78-123; sept. 1962

3.7. „Steps towards artificial intelligence".
M. Minsky
Proc. IRE vol. 49 pp. 8-30 ; januari 1960

3.8 „ Die Lernmatrix"
K. Steinbuch
Kybernetik. Band 1, Heft 1, januari 1961

3.9. „Lernende Automaten"
Bericht über die Fachtagung in Karlsruhe am 13, 14 april 1961.

3.10 „ Adaptive multiple-output threshold systems and their storage capacities."
R. J. Brown.
Stanford Electronics Labs., Stanford, Calif.
Rept. SEL-64-018 juni 1964.

4.1. „ Theory and applications of Holography"
J. B. De Velis, G. O. Reynolds.
Addison-Wesley Publishing Company.

4.2. „ L'Onde électrique"
vol. 48 no. 492 mars 1968

4.3. „Processing Information with light"
D. Redman.
Science journal february 1968

4.4. „Optical and Electro-optical Information Processing.
M. I. T. Press.

4.5. „ Analysis and synthesis of optical images"
J. E. Rhodes Jr.
Am. J. of Phys. vol. 21 pp. 337-343 mei 1953

4.6. „Spatial Filtring in Optics"
O'Neill. E. L.
I. R. E. Trans. Inform. Theory IT-2, 56-65 june 1956

4.7. „Optical Data Processing and Filtering Systems"
L. J. Cutrona.
I. R. E. Trans. Inform. Theory IT-6, 386-400 june 1960.

5.1 „An Experimental Investigation of a Mixed-Font Print Recognition System."
Liu, Shelton.
IEEE Trans. on Elec. Computers vol. EC-15 no. 6, december 1966.

5.2. „An Experimental Study of Machine Recognition of Hand-Printed Numerals."
R. Bakis, Noel M. Herbst, G. Nagy
IEEE Trans. on Systems Science and Cybernetics.
vol. SSC-4, no. 2, july 1968.

5.3. „The I.B.M. optical Page Reader"
I.B.M. Journal of Research and Development.

5.4. „Experiments in the recognition of hand-printed text."
1 „Character Recognition"
2 Context Analysis."
Afips. Conference Proceedings. Vol. 33 part 2, 1968

5.5 „Pattern Recognition and Reading by Machine"
Bledsoe, Browning.
Proc. of The Eastern Joint Computer Conference. 1959, pp. 225-232

6.1. „Recognition of Handwritten Numerical Characters for Automatic Letter Sorting"
Hiroshi Genchi, Ken-Ichi Mori, Sadakazu Watanabe.
Sumio Katsuragi.
Proc. of the IEEE. vol. 56 no. 8, august 1968.

6.2. „A System for the Automatic Recognition of Patterns"
Grimsdale, Sumner, Tunis, Kilburn.
Proceedings of the Institute of Electrical Engineers,
vol. 106. part B no. 26 , maart 1959.