

# Anwendung der optischen Filtriertechnik

von H. J. Tiziani, B. H. Beyeler und W. Witz

## Zusammenfassung

In die verschiedenen optischen Filtriermethoden wird eine kurze Einführung gegeben und auf mögliche Anwendungen im Pattern-Recognition hingewiesen. Ein konkretes Problem, nämlich die Formerkennung von Ag-Halogenid-Körnern einer photographischen Emulsion wird behandelt. Da die zu unterscheidenden Formen durch ihr Verhalten unter der zweidimensionalen Drehgruppe charakterisiert sind, wurde eine spezielle Filtriermethode entwickelt, die den gestellten Forderungen nach Lage- und Drehinvarianz in natürlicher Weise Rechnung trägt. Abschliessend werden einige experimentelle Resultate vorgestellt.

## Résumé

Tout d'abord, quelques méthodes de filtrage optique sont résumées en vue d'application à un problème particulier de «Pattern recognition»: la reconnaissance de la forme des grains d'une émulsion photographique. Puisque ces formes sont caractérisées par leur comportement dans le groupe des rotations, une méthode particulière a été développée. Elle satisfait d'une manière directe les invariances positionnelles et rotationnelles exigées. Finalement, quelques résultats pratiques sont présentés.

## Abstract

A brief introduction to the concept of optical filtering is given and possible applications in pattern recognition are pointed out. In particular, the classification of the shape of grains in photographic emulsions is discussed. Fortunately, the shapes to be distinguished differ only in their two-dimensional rotational symmetry group, so that a special optical filtering procedure was developed which takes account of the required conditions of positional and rotational invariance. Finally, some experimental results are presented.

## 1. Einleitung

Seit der Entdeckung des Lasers in den späten fünfziger Jahren steht der modernen Optik eine Quelle weitgehend monochromatischen Lichtes hervorragender räumlicher und zeitlicher Kohärenz zur Verfügung. Es ist deshalb nicht verwunderlich, dass in der letzten Zeit viele neue Teilgebiete der Forschung erschlossen worden sind. Eines davon soll hier näher beleuchtet werden, nämlich die Anwendung der optischen Filtrierung zur Informationsverarbeitung, speziell in Pattern Recognition.

Grundlegend für die optische Filtrierung ist die Existenz einer Fouriertransformation zwischen den komplexen Lichtamplituden  $A(\xi, \eta)$  in der vorderen ( $E_1$ ) bzw.  $a(x, y)$  in der hinteren ( $E_2$ ) Brennebene eines optischen Systems  $L$ . Physikalisch gesehen bildet  $L$  das (sonst im Unendlichen lokalisierte) Fraunhofersche Spektrum des Objektes  $A(\xi, \eta)$  in die Ebene  $E_2$  ab, wie Fig. 1 veranschaulicht.

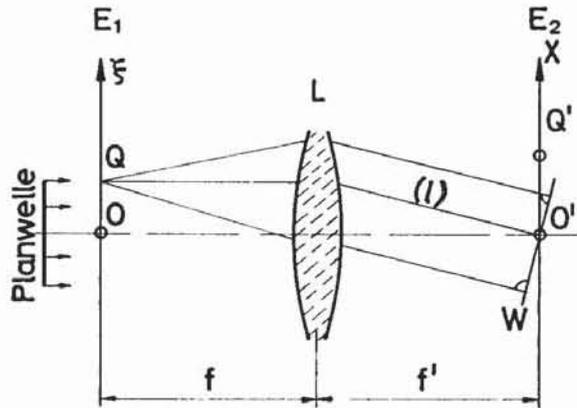


Fig. 1 Fraunhofersche Beugung.

Eine einfallende Planwelle beleuchtet die Objektebene  $E_1$ , deren Flächenelemente nach dem Prinzip von Huygens-Fresnel [1] zum Ausgangspunkt sekundärer kohärenter Wellen werden. Fällt  $E_1$  mit der objektseitigen Brennebene des optischen Systems  $L$  zusammen, so breiten sich die von Objektpunkten  $Q$  ausgehenden Kugelwellen hinter  $L$  als Planwellen aus. Der optische Weg von  $Q$  bis zur Wellenfront  $W$  ist auf allen Strahlen konstant. Ferner hängt bei einem idealen optischen System der optische Weg  $[QO]$  nicht von  $Q$  ab ([2], Abschnitte 2 und 4). Die komplexe Amplitude in  $O'$  der Wellenfront  $W$  eines Flächenelementes  $d\xi d\eta$  in  $Q(\xi, \eta)$  mit der komplexen Amplitude  $A(\xi, \eta)$  ist

$$A(\xi, \eta) \exp(-ik[QO]) d\xi d\eta$$

wobei  $\lambda$  die Wellenlänge des verwendeten Lichtes und  $k$  die Wellenzahl ( $2\pi/\lambda$ ) ist. Im allgemeinen Punkt  $Q'$  entsteht  $a(X, Y)$  durch phasengerechte Überlagerung der Wellen, die von allen Punkten der Objektebene ausgehen.

$$a(X, Y) = C \iint A(\xi, \eta) \exp\left[\frac{-i2\pi}{\lambda}(lX + mY)\right] d\xi d\eta$$

$C$  enthält die konstante Phase  $\exp(-ik[QO])$ ,  $l$  und  $m$  sind die Richtungskosinus der Brennstrahlen, welche die Verkippung der Wellenfronten charakterisieren. Durch Substitution von  $l, m$  und Normierung

$$x = X/(\lambda f), y = Y/(\lambda f) \text{ der Koordinaten erhält man}$$

$$a(x, y) = C \iint A(\xi, \eta) \exp[-i2\pi(\xi x + \eta y)] d\xi d\eta \quad (1)$$

Diese Gleichung beschreibt eine zweidimensionale Fouriertransformation, in welcher das Objekt  $A(\xi, \eta)$  in sein Raumfrequenzspektrum  $a(x, y)$  zerlegt wird. Lässt man  $E_2$  mit der vorderen Brennebene einer zweiten Linse  $L_2$  zusammenfallen, so erhält man in analoger Weise in deren hinteren Brennebene  $E_3$  die Fouriertransformierte von  $a(x, y)$ ; diese neue Funktion  $A'(\xi', \eta')$  stellt aber nichts anderes als das Bild des Objekts  $A(\xi, \eta)$  dar. Mathematisch (die F-Transformation ist bis auf die Orientierung der Koordinaten ihre eigene Inverse) kann das leicht eingesehen werden.

## 2. Optische Filtriermethoden

Die Auswertung des Raumfrequenzspektrums durch optische Filtrierung ist in den letzten Jahren auf grosses Interesse gestossen und hat zu einer Vielzahl neuer Erkenntnisse geführt. Dieses Gebiet soll hier nur gestreift werden; Näheres erfährt der Leser in [3].

### 2.1. Kohärente Filtrierung

Eine ebene Welle beleuchtet die Objektebene  $E_1$ , wo sich das durch die komplexe Transparenz  $A(\xi, \eta)$  gegebene Original befindet. In der hinteren Brennebene  $E_2$  der Linse  $L_1$  entsteht (wie im 1. Abschnitt gezeigt) das zugehörige Raumfrequenzspektrum  $a(x, y)$ , welches, sofern sich in  $E_2$  kein modulierendes Medium (Filter) befindet, durch  $L_2$  einer zweiten Fouriertransformation unterzogen wird, so dass in  $E_3$  eine Abbildung des Originals  $A(\xi, \eta)$  erscheint. Wird nun in  $E_2$  ein Filter der komplexen Transparenz  $t(x, y)$  angebracht, so geht das Frequenzspektrum über in

$$f(x, y) = t(x, y) \cdot a(x, y) \quad (2)$$

In der Bildebene  $E_3$  wird dann die Fouriertransformierte von  $f(x, y)$  beobachtet. Nach einem bekannten Satz aus der Theorie des Fourierintegrals entspricht einer Multiplikation der Transformierten die Faltung der Originale:

$$F(\xi', \eta') = \iint A(\xi, \eta) T(-\xi' - \xi, -\eta' - \eta) d\xi d\eta \quad (3)$$

$T(\xi, \eta)$  bedeutet dabei die inverse Fouriertransformierte der Filterfunktion  $t(x, y)$ :

$$T(\xi, \eta) = \iint t(x, y) \exp[+i2\pi(\xi x + \eta y)] dx dy \quad (4)$$

Definiert man  $\tilde{T}(\xi, \eta) = T^*(-\xi, -\eta)$ , so kann das Faltungsintegral (3) auch als die Kreuzkorrelationsfunktion der Funktionen  $A$  und  $\tilde{T}$  interpretiert werden.

Allgemein stellt die Kreuzkorrelation ein Mass für die Übereinstimmung zweier Funktionen dar. Dies hat für die kohärente Filtrierung folgende anschauliche Bedeutung: Befindet sich in der Objektebene  $E_1$  eine Anzahl verschiedener Figuren, so leuchten an den korrespondierenden Stellen der Bildebene  $E_3$  Lichtpunkte auf, deren Intensität ein Mass für die Übereinstimmung der betreffenden Figur mit dem in  $E_2$  angebrachten Filter ist. Unter Anwendung verschiedener mathematischer Sätze kann gezeigt werden, dass die Intensität dann maximal wird, wenn die Filterfunktion proportional dem konjugiert komplexen Objektspektrum ist («Matched Filter»).

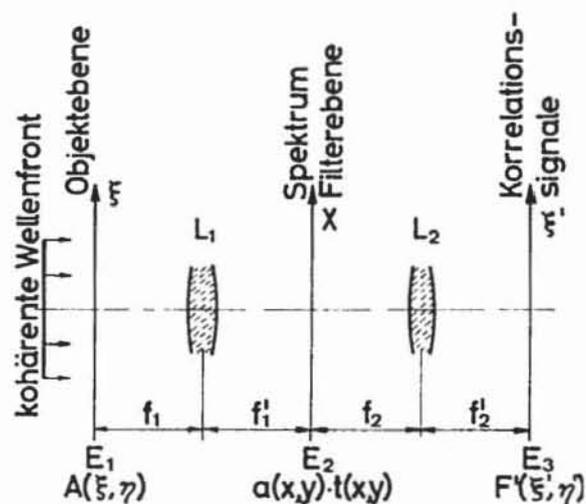


Fig. 2 Anordnung zur kohärenten optischen Filtrierung.

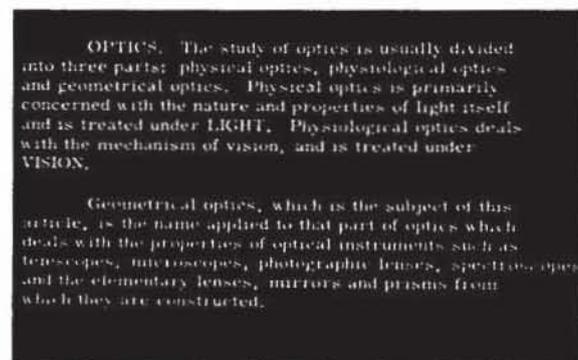


Fig. 3a Textvorlage zur kohärenten Filtrierung.



Fig. 3b Erkennung des Wortes «optics». Nach A. Vander Lugt mit dessen freundlicher Erlaubnis.

Der in Fig. 3a wiedergegebene Text wurde mittels kohärenter Filtrierung auf das Vorkommen des Wortes «optics» untersucht. Als Filter diente das konjugierte komplexe Spektrum dieses Wortes. Die hellen Punkte an den entsprechenden Stellen der Bildebene (Fig. 3b) sind deutlich erkennbar.

Die Nützlichkeit der kohärenten Filtrierung zur Lösung gewisser Pattern-Recognition-Probleme wird hier evident: Zur Erkennung eines bestimmten Patterns sind bloss Intensitätsmessungen erforderlich. Andere Anwendungen der kohärenten Filtrertechnik, zum Beispiel Unterdrückung von Rauschen, findet der Leser in den Referenzen [3], [4], wo auch auf weitere Literatur verwiesen wird.

### 2.2. Inkohärente Filtrierung

Sie stellt eine weitere Möglichkeit zur optischen Lösung gewisser Probleme des Pattern Recognition dar. Einzelheiten über die verschiedenen Methoden findet der interessierte Leser in den Literaturzitierten von Referenz [5].

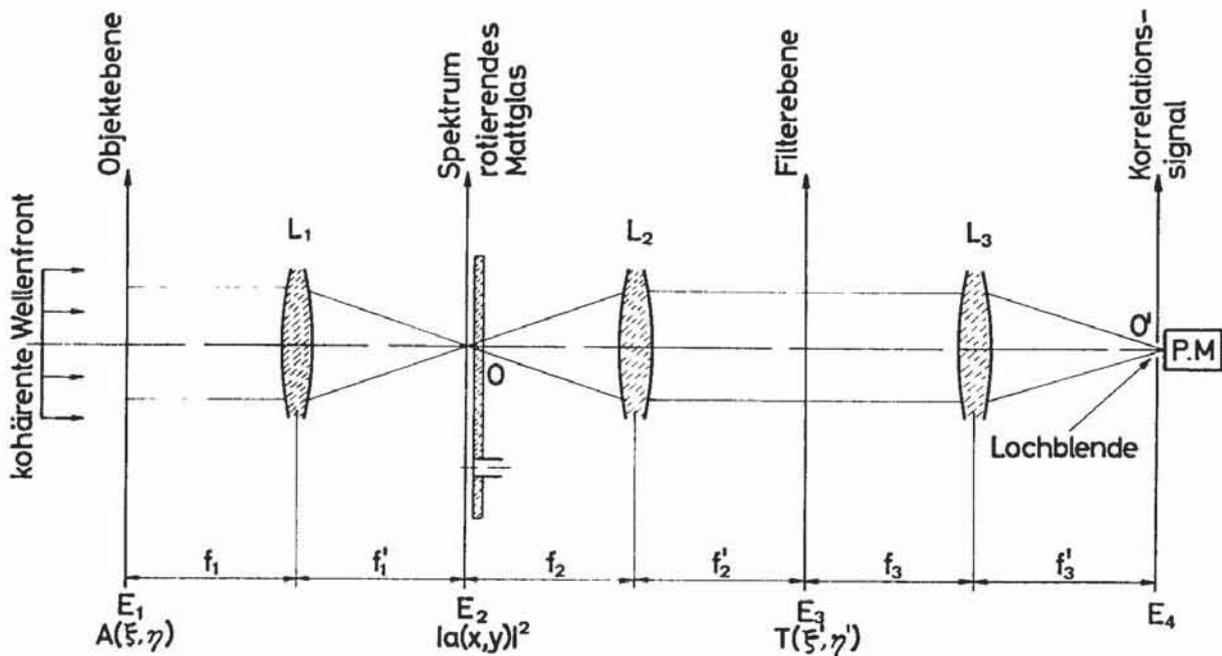


Fig. 4 Anordnung zur inkohärenten optischen Filtrierung.

Armitage und Lohmann haben die räumliche Kohärenz in der Spektrumsebene  $E_2$  (Fig. 4) künstlich zerstört, um auf diese Weise ein von der Lage des Objektes unabhängiges Intensitätsspektrum zu erhalten. Erst anschliessend erfolgt hier der eigentliche Filtrierungsprozess, bei welchem (analog wie schon in 2.1. beschrieben) die Korrelation zwischen Filter und Objekt gebildet wird. Im Gegensatz zur kohärenten Filtrierung erscheint jedoch der zu messende Lichtpunkt stets im Zentrum  $O'$  der Ebene  $E_4$ , so dass sich ein Abtasten dieser Ebene erübrigt.



Fig. 5 Beispiel einer Kornverteilung.

### 3. Erkennung einfacher geometrischer Figuren durch Abtastung ihrer Raumfrequenzspektren

Im Zusammenhang mit der Beurteilung und Kontrolle photographischer Schichten stellt sich das folgende Problem: Es soll ein Apparat gebaut werden, welcher die in einer gegebenen Vorlage (typisches Beispiel Fig. 5) enthaltenen geometrischen Figuren analysiert und in vier Klassen einteilt:

1. rhombische Figuren ( $60^\circ$ -Winkel),
2. dreieckige Figuren (gleichseitig),
3. quadratische Figuren,
4. sechseckige Figuren (regulär).

Die zu analysierenden Figuren weisen im allgemeinen mehr oder weniger abgerundete Ecken auf. Ihre Linearausdehnungen umfassen einen Bereich von ca.  $(0,2 \dots 1) \cdot$  Maximalausdehnung. Die einzelnen Figuren liegen meist ziemlich locker; es treten jedoch Berührungen oder Überlagerungen von zwei bis drei Einzelobjekten auf (Cluster).

Hier liegt also ein typisches Pattern-Recognition-Problem vor. Seine Lösung wurde zunächst unter der stark vereinfachten Annahme studiert, dass die gegebenen Vorlagen nach Fig. 5 schon in einzelne sogenannte Elementarpatterns zerlegt seien. Unter einem Elementarpattern wird eine zusammenhängende Figur (Einzelobjekt oder nicht auflösbares Cluster) verstanden, die an beliebiger Stelle und in beliebiger Lage innerhalb eines festen Gesichtsfeldes liegt, ohne jedoch seinen Rand zu berühren. Die Abmessung dieses Gesichtsfeldes ist durch die maximale Linearausdehnung der Figuren bestimmt.

Da die Klassifizierung der Form innerhalb des angegebenen Bereiches unabhängig von der Objektgrösse erfolgen soll, erscheint hier die Anwendung der optischen Filtrierung problematisch, weil ja pro Objektklasse stets der Reihe nach ein ganzer Satz von Filtern (der Objektgrösse entsprechend) verwendet werden müsste (oder dann wäre die Grössenvariation durch ein servogesteuertes Zoom-System oder durch axiale Verschiebung des Filters zu kompensieren [6], [7]). Um den Verdrehungen der Objekte Rechnung zu tragen, wäre zudem noch eine Rotation der Filter erforderlich. Es wurde deshalb nach einer Anordnung gesucht, die die Invarianzforderungen (bezüglich Translation, Verdrehung und Grösse des Objekts) in natürlicher Weise berücksichtigt. Bekanntlich ist das Intensitätsspektrum invariant gegenüber Translation des Objekts<sup>1)</sup>. Weiter erkennt man beim Betrachten der Raumfrequenzspektren der vier Grundfiguren (Fig. 6), dass die ganze Information über die Form des Objekts in der azimuthalen Verteilung der Energie im Spektrum enthalten ist.

Die dreizählige Symmetrie des Dreiecksspektrums lässt sich nur auf Grund der Filmeigenschaften er-

<sup>1)</sup> Falls  $A(\xi, \eta) \rightarrow a(x, y) = \iint A(\xi, \eta) \exp[-i 2\pi(\xi x + \eta y)] d\xi d\eta$ , dann gilt  $A'(\xi, \eta) = A(\xi - \xi_0, \eta - \eta_0) \rightarrow a'(x, y) = \exp[-i 2\pi(\xi_0 x + \eta_0 y)] \cdot a(x, y)$ . Daraus folgt  $|a'(x, y)|^2 = |a(x, y)|^2$ , d.h. Translation des Objektes beeinflusst das Intensitätsspektrum nicht.

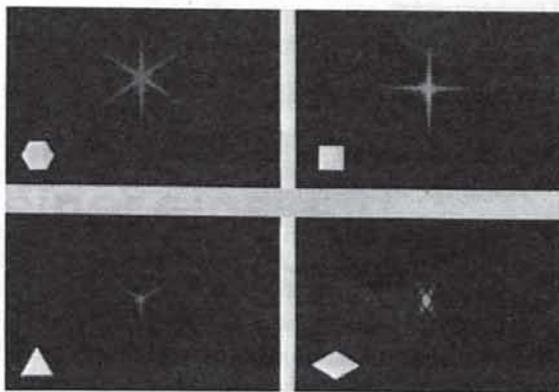


Fig. 6 Die vier Grundfiguren und ihre Raumfrequenzspektren.

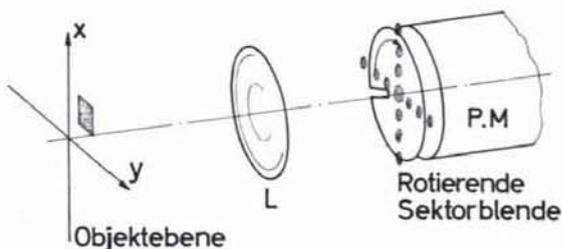


Fig. 7 Abtastung der Raumfrequenzspektren durch rotierende Blende.

klären; ein dreieckiges Loch würde nämlich ein Spektrum mit sechszähliger Symmetrie ergeben. Näheres hierzu findet sich in [8] und [9]. Dieser Effekt wird auch Gegenstand einer weiteren Publikation sein.

Während also die Form des Objekts massgebend für die Drehsymmetrie des Spektrums ist, bewirkt eine Grössenänderung lediglich eine reziproke Streckung des Spektrums; sie beeinflusst also die Winkelverteilung  $E(\varphi)$  der Energie nicht.

Fig. 7 zeigt schematisch die zur Gewinnung der Funktion  $E(\varphi)$  verwendete Apparatur. Das Raumfrequenzspektrum wird auf eine Blende mit sektorförmigem Ausschnitt projiziert (Sektorenzentrum im Ursprung der Spektrumsebene, Zentriwinkel ca.  $15^\circ$ ).

Unmittelbar hinter der Blende befindet sich die Kathode eines Photomultipliers, dessen Ausgangssignal  $U$  somit proportional zu der auf den Sektor entfallenden Energie ist. Zur Verbesserung des Signal-Rausch-Verhältnisses wurde die (energiemässig dominante) nullte Spektrumsordnung abgedeckt. Sie kann separat gemessen und zum Beispiel zur Bestimmung der Objektgrösse benützt werden. Lässt man nun den Sektor mit konstanter Winkelgeschwindigkeit  $\omega$  um sein Zentrum rotieren, so erhält man am Ausgang des PM eine periodische Zeitfunktion  $U(t)$ , welche der gesuchten Winkelverteilung der Spektrumsenergie entspricht. Jede der vier Figurenklassen ergibt eine charakteristische Signalform (Fig. 8). Es kann ferner gezeigt werden, dass die Signalamplitude ungefähr proportional zur Objektgrösse ist.

Die vier Klassen unterscheiden sich schon in der Grundfrequenz der Zeitfunktion  $U(t)$ : Diese ist nämlich im Falle des Rhombus  $2\omega$ , des Dreiecks  $3\omega$ , des Quadrates  $4\omega$  und des Sechsecks  $6\omega$ .

Im übrigen enthält das Frequenzspektrum von  $U(t)$  (infolge der Periodizität) nur ganzzahlige Vielfache der jeweiligen Grundfrequenz. Die entsprechenden Fourierkoeffizienten  $a_n$  können direkt und simultan mit Hilfe eines Satzes fest abgestimmter Frequenzfilter aus dem Signal  $U(t)$  gewonnen werden.

Die Koeffizienten  $a_n$  sind nun invariant gegen Verdrehung des Objektes; denn Verdrehung ist einer Verschiebung des Zeit-Anfangspunktes äquivalent, die natürlich ohne Einfluss auf das Frequenzspektrum bleibt. Es hat sich gezeigt, dass die Koeffizienten mit  $n > 12$  keine wesentliche Information mehr tragen. Auf Grund der Symmetrien der untersuchten Objekte verschwinden ferner  $a_1, a_5, a_7, a_{11}$  prinzipiell. Zur Auswertung benützte man deshalb nur die 8 Koeffizienten  $a_2, a_3, a_4, a_6, a_8, a_9, a_{10}, a_{12}$ . Jedem vorgelegten Pattern wurde somit ein achtkomponentiger Vektor zugeordnet, dessen Länge die Grösse und dessen Richtung die Form des Objektes angibt. Der nächste Schritt besteht nun darin, den gemessenen Vektor als Linearkombination von (bekannten) Vektoren, die den Grundfiguren ent-

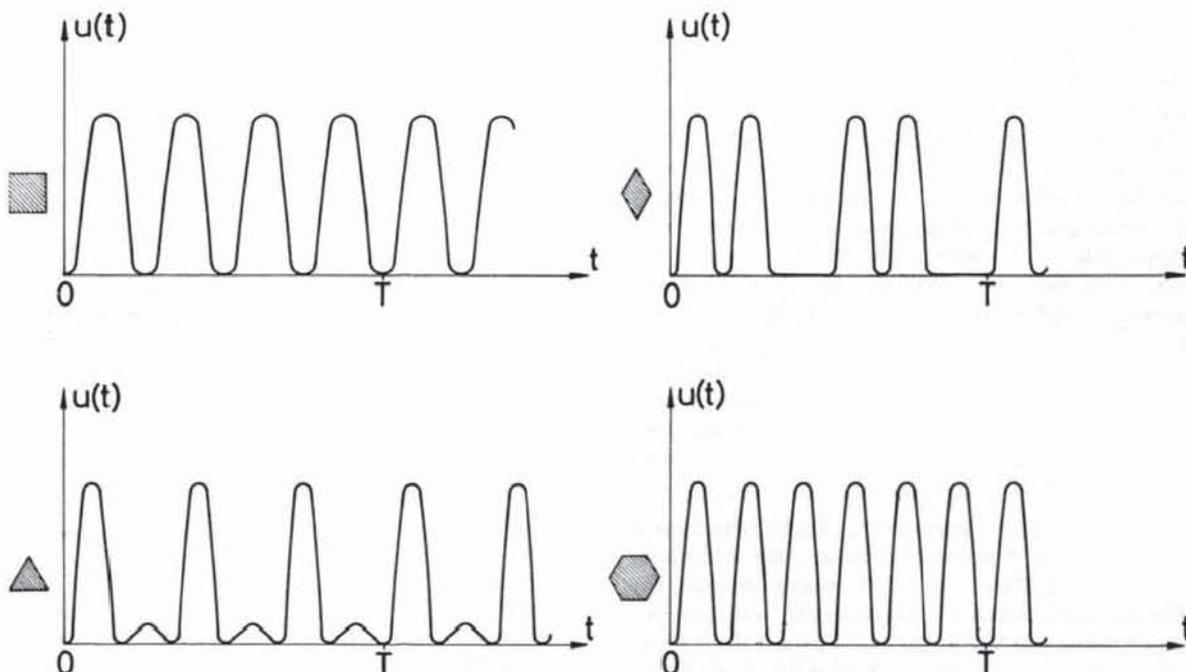


Fig. 8 Signalformen der vier Grundfiguren (schematisch).

sprechen, darzustellen. Die dabei auftretenden Koeffizienten  $p_k$  geben dann ein Mass für die Übereinstimmung des vorgelegten Elementarpatterns mit der Figurenklasse  $k$ . Die  $p_k$  gehen durch eine lineare Transformation aus den gemessenen  $a_n$  hervor:

$$p_k = \sum r_{nk} a_n \quad n = 2, 3, 4, 6, 8, 9, 10, 12$$

Die Elemente  $r_{nk}$  der Transformationsmatrix wurden in der vorliegenden Arbeit empirisch bestimmt (Vorgabe der Momente für einen Testsatz aus 37 bekannten Objekten, Ausgleichsrechnung). Da die erwähnten Frequenzfilter nicht zur Verfügung standen, wurde mittels eines Spectrum-Analyzers das Frequenzspektrum von  $U(t)$  gebildet und ausgemessen. Nach der geschilderten Methode wurden dann 42 unbekannte Einzelobjekte untersucht; davon wurden 41 richtig erkannt. Es sei hier nochmals betont, dass das Verfahren naturgemäss völlig unempfindlich gegen Translation und Rotation des Objektes sowie weitgehend unempfindlich bezüglich Grösse und Eckenabrundung arbeitet.

Bedeutend unübersichtlicher werden die Verhältnisse im Falle einfacher Cluster, das heisst Überlagerungen von zwei (eventuell drei) Einzelobjekten. Dann treten nämlich verschiedene Nichtlinearitäten auf, welche die soeben beschriebene lineare Auswertungsmethode problematisch machen (Nichtlinearität in der Objektebene bei überlappenden Figuren, nicht-lineare Bildung der Fourierkoeffizienten  $a_n$ , Filmeinflüsse usw.). Wie man beim eingehenden Studium des Problems erkennt, wirken sich diese Nichtlinearitäten in vielen praktischen Fällen nicht allzu störend aus. Es wurde deshalb auch eine Serie von 64 einfachen Clustern nach der bei Einzelobjekten bewährten Methode analysiert; davon wurden 41 richtig und 11 teilweise richtig erkannt.

#### 4. Schlussfolgerungen

Die vorliegende Arbeit gibt die Grundlagen zu einem Pattern-Recognition-System, welches sich ganz spezifisch zur Erkennung einfacher geometrischer Figuren (Kristallform-Analyse) eignet. Sämtliche gewünschten Invarianzeigenschaften ergeben sich hier auf natürliche Weise, im Gegensatz zu den in Abschnitt 2 diskutierten Methoden der optischen Filtrierung. Eine Verallgemeinerung des Systems (z. B. im Hinblick auf die Erkennung von Schriftzeichen) ist hingegen nicht denkbar, da sich Buchstaben nicht in eindeutiger Weise durch ihr Verhalten unter der zweidimensionalen Drehgruppe unterscheiden. Zur Lösung solcher Probleme wird man auf die optischen Filtrierverfahren gemäss Abschnitt 2 zurückgreifen. Möglichkeiten, die bestehende Apparatur zu verbessern, wurden untersucht. Es wäre zum Beispiel vorteilhaft, das Raumfrequenzspektrum auf elektro-optischem Wege abzutasten (mit einer Bildaufnahmeröhre oder dergleichen): Dadurch könnte die Umlaufgeschwindigkeit beträchtlich gesteigert werden. Weitere Verbesserungen betreffen die Stabilität des mechanischen Aufbaus, die Auswertung der Fourierkoeffizienten  $a_n$  und anderes mehr.

Dem Direktor des Instituts für Technische Physik und Leiter der Abteilung für industrielle Forschung an der ETH, Prof. Dr. E. Baumann, danken wir für die wohlwollende Unterstützung während der Ausführung der Versuche. Der Firma Ciba Photochemie Fribourg sei für die Präparation von Fig. 5 gedankt.

#### Literaturverzeichnis

- [1] *M. Born und E. Wolf: «Principles of Optics», Pergamon Press, 1965.*
- [2] *H. J. Tiziani: Thesis, University of London, 1967.*
- [3] *A. Vander Lugt: Optica Acta 15, 1, 1968.*
- [4] *Tippett et al.: «Optical and Electro-Optical Information Processing», MIT-Press, 1965.*
- [5] *J. D. Armitage und A. W. Lohmann: Appl. Opt. 4, 461, 1965.*
- [6] *A. W. Lohmann: Appl. Opt. 7, 561, 1967.*
- [7] *A. Vander Lugt: Appl. Opt. 5, 1760, 1966.*
- [8] *B. H. Beyeler: Diplomarbeit, Institut für Technische Physik ETH, Zürich, 1968.*
- [9] *W. Witz: Diplomarbeit, Institut für Technische Physik ETH, Zürich, 1968.*