Invariante Bildbeschreibung mit Hilfe von Autovergleichs-Funktionen*

Helmut Glünder

Der Hauptteil der Abhandlung besteht aus drei Kapiteln. Auf die Einführung folgt die umfassende Darstellung des neuen und systematischen Ansatzes zur invarianten Bildmuster-Beschreibung. Er beruht auf der Musteranalyse mit Hilfe sogenannter Autovergleichs-Funktionen, einer Verallgemeinerung der von der Signaltheorie her bekannten Autokorrelations-Analyse. Merkmale, die gegenüber allen Ähnlichkeits-Transformationen invariant sind und Verfahren zu ihrer Berechnung, werden vorgestellt und ausführlich diskutiert. Es wird nachgewiesen, daß die Merkmale in ihrer Beschreibungskraft denen des einzigen hinsichtlich der Invarianzleistungen vergleichbaren Verfahrens weit überlegen sind. Im folgenden Kapitel wird gezeigt, daß sich neuronale Netzwerke vorzüglich zur parallelen Berechnung der vorgeschlagenen Merkmale eignen. Die invariante Musteranalyse kann im selben Netzwerk erfolgen wie ein ebenfalls neuartiges Kreuzkorrelations-Verfahren zur musterunabhängigen Gewinnung von Bewegungsparametern. Dieses dualistische Analyseprinzip stellt eine wesentliche Bedingung der Möglichkeit zur Selbststrukturierung der hoch spezifischen neuronalen Verschaltung dar. Bewegungsanalysesysteme mit den geforderten Eigenschaften wurden unlängst bei Affe und Katze nachgewiesen.
Die Daten und Algorithmen finden sich in einem zweiten Teil. Mit einem eigens entwickelten opto-digitalen Prozessor wurden die vorgeschlagenen Merkmale aus 30 Zeichen ermittelt. Algorithmen zur seriellen Merkmalsberechnung werden vorgestellt und anhand zweier einfacher Beispiele erläutert. Ein Anhang zur analytischen Geometrie sowie zur Beschreibung und systemtheoretischen Analyse des verwendeten optischen Zoom-Korrelators beschließt die Abhandlung.

Einführung

Die Bedeutung invarianter Bildbeschreibungs-Merkmale für die Musteranalyse und Mustererkennung wird dargelegt. Insbesondere wird gezeigt, daß bei großen Klassenzahlen und vom Ausmaß her unbeschränkten geometrischen Varianzen – wie sie für die biologische Mustererkennung typisch sind – die Klassifikation aufgrund invarianter Mustermerkmale bezüglich des Leistung/Aufwands-Verhältnisses gegenüber den drei anderen grundlegenden Verfahren der invarianten Mustererkennung am besten abschneidet. Voraussetzung ist jedenfalls eine für

* Zusammenfassung einer von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München angenommenen Dissertation

die Klassentrennung ausreichende Beschreibungskraft der Merkmale. Bei invarianten Merkmalen ist das prinzipiell Erreichbare durch die Invarianten der in Betracht gezogenen geometrischen Transformationen (Varianzen) gegeben. Aus diesem Grund wird erstmals vorgeschlagen, Musterinvarianten selbst als Merkmale zu verwenden und sie definitionsgemäß zu ermitteln. Dazu unterzieht man das zu beschreibende Muster den geometrischen Transformationen und stellt im Vergleich mit dem untransformierten Muster die Invarianten bzw. das Ausmaß an Transformationsinvarianz fest. Im Rahmen der Abhandlung werden Merkmale untersucht, die gegenüber allen Ähnlichkeitstransformationen invariant sind (Verschiebung: ξ, η; zentrische Streckung: s; Rotation: ψ; Achsenspiegelung: σ). In der Literatur findet sich nur eine Art Ansatz zur Gewinnung ähnlichkeitsinvarianter Merkmale, der ebenfalls ganz ohne geometrische Normierungen auskommt. Er erweist sich als Spezialfall der hier vorgestellten Methode.

Theorie

Nur Muster mit nennenswerten inneren Bindungen erlauben die Extraktion brauchbarer invarianter Eigenschaften. Mit der Autovergleichs-Funktion m-ter Ordnung (AVF) wird ein universeller Ansatz zur Analyse innerer Bindungen vorzugsweise linienhafter Muster eingeführt. Die AVF ist eine Verallgemeinerung der klassischen Autokorrelations-Funktion (AKF) hinsichtlich der geometrischen Transformationen sowie der Verknüpfungsoperation und Anzahl (Ordnung) der Bildfunktionen. Die AVF 2. Ordnung wird definitionsgemäß der Invariantenbestimmung gerecht. Die im weiteren betrachtete Form ist die AVF der Ähnlichkeitstransformationen $c_ä(\xi, \eta, s, \psi, \sigma)$. Wegen der Fixpunkte der Transformationen ist diese Funktion nicht invariant. Daher werden von der AVF $c_ä$, unter anderen, drei sogenannte primäre Merkmale mit den vollständigen Invarianzen abgeleitet: die Funktionen $C_s(s)$, $C_\psi(\psi)$ und $C_\sigma(\sigma)$, jeweils gebildet aus den bezüglich der Fixpunkt-Parameter ermittelten besten Paßmaßen (z.B. Korrelationskoeffizienten) der AVF $c_ä$. Die Musterbeschreibungs-Eigenschaften werden dargelegt und verwandte oder abgeleitete Merkmalsfunktionen diskutiert, wobei der Auswertung der relativen Lage optimaler Fixpunkte besondere Bedeutung zukommt. Alternativen zur Bewältigung der Fixpunkt-Problematik werden untersucht, jedoch liefert allein der gewählte Ansatz Maße für Mustersymmetrien, Kongruenzen und Ähnlichkeiten.

Ebenso wie man die klassische AKF nicht nur über die Multiplikation zweier gleicher, aber gegeneinander verschobener (explizit transformierter) Bildfunktionen berechnen kann, sondern auch über die Gesamtheit aller Zweipunkt-Produkte (Dipol-Momente) des Bildes (impliziter Ansatz), bieten sich bei der Berechnung der AVF $c_ä$ ebenfalls beide Möglichkeiten. Neben dem expliziten Ansatz werden zwei Varianten zur impliziten Berechnung im Detail vorgestellt (verallgemeinerte Dipol-Momente). Jede der beiden ist sowohl für übliche ortsdiskrete Musterrepräsentationen ausgeführt als auch für solche, die zusätzlich die Orientierungen von Linienelementen der Muster enthalten. Für alle vier Formen werden Netzwerke zur Parallelberechnung der Merkmale angegeben. Quantitative Angaben, den Aufwand und die Eignung zur seriellen oder parallelen Berechnung betreffend, schließen das Kapitel ab.

Neurobiologie

Visuelle Musteranalyse aufgrund der vorgeschlagenen invarianten Merkmale ist aus folgenden Gründen wahrscheinlich: erstens, in Anbetracht der in natürlicher Umgebung großen zu bewältigenden Mustervarianzen und zweitens, wegen der aus (Gestalt-)psychologischen Forschungen bekannten Kategorien menschlicher Formwahrnehmung (Symmetrien, Kongruenzen, Ähnlichkeiten, etc.), also Forderungen, denen in diesem Ausmaß bislang kein anderes Verfahren gerecht wird. Elementare informations-theoretische Betrachtungen in Verbindung mit Erkenntnissen über die Signalverarbeitung in einzelnen Neuronen führen zu folgenden Einsichten:

- Die zur Informationsreduktion notwendigen Nichtlinearitäten sind vor allem als Multilinearformen sinnvoll.
- Multilinearformen von Signalwerten sind neuronal durch nichtlineare (z.B. multiplikative) Interaktionen benachbarter dendritischer Synapsen möglich, deren Resultate (Produkte) im Zellkörper summiert werden.
- In umfangreichen neuronalen Verarbeitungssystemen dominiert aus Aufwandsgründen die Bilinearform.

Der Multilinearform steht die klassische Polynomform entgegen, die durch (lineare) Summation von Signalen mit anschließender nichtlinearer Bewertung zustande kommt. Bilinearform und Polynomform werden anhand einer einfachen Bildverarbeitungsaufgabe verglichen: der Konturextraktion. Dem dafür üblicherweise angenommenen, bandbegrenzten Laplace-Operator mit Schwellenkennlinie (Polynomform) wird ein neuartiges und zumindest ebenso effektives Verfahren vom Bilineartyp entgegengestellt.

Die zur (orts)diskreten Berechnung von Funktionswerten der AVF 2. Ordnung erforderlichen Summendarstellungen sind Bilinearformen. Ihre neuronale Berechnung aus Bildmustern erfordert sehr viele gezielte Verbindungen zwischen der Musterrepräsentation und den die Vergleichskoeffizienten berechnenden Neuronen. Ohne Angabe plausibler Mechanismen zur Bildung derartiger Netzwerke ist der Ansatz für die Modellierung neuronaler Signalverarbeitung also wertlos. Es wird nachgewiesen, daß hier ebenfalls erstmals vorgeschlagene Systeme zur musterunabhängigen Extraktion von Bewegungsparametern strukturell identisch mit den geforderten Netzwerken zur Berechnung der gewünschten (gleichsinnigen) Ähnlichkeits-AVF sind. Lediglich die paarweise nichtlineare Verknüpfung geschieht bei der Bewegungsanalyse durch sogenannte Hassenstein/Reichardt-Detektoren (HRDen). In diesen wird eines der beiden Signale vor der Verknüpfung um ΔT verzögert. Für Fixationszeiten $t > \Delta T$ verlieren die HRDen naturgemäß ihre zeitlichen Eigenschaften, so daß aus ihnen aufgebaute Netzwerke sowohl der Bewegungs- als auch der Formanalyse dienen können.

Das aus der Sicht der Formanalyse schier unlösbare Problem der Systemstrukturierung erscheint vom Standpunkt der Bewegungsanalyse lösbar: Ein System zur Bewegungsanalyse hat, aufgrund der vom Lebewesen ständig ausgeführten Bewegungen, gute Strukturierungschancen. Die Eigenschaft der resultierenden Systeme, objektunabhängig zu analysieren, folgt unmittelbar aus der für die Strukturierung entscheidenden Wirksamkeit von Gemeinsamkeiten konkreter Bewegungsreize (Abstraktion): Nur häufige und gleiche Reizungen führen zur Konsolidierung neuronaler Verbindungen. Mit dieser Verarbeitungsstruktur ist es erstmals gelungen, für ein umfangreiches Modell neuronaler Verarbeitung auf Signalebene plausible Bedingungen der Möglichkeit zur Strukturierung anzugeben. Die Darstellung und Untersuchung von Mechanismen und Prozessen der Selbstorganisation aufgrund von Bewegungsreizen ist nicht Teil der Abhandlung.

Neuronale Systeme zur Analyse von Tiefen- und frontoparallelen Rotationsbewegungen der postulierten Art wurden vor kurzem bei Affe und Katze neurophysiologisch nachgewiesen. Die aufgrund des vorgeschlagenen Verfahrens zur objektunabhängigen Gewinnung von Bewegungsparametern prognostizierten Eigenschaften der rezeptiven Felder (RFer) von Netzwerk-Neuronen sind im Einklang mit den gemessenen. Darüber hinaus werden Voraussagen über RF-Eigenschaften bei statischer Reizung ($t > \Delta T$) gemacht.

Datenteil

Umfangreicher Simulationsexperimente stehen hier an zentraler Stelle. (Insgesamt wurden ca. 2000 Kreuzkorrelations-Funktionen der 30 Muster, jeweils mit einem Ortsbandbreite-Produkt von etwa 4000, berechnet.) Durch diese werden die Eigenschaften der primären Beschreibungsmerkmale veranschaulicht. Der eigens zu ihrer Berechnung über explizite Transformationen entwickelte optische Zoom-Korrelator ist bezüglich des Verhältnisses von Ortsbandbreite-Produkt zu Aufwand optimiert. Geometrisch-optische wie auch system-theoretische Analysen dieses Parallelrechners werden geliefert.

Algorithmen zur seriellen Merkmalsberechnung über verallgemeinerte Dipol-Momente, mit und ohne Orientierungsparameter, werden vorgestellt und anhand der Analysen zweier Muster erläutert.