Unscharfe Histogramm Klassifikation mit nichtlinearen Zirkulartransformationen und Potentialfunktionen für die Bildfindung und -analyse

Volker Lohweg und Dietmar Müller*

Linnenstr. 35, 33699 Bielefeld
Email: v.lohweg@owl-online.de
*Professur Schaltungs- und Systementwurf
Technische Universität Chemnitz, 09107 Chemnitz

Zusammenfassung. Basierend auf der Klasse der Zirkulartransformationen und unscharfer Histogrammerzeugung mittels Potentialfunktionen wird ein Konzept zur Klassifizierung von Bildinhalten vorgestellt. Durch eine unscharfe Gewichtung der Histogrammbereiche entsteht auch bei nicht idealen Praxissituationen ein stabiles Histogramm, dass als Ähnlichkeitsmaß verwendet werden kann. Als Auswertung wird die Methode der Histogramm Intersektion verwendet. Das vorgeschlagene Verfahren geht mit einer deutlichen Datenmengenreduktion einher und kann deshalb gut im Bereich von Image Retrieval eingesetzt werden.

1 Einleitung

Im Bereich der Bildfindung und -analyse besteht die Aufgabe anhand von Prototypen- oder Teilbildern entsprechende Informationen in großen Datenbanken zu extrahieren. Das Aufkommen an Bildmaterial ist dabei derart groß, dass eine Suche i.a. nicht mehr händisch durchgeführt werden kann. Eine automatische Bildfindung ist daher unumgänglich. Diese sollte auf das Bildmaterial direkt, also ohne Segmentierung und Festlegen von Schlüsselpunkten durchgeführt werden können. Typische Image Retrieval Verfahren beruhen auf entsprechenden Metriken zur Bestimmung von Gleichartigkeiten; z.B. Abstandsmaße für segmentierte Objekte oder der Berechnung von globalen Häufigkeitsverteilungen.

Es soll hier ein Verfahren vorgeschlagen werden, das auf nichtlinearen translationsinvarianten Zirkulartransformationen und unscharf gewichteten Häufigkeitverteilungen der Ausgangsspektren basiert. Diese Häufigkeitverteilungen werden mit Hilfe von Potentialfunktionen erzeugt.

Bei den verwendeten Zirkulartransformationen handelt es sich um Spektraltransformationen, die eine schnelle Berechnung translationsinvarianter Merkmale mit einem rechentechnischen Aufwand von $\boldsymbol{O}(N)$ bis $\boldsymbol{O}(N\mathrm{ld}(N))$ erlauben (N ist die Länge eines Dateneingangsvektors). Die Transformationen haben die Eigenschaft, dass ein *Betragsspektrum G* mit ld(N)+1 Koeffizienten (1D-Fall) definiert werden kann, welches mit absoluten Beträgen operiert und invariant bezüglich zyklischer Verschiebungen eines Eingangsvektors ist [1,2]. Erwähnenswert ist, dass auch die modifizierte Walsh-Hadamard-Transformation (MWHT) [3] und auch eine vom Pender und Covey vorgeschlagene *Square Wave Transform* (SWT) [4] ebenfalls zu der Klasse der Zirkulartransformationen gehören. Diese Eigenschaft ist von der WHT und generalisierten Transformationen (GT, bzw. MGT) [3] nicht bekannt.

2 Zirkulartransformationen

Ausgehend von einem Eingangsdatenvektor $\boldsymbol{x}^T = \{x_0, x_1, ..., x_{N-1}\}$ der ohne Einschränkung für $x_i \in \mathbb{R}$ gelte und einem transformierten Datenvektor $\boldsymbol{X}^T = \{X_0, X_1, ..., X_{N-1}\}$, sind die Transformationen sowie ihre Inversen gegeben durch (***A*** und ***B*** sind quadratische (N x N)-Transformationsmatrizen):

$$\boldsymbol{X} = \boldsymbol{A}_N \cdot \boldsymbol{x} \text{ und } \boldsymbol{x} = \tfrac{1}{N} \cdot \boldsymbol{B}_N^T \cdot \boldsymbol{X} \tag{1}$$

Es gilt für sie: $\boldsymbol{A}_N \cdot \boldsymbol{B}_N^T = \boldsymbol{B}_N^T \cdot \boldsymbol{A}_N = \boldsymbol{A}_N^T \cdot \boldsymbol{B}_N = \boldsymbol{B}_N \cdot \boldsymbol{A}_N^T = N \cdot \boldsymbol{I}_N.$ (2)

Ausgehend von einer (2x2)-Hadamard-Matrix $\boldsymbol{K} = \begin{bmatrix} +1 & -1 \\ +1 & +1 \end{bmatrix}$ [1] werden die Transformationsmatrizen rekursiv erzeugt.

$$\boldsymbol{A}_N = \boldsymbol{diag}({}^f\boldsymbol{T}_{\frac{N}{2}}, \boldsymbol{A}_{\frac{N}{2}}) \cdot \left[\boldsymbol{K} \otimes \boldsymbol{I}_{\frac{N}{2}}\right] \text{ und } \boldsymbol{B}_N = \boldsymbol{diag}({}^r\boldsymbol{T}_{\frac{N}{2}}, \boldsymbol{B}_{\frac{N}{2}}) \cdot \left[\boldsymbol{K} \otimes \boldsymbol{I}_{\frac{N}{2}}\right]. \tag{3}$$

Die *charakteristischen Matrizen* ${}^f\boldsymbol{T}$ und ${}^r\boldsymbol{T}$ besitzen die Dimension $(\frac{N}{2} x \frac{N}{2})$. Je nach Definition der Transformationskerne werden verschiedene Transformationen mit unterschiedlichen Eigenschaften möglich [1]. Alle Transformationen besitzen die Eigenschaft eines nach Perioden geordneten Spektrums. Beginnend mit den ersten $N/2$ Basisvektoren mit der Periode N, folgen $N/4$ Basisvektoren mit der Periode $N/2$, bis hin zu dem Basisvektor mit der kürzesten mögliche Periode zwei und einem Basisvektor der Periode Null. Er stellt den Mittelwert der Eingangsfolge dar. Die Transformationsmatrix besteht aus $N/2$ ungeraden und $N/2$ geraden Basisvektoren.

2.1 Betragsspektrum G

Das translationsinvariante Betragsspektrum G ist im Gegensatz zum Leistungsspektrum der DFT durch die Bildung von Periodengruppen, ähnlich dem Leistungsspektrum der WHT, definiert. Mit Hilfe des bekannten Verfahrens der Berechnung einer Shiftmatrix ${}^s\boldsymbol{S}_N := \frac{1}{N} \cdot \boldsymbol{A}_N \cdot {}^s\boldsymbol{I}_N \cdot \boldsymbol{B}_N^T$ mit $-(N-1) \le s \le (N-1)$ [3], lässt sich für alle Zirkulartransformationen zeigen, dass durch eine Summation der Beträge der Spektralkoeffizienten (jeweils innerhalb einer Periodengruppe) ein translationsinvariantes Spektrum mit ld(N)+1 Koeffizienten (1D-Fall) entsteht, dass als Merkmalvektor verwendet werden kann [1,2].

Mit ${}^{s}\boldsymbol{I}_m$ wird eine (m x m)-Einheitsmatrix bezeichnet, deren Spalten um s Stellen zyklisch verschoben werden. Hierbei gilt für $s \geq 0$, dass die Spalten um s Stellen nach rechts verschoben und für $s < 0$, dass die Spalten um s Stellen nach links verschoben sind.

3 Methode

Grundsätzlich genügen translationsinvariante Ausgangsspektren, um Bildinhalte zu beschreiben. In realen Bildszenen und Anwendungen ist jedoch ein einfacher Vergleich von invarianten Spektren nicht ohne weiteres möglich, da in der Praxis Situationen auftreten, die den o.g. einfachen Vergleich nicht möglich machen. Zu nennen sind hier: Objektverschiebungen unter dem Aufnahmesystem, nicht zyklische Verschiebungen, Aliasing-Effekte bei der Digitalisierung, verschiedene Untergründe, usw. [5]. Aus diesem Grund werden die Spektralkoeffizienten einer Histogrammanalyse unterzogen. Die Histogramme werden nicht im üblichen Sinne erzeugt, sondern durch Potentialfunktionen nach Aizerman [6], Bocklisch [7], u.a.. Durch eine unscharfe Gewichtung der Histogrammbereiche entsteht auch bei den o.g. nicht idealen Praxissituationen ein stabiles Histogramm, das als Ähnlichkeitsmaß verwendet wird. Die Grundform einer Potentialfunktion, die auch als *Zugehörigkeitsfunktion* bezeichnet wird, ist wie folgt definiert:

$$\mu(x) := \left[1 + W \cdot \left[\tfrac{|x-x_0|}{C}\right]^{D}\right]^{-1}. \tag{4}$$

Der Parameter x_0 beschreibt die Lage des Maximums der Funktion. D legt die Verteilung der Werte x fest und beschreibt den Hangabfall der Funktion und damit den Grad der Veränderung bei wachsender Entfernung vom Maximalpunkt x_0. Der Faktor W steuert eine festzulegende Randzugehörigkeit. Es sei darauf hingewiesen, dass die oben angegebene Grundform in vielfacher Weise verändert werden kann. Hier muß auf die entsprechende Literatur verwiesen werden; z. B. [8].

Die Merkmale werden innerhalb von (NxN)-Fenstern der Größe (8x8) oder (16x16) Pixel berechnet. Diese Fenster sind in Form eines Rasters über die zu analysierenden Aufnahmen gelegt, so dass die innerhalb jedes Fensters berechneten Spektralkoeffizienten als lokal angesehen werden können. Basierend auf dem Konzept der generalisierten Zirkulartransformationen [1] werden zweidimensionale Spektren G mit einer Anzahl $(\mathrm{ld}(N)+1)^2$ Ausgangskoeffizienten bestimmt. Sie dienen als Initialmerkmale zur Bestimmung eines Ähnlichkeitsmaßes pro Fenster. Dieses Maß wird als eindimensionale unscharfe Häufigkeitsverteilung definiert.

Es ist als bekannt vorauszusetzen, dass übliche Histogramme stark ihre Form ändern können falls Amplitudenwerte an den Rändern der Zuordnungsklassen schwanken. Dieses geschieht z. B. durch Rauschen. Gerade dieses Verhalten ist nicht gewünscht, da kleine Änderungen im Spektrum aus praktischen Gesichtspunkten zugelassen werden müssen. Aus diesem Grund wird jeder Spektralamplitudenwert nicht direkt einer Histogrammklasse zugeordnet, sondern über eine Potentialfunktion gewichtet auf alle Histogrammklassen verteilt. Anhand eines

Amplitudenwertes wird das Maximum x_0 der Potentialfunktion im Histogramm anhand des jeweiligen Spektralamplitudenwertes des zweidimensionalen Betragsspektrums G positioniert und in der Mitte der jeweiligen Histogrammklasse abgetastet; es findet somit in jeder Klasse ein gewichteter Eintrag pro Amplitudenwert statt. Dieses Verfahren erzeugt ein unscharfes Histogramm, das stabil bei kleinen Änderungen der Aufnahmen bleibt, jedoch die Bildinhalte genügend genau beschreibt. Auch die Forderung nach einer Summenkonstanz des Histogramms bleibt bei dem vorgeschlagene Ansatz erhalten, da die Summe der Histogrammeinträge für unterschiedliche Histogramme (Bildinhalte) gleich bleibt. Dieses Verhalten ist in Bezug auf ein Histogramm im engeren Sinne ebenso gefordert. Vorteilhaft ist die Tatsache zu werten, dass Potentialfunktionen parametrierbar sind. D. h., das Maß der Ähnlichkeit kann über entsprechende Koeffizienten eingestellt werden. Beispielsweise kann im Fall der Verwendung der o.g. Potentialfunktion das Ähnlichkeitsmaß über die Konstanten C, D und W justiert werden.

Damit besteht die Möglichkeit einerseits über die Parameter der Zirkulartransformationen und andererseits über die Parameter der unscharfen Häufigkeitsverteilung das Ähnlichkeitsmaß anzupassen. Als Abstandsmaße der Auswertung können bekannte Verfahren wie z. B. der normierte euklidische Abstand, Histogramm Intersektion, usw. dienen

4 Ergebnisse

Es wurden exemplarisch Bildmerkmale von 100 Aufnahmen verschiedener Szenen analysiert. Die Bildgröße entsprach dabei 128x128 Pixel; es wurden (16x16)-Fenster verwendet. Damit ergab sich eine Fensteranzahl von 64 Fenstern. Es wurden jeweils unscharfe Histogramme mit 4 und 8 Klassen gewählt, so dass die Bilddaten durch jeweils 256, bzw. 512 Einträge repräsentiert wurden. Dieses entspricht einer Datenmengenreduktion von 64 und 32 pro Bildaufnahme.

Untersucht wurden verschiedene Zirkulartransformationen (RMWHT, SWT, ZT(n)) [1,2] und eine Potentialfunktion mit $W = 1$, $D = 4$ und verschieden groß gewählte C. Es ist festzustellen, das Transformationen mit guten Trenneigenschaften sich günstig bei (16x16)-Fenstern auswirken, während sie bei dem (8x8)-Fenster zu viele Details abbilden. Für das (8x8)-Fenster eignen sich Transformationen deren Trenneigenschaften nicht optimal sind (z. B.: RMWHT, SWT).

Der Parameter C muss so groß gewählt werden, dass in jedem Fall ein Eintrag in alle Histogrammklassen stattfindet ($C_{\min}$). Er muss jedoch immer so klein gewählt werden, dass eine Unterscheidbarkeit der Bildinhalte gewährleistet ist. Eine empirische Abschätzung zeigt, dass der Einstellbereich etwa die Größenordnung $\mathrm{span}(C) \approx \{C_{\min} \ldots 10 \cdot C_{\min}\}$ einnimmt. Innerhalb dieses Bereichs kann C variiert werden, um den in der Praxis vorkommenden Effekten Rechnung zu tragen. Alle Aufnahmen, die für die Untersuchung genutzt wurden, konnten mit den oben genannten Randbedingungen klassifiziert werden. Als Klassifikator erbrachte das Verfahren der Histogramm Intersektion die besten Ergebnisse. Es wurde hierbei ein Ähnlichkeitsmaß durch die Bestimmung der Intersektion für jedes der 64 Fenster berechnet. Diese 64 Daten je Bild wurden als Klassifizierungsergebnis in Form der Bestimmung einer Vertauschungsmatrix ausgewertet.

5 Diskussion und Resümee

Es wurde ein Verfahren vorgestellt, dass auf einfache Weise eine Klassifikation von Bildinhalten ermöglicht. Eine Segmentierung von Objekten wird nicht benötigt. Durch das Verfahren der Rasterung ist es möglich nach charakteristischen Bereichen in A ufnahmen zu s uchen und e ine Klassifizierung nu r anhand v on Teilbildern vorzunehmen. Dieser Ansatz muss jedoch mit einer größeren Menge an Bildmaterial noch genauer untersucht werden. Die bisher erzielten Ergebnisse lassen den Schluss zu, dass der vorgeschlagene Ansatz aufgrund seiner einfachen Berechnungsstrategien auch eine einfache Hardwareimplementation zulässt [9]. Mit diesem Ansatz könnten innerhalb kurzer Zeit große Datenmengen bearbeitet werden.

6 Literatur

1. Lohweg V, Müller D: Ein generalisiertes Verfahren zur Berechnung von translationsinvarianten Zirkulartransformationen für die Anwendung in der Signal- und Bildverarbeitung, Mustererkennung 2000, 22. DAGM-Symposium,Kiel, 13. -15. Sept. 2000, pp. 213 - 220, Springer-Verlag, 2000
2. Lohweg V, Müller D: Anwendung schneller diskreter Spektraltransformationen zur transla- tionsinvarianten Merkmalgewinnung, Mustererkennung 1999, 21. DAGM-Symposium, Bonn, 15. -17. Sept. 1999, pp. 266 - 275, Springer-Verlag, 1999
3. Ahmed N, Rao K R: Orthogonal Transforms for Digital Signal Processing, Springer-Verlag, 1975
4. Covey D, Pender J: New Square Wave Transform for Digital Signal Processing, IEEE Trans. on Signal Processing, Vol. 40, No. 8, pp. 2095-2097, 1992
5. Siggelkow S, Burkhardt H: Image Retrieval Based On Local Invariant Features, Proceedings of the IASTED International Conference Signal and Image Processing, October 27-31, 1998, Las Vegas, Nevada
6. Aizerman S: Methode der Potentialfunktionen und ihre Anwendung (Russ.), Nauka 1972, Moskau
7. Bocklisch S F: Prozeßanalyse mit unscharfen Verfahren, VEB Verlag Technik, Berlin 1987
8. Eichhorn K: Entwurf und Anwendung von ASICs für musterbasierte Fuzzy-Klassifikationsverfahren am Beispiel eines Schwingungsüberwachungssystems, Technische Universität Chemnitz, Lehrstuhl Schaltungs- und Systementwurf, Dissertationsschrift 1999
9. Mauersberger H, Müller D: Effektive Entwurfsmethodik für leistungsfähige Bildverarbeitungssysteme, Technische Universität Chemnitz, Lehrstuhl Schaltungs- und Systementwurf, Dresdner Arbeitstagung Schaltungs- und Systementwurf, DASS'99, 19.-20.05.99