Bildkompression mit endlichen Automaten

A. Schrader
Universität-Gesamthochschule Siegen
Fachbereich 12 - Elektrotechnik und Informatik
Programmiersprachen
57068 Siegen
e–mail: schrader@informatik.uni-siegen.de

Das 'multimediale Zeitalter' hat zu einer ständig wachsenden Integration von Bilddaten bei der Übertragung und Speicherung von Informationen geführt. Die Begrenzungen der Kapazitäten in Übertragungskanälen und Speichermedien erforderte dabei schon früh die Entwicklung geeigneter Kompressionsalgorithmen.

In den klassischen Verfahren der ersten Generation werden Kenntnisse oder Vermutungen über die statistische Verteilung der Daten zur Konstruktion eines redundanzreduzierenden Codes verwendet. Der zweidimensionale Charakter der Bilddaten bleibt dabei im wesentlichen unberücksichtigt.

In diesem Artikel wird die Codierung von endlichen diskreten Bilddaten mit endlichen Automaten beschrieben. Ausgangspunkt unserer Überlegungen ist eine zweidimensional in einem kartesischen Raster mit äquidistanten Rasterpunkten angeordnete Menge P von $n \times n$ Bildpunkten (Pixel), denen als Intensitätswert ein Element aus der Menge $\{0, 1\}$ zugeordnet wird (Binärbild, oder Bitmap mit 2 Graustufen):

$$P := [0 \ldots n-1] \times [0 \ldots n-1] \rightarrow \{0, 1\}.$$

Die Beschreibung des Bildes P wird mit Hilfe einer endlichen Sprache $L \subset \Sigma^*$ über einem festgelegten Alphabet Σ vorgenommen. Endliche Sprache lassen sich durch reguläre Ausdrücke beschreiben und sind somit Teilmenge der (regulären) Chomsky-Typ3-Sprachklasse, so daß die Beschreibung der Sprache auch durch die Definition eines entsprechenden endlichen Automaten (FA) erfolgen kann.

Zur Erzeugung einer solchen Bildsprache L wird eine bijektive Abbildung

$$\phi := [0 \ldots n-1] \times [0 \ldots n-1] \rightarrow L$$

verwendet, die jedem Pixel des Bildes ein eindeutiges Codewort $\omega \in L$ zuordnet. Eine Möglichkeit besteht in der Verwendung der Quadtree-Zerlegung (Abb. 1), die mit Hilfe des Alphabets $\Sigma = \{0, 1, 2, 3\}$ für jeden Pixel ein Codewort $\omega \in \Sigma^k$ der Länge $|\omega| = k$ erzeugt.

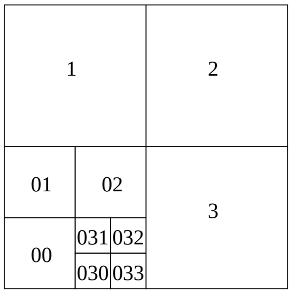

Abbildung 1: Einige Quadtree-Zuordnungen bis $k = 3$

Die Sprache L_P enthält nun alle Codewörter $\omega \in L$, deren Pixel in der Bitmap den Wert 1 besitzen:

$$L_P := \{\omega \in L | P(\phi^{-1}(\omega)) = 1\}.$$

Für diese Sprache L_P wird ein deterministischer endlicher Automat (DFA) $A = (Q, \Sigma, \delta, q_0, F)$ erzeugt, für dessen akzeptierte Sprache gilt:

$$L(A) := \{\omega \in \Sigma^* | \delta^*(q_0, \omega) \in F\} = L_P.$$

Das Bild P läßt sich aus diesem Automaten rekonstruieren, indem der Automat für alle $\omega \in L_P$ in einen akzeptierenden Endzustand übergeht und für alle $\omega \in L \setminus L_P$ in einem nichtakzeptierenden Zustand verbleibt.

Jeder Zustand im Automaten repräsentiert einen bestimmten Subquadranten des Bildes. Die Ähnlichkeit zweier Subquadranten kann auf die Äquivalenz der entsprechenden Zustände übertragen werden und führt so zu einem reduzierten minimalen Automaten mit einer geringen Zustandszahl.

Der Artikel beschreibt zwei heuristische Ansätze zur effektiven Codierung der Zustandsübergangstabelle des Automaten und vergleicht die Ergebnisse mit Werten herkömmlicher Kompressionsverfahren anhand von ausgesuchten Bildbeispielen.