DEUTSCHE DEMOKRATISCHE REPUBLIK

PATENTSCHRIFT

(12) Wirtschaftspatent

(19) DD (11) 281 069 A1

Erteilt gemäß § 17 Absatz 1 Patentgesetz

4(51) H 03 M 7/30

PATENTAMT der DDR

In der vom Anmelder eingereichten Fassung veröffentlicht

(21)	WP H 03 M / 327 218 5	(22)	03.04.89	(44)	25.07.90

(71) Akademie der Wissenschaften der DDR, Institut für Kosmosforschung, Rudower Chaussee 5, Berlin, 1199, DD

(72) Günther, Adolf, DD

(54) Verfahren zur Kompression digitaler Datenfolgen mittels Kodierung

(55) Kompression; Daten, digital; Videosignal; Datenübertragung; Datenfolge, unregelmäßig; Zerlegung; Teilfolgen; Alphabete; Kodiereinrichtung; Dekodiereinrichtung

(57) Die Erfindung betrifft ein Verfahren zur Kompression digitaler Datenfolgen mittels Kodierung und ist beispielsweise bei der Übertragung von Videosignalen mit geringer Bitrate bei Beibehaltung einer hohen Übertragungsqualität anwendbar. Aufgabe der Erfindung ist die Schaffung eines Verfahrens zur Kompression unregelmäßiger digitaler Datenfolgen mittels Kodierung, wobei sowohl innerhalb der Datenfolgen als auch von einer betrachteten Datenfolge zur nächsten Folge große inhaltliche Unterschiede in den Daten auftreten können, so daß eine einheitliche statistische Beschreibung nicht möglich ist. Erfindungsgemäß erfolgt die Lösung der Aufgabe zunächst durch Zerlegung de. gesamten Datenfolge in Teilfolgen, welche in sich weniger Variationen aufweisen als die Gesamtfolge, gegeben z. B. bei örtlichen Häufungen von großen oder kleinen Datenworten. Diese somit definierten Teilfolgen werden durch verschiedene Alphabete dargestellt, die einem gegebenen Satz von Alphabeten entnommen werden.

ISSN 0433-6461

9 Seiten

Patentansprüche:

1. Verfahren zur Kompression digitaler Datenfolgen mittels Kodierung, **gekennzeichnet dadurch,** daß zunächst die gesamte Datenfolge in Teilfolgen zerlegt wird, welche in sich weniger Variationen aufweisen als die Gesamtfolge, die somit definierten Teilfolgen durch verschiedene Alphabete dargestellt werden, welche einem gegebenen Satz von Alphabeten entnommen werden, wobei der Satz von Alphabeten ein uneingeschränktes erstes Alphabet, mit dem alle gegebenen Datenworte dargestellt werden können, sowie eine Anzahl von eingeschränkten zweiten Alphabeten mit einer geringeren Zeichenzahl enthält sowie ein drittes an sich bekanntes Alphabet der Darstellung der Länge der Teilfolge dient und dieses dritte Alphabet sowohl für die Darstellung der nichtsignifikanten Teilfolgen, d.h. Nullfolgen als auch für signifikante Datenfolgen gleicher Werte benutzt wird, wenn dies für eine Vergrößerung des Kompressionsfaktors sinnvoll ist, weiterhin wird eine Nullfolge durch ein Kennzeichen für das benutzte Alphabet und in bekannter Weise durch ein Zeichen variabler Länge für die Länge der Nullfolge dargestellt, wobei der erste Teil des Zeichens eine Information über einen konstanten Wert und über die Länge des zweiten Teiles des Zeichens enthält und der zweite Teil des Zeichens einen variablen Teil des Wertes enthält, der zu dem konstanten Teil summiert wird und eine Teilfolge von signifikanten Daten nunmehr entweder durch ein Kennzeichen für das benutzte Alphabet, gefolgt von einzelnen Zeichen für die Werte der Datenworte und mit einem Endekennzeichen, das zu dem entsprechenden Alphabet gehört oder durch ein Kennzeichen für das benutzte Alphabet, gefolgt von einem Endekennzeichen sowie gefolgt von einem Zeichen für die Größe der Datenworte und zum Ende mit einem Zeichen variabler Länge, ähnlich der Darstellung der Nullfolge, dargestellt wird.
2. Verfahren zur Kompression digitaler Datenfolgen mittels Kodierung nach Anspruch 1, **gekennzeichnet dadurch,** daß während des Kompressionsvorganges ständig das günstigste Alphabet in bezug auf eine Teilfolge ausgesucht und geprüft wird, ob bei dessen Benutzung eine Kompression gegenüber der ungeteilten Folge erreichbar ist.

Hierzu 3 Seiten Zeichnungen

Anwendungsgebiet der Erfindung

Die Erfindung betrifft ein Verfahren zur Kompression digitaler Datenfolgen mittels Kodierung und ist beispielsweise bei der Übertragung von Videosignalen mit geringer Bit-Rate bei Beibehaltung einer hohen Übertragungsqualität anwendbar.

Charakteristik des bekannten Standes der Technik

Es ist bekannt, eine Datenkompression auf der Basis einer Kodierung zu erreichen, wobei diese Kodierung die Wortlänge der Datenfolge verändert. Hierbei wird die Kompression durch die Verwendung von unterschiedlich langen Worten erreicht, d.h. häufig auftretende Werte werden mit kurzen Worten, selten auftretende Werte werden mit langen Worten kodiert.
Mit der Verwendung von unterschiedlichen Worten tritt das Problem auf, die einzelnen Worte gegeneinander abzugrenzen. Diese Abgrenzung verschlechtert die Kompression, da sie auch mit kodiert werden muß, um eine exakte Dekodierung zu gewährleisten. Die Art der Trennung der Datenworte charakterisiert hierbei die unterschiedlichen bekannten Kodierungen.
So wird z.B. bei Präfixkodes die Worttrennung durch ein Präfix vorgenommen. Dadurch können nicht alle Kombinationen der Binärdarstellung ausgenutzt werden und das bewirkt in diesem Fall einen Kompressionsverlust. Ist z.B. das erste Kodewort 0, so kann das zweite Kodewort nicht mehr einstellig sein, weil die 1 als Präfix für die nächstgrößeren Kodeworte verwendet wird.
Bei den B-Kodes wird jeder Gruppe von Informationsbits ein sogenanntes Folgebit zugeordnet, dessen Wertwechsel ein neues Datenwort anzeigt. Hierbei bewirken die Folgebits die Kompressionsverschlechterung.
Eine weitere Möglichkeit ist die Unterteilung der Datenfolge in Teilfolgen mit gleichen Werten, so daß nur der Wert und die Länge der Teilfolgen kodiert werden muß. Diese Kodierung ist besonders für zweiwertige Daten mit der Wortgröße 1 Bit günstig. Hier ergibt sich eine Kompressionsverschlechterung, insbesondere bei kurzen Teilfolgen durch die Längenangabe.
Es muß hier auch festgestellt werden, daß mit steigender Größe des Quellenwortes die Anzahl von längeren Teilfolgen mit gleichem Wert zurückgeht, so daß die Bildung von Teilfolgen keine Vorteile bringt.
Der Nachteil bekannter Kodierungen ist, daß bei Kodes mit variabler Wortlänge, bedingt durch den zusätzlichen Aufwand der Worttrennung, die Entropie 1. Ordnung (mittlere notwendige Wortgröße) nicht unterschritten werden kann.
In den der Erfindung nächstliegenden Patenten US-PS 4684923 und DE-OS 3631252 werden hybride Verfahren verwendet, bei denen signifikante Datenfolgen (Datenworte ungleich Null) und nichtsignifikante Datenfolgen (Datenworte gleich Null) getrennt behandelt werden. Es wird dabei ausgenutzt, daß eine Folge von Nullen nur noch als Lauflänge kodiert zu werden braucht, wodurch wesentlich bessere Kompressionsfaktoren erzielt werden können als z.B. bei der Huffmannkodierung, wo jeder Null einzeln ein Kode zugeordnet wird.

Für diese Lösungen ergibt sich bei den signifikanten Datenfolgen dennoch der Nachteil, daß eine Kompression nur erreicht werden kann, wenn die Werte nicht gleich verteilt sind und wenn außerdem die statistische Verteilung der Werte bekannt ist. Ein weiterer Nachteil ist, daß bei der Trennung von signifikanten Datenfolgen und Nullfolgen, speziell bei kurzen Nullfolgen, der Fall auftreten kann, daß eine Abtrennung der Nullfolge keinen Vorteil bringt.
Diese Nachteile können summarisch auftreten und ergeben, daß bei ungünstigen Datenfolgen, an die die Verfahren nicht angepaßt sind, sogar Kompressionsfaktoren < 1 auftreten können, aber auch in anderen Fällen die Kompressionsmöglichkeiten nicht genügend ausgenutzt werden.

Ziel der Erfindung

Das Ziel der vorliegenden Erfindung ist es, ein Verfahren zur Kompression digitaler Datenfolgen mittels Kodierung vorzuschlagen, wobei die Kompression von den Eigenschaften der zu komprimierenden Datenfolge, unabhängiger ist und dadurch im Durchschnitt höhere Kompressionsfaktoren, insbesondere bei unregelmäßigen Datenfolgen erreicht werden.

Darlegung des Wesens der Erfindung

Aufgabe der Erfindung ist die Schaffung eines Verfahrens zur Kompression unregelmäßiger digitaler Datenfolgen mittels Kodierung, wobei sowohl innerhalb der Datenfolgen als auch von einer betrachteten Datenfolge zur nächsten Folge große inhaltliche Unterschiede in den Daten auftreten können, so daß eine einheitliche statistische Beschreibung nicht gegeben ist.
Erfindungsgemäß erfolgt die Lösung der Aufgabe der Erfindung zunächst durch Zerlegung der gesamten Datenfolge in Teilfolgen, welche in sich weniger Variationen aufweisen als die Gesamtfolge. Gegeben z. B. bei örtlichen Häufungen von großen oder kleinen Datenworten. Diese somit definierten Teilfolgen werden durch verschiedene Alphabete dargestellt, die einem gegebenen Satz von Alphabeten entnommen werden. Der Satz von Alphabeten enthält ein uneingeschränktes erstes Alphabet, mit dem alle gegebenen Datenworte dargestellt werden können, sowie eine Anzahl von eingeschränkten zweiten Alphabeten mit einer geringeren Zeichenzahl. Ein drittes an sich bekanntes Alphabet dient der Darstellung der Länge der Teilfolge (Längenalphabet) und wird für die nichtsignifikanten Teilfolgen (Nullfolgen) sowie für signifikante Datenfolgen gleiche Werte benutzt, wenn dies für eine Vergrößerung des Kompressionsfaktors sinnvoll ist.
Weiterhin wird eine Nullfolge durch ein Kennzeichen für das benutzte Alphabet und in bekannter Weise durch ein Zeichen variabler Länge für die Länge der Nullfolge dargestellt, wobei der erste Teil des Zeichens eine Information über einen konstanten Wert und über die Länge des zweiten Teiles des Zeichens enthält und der zweite Teil des Zeichens einen variablen Teil des Wertes enthält, der zu dem konstanten Teil summiert wird.
Eine Teilfolge von signifikanten Daten wird nunmehr entweder durch ein Kennzeichen für das benutzte Alphabet, gefolgt von einzelnen Zeichen für die Werte der Datenworte und mit einem Endekennzeichen, das zu dem entsprechenden Alphabet gehört oder durch ein Kennzeichen für das benutzte Alphabet, gefolgt von einem Endekennzeichen sowie gefolgt von einem Zeichen für den Wert der Datenworte und zum Ende mit einem Zeichen variabler Länge, ähnlich der Darstellung der Nullfolge, dargestellt.
Die Aufgabe der Erfindung wird weiterhin dadurch gelöst, daß während des Kompressionsvorganges ständig das günstigste Alphabet in bezug auf eine Teilfolge ausgesucht und geprüft wird, ob bei dessen Benutzung eine Kompression gegenüber der ungeteilten Folge erreichbar ist.

Ausführungsbeispiel

Die Erfindung soll anhand einer prinzipiellen Ausgestaltungsform eines das Verfahren umsetzenden Kompressors sowie der dazugehörigen Kodier- bzw. Dekodiervorrichtung näher erläutert werden. Die Figur 1 zeigt das Blockschaltbild des Kompressors, Figur 2 die Darstellung einer Kodiereinrichtung sowie Figur 3 den Aufbau einer Dekodiereinrichtung.
Gemäß Figur 1 werden die Datenworte nacheinander aus einem Pufferspeicher 1 gelesen, in den die zu komprimierenden Quelldaten gelangen. Dabei wird zunächst die Darstellung der Datenworte in einem uneingeschränkten Alphabet angenommen. Es wird nun bei jedem gelesenen Datenwort der neue Zustand des bisher gelesenen, noch nicht kodierten Teiles der Datenfolge mit einer Prüfeinrichtung 2 geprüft. In dieser Prüfeinrichtung 2 sind Anfang, Länge und eine Kennzeichnung des bisher angenommenen Alphabetes für den noch nicht kodierten Teil der Datenfolge gespeichert (bisheriger Zustand). Außerdem ist die Prüfeinrichtung 2 in der Lage festzustellen, mit welchen Alphabeten das gelesene Datenwort dargestellt werden kann. Weiter sind in der Prüfeinrichtung 2 Informationen gespeichert, aus denen der aktuelle Zustand des nichtkodierten Teiles der Datenfolge bezüglich der eingeschränkten Alphabete und dem Längenalphabet erkennbar ist. Diese Informationen werden zum Anfang aus einem Festwertspeicher 3 gelesen und jeweils nach dem Lesen eines neuen Datenwortes auf den aktuellen Stand gebracht.
Damit ist die Prüfeinrichtung 2 in der Lage, folgende Zustände zu unterscheiden:

1. Das bisher für den nichtkodierten Teil der Datenfolge angenommene Alphabet ist weiterhin das günstigste und wird beibehalten.
2. Das bisher für den nichtkodierten Teil der Datenfolge angenommene Alphabet kann durch ein Alphabet ersetzt werden, das eine größere Kompression gegenüber der Darstellung mit dem bisherigen Alphabet erreicht.
3. Das bisher für den nichtkodierten Teil der Datenfolge angenommene Alphabet muß durch ein anderes ersetzt werden, weil mit dem bisher angenommenen Alphabet das neu hinzugekommene Datenwort nicht dargestellt werden kann.
4. Der bisher gelesene nichtkodierte Teil der Datenfolge läßt sich in zwei Teilfolgen zerlegen, die, wenn sie mit zwei verschiedenen Alphabeten dargestellt werden, eine Kompression gegenüber der ungeteilten nichtkodierten Folge erzielen. Gibt es mehrere verschiedene Alphabete mit den dazugehörigen Teilfolgen, welche eine Kompression erzielen würden, so wird jenes mit der größten Kompression ausgewählt.

Das Ergebnis der Prüfung wird der Steuerung 4 mitgeteilt, die für die verschiedenen Zustände den weiteren Ablauf veranlaßt. Wurde der Zustand 1. erkannt, so wird die Länge des nichtkodierten Teiles der Datenfolge korrigiert, und die Informationen der Prüfeinrichtung 2 bezüglich der Alphabete werden an den neuen Zustand des nichtkodierten Teiles der Datenfolge angepaßt (aktualisiert). Das Aktualisieren kann ein Lesen aus dem Festwertspeicher 3 oder eine Berechnung in der Prüfeinrichtung 2 sein. Wurde der Zustand 2. oder 3. erkannt, so werden das Kennzeichen für das verwendete Alphabet verändert, die Länge des nichtkodierten Teiles der Datenfolge korrigiert und die Informationen in der Prüfeinrichtung 2 bezüglich der Alphabete aktualisiert. Wurde der Zustand 4. erkannt, so wird Anfang und Länge der ersten Teilfolge und das Kennzeichen des für die erste Teilfolge benutzten Alphabetes in einen Zwischenspeicher 5 geschrieben und in der Steuerung 4 ein Kennzeichen gesetzt, daß eine Teilfolge kodiert werden kann. Anschließend werden Anfang und Länge der zweiten Teilfolge korrigiert, welche von diesem Zeitpunkt an den nichtkodierten Teil der Datenfolge darstellt. Die Informationen in der Prüfeinrichtung 2 bezüglich der Alphabete werden aktualisiert. Die Kodierung der freigegebenen ersten Teilfolge geschieht in der Kodiereinrichtung 6 zeitparallel zum Kompressionsvorgang, der mit dem Lesen des nächsten Datenwortes weiterläuft. In dieser Weise wird die gesamte Datenfolge in Teilfolgen zerlegt und kodiert.
Zweckmäßigerweise ist der Pufferspeicher 2 als Ringspeicher gestaltet, d.h., in der Adressierung folgt auf das letzte Wort automatisch das erste. Auf den Pufferspeicher 2 greifen zeitmultiplex drei Einheiten zu. Dabei hat die Abspeicherung aus der Datenquelle die höchste Priorität. Zwischen den Abspeicherzyklen kann der Kompressor auf den Pufferspeicher 2 zugreifen. Der Zugriff des Kompressors wird nur zugelassen, wenn die Quelladresse für die Kompression und die Abspeicheradresse einen festgelegten Abstand überschreiten. Nach Abtrennung einer Teilfolge greift auch die Kodiereinrichtung 6 auf den Pufferspeicher 2 zu.
Die spezielle Ausführung insbesondere der Kodiereinrichtung 6 und der Prüfeinrichtung 2 hängt von der Datenwortgröße (Bit-Anzahl) und von den verwendeten Alphabeten ab. Eine besonders günstige Ausgestaltung der Erfindung ergibt sich, wenn der Satz von Alphabeten so ausgewählt wird, daß die Wortlänge innerhalb jedes eingeschränkten Alphabetes konstant ist und die einzelnen Zeichen des Alphabetes der dualen Zahlendarstellung entsprechen. In diesem Fall kann der Teil der Prüfeinrichtung, welcher den Zustand 4. erkennt, aus einem Satz von Rückwärtszählern (pro eingeschränktem Alphabet ein Zähler) bestehen. In diese Zähler wird jeweils eine Zahl aus dem Festwertspeicher 3 geladen, die die Anzahl der Zeichen des entsprechenden Alphabetes angibt, welche mindestens für eine Zerlegung mit Kompression nötig ist. Existiert also in den Daten eine ununterbrochene Folge von Datenworten der entsprechenden Größe, so erreicht der entsprechende Zähler durch Rückwärtszählen den Wert Null, d.h., die Prüfeinrichtung hat den Zustand 4. erkannt.
Auch die Prüfung, mit welchen Alphabeten das aktuelle Datenwort dargestellt werden kann, ist in diesem Fall einfach. Das neu gelesene Datenwort muß dazu nur nacheinander mit allen Potenzen von 2, die in den Daten auftreten können, angefangen mit der größten, verglichen werden.
Die Kodierung nach Figur 2 geschieht bei signifikanten Teilfolgen durch paralleles Lesen des Alphabetkennzeichens aus dem Zwischenspeicher 5, der Datenworte aus dem Pufferspeicher 1 und des Endekennzeichens aus dem Festwertspeicher 3 jeweils in das erste Schieberegister 7. Von dort aus wird die der Wortgröße entsprechende Anzahl von Bits nach rechts in das zweite Schieberegister 8 und aus dem zweiten Schieberegister 8 auf den Ausgang der Kodiereinrichtung 6 gegeben. Bei Nullfolgen wird auch das Alphabetkennzeichen aus dem Zwischenspeicher 5 in das erste Schieberegister 7 geladen.
Der erste Teil des Zeichens für die Länge wird mittels des Komparators 9 durch Vergleich der Länge mit einer Tabelle im Festwertspeicher 3 ermittelt und aus dem Festwertspeicher 3 in das erste Schieberegister 7 geladen. Der zweite Teil des Zeichens für die Länge wird aus der Länge der Teilfolge entnommen, die aus dem Zwischenspeicher 5 in das erste Schieberegister 7 geladen wird. Das Schieben in das erste Schieberegister 7 und aus dem zweiten Schieberegister 8 auf den Ausgang erfolgt in gleicher Weise wie bei den signifikanten Datenfolgen.
Die Dekodierung wird zweckmäßigerweise mit einer Dekodiereinrichtung nach Figur 3 vorgenommen, wobei sich deren Funktionsweise wie folgt darstellt.
Aus dem ersten Pufferspeicher 10 wird parallel das erste Schieberegister 11 geladen. Zunächst wird das Alphabetkennzeichen mit fester Länge in das zweite Schieberegister 12 geschoben. Aus dem zweiten Schieberegister 12 wird durch Vergleich im Komparator 13 mit einem Vergleichskennzeichen aus dem Festwertspeicher 15 erkannt, ob es sich um eine signifikante Teilfolge oder Nullfolge handelt.
Bei der Nullfolge wird der erste Teil der Nullfolgenlänge mit fester Länge aus dem ersten Schieberegister 11 in das zweite Schieberegister 12 geschoben und in die Steuerung 16 übernommen. Darin ist die Länge des variablen Teiles enthalten. Die entsprechende Bit-Anzahl wird wiederum aus dem ersten Schieberegister 11 in das zweite Schieberegister 12 geschoben. Fester und variabler Teil der Länge werden in der Steuerung 16 kombiniert. Das zweite Schieberegister 12 wird gelöscht. Die der Länge entsprechende Anzahl von Nullen wird in den zweiten Pufferspeicher 14 geschrieben.
Bei einer signifikanten Teilfolge wird die Wortlänge, die aus dem Alphabetkennzeichen abgeleitet wird, in der Steuerung 16 abgelegt. Dann werden die einzelnen Datenworte entsprechend der Wortlänge aus dem ersten Schieberegister 11 in das zweite Schieberegister 12 geschoben und von dort in den zweiten Pufferspeicher 14 geschrieben. Wird ein Datenwort durch Vergleich als Endekennzeichen erkannt, so wird das Schreiben in den zweiten Pufferspeicher 14 unterdrückt und die Steuerung 16 erwartet das nächste Alphabetzeichen.
Das zweite Schieberegister 12 wird jeweils vor dem Lesen eines Wortes gelöscht und das erste Schieberegister 11 wird neu aus dem ersten Pufferspeicher 10 geladen, wenn alle Bits herausgeschoben wurden.
Die vorliegende Ausgestaltung der Erfindung wird im folgenden anhand eines Beispieles erläutert.
Für die Quelldatenfolge soll angenommen werden, daß sie das Ergebnis einer Vorverarbeitung durch Differenzbildung zwischen jeweils 2 aufeinanderfolgenden Worten ist. Die Subtraktionsergebnisse sollen so dargestellt sein, daß die Zahlenwerte mit 2 multipliziert werden und bei negativen Zahlenwerten als letztes Bit eine 1 hinzugefügt wird (z.B. $1 \triangleq 2$, $2 \triangleq 4$, $-2 \triangleq 5$). Durch diese Darstellung lassen sich besonders gut Alphabete aufbauen, in denen das Endekennzeichen kein zusätzliches Bit benötigt, da ja die Subtraktion niemals den Wert −0, was der 1 im Alphabet entspricht, liefert, so daß die 1 als Endekennzeichen verwendet werden kann. Das uneingeschränkte und die eingeschränkten Alphabete sind in Tab. 1 gezeigt.

Tabelle 1:

	Alphabet-kennzeichen	dargestellter Zahlenbereich	Darstellung
uneingeschränktes Alphabet		−127 ... +127 + Endekennzeichen	0 ... 255
	6	−63 ... +63 + Endekennzeichen	0 ... 127
	5	−31 ... +31 + Endekennzeichen	0 ... 63
	4	−15 ... +15 +Endekennzeichen	0 ... 31
eingeschränkte Alphabete	3	−7 ... +7 + Endekennzeichen	0 ... 15
	2	−3 ... +3 + Endekennzeichen	0 ... 7
	1	−1 ... +1 + Endekennzeichen	0 ... 3

Das Längenalphabet wird in Tab. 2 gezeigt.

Tabelle 2:

Alphabetkennzeichen	dargestellte Länge	Darstellung 1. Teil	2. Teil
0	1	000	–
	2 ... 3	001	x
	4 ... 7	010	xx
	8 ... 15	011	xxx
	16 ... 31	100	xxxx
	32 ... 63	101	xxxxx
	64 ... 127	110	xxxxxx
	128 ... 255	111	xxxxxxx

x ist dabei eine binäre 0 oder 1

Es sei nun beispielsweise die Quelldatenfolge 72,3,3,3,4,7,2,2,3,0,3,0 als Ergebnis der Differenzbildung gegeben. Zunächst wird ein Anfangszustand durch die Steuerung eingestellt. Der Anfang der nichtkodierten Datenfolge wird gespeichert, deren Länge auf Null gesetzt, und in die Prüfeinrichtung wird eine Tabelle aus dem Festwertspeicher geladen, die sich auf das uneingeschränkte Alphabet bezieht. Die Tabellen für die Prüfeinrichtung sind eine Funktion der zur Kodierung benutzten Alphabete. Die Anzahl der benutzten Alphabete bestimmt die Wortgröße des Kennzeichens für das benutzte Alphabet. Bei dem behandelten Beispiel ist die Anzahl der Alphabete gleich 8, d. h., das Kennzeichen muß 3 Bit lang sein. In der Tabelle für die Prüfeinrichtung sind für die einzelnen eingeschränkten Alphabete und das Längenalphabet diejenigen Längen enthalten, die jeweils für die 2. Teilfolge notwendig sind, um eine Kompression bei Zerlegung in 2 Teilfolgen zu erreichen. Bei diesem Beispiel ergeben sich die Längen aus dem Mehraufwand von 3 Bit für das Alphabetkennzeichen plus N Bit für das Endekennzeichen der 2. Teilfolge. Wenn z. B. das bisherige Alphabet mit 8 Bit dargestellt werden kann und das der 2. Teilfolge mit 7 Bit, so muß die zweite Teilfolge mindestens 3 + 7 + 1 = 11 Worte bei Einsparung von 1 Bit pro Wort der zweiten Teilfolge sein, um eine Kompression zu erzielen. Die erste in die Prüfeinrichtung gelesene Tabelle für das Alphabet 7 hätte also folgendes Aussehen (die Tabellen bestehen jeweils nur aus der 1. Spalte, die 2. Spalte dient zur Erläuterung):

nötige Länge der 2. Teilfolge	Alphabet-Nr. der 2. Teilfolge
11	6
6	5
4	4
3	3
2	2
2	1
1	0

Darin bedeutet z. B. die 11, daß für die Zerlegung in 2 Teilfolgen die 2. Teilfolge 11 Worte lang sein muß, wenn das Alphabet 6 für die 2. Teilfolge gültig ist.

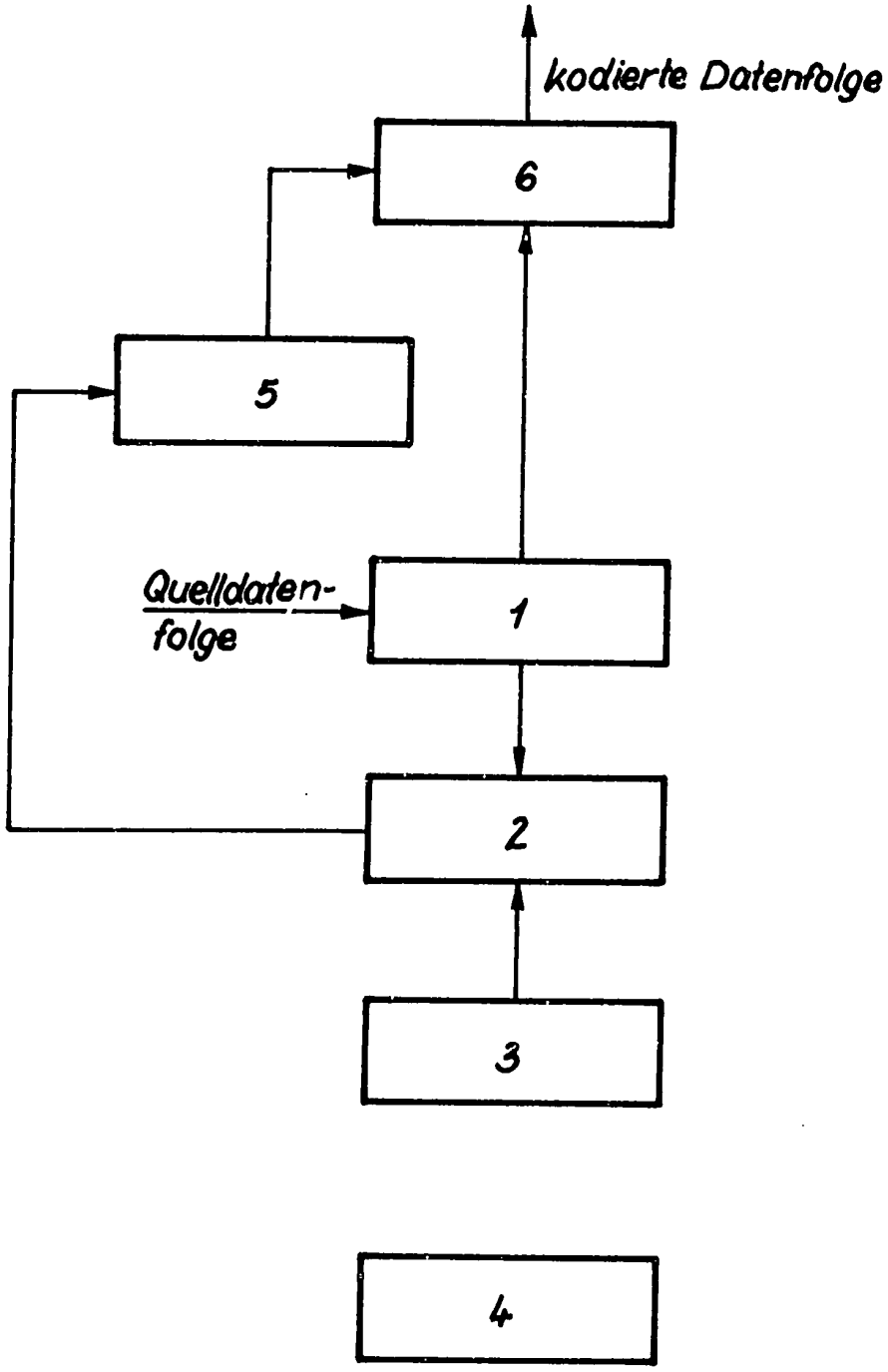
Nach dem Einlesen des 1. Datenwortes „72" wird der Tabellenplatz mit der 11 um 1 erniedrigt, da sich die 72 mit einem 6 Bit großen Alphabet darstellen läßt. Da die Folge erst einen Wert enthält, ist keine Zerlegung möglich, aber die Verwendung eines eingeschränkten Alphabetes. Das heißt, die Prüfeinrichtung erkennt den Zustand 2, und die Steuerung veranlaßt die Erhöhung der Länge des nichtkodierten Folgeteiles, die Änderung des Alphabetzeichens und das Lesen einer neuen Tabelle aus dem Festwertspeicher, die sich nun auf das eingeschränkte Alphabet 6 bezieht. Die neue Tabelle hat dann folgendes Aussehen:

nötige Länge der 2. Teilfolge	Alphabet-Nr. der 2. Teilfolge
10	5
5	4
3	3
2	2
2	1
1	0

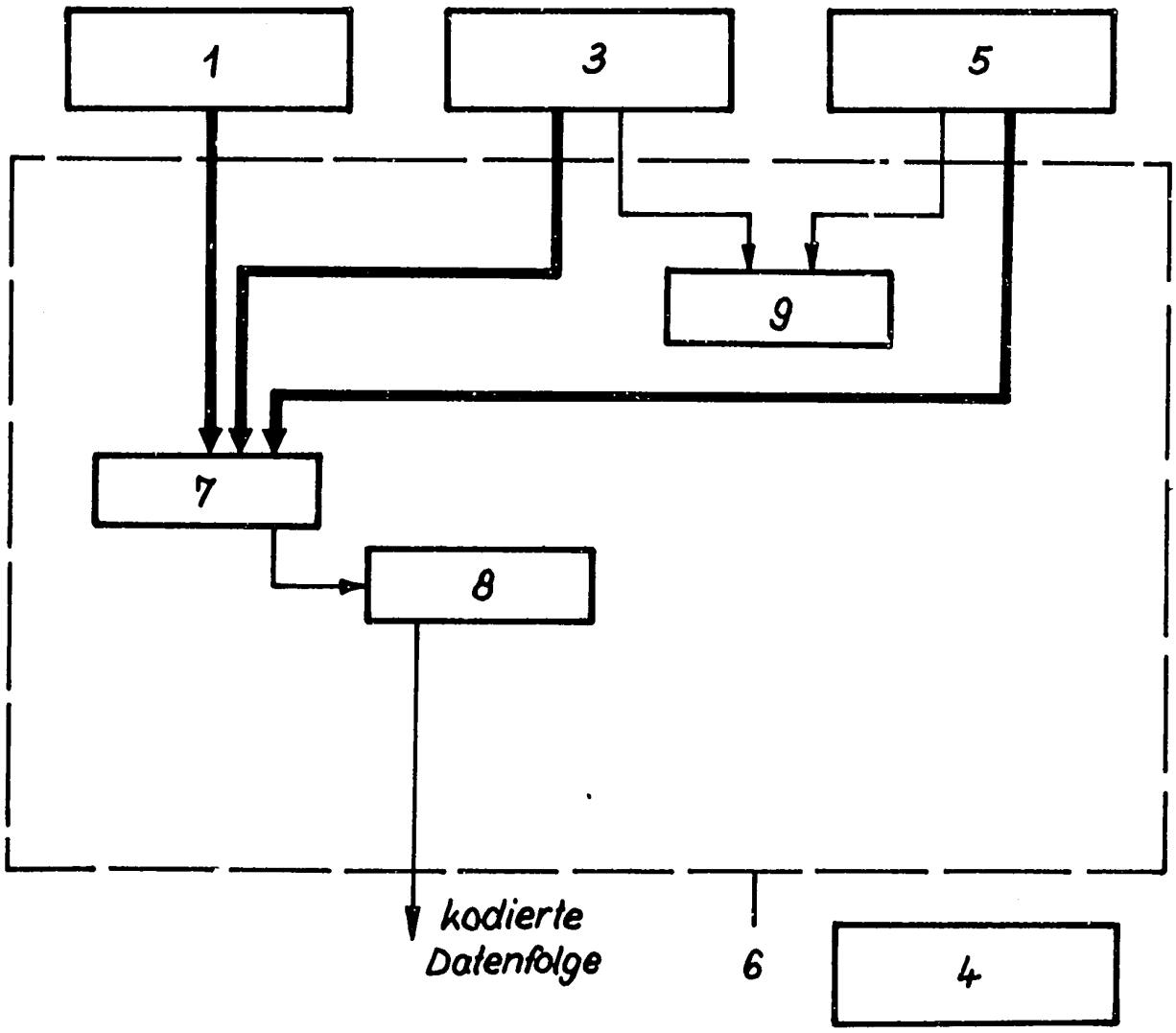
Nach dem Einlesen des 2. Datenwortes „3" werden in der Prüfeinrichtung die ersten 5 Tabellenplätze um 1 erniedrigt, weil die entsprechenden Alphabete für das neue Datenwort verwendet werden könnten, so daß jetzt die Tabelle den folgenden Zustand hat:

nötige Länge der 2. Teilfolge	Alphabet-Nr. der 2. Teilfolge
9	5
4	4
2	3
1	2
1	1
1	0

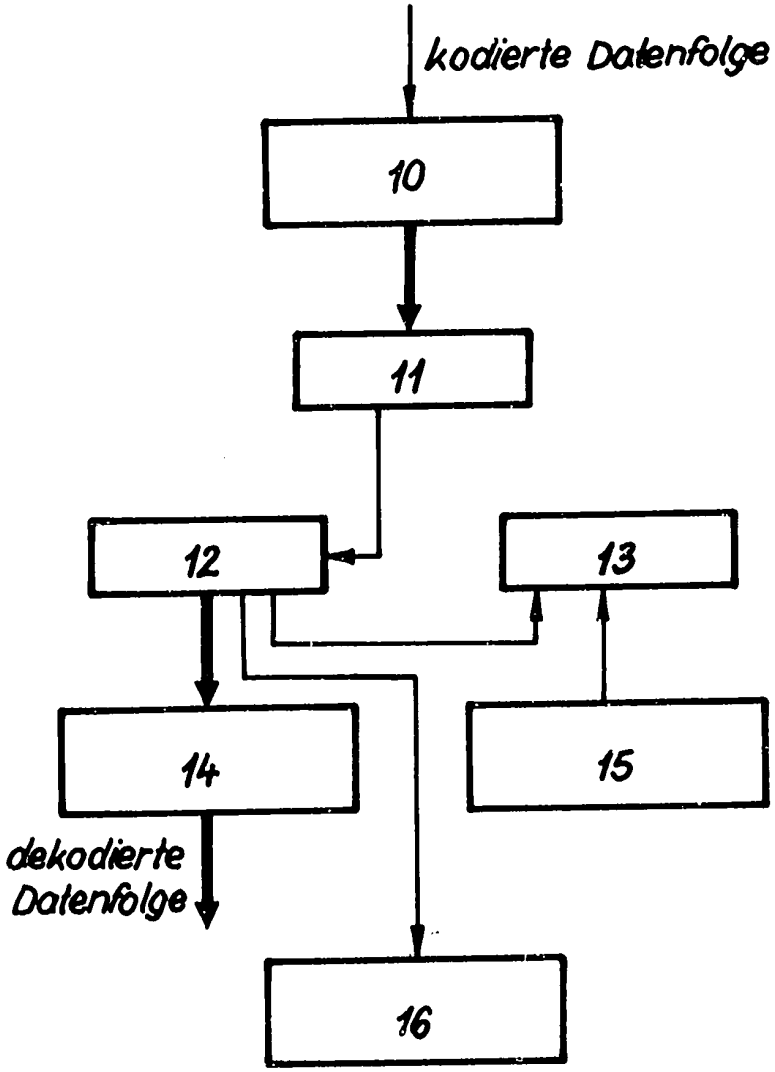
Da aber keiner der erniedrigten Tabellenplätze 0 wurde, was bedeutet, daß bei keinem der möglichen Alphabete die für die Zerlegung erforderliche Länge erreicht wurde, erkennt die Prüfeinrichtung den Zustand 1, und die Steuerung veranlaßt die nötigen Korrekturen. Erst nach dem Lesen des 3. Datenwortes „3" werden infolge der nochmaligen Erniedrigung der ersten 5 Tabellenplätze 2 Tabellenplätze Null. Es gibt also 2 Alphabete, die bei der Zerlegung eine Kompression erzielen würden. Das Alphabet Nr. 1 erzielt die größere Kompression und wird deshalb für die 2. Teilfolge benutzt. Durch das Erkennen des Zustandes 4 wird nun die Zerlegung in 2 Teilfolgen veranlaßt. Aus dem ursprünglichen Wert des Tabellenplatzes, welcher Null wurde, kann die Trennstelle für die Zerlegung berechnet werden. In diesem Fall ist die 2. Teilfolge 2 Worte lang, d.h., die bisher gelesene Folge wird in die Teilfolge 1 (enthält das 1. Datenwort) und die Teilfolge 2 (enthält das 2. und 3. Datenwort) zerlegt. Anfang, Länge und Alphabetkennzeichen der 1. Teilfolge werden nun in den Zwischenspeicher geschrieben, und die Kodierung der 1. Teilfolge wird veranlaßt. Länge und Anfang der 2. Teilfolge werden abgespeichert.
Durch die Arbeit mit den Tabellen bestimmt also die Prüfeinrichtung den aktuellen Zustand. So wird z. B. nach dem Lesen des 5. Datenwortes „4" der Zustand 3 eingestellt, da in dem Alphabet 1 die 4 nicht enthalten ist und die Trennung der Teilfolge 3,3,3 von der Teilfolge 4 keine Kompression erzielt. Auf diese Weise wird die ganze Datenfolge zerlegt und kodiert. Andere Ausgestaltungen der Erfindung, wie z. B. die Benutzung des Längenalphabetes auch für signifikante Teilfolgen oder wie die Benutzung von eingeschränkten Alphabeten mit nichtkonstanter Wortlänge, bedingen einen höheren Hardwareaufwand, erzielen dafür aber auch größere Kompressionsfaktoren.



Figur 1 Kompressor



Figur 2 Kodiereinrichtung



Figur 3 Dekodiereinrichtung