



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 698 20 264 T2 2004.09.09**

(12)

Übersetzung der europäischen Patentschrift

(97) **EP 0 993 649 B1**

(21) Deutsches Aktenzeichen: **698 20 264.3**

(86) PCT-Aktenzeichen: **PCT/US98/14491**

(96) Europäisches Aktenzeichen: **98 935 635.7**

(87) PCT-Veröffentlichungs-Nr.: **WO 99/003058**

(86) PCT-Anmeldetag: **09.07.1998**

(87) Veröffentlichungstag
der PCT-Anmeldung: **21.01.1999**

(97) Erstveröffentlichung durch das EPA: **19.04.2000**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **03.12.2003**

(47) Veröffentlichungstag im Patentblatt: **09.09.2004**

(51) Int Cl.7: **H03M 7/30**
G06T 9/00

(30) Unionspriorität:

52144 P	09.07.1997	US
112668	09.07.1998	US

(73) Patentinhaber:

Quvis, Inc., Topeka, Kan., US

(74) Vertreter:

BOEHMERT & BOEHMERT, 80336 München

(84) Benannte Vertragsstaaten:

**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,
LI, LU, MC, NL, PT, SE**

(72) Erfinder:

GOERTZEN, Kenbe, Topeka, US

(54) Bezeichnung: **GERÄT UND VERFAHREN ZUR ENTROPIEKODIERUNG**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

Gebiet der Erfindung

[0001] Die vorliegende Erfindung betrifft die Signalverarbeitung und insbesondere eine Vorrichtung und ein Verfahren zur statistischen Kodierung bei der Signalverarbeitung.

Hintergrund der Erfindung

[0002] US-A-5,357,250 betrifft ein Verfahren und eine Vorrichtung zum adaptiven Bestimmen von Wahrscheinlichkeiten eines Auftretens jedes Zeichens eines endlichen Alphabets innerhalb einer Zeichenkette. Gemäß diesem Dokument wird eine einzige Wahrscheinlichkeitstabelle auf einer Pro-Zeichen-Basis durch Zählen der historischen Auftritte eines Zeichens aktualisiert.

[0003] WO-A-9 619 045 betrifft eine auf der Syntax basierende arithmetische Kodierungseinheit, die eine Kette von Zeichenwerten unter Verwendung eines arithmetischen Kodierungsmodells, das gemäß der Syntax der Kette ausgewählt wird, kodiert. Dieses Dokument offenbart ein System, bei dem verschiedene Datensätze mit einem unterschiedlichen statistischen Verhalten effektiver kodiert werden können, wenn sie getrennt sind oder so behandelt werden, als ob sie getrennt wären. Es stützt sich auf zuvor bekannte Regeln bezüglich der Syntax bei der Organisation bezüglich der Darstellung von Daten.

[0004] EP-A-0 600 646 betrifft ein Verfahren und eine Vorrichtung zum Herstellen und Authentifizieren eines geheimen Dokuments. Diese Druckschrift lehrt die direkte Verwendung eines adaptiven statistischen Kodierers als eine Verschlüsselungs-Entschlüsselungsvorrichtung.

[0005] Eine Methode zur Durchführung einer sicheren Datenkompression wird von A. Barbir in PROCEEDINGS OF THE TWENTY-NINTH SOUTHEASTERN SYMPOSIUM ON SYSTEM THEORY, 1997, Cookeville, TN, 09.-11. März 1997, Seiten 266 bis 270 beschrieben.

[0006] Ein Nachteil bei der herkömmlichen Verarbeitung elektronischer Signale betrifft die Techniken, die verwendet werden, um die Datenmenge zu reduzieren, die verwendet wird, um das Signal darzustellen. Eine Art, in der die Datenmenge, die erforderlich ist, um ein Signal darzustellen, reduziert wird, ist die Verwendung von Datenkompressionstechniken.

[0007] Die statistische Kodierung ist eine Technik zur Reduzierung der Datenmenge, die zur Darstellung von Signalen verwendet wird. Die statistische Kodierung ändert die Art, in der tatsächliche Werte gespeichert werden, oder definiert sie neu, um die Informationsmenge zu reduzieren, die verwendet wird, um ein Signal darzustellen. Eine einfache Art einer Kodierung, die beispielsweise zur Bildkodierung verwendet wird, verwendet einen Differenzwert anstelle eines tatsächlichen Wertes, um Pixel darzustellen. Insbesondere wird bei der Differenzkodierung die Differenz zwischen jedem Pixel und dem vorhergehenden Pixel in einem Bild bewahrt. Diese Art einer Kodierung bietet eine bessere Kompression, wenn die Bilder stark korreliert sind, da die Pixeldifferenzen dann niedrig sind. Jedoch ist die Differenzkodierung hinsichtlich ihres Mangels viele Bilder oder andere Arten von Datensignalen, die nicht stark korreliert sind, passend zu komprimieren, unzureichend.

[0008] Viele herkömmliche Prozessoren verwenden feste Wahrscheinlichkeiten, um den Zeichenstrom darzustellen. Eine Art einer Kodierung mit einer festen Wahrscheinlichkeit ist die Kodierung mit variabler Länge, wobei Werte basierend auf ihrer erwarteten Auftrittshäufigkeit kodiert werden. Insbesondere wird, falls von einem bestimmten Wert erwartet wird, daß er in den zu verarbeitenden Bildern häufig auftritt, ein kurzer Code (z. B. 1) verwendet, um diesen Wert darzustellen. Falls jedoch von einem Wert erwartet wird, daß er selten auftritt, wird er durch einen langen Code (z. B. 000001) dargestellt. Eine Kodierung mit fester Wahrscheinlichkeit, wie beispielsweise beim Huffman-Typ, bietet die Kompression von Signalen mit Erwartungswertverteilungen. Sie ist jedoch nicht so wirkungsvoll, wenn die Werte stark variieren, was häufig der Fall ist, da Daten schnellen und extrem lokalisierten Änderungen bezüglich der Wahrscheinlichkeitsverteilungen unterworfen sein können.

[0009] Bei einer Verbesserung dieser Art einer Signalverarbeitung wird ein Kodierungssystem verwendet, das historische Kodierung genannt wird, bei der das Signal unter Verwendung einer Statistik kodiert wird, die bezüglich den neuesten bzw. kürzlichen Zeichen im Signal aufgestellt wird. Gemäß einer Umsetzung dieses Systems wird eine Tabelle verwendet, um eine Strichliste der Zeichen im Signal zu führen. Die für jedes Zeichen geführte Strichliste bestimmt, wie jedes Zeichen kodiert wird. Die historischen Kodierungssysteme unterliegen jedoch mehreren Problemen. Zuerst hängt der Verarbeitungsaufwand von der Auflösung der zu verarbeitenden Daten ab. Relativ große Zeichensätze erfordern einen größeren rechnerischen Aufwand. Zusätzlich sind größere Datenumgebungen erforderlich, um eine aussagekräftige Statistik aufzustellen, wenn die Zeichensätze vergrößert werden, um die Signalauflösung zu erhöhen. Somit kann eine große Probengröße einer raschen und genauen Bestimmung der lokalen Wahrscheinlichkeit zuwider laufen, wenn sich bei großen Zeichensätzen die lokale Wahrscheinlichkeit rasch ändert. Da der Zeichensatz groß wird, wird es stufenweise weniger wahrscheinlich, daß ein bestimmtes Zeichen in einer lokalen Umgebung wiederholt wird.

[0010] Anstrengungen, die Probleme der Kodierung basierend auf der Sammlung der Datenverteilung zu lö-

sen, führten zu komplizierten Systemen, die mit Hardware schwierig zu implementieren sind, eine rasche Ausführung oder vernünftige Kosten nicht ohne weiteres zulassen und eine Näherung der lokalen Wahrscheinlichkeit basierend auf geringer Häufigkeit liefern. Diese Kodierverfahren sind ferner nicht dazu in der Lage, einen effektiven Verschlüsselungsschutz für den übertragenen kodierten Strom umzusetzen. Die Kombination der statistischen Kodierung und der Verschlüsselung verbessert sowohl die Sicherheit als auch die Effizienz der Kommunikation.

[0011] Demgemäß besteht ein Bedürfnis für eine Signalverarbeitung weiter fort, bei der ein statistisches Kodierungsschema umgesetzt wird, das Datenmengen wirkungsvoller reduziert und wobei dies in einer Echtzeitumgebung geschieht. Zusätzlich besteht das Bedürfnis, die Verschlüsselungstechnologie während der Signalverarbeitung wirkungsvoll umzusetzen, weiter fort, so daß ein nicht autorisierter Empfang oder eine Verwendung des Zeichenstroms verhindert werden kann.

Abriss der Erfindung

[0012] Die vorliegende Erfindung ist ein System und ein Verfahren zum Kodieren von Zeichen in einem Zeichenstrom, um die Datenmenge zu komprimieren, die erforderlich ist, um ein Signal darzustellen, das dem Zeichenstrom entspricht. Wenn ein aktuelles Zeichen kodiert wird, wird ein Modellierungswert durch ein Modellierungsmodul in Abhängigkeit von einer Abtastung von Zeichen berechnet, die für ein aktuelles Zeichen lokal sind. Das Klassifizierungsmodul verwendet den Modifizierungswert, um eine Wahrscheinlichkeitsfunktion, die dem Strom entspricht, zu erzeugen oder abzurufen. Die Wahrscheinlichkeitsfunktion wird verwendet, um die Wahrscheinlichkeit zu berechnen, daß das aktuelle Zeichen das nächste Zeichen im Strom gewesen wäre. Die Kodierung des aktuellen Zeichens hängt von der Wahrscheinlichkeit ab, daß das aktuelle Zeichen das nächste Zeichen im Strom gewesen wäre. Ein Kodierer verwendet mehr oder weniger Bits, um das Zeichen basierend auf der zur Verfügung gestellten Wahrscheinlichkeit zu kodieren.

[0013] Das System und das Verfahren können des weiteren dazu verwendet werden, die Zeichen unter Verwendung desselben Prozesses zur Bestimmung eines Modellierungswertes und zum Abbilden des Modellierungswertes auf eine Wahrscheinlichkeitsfunktion, wie vom Kodierer durchgeführt, zu dekodieren. Die Eigenschaften der Bits in dem kodierten Zeichen werden vom Dekodierer verwendet, um die vom Kodierer abgeschätzte Wahrscheinlichkeit zu bestimmen. Die Wahrscheinlichkeitsfunktion, die vom Dekoder wiedererlangt wird, wird in einem umgekehrten Look-Up verwendet, um den ursprünglichen Zeichenwert, der kodiert wurde, zu bestimmen. Des weiteren kann ein Verschlüsselungswert verwendet werden, um die geeignete Modellierungs-, Klassifizierungs- oder Wahrscheinlichkeitsfunktion auszuwählen, um ein nicht autorisiertes Dekodieren des Stroms zu verhindern.

[0014] Insbesondere stellt die vorliegende Erfindung ein Verfahren zum Modellieren von Zeichen in einem Zeichenstrom für eine Kompression nach Anspruch 1, ein Verfahren zum Dekodieren von Zeichen nach Anspruch 8, ein System zum Kodieren von Zeichen und ein System zum Dekodieren von Zeichen nach Anspruch 14 bzw. 17 und ein computerlesbares Medium nach Anspruch 19 bereit.

Kurze Beschreibung der Zeichnungen

[0015] Diese und andere detailliertere und speziellere Gegenstände und Merkmale der vorliegenden Erfindung werden umfassender in der folgenden Beschreibung offenbart, die in Bezug zu den begleitenden Zeichnungen steht, in welchen:

[0016] **Fig. 1** ein Blockdiagramm des Kodierungssystems der vorliegenden Erfindung ist,

[0017] **Fig. 2** ein Schema einer Kodierungsvorrichtung darstellt, mit der ein Zeichenstrom gemäß der vorliegenden Erfindung kodiert wird,

[0018] **Fig. 3** ein Schema einer Dekodierungsvorrichtung darstellt, mit der ein Zeichenstrom gemäß der vorliegenden Erfindung dekodiert wird,

[0019] **Fig. 4** ein Probensatz von Wahrscheinlichkeitsfunktionen ist, die verwendet werden, um den Zeichenstrom zu kodieren und dekodieren;

[0020] **Fig. 5** ein Flußdiagramm ist, das das Verfahren zum Kodieren von Daten gemäß der vorliegenden Erfindung veranschaulicht,

[0021] **Fig. 6** eine Tabelle ist, die Probenwerte aufweist, welche den Prozeß zur Kodierung eines Zeichenstroms veranschaulichen.

Beschreibung der bevorzugten Ausführungsformen

[0022] Die vorliegende Erfindung ist ein System und ein Verfahren zum Komprimieren und Kodieren von Signalen in einem Zeichenstrom. Die Signale können unter Verwendung von Verschlüsselungscodes kodiert werden, die ein nicht autorisiertes Dekodieren des Stroms verhindern. Gemäß der Erfindung werden sowohl

Zeichenströme komprimiert, die in Echtzeit übertragen werden, als auch Daten in einem diskreten Datensatz. Auf den Zeichenstrom werden Bearbeitungsschritte ausgeführt, die eine Datenmodellierung und -klassifizierung ermöglichen, so daß die geeignete Kodierung der Zeichen im Zeichenstrom ausgeführt werden kann. Vorzugsweise ist der dem statistischen Kodierer zugeführte Datenstrom relativ null-zentriert, was immer dann der Fall ist, wenn eine geeignete Transformation, wie z. B. eine Elementarwellentransformation auf den Zeichenstrom ausgeführt wird.

[0023] **Fig. 1** ist ein Blockdiagramm des erfindungsgemäßen Kodierungssystems. Das System umfaßt: ein Modellierungsmodul **110**, ein Klassifizierungsmodul **120**, ein Wahrscheinlichkeitsabbildungssystem **130** und einen statistischen bzw. Entropiekodierer **140**. Der statistische Kodierer **140** empfängt eine Wahrscheinlichkeitsfunktion basierend auf den Ergebnissen von auf den Zeichenstrom durch das Modellierungsmodul **110**, das Klassifizierungsmodul **120** und das Wahrscheinlichkeitsabbildungssystem **130** ausgeführten Bearbeitungsschritten. Der statistische Kodierer **140** kann des weiteren einen Verschlüsselungsalgorithmus verwenden, um einen komprimierten und verschlüsselten Zeichenstrom zu liefern.

[0024] Das Modellierungsmodul **110** empfängt den quantifizierten Zeichenstrom. Das Modellierungsmodul **110** verwendet die quantifizierten Zeichen, um einen Modellierungswert oder einen Satz von Modellierungswerten, die den Strom charakterisieren, zu berechnen. Bei der bevorzugten Ausführungsform ist der Modellierungswert eine Zahl, die basierend auf den am kürzlichsten quantifizierten Zeichenwerten berechnet wird. Eine lokale Abschätzung des Modellierungswertes wird unter Verwendung einer Funktion $B[t]$ ausgeführt, welche den lokalen Modellierungswert an einem Punkt t im Zeichenstrom annähert. Beispielsweise kann der Modellierungswert unter Verwendung einer Funktion bestimmt werden, die auf eine lokale Probe des Zeichenstroms wirkt, wie beispielsweise:

$$\text{Modellierungswert} = B [s(t - 1), s(t - 2) \dots s(t - n)], \quad (\text{Gl. 1})$$

wobei der Modellierungswert zur Zeit t durch die Funktion $B[s(t)]$ angenähert wird, die auf die letzten n Zeichen wirkt. Eine bevorzugte Funktion, die in dem Modellierungsmodul **110** verwendet werden kann, um die Erzeugung eines Modellierungswertes zu unterstützen und die in einem breiten Bereich von Bildtypen gut funktioniert, ist ein kurzfristig gewichteter Mittelwert der durch die Zeichen dargestellten Größen. Die folgende Gleichung kann im Modellierungsmodul **110** verwendet werden, um einen gewichteten Mittelwert zu erzeugen:

$$ws = \sum_{i=1}^n \frac{|s(t - n)|}{n} \quad (\text{Gl. 2})$$

wobei ws die gewichtete Summe der Zeichenwerte zur Zeit $t - 1$ bis $t - n$ ist und n die Anzahl von Zeichen ist, die in der gewichteten Summe umfaßt sind. Die Gleichung mit der gewichteten Summe gibt dem kürzlichsten bzw. neuesten Zeichen das größte Gewicht, da das letzte Objekt häufig als das beste Vorhersagemittel für die Funktion dient, mit der die tatsächliche Wahrscheinlichkeit des nächsten Wertes im Strom angenähert wird. Gemäß einer Ausführungsform ist die gewichtete Summe der Modellierungswert **115**, der dem Klassifizierungsmodul **120** zugeführt wird. Alternativ könnte die gewichtete Summe ebenso in einer weiteren Berechnung verwendet werden, um einen Modellierungswert **115** zu erzeugen. Dies umfaßt die Verwendung einer speziellen Funktion, um den Modellierungswert **115** für einige Werte zu erzeugen, während eine unterschiedliche Funktion verwendet wird, um den Modellierungswert **115** für andere Werte zu berechnen. Beispielsweise könnte eine logarithmische Funktion auf die gewichtete Summe angewandt werden, falls sie mehr als 1 beträgt, während eine direkte Abbildung zwischen der gewichteten Summe und dem Modellierungswert **115** verwendet werden könnte, wenn die Summe weniger als 1 beträgt. Darüber hinaus könnte der Modellierungswert **115**, obwohl er als ein einzelner numerischer Wert beschrieben wurde, mehrere Werte umfassen, die Informationen über die Objekte im Zeichenstrom liefern. Der Modellierungswert **115** könnte einen Abweichungswert umfassen, der beispielsweise die Abweichung zwischen den Werten der quantifizierten Zeichen anzeigt. Sobald der Modellierungswert **115** erzeugt wurde, wird er zum Klassifizierungsmodul **120** weitergegeben.

[0025] Das Klassifizierungsmodul **120** empfängt den Modellierungswert **115** vom Modellierungsmodul **110**. Das Klassifizierungsmodul **120** verwendet den Modellierungswert **115**, um den Zeichenstrom auf eine Wahrscheinlichkeitsfunktion zu klassifizieren oder abzubilden. Das Klassifizierungsmodul **120** könnte eine Wahrscheinlichkeitsfunktion basierend auf dem Modellierungswert **115** erzeugen, wie beispielsweise unter Verwendung des Wertes als einen Koeffizienten in der Funktion oder könnte den Wert einfach dazu verwenden, um den Zeichenstrom auf eine Wahrscheinlichkeitsfunktion abzubilden. Falls der Wert dazu verwendet wird, den Zeichenstrom auf eine Wahrscheinlichkeitsfunktion abzubilden, kann die Wahrscheinlichkeitsfunktion in einer kontinuierlichen Weise unter Verwendung eines jeweiligen neuen Modellierungswertes **115** erzeugt werden, um Anpassungen an die Wahrscheinlichkeitsfunktion durchzuführen. Darüber hinaus könnte das Klassifizierungsmodul **120** den Zeichenstrom auf eine von mehreren diskreten Wahrscheinlichkeitsfunktionen abbilden.

[0026] Beispielsweise könnte ein Satz von Glockenfunktionen mit unterschiedlichen Formen bereitgestellt

werden, falls die Daten sich einer Glockenfunktion annähern. Der Modellierungswert **115** könnte basierend auf dem Betrag, um welchen sich die quantifizierten Zeichenwerte voneinander unterscheiden, erzeugt werden, und könnte verwendet werden, um die Form der Funktion zu definieren. Beispielsweise könnte der Zeichenstrom auf eine Glockenfunktion abgebildet werden, die groß und schmal ist, falls die meisten Werte einander ähnlich sind. Besteht eine starke Abweichung unter den Werten könnte der Zeichenstrom auf eine Glockenfunktion abgebildet werden, die klein und breit ist. In einem derartigen Fall könnte der Modellierungswert **115** ein Maß der Änderung zwischen den Zeichenwerten im Zeichenstrom unter Verwendung eines berechneten Wertes, wie beispielsweise der Standardabweichung der Zeichenwerte sein.

[0027] Die Wahrscheinlichkeitsfunktion sowie Parameter, die die Form der Funktion definieren, werden an das Wahrscheinlichkeitsabbildungssystem **130** weitergegeben. Das Wahrscheinlichkeitsabbildungssystem **130** verwendet die Wahrscheinlichkeitsfunktion, um die Wahrscheinlichkeit eines oder mehrerer Zeichen im Zeichensatz zu bestimmen. Die Wahrscheinlichkeit eines Objektes bestimmt, wie viele Bits verwendet werden, um das Objekt im Zeichenstrom zu kodieren. Falls das Objekt eine niedrige Auftrittswahrscheinlichkeit aufweist, können mehr Bits verwendet werden, um das Objekt zu kodieren. Falls das Wahrscheinlichkeitsabbildungssystem **130** ermittelt, daß das Objekt hochwahrscheinlich ist, können andererseits wenige oder sogar weniger als 1 Bit verwendet werden, um die Information über den Zeichenstrom zu übertragen. In beiden Fällen werden die unter Verwendung der Wahrscheinlichkeitsfunktion generierten Wahrscheinlichkeitswerte vom Wahrscheinlichkeitsabbildungssystem **130** zum Kodierer **140** weitergegeben.

[0028] Der Kodierer **140** kodiert oder dekodiert die Zeichen im Zeichenstrom basierend auf den Wahrscheinlichkeiten, die durch das Wahrscheinlichkeitsabbildungssystem **130** geliefert werden. Der Kodierer **140** kann irgendein parametrisierter Kodierer sein, der einen Eingangszeichenstrom basierend auf einer Wahrscheinlichkeit oder einer Wahrscheinlichkeitsfunktion kodieren kann. Der Kodierer **140** empfängt den Zeichenstrom und kodiert oder dekodiert die Zeichen im Strom entsprechend.

[0029] Im folgenden ist unter Bezugnahme auf **Fig. 2** ein Schema einer Kodiervorrichtung gemäß der vorliegenden Erfindung gezeigt, mit der ein Zeichenstrom kodiert wird. Die Kodiervorrichtung umfaßt ein Modellierungssystem **210**, einen Klassifizierer **220**, eine Wahrscheinlichkeitskarte **230**, einen statistischen Kodierer **240** und eine Steuerungsvorrichtung **285**. Ein Zeichenstrom wird vom Modellierungssystem **210** empfangen und die quantifizierten Zeichenwerte werden verwendet, um eine gewichtete Summe zu erzeugen. Die gewichtete Summe wird nur als ein Beispiel der Funktion verwendet, die verwendet werden kann, um einen Modellierungswert **115** zu erzeugen und es könnten wechselnde oder zusätzliche Funktionen verwendet werden. Bei diesem Beispiel wird eine weitere Berechnung unter Verwendung des Logarithmus zur Basis **2** der gewichteten Summe verwendet, um den Modellierungswert **115** zu erzeugen, der zum Klassifizierer **220** weitergegeben wird.

[0030] Das Modellierungssystem **210** weist auch einen Nullzählungskodierer **260** auf. Der Nullzählungskodierer **260** überwacht den Strom nach Ketten von Nullen. Falls ein triggerndes Ereignis auftritt, wie die Berechnung der gewichteten Summe, die gleich 0 ist, beginnt der Nullzählungskodierer **260** mit der Verfolgung der Anzahl von Nullen. Die erste nach dem auslösenden Ereignis auftretende 0 bewirkt, daß der Nullzählungskodierer **260** einen Modellierungswert **115** überträgt, der anzeigt, daß der Nullzählungskodierer **260** einen Modellierungswert **115** sowie die gewichtete Summenfunktion **250** liefert wird. Falls das nächste Zeichen ebenso eine 0 ist, wartet der Nullzählungskodierer bis das nächste Zeichen im Strom empfangen wurde, bevor er einen Modellierungswert **115** überträgt. Falls das Zeichen keine Null ist, wird der Modellierungswert **115**, der unter Verwendung des Logarithmus der gewichteten Summe erzeugt wurde, wiederverwendet, um den Strom zu kodieren. Falls der nächste Wert ebenso Null ist, überträgt der Nullzählungskodierer **260** jedoch einen -1 Modellierungswert **115**. Dieser Modellierungswert **115** zeigt an, daß zwei Nullwerte aufgetreten sind. Mit dem Fortdauern des Stroms von Nullen fährt der Nullzählungskodierer **260** fort, anwachsend große Sätze von Nullen zu zählen, bevor er einen Modellierungswert **115** überträgt. Beispielsweise überträgt der Nullzählungskodierer **260** den nächsten Modellierungswert **115** von -2 solange nicht, bis vier Nullen aufgetreten sind, nachdem ein Modellierungswert **115** von -1 basierend auf einer Kette von zwei Nullen in dem Strom übertragen wurde. Dieser Prozeß dauert an, wobei die Anzahl von Nullen in dem Satz um Potenzen von zwei anwachsen, bevor der nächste Modellierungswert **115** übertragen wird. Falls ein Wert, der nicht Null ist, sich im Strom befindet, bevor der nächste Satz von Nullen vollständig ist, wird der Modellierungswert **115** verwendet, der gemäß der Berechnung der gewichteten Summe berechnet wurde.

[0031] Der Klassifizierer **220** verwendet den (die) Modellierungswert(e) **115**, der (die) vom Modellierungssystem **210** geliefert werden, um eine Wahrscheinlichkeitsfunktion zu generieren oder auszuwählen. Der Klassifizierer **220** führt die Funktionen des Klassifizierungsmoduls **120** und des Wahrscheinlichkeitsabbildungssystems **130**, wie mit Bezugnahme auf **Fig. 1** beschrieben wurde, aus. Bei dieser Ausführungsform ruft der Klassifizierer **220** eine der gespeicherten Wahrscheinlichkeitsfunktionen, die auf dem Modellierungswert **115** basieren, ab und überträgt die Wahrscheinlichkeitskarte **230** zur Code-Look-Up-Tabelle **275** des statistischen Kodierers. Die Wahrscheinlichkeit des aktuellen Zeichens wird aus der Code-Look-Up-Tabelle **275** abgerufen und zum Codeassembler **280** übertragen. Der Codeassembler **280** empfängt die Wahrscheinlichkeit und andere

- Parameter, die verwendet werden, um das Zeichen zu kodieren und gibt das Zeichen in den kodierten Strom **150** aus.
- [0032] Eine Nulldetektionsschaltung **265** ist ebenso im statistischen Kodierer **240** vorgesehen, um die Ketten von Nullen zu kodieren. Wenn der Nullzählungskodierer **240** einen Modellierungswert **115** zum Klassifizierer **220** überträgt, kodiert der statistische Kodierer **240** den Satz von Nullen entsprechend diesem Modellierungswert **115** als ein einzelnes kodiertes Zeichen.
- [0033] Der Modellierungswert **115** bestimmt die Anzahl von Nullen in dem durch das kodierte Zeichen dargestellten Satz. Beispielsweise zeigt ein Modellierungswert **115** von -3 an, daß acht Nullen (2^3) durch eine einzelne kodierte Null im Strom **150** dargestellt sind. Somit zählt die Nulldetektionsschaltung **265** die Nullen entsprechend dem Satz und entfernt sie und gibt eine einzelne nicht kodierte Null für jeden kompletten Satz an den Codeassembler weiter, um kodiert zu werden. Eine einzelne durch den Codeassembler **280** kodierte Null wird in dem Strom **150** eingefügt und stellt alle Nullen im Satz entsprechend dem letzten Modellierungswert **115** dar. Alle Werte, die nicht Null sind, oder Nullwerte, die normal kodiert wurden, werden von der Nulldetektionsschaltung **265** ignoriert.
- [0034] Die Kodiervorrichtung umfaßt des weiteren eine Steuerungsvorrichtung **285**. Die Steuerungsvorrichtung **285** überwacht die Übertragungen zwischen dem Modellierungssystem **210**, dem Klassifizierer **220** und dem Kodierer **240**, um sicherzustellen, daß die Signale sich in der richtigen zeitlichen Phase befinden. Die veranschaulichte Kodiervorrichtung ist lediglich eine einzelne Ausführung der vorliegenden Erfindung, da die speziellen im System beinhalteten Komponenten und die von jeder Komponente verwendeten Funktionen zwischen den Ausführungen variieren können und werden.
- [0035] Im folgenden wird auf **Fig. 3** Bezug genommen. Dort ist eine schematische Zeichnung einer Vorrichtung gezeigt, mit der ein Zeichenstrom gemäß der vorliegenden Erfindung dekodiert wird. Die Dekodiervorrichtung umfaßt ein Modellierungssystem **210**, einen Klassifizierer **220**, eine Wahrscheinlichkeitskarte **230** und einen statistischen Dekodierer **340**. Der kodierte Strom **150** wird am Anfang vom statistischen Dekodierer **340** empfangen. Der Dekodierer **340** entnimmt eine Wahrscheinlichkeitskarte **230** und verwendet die Charakteristika des kodierten Zeichens, um das Zeichen zu dekodieren. Der Kodierprozeß umfaßt das Abschätzen der Wahrscheinlichkeit des Zeichens und das Kodieren basiert auf dieser Wahrscheinlichkeit. Beim Dekodierungsprozeß wird der Prozeß umgekehrt durchgeführt. Die Charakteristika der Bits, die verwendet werden, um die Daten zu kodieren, ermöglichen, daß der Dekodierer **340** die Wahrscheinlichkeit berechnet, die verwendet wurde, um die Daten zu kodieren. Der Dekodierer **340** verwendet die Wahrscheinlichkeitskarte **230**, um die berechnete Wahrscheinlichkeit mit dem ursprünglichen Zeichen abzugleichen. Das ursprüngliche Zeichen ersetzt das kodierte Zeichen im resultierenden dekodierten Strom **105**.
- [0036] Wenn der Strom **150** am Anfang vom Dekodierer **300** empfangen wird, haben der Klassifizierer **220** und das Modellierungssystem **210** nicht genug Werte empfangen, um eine Wahrscheinlichkeitskarte **230** zu erzeugen, die dem Strom **105** entspricht, der empfangen wurde. Um die ersten wenigen Zeichen im Strom **150** zu dekodieren, kann eine Voreinstellungswahrscheinlichkeitskarte **230** oder -funktion **250** verwendet werden. Dies gewährleistet, daß der Kodierer **240** und der Dekodierer **340** dieselben Wahrscheinlichkeiten für dieselben Zeichen berechnen. Bei vielen Anwendungen, wie beispielsweise einem Streamer-Video, beinhaltet der Zeichenstrom einen Anfangsrahmen, der verwendet werden kann, um spezielle Zeichen zu speichern, die die Funktion kennzeichnen, die verwendet werden soll. Beispielsweise könnte ein Anfangszeichen, wie beispielsweise eine Null dazu verwendet werden, den Strom zu beginnen, und da der Dekodierer weiß, daß der kodierte Wert eine Null ist, kann der Dekodierer die Anzahl von Bits, die verwendet wird, um die Null zu kodieren, als einen Indikator der Funktion verwenden, die verwendet wird, um die Wahrscheinlichkeitskarte **230** zu generieren. Das Dekodierungsschema kann ebenso einen konvergierenden Fehlerwert bereitstellen, so daß die Werte, die unter Verwendung der nicht korrekten Wahrscheinlichkeitskarten dekodiert wurden, mit der Zeit einen Modellierungswert **115** generieren, der die richtige Wahrscheinlichkeitskarte **230** erzeugt.
- [0037] Des weiteren kann die Auswahl der Anfangswahrscheinlichkeitsfunktion auf einem Verschlüsselungswert basieren. Der Verschlüsselungswert kann durch irgendein Verschlüsselungsmittel unter Verwendung verschiedener Variablen erzeugt werden, um den Wert zu berechnen. Beispielsweise könnten die Seriennummer und der Benutzername in einem Softwareprogramm in einem Verschlüsselungssystem mit einem öffentlichen Schlüssel verwendet werden. Verschlüsselungssysteme mit öffentlichem Schlüssel setzen voraus, daß jede Person zwei Schlüssel, einen öffentlichen Schlüssel und einen privaten Schlüssel, erhält. Die Schlüssel ermöglichen es einer Person entweder eine Nachricht zu verschlüsseln (sperrern) oder eine kodierte Nachricht zu entschlüsseln (entsperren). Der öffentliche Schlüssel jeder Person wird veröffentlicht und der private Schlüssel wird geheim gehalten. Die Nachrichten werden unter Verwendung des öffentlichen Schlüssels des gewünschten Empfängers verschlüsselt und können und unter Verwendung des privaten Schlüssels entschlüsselt werden, der unter Verwendung von Variablen, wie beispielsweise der Seriennummer und des Benutzernamens erzeugt wird. Da die Bestimmung der Wahrscheinlichkeitsfunktion, die verwendet wird, um die Werte zu dekodieren auf der richtigen Entschlüsselung der Zeichen basiert, kann dieses Verfahren verwendet werden, um einen nicht autorisierten Empfang des Zeichenstroms zu verhindern. Die Verschlüsselung mit öffentlichem

Schlüssel wird als ein Beispiel eines Verschlüsselungsverfahrens verwendet, jedoch können auch andere Verschlüsselungsverfahren verwendet werden.

[0038] Sobald eine Anfangswahrscheinlichkeitsfunktion abgerufen wurde und die entsprechende Wahrscheinlichkeitskarte **230** zum Dekoder **340** übertragen wurde, wird der kodierte Zeichenstrom **150** unter Verwendung desselben Prozesses wie bei der Kodierung dekodiert. Die quantifizierten Werte der dekodierten Zeichen werden zum Modellierungssystem **210** übertragen und es wird ein Modellierungswert **115** in derselben Weise wie auf der Kodierungsseite erzeugt. Der Modellierungswert **115** wird dann zum Klassifizierer **220** weitergegeben und die entsprechende Wahrscheinlichkeitsfunktion wird verwendet, um eine Wahrscheinlichkeitskarte **230** zu generieren. Die Charakteristika der Bits, die verwendet werden, um die Daten zu kodieren, ermöglichen es, daß der Dekodierer **340** die Wahrscheinlichkeit berechnet, die für das Zeichen berechnet wurde, als es kodiert wurde. Die Wahrscheinlichkeitskarte **230** wird vom Dekodierer empfangen und wird in die Code-Lookup-Tabelle **275** eingegeben. Der Code-Disassembler **340** verwendet dann die Code-Lookup-Tabelle **275**, um die berechnete Wahrscheinlichkeit mit dem ursprünglichen Zeichen abzugleichen. Die resultierenden dekodierten Werte werden im dekodierten Zeichenstrom **105** übertragen. Zusätzlich wird ein Nullzählungsmechanismus **270** vorgesehen, um Nullen zu dekodieren, die unter Verwendung des Nullzählungssystems kodiert wurden. Wenn im dekodierten Strom **105**, der vom Modellierungssystem **210** empfangen wird, ein triggerndes Ereignis auftritt, werden die Werte zum Code-Disassembler **380** übertragen. Der Nullzählungsmechanismus liefert dem Code-Disassembler **380** einen Nullzählimpuls, der die Anzahl von Nullen kennzeichnet, die in dem Strom **105** eingegeben werden.

[0039] Im folgenden wird auf **Fig. 4** Bezug genommen. Dort ist ein Beispielsatz von Wahrscheinlichkeitsfunktionen gezeigt. Das Klassifizierungsmodul **120** kann die Funktionen basierend auf dem empfangenen Modellierungswert **115** erzeugen oder kann Funktionen speichern, die auf spezielle Modellierungswerte **115** abgebildet sind. Der Modellierungswert **115** charakterisiert den Strom, der die Identifizierung der besten entsprechenden Funktion ermöglicht.

[0040] Wie die Darstellung zeigt, reflektieren die Funktionen **410**, **420**, **430** verschiedene Wahrscheinlichkeiten für identische quantifizierte Werte. Falls der Modellierungswert **115** den Zeichenstrom auf die Funktion **420** abbildet, wird beispielsweise die Wahrscheinlichkeit eine Null zu erhalten mit ungefähr 0,01 berechnet. Falls mehr Nullen empfangen werden, wird der Modellierungswert **115** abfallen und der neue Modellierungswert **115** wird auf die Funktion **410** abbilden, die eine ansteigende Wahrscheinlichkeit (ungefähr 0,02) dafür anzeigt, daß ein Nullzeichen das nächste Zeichen im Strom ist. Dies führt zu weniger Bits, die verwendet werden, um die Nullen im Strom zu kodieren und somit steigt die Kodierungseffektivität an.

[0041] Ein Flußdiagramm, das das Verfahren zum Kodieren von Daten gemäß der vorliegenden Erfindung veranschaulicht, ist in **Fig. 5** gezeigt. Das Verfahren beginnt, wenn ein Zeichenstrom vom Modellierungsmodul **110** empfangen wird. Das Modellierungsmodul **110** empfängt **505** ein aktuelles Zeichen. Das Modellierungsmodul **110** verwendet den quantifizierten Wert, um einen Modellierungswert **115** zu berechnen **510**. Eine Funktion, wie z. B. eine gewichtete Mittelwertfunktion, wird verwendet, um den Modellierungswert **115** zu berechnen. Die Art der verwendeten Funktion und die verwendeten Parameter der Funktion können abhängig von den Charakteristika der Daten, die kodiert werden, variieren. Beispielsweise kann der Parameter, der die Anzahl von Elementen steuert, die verwendet werden, um den gewichteten Mittelwert zu berechnen, basierend auf dem Zeichenstrom variieren, falls die Funktion einen gewichteten Mittelwert berechnet. Ein gewichteter Mittelwert kann unter Verwendung von mehr Werten berechnet werden, wenn ein Zeichenstrom gelegentlich Spitzen aufweist, und es können weniger Werte zur Berechnung des gewichteten Mittelwertes für einen Strom verwendet werden, der dazu tendiert, stabil zu sein. Falls ein gewichteter Mittelwert verwendet wird, ist der berechnete Modellierungswert **115** ein einzelner numerischer Wert. Es können jedoch andere Funktionen einen oder mehrere Werte erzeugen, die den Zeichenstrom modellieren. Der gewichtete Mittelwert könnte des Weiteren in einer zweiten Funktion verwendet werden, um den Modellierungswert **115** zu berechnen.

[0042] Bei der bevorzugten Ausführungsform wird ein Modellierungswert **115** berechnet **510**, indem der Logarithmus zur Basis **2** eines gewichteten Mittels der quantifizierten Zeichen genommen wird und der logarithmische Wert auf die nächste ganze Zahl gerundet wird. Der berechnete Modellierungswert **115** wird zum Klassifizierungsmodul **120** übertragen. Andere Funktionen könnten jedoch ebenso verwendet werden, um die Modellierungswerte **115** basierend auf speziellen Charakteristika des Stroms zu berechnen. Dies ist insbesondere der Fall, wenn eine Modellierungsfunktion nicht über bestimmte Wertebereiche verwendet werden kann. Beispielsweise liefert eine Logarithmusfunktion keine Information darüber, wann das gewichtete Mittel gleich 0 ist. Unter diesen Umständen kann das Modellierungsmodul **110** eine zusätzliche Modellierungsfunktion liefern, um einen weiteren Modellierungswert **115** zu erzeugen. Beispielsweise wird ein zweiter Modellierungswert **115** unter Verwendung der Anzahl von Nullen im Strom als ein korrigierter Modellierungswert **115** berechnet. Falls das gewichtete Mittel gleich 0 ist und das nächste Zeichen ebenfalls eine Null ist, kann das Modellierungsmodul **110** einen Modellierungswert **115** von Null zum Klassifizierungsmodul **120** weitergeben. Der Modellierungswert **115**, der unter Verwendung der alternierenden Funktion erzeugt wurde, kann solange verwendet werden, wie die gewichtete Summe Null ist. Dies ermöglicht eine größere Flexibilität bei der Kompression, da bestimmte

Teile des Zeichenstroms in einer Weise variieren können, die mit einer verbesserten Modellierungswert-**115**-Funktion antizipiert werden kann. Weitere Einzelheiten des Prozesses der Bewegung zwischen Modellierungsfunktionen werden mit Bezugnahme auf die Tabelle in **Fig. 6** beschrieben.

[0043] Das Klassifizierungsmodul **120** bildet den Modellwert auf eine Wahrscheinlichkeitsfunktion ab **515**. Die Wahrscheinlichkeitsfunktion könnte unter Verwendung des Modellwertes wie beispielsweise unter Verwendung des Modellwertes als die Basis einer Potenzfunktion erzeugt werden, oder kann verwendet werden, um eine Wahrscheinlichkeitsfunktion aus einem Satz von Funktionen, die im Klassifizierungsmodul **120** gespeichert sind, auszuwählen. Beispielsweise könnte das Klassifizierungsmodul **120** eine Auswahl von sechzehn Funktionen speichern, die verschiedene Typen von Zeichenströmen reflektieren und der Modellierungswert **115** könnte dazu verwendet werden, um unter diesen 16 Wahrscheinlichkeitsfunktionen auszuwählen. Alternativ könnte die Funktion eine Funktion sein, die kontinuierlich mit der Änderung des Modellierungswertes **115** angepaßt wird. Die ausgewählte Wahrscheinlichkeitsfunktion oder Funktion wird zum Wahrscheinlichkeitsabbildungssystem **130** übertragen **520**.

[0044] Das Wahrscheinlichkeitsabbildungssystem **130** empfängt die Wahrscheinlichkeitsfunktion. Falls der Strom kodiert wird **530**, berechnet **525** das Wahrscheinlichkeitsabbildungssystem **130** eine Wahrscheinlichkeitskarte **230**, die die Wahrscheinlichkeit beinhaltet, daß das aktuelle Zeichen das nächste Zeichen im Strom gewesen wäre. Die Wahrscheinlichkeiten werden der Code-Lookup-Tabelle **275** zugeführt, die die Wahrscheinlichkeiten für einen Abruf durch den Code-Assembler **280** speichert. Die Wahrscheinlichkeit des aktuellen Zeichens bestimmt die Anzahl von Bits, die verwendet wird, um das Zeichen zu kodieren. Das Zeichen wird auf die richtige Wahrscheinlichkeit abgebildet **550**. Das Zeichen wird mit nur einigen Bits kodiert **555**, falls die Wahrscheinlichkeit des aktuellen Zeichens, daß es das nächste Zeichen im Strom ist, im Vergleich zur Wahrscheinlichkeit der anderen Zeichen im Zeichensatz sehr wahrscheinlich ist. Mehr Bits werden verwendet, um das Zeichen zu kodieren, wenn die relative Wahrscheinlichkeit des aktuellen Zeichens abnimmt. Das dekodierte Zeichen wird dann im kodierten Zeichenstrom **150** übertragen **560**.

[0045] Falls der Strom dekodiert wird **530**, wird die Wahrscheinlichkeitskarte zur Code-Look-Up-Tabelle **275** weitergegeben, die ein Mittel zum Abbilden **535** der Anzahl von Bits in dem kodierten Zeichen mit der Wahrscheinlichkeit, die vom Kodierer verwendet wurde, um das Zeichen zu kodieren, bereitstellt. Beispielsweise, falls nur ein einziges Bit verwendet wurde, um das Zeichen zu kodieren, war die kodierseitig berechnete Wahrscheinlichkeit ziemlich hoch. Die Charakteristika des Bits, wie beispielsweise, ob eine 1 oder 0 verwendet wurde, werden ebenso verwendet, um diese Wahrscheinlichkeit zu verfeinern. Falls beispielsweise eine einzelne Null verwendet wurde, wird das Zeichen mit der höchsten Wahrscheinlichkeit in der Wahrscheinlichkeitsfunktion abgeglichen.

[0046] Der Dekoder bildet dann die Wahrscheinlichkeit mit dem richtigen quantifizierten Zeichen, das das kodierte Zeichen im dekodierten Zeichenstrom ersetzt, ab **540**. Kodierseitig wird jedem Zeichen eine Wahrscheinlichkeit basierend auf dem Modellierungswert **115** und der entsprechenden Wahrscheinlichkeitsfunktion zugeteilt. Dieser Prozeß wird nun umgekehrt. Beispielsweise, falls die Zahl "N" das Zeichen war, das gemäß der Wahrscheinlichkeitsfunktion am wahrscheinlichsten das nächste Zeichen im Strom war, würde der Kodierer ein Zeichen mit der geringsten Anzahl von Bits als "N" dekodieren. Dieser Prozeß wird wiederholt und ein neuer Modellierungswert **115** und die entsprechende Wahrscheinlichkeitsfunktion wird verwendet, um jedes Zeichen dynamisch zu dekodieren. Das dekodierte Zeichen wird dann in den Strom **545** übertragen.

[0047] Die folgende Beschreibung liefert ein Beispiel eines Zeichenstroms, der gemäß dem vorliegenden Verfahren kodiert wird. Für Veranschaulichungszwecke wurde die Wahrscheinlichkeit eine Null als das nächste Zeichen im Zeichenstrom zu erhalten, wie durch die Wahrscheinlichkeitsfunktion erhalten, die mit dem Modellierungswert **115** und anderen relevanten Faktoren zusammenpaßt, in einer Tabelle in **Fig. 6** bereitgestellt. In dieser beispielhaften Ausführungsform bestimmt das Modellierungsmodul **110** den Modellierungswert **115** durch Berechnen eines angepaßten Logarithmus zur Basis **2** des gewichteten Mittels der letzten vier Zeichen im Strom. Der berechnete Wert wird auf die nächstliegende ganze Zahl gerundet und dieser Modellierungswert **115** wird zum Klassifizierungsmodul **120** weitergegeben. Zusätzlich speichert bei diesem Beispiel das Klassifizierungsmodul **120** eine Auswahl von Wahrscheinlichkeitsfunktionen, die auf jeden Modellierungswert **115** abgebildet werden. Schließlich wird eine Zeichensatzgröße von 4096 verwendet und 24 Funktionen im Klassifizierungsmodul **120** gespeichert. Die Verwendung dieser Parameter in diesem Beispiel soll die praktische Anwendung der Erfindung bei der Datenübertragung demonstrieren und soll nicht als Beschränkung des Umfangs der Ansprüche verstanden werden. Irgendeine Anzahl von unterschiedlichen Berechnungen des Modellierungswertes **115**, Wahrscheinlichkeitsfunktionen und Zeichensätzen kann verwendet werden, um die Erfindung wirkungsvoll umzusetzen.

[0048] Es sei angenommen, daß die letzten vier Zeichen im Strom vor dem aktuellen Zeichen, vorausgesetzt, daß es sich um hohe Zeichenwerte handelt, in einem gewichteten Mittel von 768 resultieren. Der angepaßte Logarithmus zur Basis **2** des gewichteten Mittelwerts ist gleich 10 und dieser Modellierungswert **115** wird zum Klassifizierungsmodul **120** weitergegeben. Die Wahrscheinlichkeitsfunktion, die einem Modellierungswert **115**, der gleich 10 ist, entspricht, wird durch das Klassifizierungsmodul **120** abgerufen und zum Wahrscheinlich-

keitsabbildungssystem **130** weitergegeben. Die durch das Wahrscheinlichkeitsabbildungssystem **130** bereitgestellten Wahrscheinlichkeiten werden vom Kodierer verwendet, um das Zeichen zu kodieren. Ferner sei angenommen, daß das aktuelle Zeichen, das übertragen wird, eine Null ist. Mit Bezugnahme auf die Tabelle wird die Wahrscheinlichkeit, daß das erste Zeichen gleich Null ist, in 1535 mit 1 berechnet. Da die Null äußerst unerwartet war, würde die Null mit ungefähr 12 Bits kodiert werden. Diese **12** Bits würden im Strom kodierter Werte **150** vom statistischen Kodierer **150** bereitgestellt werden.

[0049] Im folgenden werde angenommen, daß das nächste Zeichen (das zweite Zeichen) ebenfalls Null ist. Dies resultiert in einer Reduzierung des Modellierungswertes **115**, möglicherweise auf neun, abhängig von den Werten der Zeichen vor dem ersten Zeichen. Die Wahrscheinlichkeitsverteilung entsprechend einem Modellierungswert **115**, der gleich neun ist, würde dann verwendet werden, um das zweite Zeichen zu kodieren. Das zweite Zeichen würde wie das erste Zeichen kodiert und durch den Kodierer **250** dem Strom kodierter Werte **150** zugeführt. In diesem Fall würde das Zeichen jedoch nur 11 Bits zur Kodierung erfordern, da die Wahrscheinlichkeitsverteilung eine größere Wahrscheinlichkeit für Null angezeigt hat. Da niedrige Werte (wie beispielsweise Nullen) angetroffen werden, gibt das Klassifizierungsmodul **120** Wahrscheinlichkeitsfunktionen, die eine größere Wahrscheinlichkeit für derartige Werte angeben, weiter und die Kodierungseffektivität steigt an. Zusätzlich wird, insbesondere dort, wie im vorliegenden Fall, wo eine gewichtete Summe unter Verwendung einer begrenzten Anzahl von Werten, wie beispielsweise 4 verwendet wird, die geeignete Wahrscheinlichkeitsverteilung rasch zu einer übergehen, die für den Strom geeignet ist. Dieses Verfahren wird durch das Vorsehen einer alternierenden Funktion verstärkt, die einen Modellierungswert **115** gemäß der Nullzählung berechnet. Falls beispielsweise vier aufeinanderfolgende Nullen angetroffen werden, könnte für das fünfte Zeichen die Funktion entsprechend dem Modellierungswert **115** von Null verwendet werden. Das folgende Beispiel veranschaulicht den Übergang zwischen zwei Funktionen, die verwendet werden, um den Modellierungswert **115** für verschiedene Typen von Strömen zu berechnen. Für Veranschaulichungszwecke werde angenommen, daß wenn der Strom einen Mittelwert von Null aufweist, die alternierende Funktion implementiert wird.

[0050] Unter Beibehaltung derselben Parameter, die im vorhergehenden Beispiel verwendet wurden, werde angenommen, daß, nachdem die Funktion gemäß einem Modellierungswert **115** von 10 wegen der früheren Historie mit hohem Wert verwendet wurde, die nächsten vier Zeichen Null wären (Zeichen eins bis vier). Die Zeichen wären gemäß den auf den berechneten Modellierungswert **115** abgebildeten Wahrscheinlichkeitsfunktionen kodiert worden und würden in den Strom kodierter Werte **150** zugeführt werden, die durch den statistischen Kodierer **240** produziert werden. Für das fünfte Zeichen wäre die gewichtete Summe gleich Null, so daß die Wahrscheinlichkeitsverteilung, bei der der Modellierungswert **115** gleich Null ist, verwendet wird.

[0051] Wenn der Modellierungswert **115** Null ist, wird jedoch eine alternierende Funktion verwendet, um einen zusätzlichen Modellierungswert **115** zu liefern. In diesem Fall wird vorzugsweise das Nullzählungsverfahren implementiert. Falls das fünfte Zeichen nicht gleich Null ist, wird es unter Verwendung der Verteilung gemäß dem vorhergehenden Modellierungswert **115** unter Verwendung des Standardkodierungsprozesses, der oben beschrieben wurde, kodiert. Man nehme jedoch an, daß das fünfte Zeichen gleich Null ist. Diese Null würde nicht unmittelbar verwendet, um einen neuen Modellierungswert **115** zu generieren. Wie oben erläutert, wird die Wahrscheinlichkeitsfunktion, die verwendet wird, um das Zeichen zu kodieren, von den nachfolgenden Zeichenwerten abhängen. Es werde angenommen, daß das sechste Zeichen ebenfalls Null ist. Da zwei aufeinander folgende Nullen angetroffen worden wären, wird ein neuer Modellierungswert **115** von -1 berechnet. Wenn der Modellierungswert **115** -1 ist, wird eine Null, die durch den Kodierer in den Strom kodierter Werte **150** eingefügt wird, das Äquivalent von zwei (2^1) Nullen. Somit würde tatsächlich eine einzige kodierte Null die zwei Nullen entsprechend dem fünften und sechsten Zeichen darstellen.

[0052] Falls vier weitere Nullen (für die Zeichen 7–10) in den Strom **105** eingegeben würden, würde ein Modellierungswert **115** von -2 berechnet und die vier gesetzten Nullen würden durch eine einzige Null im Strom kodierter Werte dargestellt werden. Somit würde der Kodierer **240** "00" in den kodierten Strom für die tatsächlichen Werte "000000" für die Zeichen 5–10 einfügen. Acht weitere Nullen (für die Zeichen 11–18) müssen in den Strom **105** eingeführt werden, um den Satz von Nullen entsprechend einen Modellierungswert **115** von -3 zu vervollständigen. An diesem Punkt würden acht aufeinanderfolgende Nullen als eine Null gespeichert werden. Somit könnte "000" der Strom kodierter Werte für die tatsächlichen Werte "0000000000000000" der Zeichen 5–18 sein. Dieser Prozeß wird durch Verdoppeln der Anzahl von Nullen, die in einem Satz erforderlich sind, bevor der nächste Modellierungswert **115** berechnet wird, fortgesetzt.

[0053] Man nehme jedoch an, daß nach einem Anfangszustand mit einem Modellierungswert **115**, der gleich 10 ist, nur die Zeichen 1–14 Null wären und daß die Zeichen **15** nicht Null wären. Der Strom kodierter Werte **150** würde eine geeignete Kodierung für die Nullen der Zeichen 1–4 und "00" für die Zeichen 5–10 umfassen. Eine dritte nachfolgende Null, die acht 0en (11–18) entsprechen würde, wäre nicht im Strom **150** kodierter Werte beinhaltet. Jedoch würde, sobald ein Wert 15, der nicht Null ist, angetroffen wird, dieser unter Verwendung des herkömmlichen Kodierungsschemas kodiert werden (gemäß der b-gleich-minus-drei-Funktion) und in den Strom kodierter Werte **150** eingefügt werden. Zusätzlich würden die vier Nullen für die Zeichen 11–14 durch

ein Codewort dargestellt werden, das anzeigt, daß der nächste Satz von Nullen ebenfalls direkt in den Strom **150** der kodierten Werte eingefügt würde.

[0054] Unter der Annahme, daß die Bezeichnung (EC_n) der kodierte Wert für das me Zeichen ist (z. B. für das erste Zeichen, wie oben beschrieben, wäre EC₁ 11 Bit an Information). Ein resultierender beispielhafter Strom **150** kodierter Werte wäre aufeinanderfolgend (EC₁), (EC₂), (EC₃), (EC₄), 0, 0, (EC₁₅), (Codewort). Alternativ könnte das (Codewort) im Strom kodierter Werte vorangehen (V 15). In diesem Beispiel wäre die Abfolge von Nullen im Strom kodierter Daten **150** kompakt vorgesehen. Beispielsweise würde "00000" in einem Strom kodierter Daten **150** 44 Nullen des ursprünglichen Zeichenstroms **105** darstellen. Tatsächlich würde, wenn der Modellierungswert **115** gleich -12 ist, eine einzelne Null in dem Strom kodierter Daten **150** 4096 (2¹²) Nullen in dem ursprünglichen Zeichenstrom **105** darstellen. 4096 ist lediglich eine beispielhafte Grenze und die Erfindung kann mit höheren Darstellungen aufeinanderfolgender Nullen implementiert werden. Zusätzlich wird Null verwendet, um das Verfahren zu demonstrieren und es könnten andere nicht null-zentrierte Systeme implementiert werden. Des weiteren muß das Verfahren nicht auf der Basis von 2 betrieben werden, obwohl Potenzen von 2 bei der bevorzugten Ausführungsform verwendet werden.

Patentansprüche

1. Verfahren zum Modellieren von Zeichen in einem Zeichenstrom für eine Kompression, um die Datenmenge zu komprimieren, die erforderlich ist, um ein Signal darzustellen, das dem Zeichenstrom entspricht, wobei das Verfahren aufweist:

Erhalten (**505**) eines aktuellen Zeichens in dem Zeichenstrom; und Modellieren des aktuellen Zeichens, wobei dieses Modellieren das Ergebnis ist des:

Verwendens einer Modellierungsfunktion, um für Zeichenwerte, die für das aktuelle Zeichen lokal sind, einen Modellierungswert basierend auf Zeichengrößen der lokalen Zeichenwerte zu bestimmen (**510**);

Verwendens des Modellierungswertes, um eine Wahrscheinlichkeitsfunktion zu bestimmen (**515**), die jedem Zeichen in dem Strom einen Wahrscheinlichkeitswert dafür zuordnet, dass das aktuelle Zeichen das nächste Zeichen in dem Strom gewesen wäre; und

Verwendens (**525**) der Wahrscheinlichkeitsfunktion, um die Wahrscheinlichkeit für das aktuelle Zeichen zu bestimmen.

2. Verfahren nach Anspruch 1, wobei das Verwenden des Modellierungswertes zum Bestimmen der Wahrscheinlichkeitsfunktion (**515**) das Verwenden des Modellierungswertes als eine Adresse für ein Lokalisieren der Wahrscheinlichkeitsfunktion in einem definierten Satz von Funktionen beinhaltet.

3. Verfahren nach Anspruch 1, wobei das Verwenden des Modellierungswertes zum Bestimmen der Wahrscheinlichkeitsfunktion (**515**) das Verwenden des Modellierungswertes als eine Variable in einer Wahrscheinlichkeitsfunktion beinhaltet.

4. Verfahren nach Anspruch 1, wobei das Verwenden des Modellierungswertes zum Bestimmen der Wahrscheinlichkeitsfunktion (**515**) das Verwenden eines Verschlüsselungswertes beinhaltet.

5. Verfahren nach Anspruch 4, wobei der Verschlüsselungswert ein öffentlicher Schlüssel in einem öffentlichen Verschlüsselungssystem ist.

6. Verfahren nach einem der Ansprüche 1 bis 5, welches ferner umfaßt:

Wiederholen der Schritte von Anspruch 1 für jedes Zeichen, das dem aktuellen Zeichen in dem Zeichenstrom nachfolgt.

7. Verfahren nach einem der Ansprüche 1 bis 6, welches ferner umfaßt:

Kodieren (**555**) des aktuellen Zeichens basierend auf der bestimmten Wahrscheinlichkeit.

8. Verfahren zum Dekodieren von Zeichen in einem Datenstrom, welches umfaßt:

Erhalten (**505**) eines aktuellen, kodierten Zeichens in einem Strom kodierter Werte und

Dekodieren des aktuellen kodierten Zeichens, wobei dieses Dekodieren das Ergebnis ist des:

Verwendens einer Modellierungsfunktion, um für Zeichenwerte, die für das aktuelle, kodierte Zeichen lokal sind, einen Modellierungswert basierend auf Zeichengrößen derartiger Zeichenwerte zu bestimmen (**510**);

Verwendens des Modellierungswertes, um eine Wahrscheinlichkeitsfunktion zu bestimmen (**515**), die jedem Zeichen in dem Strom einen Wahrscheinlichkeitswert dafür zuordnet, dass das aktuelle Zeichen das nächste Zeichen in dem Strom gewesen wäre; und

Verwendens der Wahrscheinlichkeitsfunktion beim Dekodieren (**540**) des aktuellen, kodierten Zeichens.

9. Verfahren nach Anspruch 8, das in Reaktion auf Empfangen eines ersten Zeichens in dem Zeichenstrom aufweist:
Erzeugen einer Anfangswahrscheinlichkeitsfunktion zum Dekodieren der Zeichen in dem Strom; und
Dekodieren (**540**) der Zeichen in dem Strom unter Verwendung der Anfangswahrscheinlichkeitsfunktion, bis ein Modellierungswert berechnet ist.
10. Verfahren nach Anspruch 9, wobei das Erzeugen einer Anfangswahrscheinlichkeitsfunktion das Verwenden eines Verschlüsselungswertes beinhaltet, um einen Anfangsmodellierungswert zu dekodieren.
11. Verfahren nach Anspruch 8, wobei das Verwenden des Modellierungswertes zum Bestimmen der Wahrscheinlichkeitsfunktion das Verwenden des Modellierungswertes als Adresse für ein Lokalisieren der Wahrscheinlichkeitsfunktion in einem definierten Satz von Funktionen beinhaltet.
12. Verfahren nach Anspruch 8, wobei das Verwenden des Modellierungswertes zum Bestimmen der Wahrscheinlichkeitsfunktion das Verwenden des Modellierungswertes als eine Variable in einer Wahrscheinlichkeitsfunktion beinhaltet.
13. Verfahren nach Anspruch 8, wobei das Verfahren des weiteren das Ersetzen des aktuellen, kodierten Zeichens durch mehr als ein Zeichen aufweist.
14. System zum Kodieren von Zeichen in einem Zeichenstrom, das aufweist:
ein Modellierungsmodul (**110**) zum Bestimmen eines Modellierungswertes für Zeichenwerte, basierend auf Zeichengrößen derartiger Zeichenwerte, die für ein aktuelles Zeichen in dem Zeichenstrom lokal sind;
ein mit dem Modellierungsmodul (**110**) gekoppeltes Klassifizierungsmodul (**120**) zum Verwenden des Modellierungswertes zum Bestimmen einer Wahrscheinlichkeitsfunktion, die jedem Zeichen in dem Zeichenstrom einen Wahrscheinlichkeitswert dafür zuordnet, dass das aktuelle Zeichen das nächste Zeichen in dem Strom gewesen wäre;
ein mit dem Modellierungsmodul (**110**) gekoppeltes Wahrscheinlichkeitsabbildungssystem (**130**) zum Bestimmen der Wahrscheinlichkeit des aktuellen Zeichens in dem Zeichenstrom; und
einen mit dem Wahrscheinlichkeitsaufzeichnungssystem (**130**) gekoppelten Entropie-Kodierer (**140**) zum Kodieren des aktuellen Zeichens basierend auf der Wahrscheinlichkeit des aktuellen Zeichens.
15. System nach Anspruch 14, wobei das Klassifizierungsmodul (**120**) ferner eine Tabelle aufweist, die Wahrscheinlichkeitsfunktionen beinhaltet.
16. System nach Anspruch 14, wobei das Modellierungsmodul (**110**) ferner einen Zeichenzähler zum Zählen der Anzahl aufeinanderfolgender identischer Zeichen und zum Erzeugen eines Modellierungswertes basierend auf der Zählung aufweist.
17. System zum Decodieren von Zeichen in einem Zeichenstrom, welches aufweist:
Ein Modellierungsmodul (**110**) zum Bestimmen eines Modellierungswertes basierend auf Zeichengrößen derartiger Zeichenwerte für Zeichenwerte, die für ein aktuelles, kodiertes Zeichen lokal sind;
ein mit dem Modellierungsmodul (**110**) gekoppeltes Klassifizierungsmodul (**120**) zum Verwenden des Modellierungswertes, um eine Wahrscheinlichkeitsfunktion zu bestimmen, die jedem Zeichen in dem Zeichenstrom einen Wahrscheinlichkeitswert dafür zuordnet, dass das aktuelle Zeichen das nächste Zeichen in dem Strom gewesen wäre;
ein mit dem Modellierungsmodul (**110**) gekoppeltes Wahrscheinlichkeitsabbildungssystem (**130**) zum Bestimmen der Wahrscheinlichkeit des aktuellen, codierten Zeichens; und
einen mit dem Wahrscheinlichkeitsabbildungssystem (**130**) gekoppelten Entropie-Kodierer (**140**) zum Kodieren des aktuellen, kodierten Zeichens basierend auf der Wahrscheinlichkeit des aktuellen, kodierten Zeichens.
18. System nach Anspruch 17, wobei das Klassifizierungsmodul (**120**) ferner eine Tabelle aufweist, die Wahrscheinlichkeitsfunktionen enthält.
19. Computer-lesbares Medium, das ein Computerprogramm enthält, das:
ein aktuelles Zeichen in einem Zeichenstrom erhält; und
das aktuelle Zeichen kodiert, wobei dieses Kodieren das Ergebnis ist des:
Verwendens einer Modellierungsfunktion, um für Zeichenwerte, die für das aktuelle Zeichen lokal sind, einen Modellierungswert basierend auf Zeichengrößen für derartige Zeichenwerte zu bestimmen (**510**); und
Verwendens des Modellierungswertes, um eine Wahrscheinlichkeitsfunktion zu bestimmen (**515**), die jedem

Zeichen in dem Strom einen Wahrscheinlichkeitswert dafür zuordnet, dass das aktuelle Zeichen das nächste Zeichen in dem Zeichenstrom gewesen wäre; und Verwendens der Wahrscheinlichkeitsfunktion bei einem Kodieren (**555**) des aktuellen Zeichens.

Es folgen 6 Blatt Zeichnungen

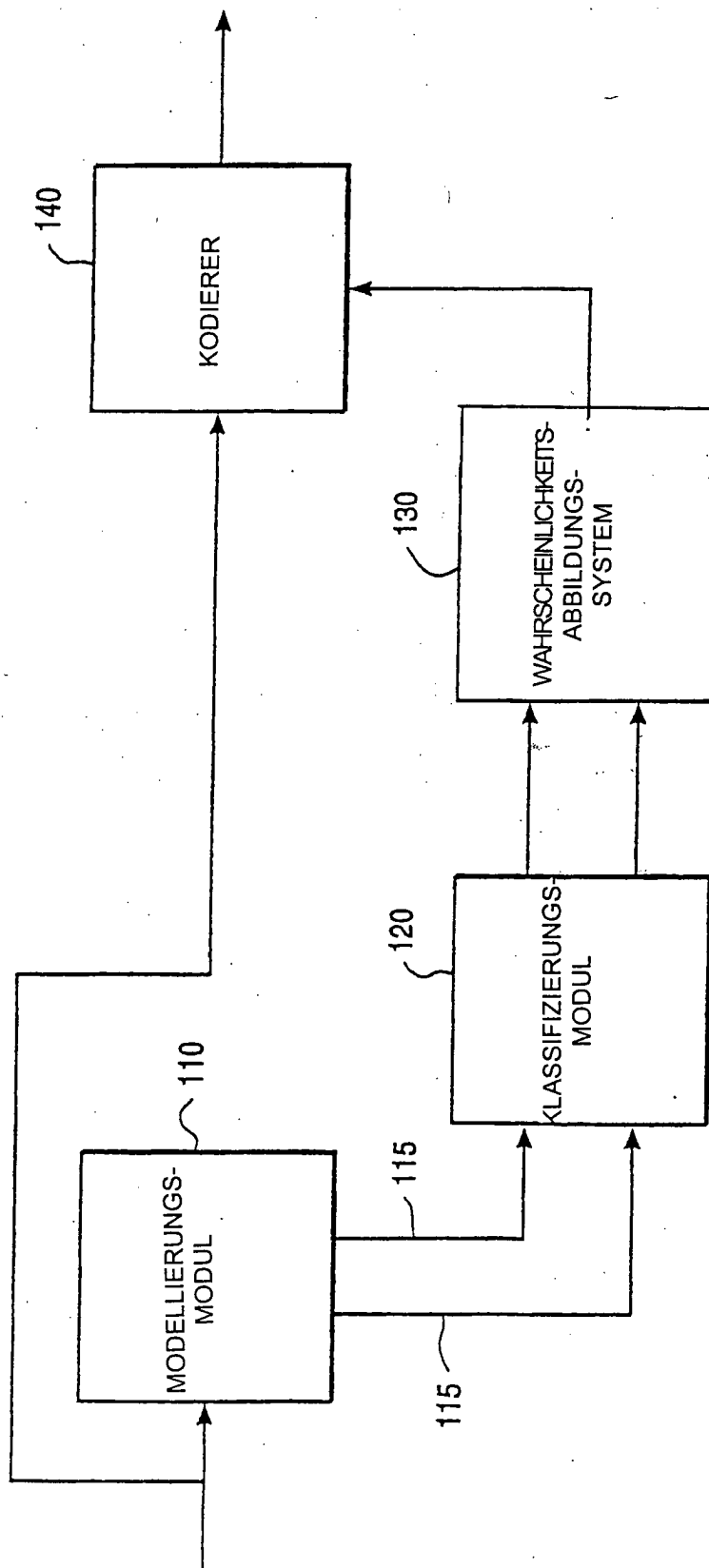


FIG. 1

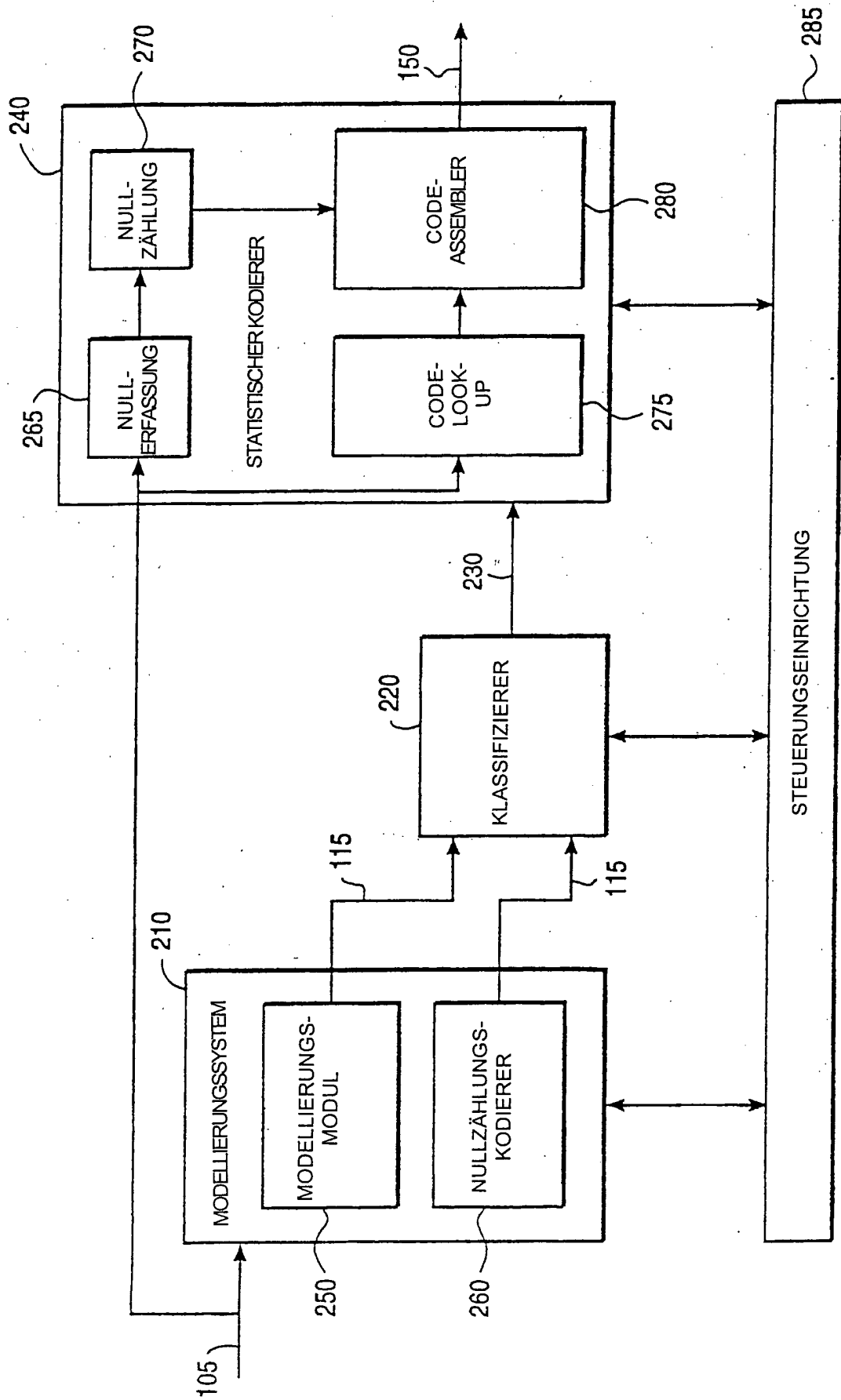


FIG. 2

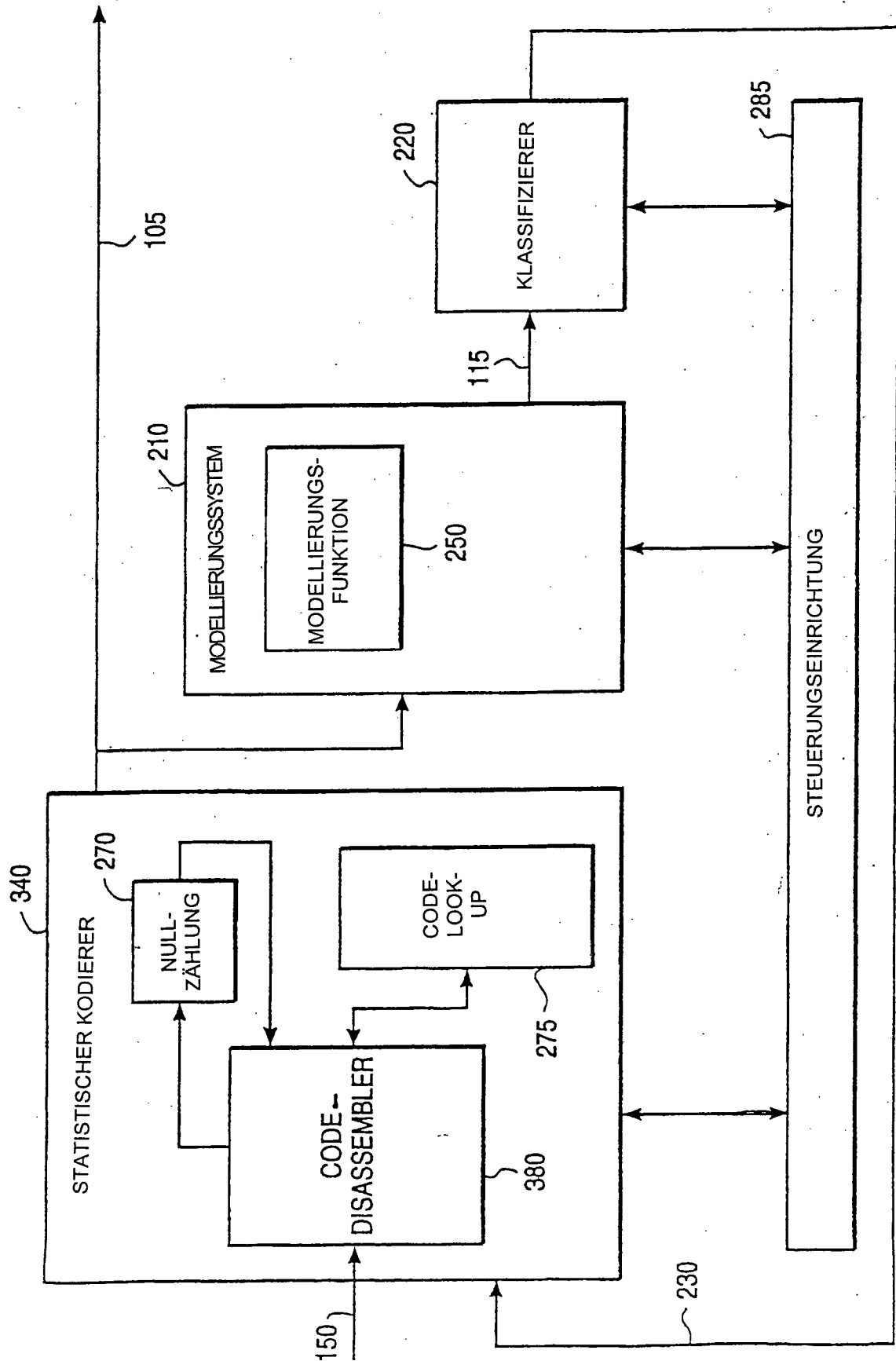


FIG. 3

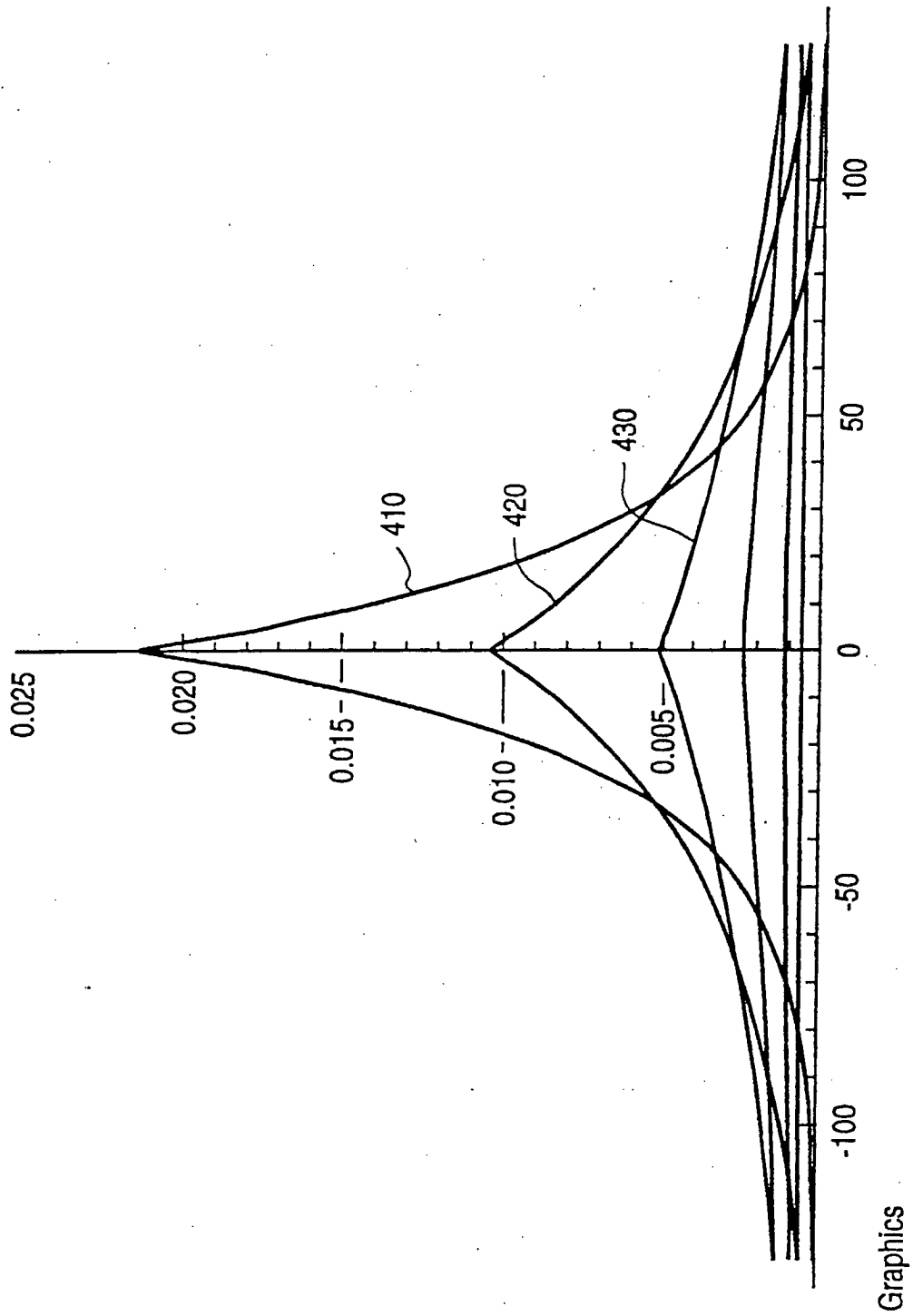


FIG. 4

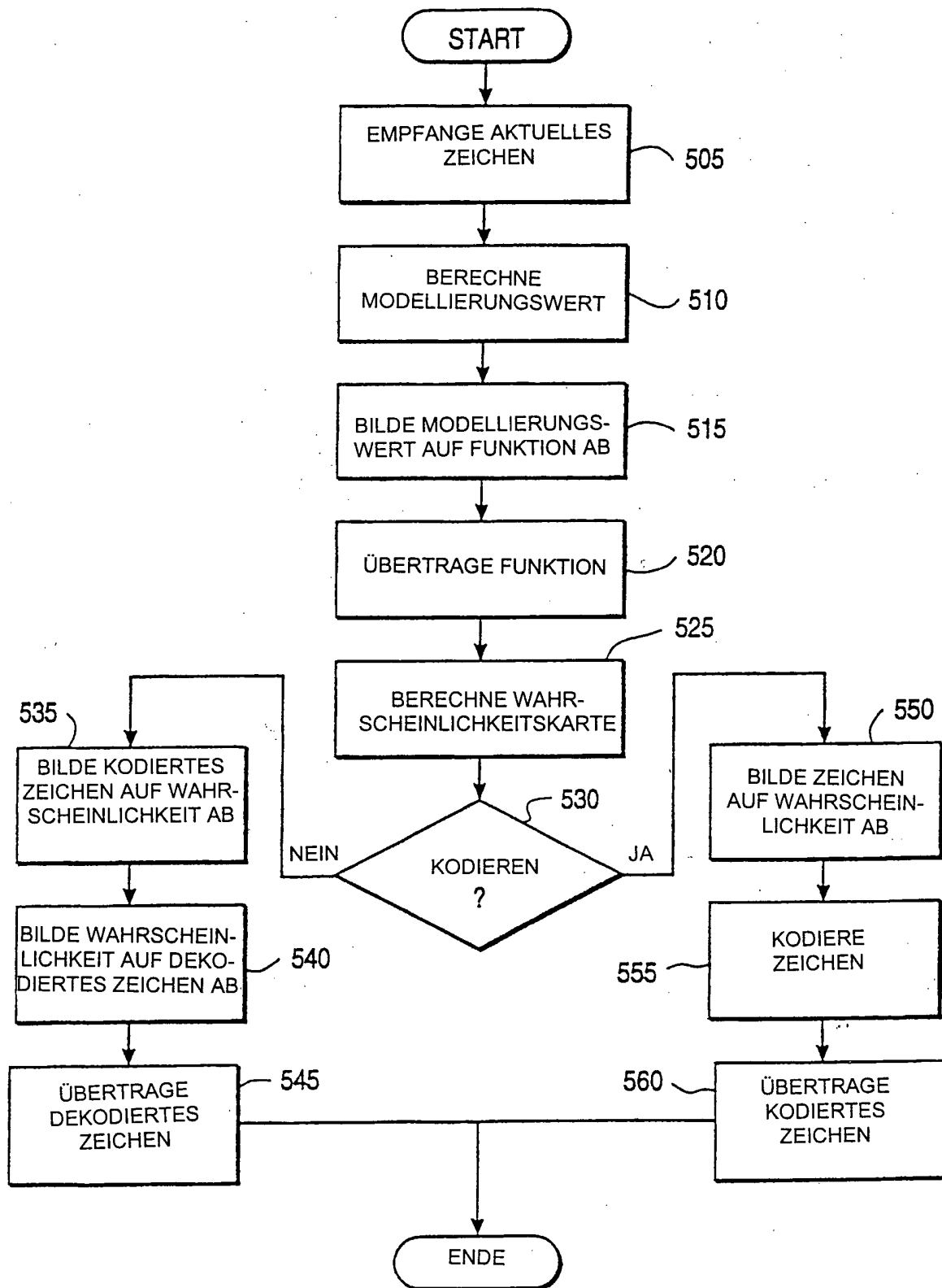


FIG. 5

GEWICHTETE SUMME	MODELLIERUNGS- WERT	P[0]	ZCNT
768	10	1/1535	1
384	9	1/767	1
192	8	1/383	1
96	7	1/191	1
48	6	1/95	1
24	5	1/47	1
12	4	1/23	1
6	3	1/11	1
3	2	1/5	1
4/3	1	1/3	1
3/4	0	$(1/2)^1$	1
1/3	-1	$(1/2)^{(1/2)}$	2
1/6	-2	$(1/2)^{(1/4)}$	4
1/12	-3	$(1/2)^{(1/8)}$	8
1/24	-4	$(1/2)^{(1/16)}$	16
1/48	-5	$(1/2)^{(1/32)}$	32
1/96	-6	$(1/2)^{(1/64)}$	64
1/192	-7	$(1/2)^{(1/128)}$	128
1/384	-8	$(1/2)^{(1/256)}$	256
1/768	-9	$(1/2)^{(1/512)}$	512
1/1536	-10	$(1/2)^{(1/1024)}$	1024
1/3072	-11	$(1/2)^{(1/2048)}$	2048
1/6144	-12	$(1/2)^{(1/4096)}$	4096

FIG. 6