

PCT

世界知的所有権機関
国際事務局

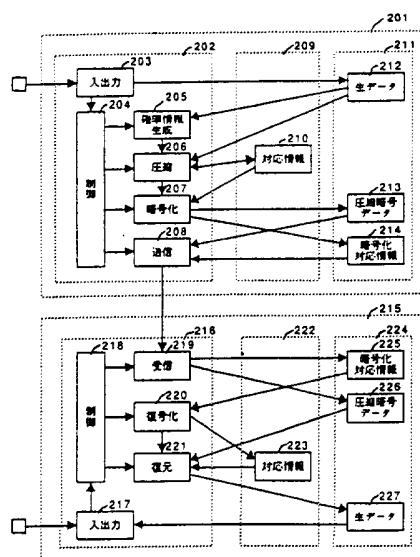
特許協力条約に基づいて公開された国際出願



(51) 国際特許分類6 H04L 9/06, 9/16	A1	(11) 国際公開番号 WO97/10659
(21) 国際出願番号 PCT/JP95/01815	(43) 国際公開日 1997年3月20日(20.03.97)	
(22) 国際出願日 1995年9月13日(13.09.95)	(74) 代理人 弁理士 小川勝男(OGAWA, Katsuo) 〒100 東京都千代田区丸の内一丁目5番1号 株式会社 日立製作所内 Tokyo, (JP)	(81) 指定国 AU, CN, JP, KR, SG, US, 歐州特許 (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). 添付公開書類 国際調査報告書

(54) Title: METHOD AND DEVICE FOR COMPRESSING AND CIPHERING DATA

(54) 発明の名称 データの圧縮・暗号化方法および装置



203, 217 ... input/output means
 204, 218 ... control means
 205 ... probability information generating means
 206 ... compressing means
 207 ... ciphering means
 208 ... transmitting means
 210, 223 ... corresponding information
 212, 227 ... raw data
 213, 226 ... compressed and ciphered data
 214, 225 ... ciphered corresponding information
 219 ... receiving means
 220 ... deciphering means
 221 ... restoring means

(57) Abstract

In an information processing system provided with data inputting means (203, 208, 217, and 219) which input or receive data, a data compressing means which compresses the data, a restoring means (221) which restores compressed data, and main storage devices (209 and 222); a ciphering means (207) which ciphers data, and deciphering means (220) which deciphers ciphered data are provided. At a compressing/ciphering step where the compressing means (206) and ciphering means (207) respectively compresses and ciphers part of the data and at a restoring/deciphering step where the restoring means (221) and deciphering means (220) respectively restores and deciphers the compressed and deciphered data, the amount of data to be processed is determined so that the quality of data does not exceed the capacities of the main storage devices (209 and 222). Such compressing/ciphering step and restoring/deciphering step are performed once or more times.

(57) 要約

データを入力または受信するデータ入力手段（203、208、217、219）と、このデータを圧縮するデータ圧縮手段（206）と、圧縮されたデータを復元する復元手段（221）と、主記憶装置（209、222）とを有する情報処理システムにおいて、データを暗号化する暗号化手段（207）と、暗号化されたデータを復号化する復号化手段（220）とを備える。そして、圧縮手段（206）と暗号化手段（207）とが前記データの一部について圧縮と暗号化を行う圧縮暗号化ステップと、復元手段（221）と復号化手段（220）とが圧縮かつ暗号化されたデータの一部について復元と復号化を行う復元復号化ステップの実行において、圧縮暗号化、および復元復号化ステップが処理する一連のデータの量を、この処理に必要なメモリの量が主記憶装置の容量を越えないように設定し、圧縮暗号化、および復元復号化ステップを1回以上繰り返す。

情報としての用途のみ

PCTに基づいて公開される国際出願をパンフレット第一頁にPCT加盟国を同定するために使用されるコード

AL	アルバニア	EE	エストニア	LR	リベリア	RU	ロシア連邦
AM	アルメニア	ES	スペイン	LS	レソト	SD	スーダン
AT	オーストリア	FI	フィンランド	LT	リトアニア	SE	スウェーデン
AU	オーストラリア	FR	フランス	LU	ルクセンブルグ	SG	シンガポール
AZ	アゼルバイジャン	GA	ガボン	LV	ラトヴィア	SI	スロヴェニア
BB	バルバドス	GB	イギリス	MC	モナコ	SK	スロヴァキア共和国
BE	ベルギー	GE	グルジア	MD	モルドバ	SZ	セネガル
BF	ブルキナ・ファソ	GH	ガーナ	MG	マダガスカル	TZ	スワジランド
BG	ブルガリア	GN	ギニア	MK	マケドニア旧ユーゴスラ	TD	チャード
BJ	ベナン	GR	ギリシャ	VU	ヴィア共和国	TG	トーゴ
BR	ブラジル	HU	ハンガリー	ML	マリ	TJ	タジキスタン
BY	ベラルーシ	IE	アイルランド	MN	モンゴル	TM	トルクメニスタン
CA	カナダ	IS	イスランド	MR	モーリタニア	TR	トルコ
CF	中央アフリカ共和国	IT	イタリー	MW	マラウイ	TT	トリニダード・トバゴ
CG	コンゴー	JP	日本	MX	メキシコ	UA	ウクライナ
CH	スイス	KE	ケニア	NE	ニジェール	UG	ウガンダ
CI	コート・ジボアール	KG	キルギスタン	NL	オランダ	US	米国
CM	カメルーン	KP	朝鮮民主主義人民共和国	NO	ノルウェー	UZ	ウズベキスタン共和国
CN	中国	KR	大韓民国	NZ	ニュージーランド	VN	ヴィエトナム
CZ	チェコ共和国	KZ	カザフスタン	PL	ポーランド	YU	ユーゴスラビア
DE	ドイツ	LI	リヒテンシュタイン	PT	ポルトガル		
DK	デンマーク	LK	スリランカ	RO	ルーマニア		

明細書

データの圧縮・暗号化方法および装置

技術分野

本発明は、データの圧縮および暗号化に関し、特に、データに圧縮と暗号化の両方を施す場合の処理効率の向上および電力消費量の削減を図ったデータの圧縮・暗号化方法および装置に関する。

背景技術

通信の利用の急激な増加に伴い、通信処理の効率化、通信データの盗聴や改竄防止のために、データを圧縮し、かつ暗号化する（ここでは圧縮暗号化と略する）機会が多くなっている。特に、無線通信の場合、帯域が小さく、傍受が容易であるため、圧縮暗号化が必須である。一方、近年急速に普及しつつある携帯型計算機では、バッテリーの持続が重要な問題であり、電力の節約に適した処理方法が重要になっている。以上から、効率的で電力消費の小さい圧縮暗号化技術が必須である。

従来の圧縮暗号化技術では、まず、データを圧縮してハードディスク等の二次記憶装置に書き込み、その後、圧縮データを二次記憶装置から読み出して暗号化していた。暗号化処理では、データが圧縮されていることを考慮せず、圧縮されていないデータに対する場合と同じ暗号化を行っていた。また、圧縮暗号化されたデータを元のデータに復元する（ここでは復号化復元と呼ぶ）処理では、まず、データを復号化して二次記憶装置に書き込み、その後、復号化されたデータを二次記憶装置から読み出して復元していた。

上記従来の圧縮暗号化技術では、圧縮と暗号化を独立して実行しており、両者を融合することによる処理効率向上の可能性を生かしていなかった。また、圧縮処理と暗号化処理、および、復号化処理と復元処理を二次記憶装置経由で連動していたので、二次記憶装置の読み書きのための時間と消費電力を要して

いた。

本発明の目的は、圧縮と暗号化の処理を融合し、圧縮と暗号化および復号化と復元の二次記憶装置経由による運動を不要化することにより、圧縮暗号化および復号化復元の処理効率を向上し、かつ、電力消費を削減することのできるデータの圧縮・暗号化方法および装置を得ることにある。

発明の開示

二次記憶装置経由による圧縮処理と暗号化処理および復号化処理と復元処理の運動を不要化するためには、圧縮暗号化処理および復号化復元処理に必要なメモリ量が当該情報処理システムの主記憶装置の容量を越えないように、一度に処理するデータの量を制限すればよい。換言すれば、圧縮暗号化処理および復号化復元処理に必要なメモリ量が当該情報処理システムの主記憶装置の容量を越えないように、一度に処理するデータの量を区分し、この区分した単位でデータの圧縮・暗号化を順次繰り返して実行し、1組のデータ全体を圧縮・暗号化する。

すなわち、本発明の第一の方法は、データの圧縮手段と復元手段を有する情報処理システムにおいて、暗号化手段および復号化手段を設け、データに圧縮手段と暗号化手段を適用するときに必要なメモリ量、および、データに復号化手段と復元手段を適用するときにメモリ量が、当該情報処理手段の有する主記憶装置の容量を越えないように、圧縮と暗号化あるいは復号化と復元を行うデータの量を設定し、設定した量のデータに対する圧縮手段と暗号化手段の適用あるいは復号化手段と復元手段の適用を繰り返すことにより、データ全体を圧縮暗号化あるいは複合化復元する。

次に、圧縮処理と暗号化処理の融合による処理効率の向上について述べる。データ圧縮ハンドブック、トッパン（1994年）、21頁から247頁に述べられているように、従来の圧縮の方法には、以下の（1）～（3）がある。実用的なデータ圧縮プログラムでは、（1）と（3）あるいは（2）と（3）を組合せる場合が多い。

- (1) 固定型確率モデルに基づく方法。例えば、固定型ハフマン符号化。
- (2) 適応型確率モデルに基づく方法。例えば、適応型ハフマン符号化。
- (3) 辞書ベースの方法。たとえば、LZ77、LZ78。

本発明の圧縮暗号化方法は、上記(1)と(2)に適用できる。また、本発明は、(3)には直接は適用できないが、(1)と(3)あるいは(2)と(3)の組合せには適用できる。

第一に、固定型確率モデルに基づく圧縮処理への暗号化処理の組込みについて述べる。固定型確率モデルに基づく圧縮方法では、まず、データ中の各記号の頻度を調べるといった方法により、各記号の出現確率を求める。この記号の出現確率に基づいて、記号とビット列の対応情報（ハフマン符号化の場合はハフマン木）を生成する。このとき、出現確率の大きい記号ほど、短いビット列に対応付ける。その後、データ中の記号を、対応するビット列に変換することによりデータを圧縮する。復元処理は、圧縮後データと、記号一ビット列対応情報（または、それを生成するための記号の出現頻度や出現確率）を受取り、ビット列を記号に逆変換することにより元のデータを復元する。

上記固定型確率モデルに基づく圧縮方法では、記号一ビット列対応情報が利用できなければ、データを復元できない。そこで、圧縮されたデータ自体を暗号化しなくても、記号一ビット列対応情報のみ暗号化すれば、暗号の目的を達成できる。記号一ビット列の対応情報は、圧縮データに比べ、極めて量が少ないので、従来の圧縮データ自体を暗号化する方法に比べ、暗号化の処理量および対応する復号化の処理量を大幅に削減できる。すなわち、本発明の第2の方法は、固定型確率モデルに基づく圧縮において、記号一ビット列対応情報を暗号化する。

記号一ビット列対応情報の暗号化のみでは、暗号の強度（解読の難しさ）が不十分な場合がある。すなわち、上記暗号化に対し、以下の解読方法を考えられる。

- (a) 複数の類似したデータと対応する圧縮データ入手し、それらの比較により、本方式による記号とビット列の対応付けの傾向を推定する。

(b) 一つのデータの中で同じ記号や記号列が繰り返し現れる場合、対応する圧縮データ中の繰り返しパターンの分析により、記号とビット列の対応付けを推定する。

上記(a)の解読を防止するには、データが少しでも異なれば、記号とビット列の対応が全く異なるようにすればよい。前述のように、固定型確率モデルに基づく圧縮では、記号の出現確率に依存して対応するビット列の長さが決まるが、その長さでの0と1の配列には、自由度がある。例えば、記号aに対応するビット列の長さが4に決まったとしても、対応するビット列には、0000、0101、1101等の自由度がある。そこで、記号に対応するビット列が、上記複数の可能性の中から、乱数などの偶然あるいは確率的な要因に依存して、選択されるようにする。この方法によれば、データが少しでも異なれば（全く同じ入力データであっても）、圧縮計算の度に、記号とビット列の対応が全く異なるので、対応付けの傾向を推定することが推定困難となる。そこで、本発明の第3の方法は、上記第2の方法において、記号とビット列の対応付けを、偶然あるいは確率に基づく計算を用いて行う。

上記(b)の解読を防止するには、データとビット列の対応付けをデータ処理の途中で変更すればよい。そこで、本発明の第4の方法は、上記第2の方法において、記号とビット列の対応付けを、データ処理の途中で変更する。

以上の暗号化の強度をさらに増加するためには、第2の暗号化手段を設け、以上的方法で得られた暗号データをさらに暗号化すればよい。ただし、以下の理由から、第2の暗号化手段は、従来の単純な圧縮データに対する暗号化手段より簡易な処理で十分である。

- (a) すでに暗号化されたデータをさらに暗号化するため。
- (b) 類似データの比較や、繰り返しパターンの分析などの解読方法への対処がすでになされているため。

そこで、本発明の第5の方法は、上記2～4の方法における暗号データに対して、簡易な暗号化をさらに加える。

次に、前記(2)の適応型確率モデルへの暗号化処理の組込みについて述べ

る。適応型確率モデルに基づく圧縮では、固定型モデルのように予めデータを調べて記号の出現確率を求めるのではなく、出現確率の予想値を用いる。出現確率の初期予想値として、例えば、全ての記号の出現確率は等しいと予想する。この予想値に基づいて、記号をビット列に変換すると共に、予想値を修正する（処理した記号の出現確率予想値を増やし、他の記号の出現確率予想値を減らす）。上記記号からビット列への変換と出現確率予想値の修正の繰返しにより、データ全体を圧縮する。復元処理は、圧縮データを受取り、圧縮処理と同じ出現確率の初期予想値を用い、ビット列から記号への逆変換と、出現確率予想値の修正の繰返しによりデータを復元する。

上記適応型確率モデルに基づく圧縮方法では、圧縮データのある部分を復元するためには、それより前の全ての部分を復元し、その時点での出現確率の予想値を求めておく必要がある。そこで、圧縮データの先頭部分を暗号化しておけば、それ以降の部分を暗号化しなくても、全体を暗号化したことになる。そこで、本発明の第6の方法は、適応型確率モデルに基づく圧縮において、圧縮データの先頭部分を圧縮する。

固定型確率モデルの場合と同様に、上記方法に対して、以下の解読方法が考えられる。

- (a) 複数の類似したデータと対応する圧縮データ入手し、それらの比較により、本方式による記号とビット列の対応付けの傾向を推定する。
- (b) 一般に、記号の出現確率予想値はデータ処理の過程で急激には変化しないので、記号とビット列の対応付けも、急激には変化しない。そこで、一つのデータの中で同じ記号や記号列が繰り返し現れる場合、圧縮データ中に類似パターンが繰り返し現れる。その分析により、記号とビット列の対応付けを推定する。

上記(a)に対しては、固定型確率モデルの場合と同様に、記号とビット列の対応付けを、偶然あるいは確率に基づく計算を用いて行えばよいので、上記第3の方法により対処できる。上記(b)については、記号のビット列の対応付けを、記号の出現確率予想値以外の情報に基づいて、変更すればよい。固定

型確率モデルの場合と同様に、記号の出現確率予想値に依存して対応するビット列の長さが決まるが、その長さでの0と1の配列には、自由度がある。そこで、データ処理の途中で、記号の出現確率とは無関係に、記号とビット列の対応付けを変更すればよい。以上から、本発明の第7の方法は、データ処理の途中で、記号の出現確率以外の情報に基づいて、記号とビット列の対応付けを変更する。

上記本発明の第1の方法では、データに圧縮手段と暗号化手段を適用するときに必要なメモリ量、および、データに復号化手段と復元手段を適用するときにメモリ量が、当該情報処理システムの主記憶装置の容量を越えないように、圧縮と暗号化あるいは復号化と復元を行うデータの量を設定するので、圧縮暗号化および復号化復元を主記憶装置内で実行でき、中間結果の二次記憶装置への読み書きが不要となる。そこで、効率の向上と電力消費量の削減が達成できる。

第2の方法では、固定型確率モデルに基づく圧縮処理において、記号一ビット列対応情報を暗号化することにより、圧縮データを復元不能とし、圧縮データ自身の暗号化と同じ効果を得ることができる。圧縮データの量が通常数キロから数メガバイトであるのに対し、記号一ビット列対応情報の量は、記号の種類数×1バイト程度であり、無視できるほど小さい。そこで、本方法によれば、暗号化および対応する復号化の効率を向上できる。

第3の方法では、記号とビット列の対応付けを偶然あるいは確率に基づく計算を用いて決定するので、本方法の処理を起動する度に、記号とビット列の対応が全く異なる。そこで、記号とビット列の対応付けの傾向を推定することが困難となり、暗号の強度が増加する。

第4の方法では、記号とビット列の対応付けをデータ処理の途中で変更するので、データ中に繰り返しパターンが暗号データには現れない。そこで、暗号の強度が増加する。

第5の方法では、以上の方法による暗号データを、第2の暗号化手段がさらに暗号化するので、暗号強度が増加する。二重の暗号化であるので、第2の暗

号化手段は簡易なもので十分であり、暗号化の効率を向上できる。また、一般に、暗号化が簡易であれば、対応する復号化も簡易になるので、復号化の効率も向上できる。

第6の方法では、適応型確率モデルに基づく圧縮処理において、圧縮データの先頭部分を暗号化することにより、後続部分における記号の出現確率の予想を困難にし、結果的に、圧縮データ全体を暗号化する。先頭部分のみの暗号化であるため、従来の圧縮データ全体を暗号化する方法に比べ、効率を向上でき、対応する復号化の効率も向上できる。

第7の方法では、記号の出現確率の予想値以外の情報に基づいて、記号とビット列の対応付けをデータ処理の途中で急激に変更するので、データ中の繰り返しパターンが暗号化データに現れない。そこで、暗号の強度が増加する。

図面の簡単な説明

第1図は本発明の第1の実施例における圧縮処理の方法を示すフローチャート、第2図は本発明の第1の実施例の機能構成を示す図、第3図はハフマン符号化における記号とビット列の対応情報（ハフマン木）の一例を示す図、第4図は本発明の第1の実施例における乱数に基づく記号とビット列の対応付けの方法を示すフローチャート、第5図は本発明の第1の実施例における記号とビット列の対応情報を複数通り生成する方法を示すフローチャート、第6図は第3図の記号とビット列の対応情報を変形した例を示す図、第7図は本発明の第1の実施例における復元処理の方法を示すフローチャート、第8図は記号とビット列の対応情報（ハフマン木）の生成の中間結果の例を示す図、第9図は記号とビット列の対応情報（ハフマン木）の生成の中間および最終結果の例を示す図、第10図は第9図の記号とビット列の対応情報を変形した例を示す図、第11図は圧縮かつ暗号化されたデータの例を示す図、第12図は本発明の第2の実施例における記号とビット列の対応情報を複数通り生成する方法を示すフローチャート、第13図は本発明の第2の実施例における記号とビット列の対応情報を複数通り記憶するデータ形式の例を示す図、第14図は本発明の第2の実施例における記号とビット列の対応情報の切換のための式を複数通り生

成する方法を示すフローチャート、第15図は本発明の第3の実施例の機能構成を示す図、第16図は本発明の第3の実施例の圧縮処理の方法を示すフローチャート、第17図は本発明の第3の実施例における関数値に基づく記号とビット列の対応付けの方法を示すフローチャート、第18図は本発明の第3の実施例の復元処理の方法を示すフローチャートである。

発明を実施するための最良の形態

本発明をより詳細に説述するために、添付の図面に従ってこれを説明する。また、以下では、本発明の3つの実施例を述べる。

まず、第1図～第11図を用いて、第1の実施例を説明する。第1の実施例は、前記第1～5の方法の具体例であり、当該情報処理システムの主記憶容量を越えない量のデータに対する圧縮暗号化処理を繰り返す。その圧縮暗号化処理は、固定型確率モデルに基づく圧縮の代表的手法である固定型ハフマン圧縮に、暗号化処理を組込んだものである。第2図は、第1の実施例の機能構成図を示す。ブロック201とブロック215は、それぞれ、完結した情報処理システムであり、通信回線により、ブロック201からブロック215にデータを送ることができる。

ブロック202は、中央処理装置、入出力装置により実現される処理であり、入出力203、制御204、確率情報生成205、圧縮206、暗号化207、送信208の各処理から構成される。ブロック209は、RAM (random access memory) 等により実現される主記憶であり、出現確率情報210を記憶する。ブロック211は、ハードディスク等により実現される二次記憶であり、生データ212、圧縮暗号データ213、暗号化対応情報214を記憶する。

ブロック216は、中央処理装置、入出力装置により実現される処理であり、入出力217、制御218、受信219、復号化220、復元221の各処理から構成される。ブロック222は、RAM (random access memory) 等により実現される主記憶であり、対応情報223を記憶する。ブロック224は、ハードディスク等により実現される二次記憶であり、暗号化対応情報225、圧縮

暗号データ 226、生データ 227 を記憶する。

入出力 203 は、生データを入力し、二次記憶 211 に格納する。また、圧縮暗号化コマンドおよびデータの送信コマンドを入力し、制御 204 に渡す。制御 204 は、圧縮暗号化コマンドを受け取ったときに、確率情報生成 205、圧縮 206、暗号化 207 を順次起動して生データの一部を圧縮暗号化するステップを繰り返すことにより、生データ全体を圧縮暗号化し、二次記憶 211 に格納する。また、その過程で得られた記号とビット列の対応情報（記号一ビット列対応情報と呼ぶ）を暗号化し、二次記憶 211 に格納する。このとき、1 回のステップで圧縮暗号化する生データの量は、処理に必要なメモリ量が当該情報処理システム 201 の主記憶容量を超えないように設定する。また、制御 204 は、圧縮暗号データの送信コマンドを受け取ったときに、送信 208 を起動して、圧縮暗号データと暗号化された記号一ビット列対応情報を情報処理システム 215 に送信する。

確率情報生成 205 は、生データ中の記号の出現頻度を数えることにより、生データにおける記号の出現確率を求め、圧縮 206 に渡す。圧縮 206 は、記号の出現確率情報に基づいて、記号とビット列の対応情報を生成し、主記憶 209 に格納する。また、この記号とビット列対応情報を参照して、生データ 212 の一部を圧縮暗号化し、暗号化 207 に渡す。この処理については、後に第 1、第 3、第 4、第 5、第 6 図を用いて詳述する。

暗号化 207 は、圧縮 206 から圧縮データを受け取って、これを暗号化し、二次記憶 211 に格納する。また、主記憶 210 から記号とビット列の対応情報を読み出し、これを暗号化して、二次記憶 211 に格納する。ここでは、データ保護と暗号化の研究、日本経済新聞社（1983 年）、73 頁から 153 頁に述べられるような従来の対称鍵暗号または非対称鍵暗号を用いる。

送信 208 は、圧縮暗号データおよび暗号化された記号一ビット列対応情報を情報処理システム 215 に送信する。

入出力 217 は、復号化復元コマンドを入力し、制御 218 に渡す。また、二次記憶 224 中の生データ 227 を出力する。制御 218 は、復号化 220、

復号化復元 221 を順次起動して圧縮暗号データ 226 を復号化復元するステップを繰り返すことにより、圧縮暗号化データを復号化復元し、得られた生データを二次記憶 224 に格納する。このとき、1 回のステップで復号化復元する圧縮暗号データの量は、処理に必要なメモリ量が当該システム 215 の主記憶容量を超えないように設定する。

受信 219 は、送信 208 から、暗号化された記号一ビット列対応情報と圧縮暗号データを取り、二次記憶 224 に格納する。復号化 220 は、暗号化 207 の逆変換により、暗号化された記号一ビット列対応情報を復号化し、主記憶 222 に格納する。また、圧縮暗号データを復号化し、復元 221 に渡す。

復元 221 は、復号化 220 から圧縮データを受け取ると、主記憶 222 中の記号一ビット列対応情報に基づいて、圧縮データを復元し、二次記憶 224 に格納する。この処理については、後に第 7 図を用いて詳述する。

第 1 図は、第 2 図の圧縮 206 の動作を示すフローチャートである。ステップ 101 は、乱数を発生し、その値に基づいて、記号の出現確率情報から、記号一ビット列対応情報を一通り生成し、主記憶 209 に格納する。この記号一ビット列対応情報を初期記号一ビット列対応情報と呼ぶこととする。記号一ビット列対応情報は、第 3 図のような二分木（ハフマン木と呼ばれる）により表される。ハフマン木では、各記号が葉ノード（末端のノード）に配置され、根ノードから記号までの枝に付加された 0 と 1 の並びが、記号に対応するビット列を表す。第 3 図の場合は、a が 000、b が 011、e が 1 に対応する。ステップ 101 の詳細については、第 4 図を用いて後に説明する。

ステップ 102 は、初期記号一ビット列対応情報から、他の記号一ビット列対応情報を生成する。この処理の詳細については、第 5 図を用いて後に説明する。ステップ 103 は、上記の初期記号一ビット列対応情報および他の記号一ビット列対応情報の中から一つを選択し、利用対応情報とする。

ステップ 104 は、対象とする生データ（生データ全体のうち制御 204 の指定した部部）の先頭の記号を現記号とする。ステップ 105 は、利用対応情報に基づいて、現記号をビット列に変換する。

ステップ106は、対象とする生データを全て処理したかどうかを判定する。全て処理した場合にはリターンする。そうでない場合には、ステップ107に進み、現記号の次の記号を、新たに現記号とする。ステップ108は、利用対応情報を、現在の利用対応情報以外の記号一ビット列対応情報に変更し、ステップ105に戻る。この変更は、例えば、次の方法による。

記号一ビット列対応情報がL通りあるとすると、それぞれの記号一ビット列情報に0からL-1までの番号を付けておく。ステップ104で求めたビット列を数字とみなした場合の値をv、現在利用対応情報となっている記号一ビット列対応情報の番号をwとする。 $(v + w)$ をLで割った余りyをとると、yは0以上L-1以下の整数となる。新たな利用対応情報として、y番目の記号一ビット列対応情報を選択する。この方法によると、直前に処理した記号と現在利用している記号一ビット列対応情報に依存して、次に利用する記号一ビット列対応情報が選ばれる。

次に、第4図を用いて、第1図のステップ101の詳細を示す。まず、枝に0および1の付加されていないハフマン木を、データ圧縮ハンドブック、トップン、21頁から60頁に示されるような従来のハフマン木生成方法により生成する。次に、そのハフマン木の根ノードに対して第4図に示す処理を適用することにより、ハフマン木の枝に0または1を付加する。

ステップ1011は、当該ノードが葉ノードかどうかを判定する。葉ノードの場合はリターンする（最初はそうでない）。そうでない場合には、ステップ1012にすすみ、乱数を発生する。ステップ1013は、発生した乱数が偶数か奇数かを判定する。偶数の場合には、ステップ1014、奇数の場合には、ステップ1015に進む。ステップ1014は、当該ノードの左の枝に0、右の枝に1を付加する。ステップ1015は、ステップ1014とは逆の値を付加する。ステップ1016は、当該ノードの左の子ノードに対して本処理を再帰的に適用する。ステップ1017は、当該ノードの右の子ノードに対して本処理を再帰的に適用する。

次に、第5図を用いて、図1のステップ102の詳細を説明する。ステップ

1021は、前記第4図の方法で求めた初期記号一ビット列対応情報の段数を求め、変数nに代入する。例えば、第3図のハフマン木の場合には、 $n = 3$ である。ステップ1022は、n桁の0から鳴る2進数を変数aに代入する。ステップ1023は、aの値を1だけ増加する。

ステップ1024は、aのうち値が1の桁を求め、 m_1, m_2, \dots, m_k とする。例えば、aが011の場合には、値が1の桁は（右端から）1と2である。ステップ1025は、初期記号一ビット列対応情報の m_1, m_2, \dots, m_k 段の枝について0と1を交換したもの求め、別の記号一ビット列対応情報とする。初期記号一ビット列対応情報が第3図のハフマン木で、aが011の場合には、第3図の1段および2段の枝について0と1を交換したハフマン木（第6図）を別の記号一ビット列対応情報とする。

ステップ1026は、 $a = 2$ の n 乗-1かどうかを判定する。そうであればリターンし、そうでなければステップ1023に戻る。本処理により、2の n 乗-1通りの記号一ビット列対応情報が生成され、初期記号一ビット列対応情報と合わせて、2の n 乗通りの対応付けが得られる。

次に、第7図を用いて、第2図の復号化220の詳細を説明する。ステップ701は、初期記号一ビット列対応情報から他の記号一ビット列対応情報を生成する。ここでは、第1図のステップ102と同じ処理を用いる。そこで、初期記号一ビット列対応情報が圧縮206の初期記号一ビット列対応情報と等しければ、全ての記号一ビット列対応情報が圧縮の記号一ビット列対応情報と等しくなる。

ステップ702は、一つの記号一ビット列対応情報を選択し、利用対応情報とする。この処理も、圧縮のステップ103と同じ処理を用いる。そこで、圧縮時と同じ利用対応情報が選ばれる。ステップ703は、利用対応情報に基づいて、圧縮データの先頭から記号に対応するビット列を切り出す。ステップ704は、上記ビット列を記号に変換することにより、元の生データを求める。

ステップ705は、圧縮データを全て処理したかどうかを判定する。全て処理した場合にはリターンする。そうでない場合には、ステップ706に進み、

圧縮データから処理済みのビット列を除去する。ステップ 707 は、利用対応情報を他の記号一ビット列対応情報に変更し、ステップ 703 に戻る。この処理は、第 1 図のステップ 108 と同じ処理を用いる。そこで、圧縮時と同じ利用対応情報が選ばれる。

以下、第 8 図～第 11 図を用いて、本実施例により実際のデータが圧縮暗号化できることを示す。実用時には、大量の生データを圧縮暗号化するが、ここでは、説明を簡単にするため、生データとして、"this theater"を考える。まず、入出力 203 が、このデータを入力し、二次記憶 211 に格納する。次に、入出力 203 が、圧縮暗号化コマンドを入力すると、これを制御 204 に渡す。一般には、生データを一度に処理するのではなく、当該情報処理システム 201 の主記憶内で処理可能な量に分割して処理するが、ここでは、生データの量が少ないので一度に処理する。

制御 204 は、確率情報生成 205 を起動し、生データ中の記号の出現確率を求め、圧縮 206 に渡す。ここでは、記号の総数は、s と t の間の空白を含めて 12 である。そこで、例えば、t の出現確率は 3 / 12 であり、h の出現確率は 2 / 12 である。

圧縮 206 は、第 1 図の処理により、記号の出現確率情報に基づいて、二次記憶 211 中の生データを圧縮する。ステップ 101 は、まず、従来の方法により、記号の出現確率に基づいて、0、1 の付加されていないハフマン木を生成する。ここでは、第 8 図のハフマン木を得る。次に、このハフマン木の根ノードに第 4 図の処理を適用することにより、ハフマン木の枝に 0、1 を付加する。

ステップ 1011 は、当該ノードが葉ノードかどうかを判定する。ここでは、根ノードなので No となる。ステップ 1012 は、乱数 r を発生する。ここでは、例えば、r = 649023 となるとする。ステップ 1013 は r が偶数かどうかを判定する。ここでは、No となる。ステップ 1015 は、当該ノードの左の枝に 1、右の枝に 0 を付加する。

ステップ 1016 は、当該ノードの左の子ノードに本処理を再帰的に適用す

る。左の子ノードに対して、ステップ1011はN_oとなる。ステップ1012では、r = 89024となるとする。ステップ1013はY e sとなる。ステップ1014は、当該ノードの左の枝に0、右の枝に1を付加する。この時点でのハフマン木を第9図ブロック901に示す。

次に、ステップ1016は、当該ノードの左の子ノードに本処理を再適用する。同様の処理の繰り返しにより、最終的に、ブロック902のハフマン木が得られたとする。これが、初期記号一ビット列対応情報となる。

ステップ102は、第5図の処理により、ブロック902の初期記号一ビット列対応情報から、他の記号一ビット列対応情報を生成する。ステップ1021は、ブロック902のハフマン木の段数を求め、変数nに代入する。ここでは、n = 4となる。ステップ1022は、4桁の0から成る2進数0000を変数aに代入する。ステップ1023は、aに0001を代入する。ステップ1024は、aのうち値の1の桁を求める。ここでは、1桁目が1である。ステップ1025は、ブロック902の1段目の枝の0と1を交換したものを、別の記号一ビット列対応情報として記憶する。ここでは、第10図ブロック1001を記憶する。

ステップ1026は、a = 2の4乗 - 1かどうかを判定する。ここでは、N_oとなる。ステップ1023は、aに0010を代入する。ステップ1024は、aのうち値が1の桁が、2桁目であることを認識する。ステップ1025は、ブロック902の2段目の枝を変更し、ブロック1002を得る。同様に次にループでは、ブロック902の1及び2段目の枝を変更し、ブロック1003を得る。同様の処理の繰り返しにより、15個(2の4乗 - 1個)の記号一ビット列対応情報を求め、初期記号一ビット列対応情報を含め、16通りの記号一ビット列対応情報を得る。

ブロック103は、上記16通りの記号一ビット列対応情報から一つを選択し、利用対応情報とする。ここでは、初期記号一ビット列対応情報すなわちブロック902を選択するとする。ブロック104は、生データの先頭の記号を現記号とする。ここでは、tが現記号となる。ブロック105は、利用対応情

報に基づいて、tを対応するビット列に変換する。ここでは、01に変換する。ブロック106は、生データを全て処理したかどうかを判定する。ここでは、Noとなる。ブロック107は、現記号の次の記号を現記号とする。ここでは、hが現記号となる。

ブロック108は、前述の方法により利用対応情報を他の記号一ビット列対応情報に変更する。現在の利用対応情報は、初期記号一ビット列対応情報すなわち0番目の記号一ビット列対応情報である。上記tに対応するビット列01を2進数とみなすと、その値は1である。そこで、現在の記号一ビット列対応情報の番号に1を加え、1番目の記号一ビット列対応情報を新たな利用対応情報とする。これは、ブロック1001のハフマン木である。

次に、利用対応情報に基づいて、現記号hをビット列010に変換する。ブロック106はNoとなる。ブロック107は、iを現記号とする。ブロック108は、現在の利用対応情報が1番目の記号一ビット列対応情報であり、hに対応するビット列010の表す数値が2なので、3番目の記号一ビット列対応情報を利用対応情報とする。以下同様の処理により、生データを圧縮データに変換する。

第11図は、圧縮データを示す。ただし、3番目の記号以下のビット列は省略している。

次に、暗号化207は、上記圧縮データを従来の暗号方法により、さらに暗号化し、二次記憶211に格納する。また、初期記号一ビット列対応情報を暗号化し、二次記憶211に格納する。

次に、入出力203が送信コマンドを入力し、制御204に渡す。制御204は、送信208を起動して、圧縮暗号データと暗号化された初期記号一ビット列対応情報を情報処理システム215に送信する。

受信219は、上記の圧縮暗号データと暗号化された初期記号一ビット列対応情報を受け取り、二次記憶224に格納する。入出力217は、復号化復元コマンドを入力し、これを制御218に渡す。制御218は、復号化220、復元221を起動して圧縮暗号データを復号化復元する。一般には、圧縮暗号

データを一度に処理するのではなく、当該情報処理システム215のメモリ量を超えないようにデータを分割して処理するが、ここでは、データ量が少ないので一度に処理する。

復号化220は、従来の復号化方法により、暗号化された初期記号一ビット列対応情報を復号化し、その結果すなわちブロック902を主記憶222に格納する。また、圧縮暗号データを前記暗号化207を施す前のデータ、すなわち、第11図の圧縮データに戻し、復元221に渡す。

復元221は、第7図の処理により、主記憶中の初期記号一ビット列対応情報に基づいて、第11図の圧縮データを復元する。ステップ701は、ステップ102と同じ処理により、ブロック902の初期記号一ビット列対応情報から他の記号一ビット列対応情報を生成する。ステップ702は、ステップ103と同じ処理により、一つの記号一ビット列対応情報を選択し、利用対応情報とする。ここでは、初期記号一ビット列対応情報（ブロック902）を選択する。

ステップ703は、利用対応情報に基づいて、第11図の圧縮データの先頭から、01を切り出す。ステップ704は、01をtに変換する。ステップ705はNoとなる。ステップ706は、第11図の圧縮データの先頭から01をのぞく。ステップ707は、ステップ108と同じ処理により、2番目の記号一ビット列対応情報を新たな利用対応情報とする。以下、同様の処理により、圧縮データを生データに変換し、二次記憶224に格納する。

最後に、入出力217が出力コマンドを受け取り、二次記憶224中の生データを出力する。

以上述べたように、本実施例によれば、生データを圧縮かつ暗号化し、また、圧縮暗号データを復号化かつ復元して元の生データを復元することができる。

データを圧縮暗号化するときに必要なメモリ量、および、データを復号化復元するときに必要なメモリ量が、当該情報処理システムの有する主記憶容量を超えないように、一度に圧縮暗号化するデータの量および一度に復号化復元するデータの量を設定するので、圧縮暗号化および復号化復元を主記憶内で実行

できる。そこで、中間結果の二次記憶への読み書きが不要となり、効率の向上と電力消費量の削減が達成できる。

また、従来の圧縮暗号化方法における暗号化処理では、圧縮データおよび記号一ビット列対応情報に対して暗号化処理を適用していた。これに対し、本実施例では、（1）乱数に基づく記号一ビット列対応情報の生成、（2）記号一ビット列対応情報の変更、（3）記号一ビット列対応情報への暗号化処理の適用、（4）圧縮データに対する暗号化処理の適用により生データを暗号化する。

（1）の処理は、ハフマン木の枝への0、1の付加である。ハフマン木は、記号の種類だけの葉ノードを有する木であり、枝の個数は、（記号の種類－1）×2である。記号の種類は高々300程度であるので、（1）の処理は、高々600程度の枝に0または1を付加する処理であり、数キロから数メガバイトの圧縮データに複雑な暗号化を加える従来処理に比べて、無視できるほど小さい時間で実行できる。同様に、（2）の処理時間も無視できるほど小さい。（3）については、記号一ビット列対応情報は、圧縮データに比べて量が極めて少ないので、その暗号化の時間は無視できるほど小さい。（4）における暗号化処理は、以下の理由から簡易なもので十分である。

- （a）上記（1）～（3）により、生データが既に暗号化されている。
- （b）（1）により、記号とビット列の対応付けを乱数に基づいて行うので、生データが類似あるいは同じであっても、本暗号化処理を起動する度に、記号とビット列の対応が全く異なる。また、（2）により、処理の途中で、記号とビット列の対応を変更するので、生データ中の繰り返しパターンが暗号データに現れない。そこで、類似データの比較や、繰り返しパターンの分析などに解読法への対処がすでになされている。

そこで、（4）の暗号化処理の時間は、従来の暗号化処理に比べて小さくすることができる。以上から、本実施例によれば、圧縮暗号化の時間を短縮することができる。

一方、従来の復号化復元方法における復号化処理では、圧縮暗号データおよび暗号化された記号とビット列の対応情報に対して復号化処理を適用していた。

これに対し、本実施例では、（1）記号一ビット列対応情報への復号化処理の適用、（2）記号一ビット列対応情報の変更、（3）圧縮暗号データに対する復号化処理の適用により、圧縮暗号データを復号化する。

このうち、（1）、（2）の処理時間は、暗号化の場合と同様に、無視できるほど小さい。一般に、暗号化処理が簡易な場合、復号化処理も簡易になる。そこで、（3）の復号化処理の時間も、従来の復号化処理に比べて小さい。以上から、本実施例によれば、復号化復元の時間を短縮することができる。

次に、第12図～第14図を用いて、本発明の第2の実施例を述べる。第2の実施例は、前記第1の実施例の変形例である。前記第1の実施例では、生データにおける繰り返しパターンを隠蔽するために、以下の処理により、複数通りの記号一ビット列対応情報を用いて、記号をビット列に変換していた。

（1）初期記号一ビット列対応情報（初期のハフマン木）の段数をnとしたときに、第1図ステップ102において、2のn乗通りの記号一ビット列対応情報を生成する。

（2）直前の記号から求めたビット列を数字とみなした値をv、現在利用中の記号一ビット列対応情報の番号をw、記号一ビット列対応情報の総数をLとしたときに、ステップ108において、 $(v + w)$ をLで割った余りを、次に利用する記号一ビット列対応情報の番号とする。

本実施例では、以下の方法により、繰り返しパターンの隠蔽をより完全にする。

（a）上記（1）に関しては、ステップ102において、より多数の記号一ビット列対応情報を生成する。

（b）上記（2）に関しては、次に利用する記号一ビット列対応情報の番号を求めるための式を複数通り用意しておき、ステップ108において、その中の一つを用いて、次に利用する記号一ビット列対応情報の番号を決める。どの式を用いるかは、データの送信者と受信者以外には秘密となっている共通パラメータを圧縮206と復元221に設定しておき、そのパラメータに基づいて決定する。

上記(a)は、ステップ102の詳細として、第1の実施例における第5図の方法の代わりに、第12図の方法を用いることにより実現できる。ステップ10201は、初期記号一ビット列対応情報の根および中間ノードの個数を、変数mに代入する。ステップ10202は、初期記号一ビット列対応情報の根および中間ノードに1からmまでの番号を付ける。ステップ10203は、m桁の0から成る2進数を変数aに代入する。ステップ10204は、aを1だけ増加する。ステップ10205は、aのうち値が1の桁を求め、m1, m2, ..., mkとする。

ステップ10206は、初期記号一ビット情報のm1, m2, ..., mk番目のノードから出る左右の枝の値(0または1)を交換したものを、別の記号一ビット列対応情報とする(初期記号一ビット列対応情報は2進木であり、各々の根および中間ノードから二本ずつの枝が出ている)。ステップ10207は、(a = 2のm乗 - 1)かどうかを判定し、そうであれば、リターンし、そうでなければ、ステップ10204に戻る。

上記の処理によると、初期記号一ビット列対応情報を含め、2のm乗通りの記号一ビット列対応情報が得られる。これは、前記第1の実施例における2のn乗通りの記号一ビット列対応情報よりも多数である。例えば、第3図の初期記号一ビット列対応情報の場合、m = 4, n = 3なので、本実施例の記号一ビット列対応情報は16通り、第1の実施例の記号一ビット列対応情報は8通りとなる。また、第8図の初期記号一ビット列対応情報の場合には、m = 7, n = 4なので、本実施例の記号一ビット列対応情報は128通り、第1の実施例の記号一ビット列対応情報は16通りとなる。

上記のような多数の記号一ビット列対応情報を個別に記憶すると多量のメモリを要する。そこで、次の方法により、メモリを節約する。同じ生データに対する複数の記号一ビット列対応情報は、2進木の構造が等しく、その枝に附加された0または1の値のみが異なる。そこで、複数の記号一ビット列対応情報を、一本の2進木の枝に複数の値を附加することにより、表すことにする。例えば、第10図の三つの2進木をまとめて、第13図のように表す。

次に、上記（b）において複数通りの式を求める方法を、第14図により説明する。ここでは、前記vとwから値を計算する文字数10の式を複数通り求めることにする。ステップ1401は、変数SYM-SETに、v, w, +, -, ^, ∨, ◇, (,)の9個の文字から成る集合を代入する（ただし、+および-は、v, wを数値とみなした加算および減算であり、^, ∨, ◇は、v, wをビット列とみなした論理積、論理和、排他的論理和であり、(,)は、計算の順序を指定するための括弧を表す）。ステップ1402は、EXP-SETに、空集合を代入する。ステップ1403は、SYM-SET中の文字の長さ10の配列（ただし同じ文字を二回以上使ってもよい）を求め、変数EXPに代入する。

ステップ1404は、EXPが数式として意味があるかどうかを判定する。この判定は、コンパイラ、産業図書（1981年）、41頁から140頁に述べられているような構文解析技術により実現できる。意味がある場合には、ステップ1405に進み、EXPをEXP-SETに追加する。意味がない場合には、ステップ1406に進む。

ステップ1406は、SYM-SETから生成可能な長さ10の配列を全て処理し終わったかどうかを判定する。処理し終わった場合には、リターンする。そうでない場合には、ステップ1407に進む。ステップ1407は、SYM-SET中の文字の長さ10の新たな配列をEXPに代入した後、ステップ1404に戻る。

以上述べた第2の実施例によると、第1の実施例に比べ、記号とビット列の対応付けがより多様になるため、生データ中の繰り返しパターンがより完全に隠蔽される。その結果、暗号強度がより大きくなる。

次に、第15図～第18図を用いて、本発明の第3の実施例を説明する。第3の実施例は、前記で述べた第6および7の方法の具体例であり、当該情報処理システムの主記憶容量を超えない量のデータに対する圧縮暗号化処理を繰り返す。その圧縮暗号化処理は、適応型確率モデルに基づく圧縮の代表的手法である適応型ハフマン圧縮に、暗号化処理を組み込んだものである。

まず、本実施例の説明の準備として、適応型ハフマン符号化について簡単に説明する。前記第1の実施例で用いた固定型ハフマン符号化では、生データ中の記号の頻度を調べることにより、記号の出現確率を求め、その確率に基づいて、記号とビット列の対応付け（ハフマン木の構造）を決定した後、生データ中の記号を対応するビット列に変換していた。これに対し、本実施例で用いる適応型ハフマン符号化では、記号の出現確率の予想値を用い、以下のステップにより記号をビット列に変換する。

- (1) 記号の出現確率予想値の初期値を設定する。例えば、全ての記号の出現確率を $1/p$ とする（ただし、 p は記号の種類数。予想値に基づいて、記号一ビット列対応情報（ハフマン木）の初期値を生成する。
- (2) 生データから次の記号を読み込み、記号一ビット列対応情報に基づいて、記号を対応するビット列に変換する。
- (3) 生データを全て処理した場合には終了。そうでない場合には、記号の出現確率予想値を更新し、それに伴って、記号一ビット列対応情報を更新（具体的には、ハフマン木を変更）した後、(2)に戻る。

上記記号一ビット列対応情報の更新方法は様々であるが、ここでは、以下の方法を採用する。すなわち、各記号の出現頻度の初期値を全て 1 としておく。

(3) で、直前に処理した記号の頻度を一つ増やした後、頻度に比例した出現確率を計算し、更新後の出現確率とする。

復元時には、出現確率予想値の初期値および修正方法を、圧縮時と同じ値および方法に設定し、ビット列から記号への変換、出現確率予想値の修正の繰り返しにより、生データを復元する。

第15図は、本実施例の機能構成を示す図である。ブロック 1501 とブロック 1513 は、それぞれ、完結した情報処理システムであり、通信回線により、ブロック 1501 からブロック 1513 にデータを送ることができる。

ブロック 1502 は、中央処理装置、入出力装置により実現される処理であり、入出力 1503、制御 1504、圧縮 1505、暗号化 1506、送信 1507 の各処理から構成される。ブロック 1508 は、RAM (random access

memory) 等により実現される主記憶であり、出現確率予想情報 1509 を記憶する。ブロック 1510 は、ハードディスク等により実現される二次記憶であり、生データ 1511、圧縮暗号データ 1512 を記憶する。

ブロック 1514 は、中央処理装置、入出力装置により実現される処理であり、入出力 1515、制御 1516、受信 1517、復号化 1518、復元 1519 の各処理から構成される。ブロック 1520 は、RAM (random access memory) 等により実現される主記憶であり、出現確率予想情報 1521 を記憶する。ブロック 1522 は、ハードディスク等により実現される二次記憶であり、圧縮暗号データ 1523、生データ 1524 を記憶する。

入出力 1503 は、生データを入力し、二次記憶 1510 に格納する。また、圧縮暗号化コマンドおよびデータの送信コマンドを入力し、制御 1504 に渡す。制御 1504 は、圧縮暗号化コマンドを受け取ったときに、圧縮 1505、暗号化 1506 を順次起動して生データの一部を圧縮暗号化するステップを繰り返すことにより、生データ全体を圧縮暗号化し、二次記憶 1510 に格納する。このとき、1 回のステップで圧縮暗号化する生データの量は、処理に必要なメモリ量が当該情報処理システム 1501 の主記憶容量を超えないように設定する。また、制御 1504 は、圧縮暗号データの送信コマンドを受け取ったときに、送信 1507 を起動して、圧縮暗号データを情報処理システム 1513 に送信する。

圧縮 1505 は、生データ 1511 の一部を圧縮暗号化し、暗号化 1506 に渡す。その過程で、記号の出現確率予想情報 1509 の初期値を、主記憶 1508 に設定し、その値を更新する。この処理については、後に第 16 図および第 17 図を用いて詳述する。

暗号化 1506 は、圧縮 1505 から圧縮データを受け取って、その先頭から指定された量だけを暗号化し、二次記憶 1510 に格納する。ここでは、第 1 の実施例と同様に、データ保護と暗号化の研究、日本経済新聞社（1983 年）、73 頁から 153 頁に述べられるような従来の対称鍵暗号または非対称鍵暗号を用いる。送信 1507 は、圧縮暗号データを情報処理システム 151

3に送信する。

入出力1515は、復号化復元コマンドを入力し、制御1516に渡す。また、二次記憶1522中の生データ1524を出力する。制御1516は、復号化1518、復元1519を順次起動して圧縮暗号データ1523を復号化復元するステップを繰り返すことにより、圧縮暗号データを復号化復元し、得られた生データを二次記憶1524に格納する。このとき、1回のステップで復号化復元する圧縮暗号データの量は、処理に必要なメモリ量が当該システム1513の主記憶容量を超えないように設定する。

受信1517は、送信1507から、圧縮暗号データを受取り、二次記憶1522に格納する。復号化1518は、暗号化1506の逆変換により、圧縮暗号データの先頭から指定された量だけを復号化し、復元1519に渡す。

復元1519は、復号化1518から圧縮データを受け取ると、これを復元し、二次記憶1522に格納する。その過程で、記号の出現確率予想情報1521の初期値を、主記憶1520に設定し、その値を更新する。この処理については、後に第18図を用いて詳述する。

第16図は、圧縮1505の詳細な処理方法を示す。ステップ1601は、生データのうち先頭から指定された個数の記号に対して、従来の適応型ハフマン圧縮を適用する。本ステップが終了した時点では、ハフマン木ができている。ステップ1602は、そのハフマン木の枝の1、0の値を、圧縮1505と復元1519の両者に共通な秘密パラメータ（鍵と呼ばれる）およびこれまでに処理した生データに依存して変更し、その結果を（これからの圧縮暗号化処理における）初期記号一ビット列対応情報とする。本処理の詳細については、第17図を用いて後に詳述する。

ステップ1603は、初期記号一ビット列対応情報から他の記号一ビット列対応情報を求める。その方法は、前記第1図のステップ102と同様である。ここで求めた複数のハフマン木は、個別に記憶するのではなく、前記第13図で説明したように、一本の木構造の枝に複数の1、0の値を付加することにより記憶する。

ステップ1604は、前記ステップ103と同様の方法により、一つの記号一ビット列対応情報を選択し、利用対応情報とする。ステップ1605は、生データの次の記号を現記号とする。ステップ1606は、利用対応情報に基づいて、現記号をビット列に変換する。ステップ1607は、生データを全て処理したかどうかを判定する。処理した場合には、リターンし、そうでない場合には、ステップ1608に進む。ステップ1608は、直前に処理した記号の頻度を1だけ増やし、それに伴って各記号の出現確率予想情報を更新し、その値に基づいて、全ての記号一ビット列対応情報を変更する。実際の計算としては、全ての記号一ビット列対応情報を表す一本のハフマン木の構造を変更する。ステップ1609は、ステップ108と同様の方法により、利用対応情報を他の記号一ビット列対応情報を変更する。

次に、上記ステップ1602における初期記号一ビット列対応情報の生成の詳細を説明する。まず、これまでに生成されたハフマン木の枝の0、1の値を削除し、次に、第17図の処理をそのハフマン木の根ノードに適用することより、ハフマン木の枝に新たに値を付加する。第17図の処理は、前記第1の実施例第4図の処理と基本的には同じであり、乱数 r の代わりに関数値 h に基づいて枝に値を付加する点のみが異なる。

ステップ1011は、当該ノードが葉ノードかどうかを判定する。ステップ15022は、圧縮1505と復元1519に共通の秘密パラメータ（鍵）およびこれまでに数えた記号の頻度を入力とし、整数値を出力とし、入力が少しでも変われば出力が全く変わらるような関数 f により、整数値 h を算出する。そのような関数は、アイ・イー・イー・トランザクション オン インフォーメーション セオリー、30巻、5号、776頁から780頁（IEEE Transaction on Information Theory, Vol. 30, No. 5, pp. 776-780）に述べるような方法により実現できる。上記の f の性質により、この時点までの生データが少しでも異なれば（記号の頻度が一つでも異なれば）、全く異なる記号一ビット列対応情報が生成されるので、類似の生データを用いた暗号の解析を防止することができる。

ステップ15023は、hが偶数か奇数かを判定する。偶数の場合には、ステップ1014に進み、左右の枝にそれぞれ0および1を付加し、奇数の場合には、ステップ1015に進み、左右の枝にそれぞれ1および0を付加する。ステップ1016は、左の子ノードに本処理を再帰的に適用する。ステップ1017は、右の子ノードに本処理を再帰的に適用する。

最後に、第18図を用いて、復元1519の処理の詳細を説明する。ステップ1801は、指定された個数の記号を、従来の適応型ハフマン圧縮の復元方法により復元する。ここで個数の指定値は、前記第16図ステップ1601の指定値と等しくする。ステップ1802は、前記ステップ1602と同じ方法により、初期記号一ビット列対応情報を生成する。ステップ1803は、前記ステップ1603と同じ方法により、初期記号一ビット列対応情報から他の初期記号一ビット列対応情報を生成する。ステップ1804は、前記ステップ1604と同じ方法により、一つの記号一ビット列対応情報を選択し、利用対応情報をとする。ステップ1805は、利用対応情報に基づいて、圧縮データの先頭から記号に対応するビット列を切り出す。このステップからステップ1808までは、前記第1の実施例のステップ703から706と同様である。

ステップ1809は、直前に復元した記号に依存して、記号の出現確率予想情報を修正し、修正値に基づいて、全ての記号一ビット列対応情報を変更する。実際の計算としては、全ての記号一ビット列対応情報を表す一本のハフマン木の構造を変更する。ステップ1810は、ステップ1609と同様の方法により、利用対応情報を他の記号一ビット列対応情報を変更した後、ステップ1805に戻る。

以上述べたように、本実施例によれば、生データを圧縮かつ暗号化し、また、圧縮暗号データを復号化かつ復元して元の生データを復元することができる。

データを圧縮暗号化するときに必要なメモリ量、および、データを復号化復元するときに必要なメモリ量が、当該情報処理システムの有する主記憶容量を超えないように、一度に圧縮暗号化するデータの量および一度に復号化復元す

るデータの量を設定するので、圧縮暗号化および復号化復元を主記憶内で実行できる。そこで、中間結果の二次記憶への読み書きが不要となり、効率の向上と電力消費量の削減が達成できる。

また、従来の圧縮暗号化方法における暗号化処理では、圧縮データ全体に対して暗号化処理を適用していた。これに対し、本実施例では、（1）関数値に基づく記号一ビット列対応情報の生成、（2）記号一ビット列対応情報の変更、（3）圧縮データの先頭部分に対する暗号化処理の適用により生データを暗号化する。

（1）の処理は、ハフマン木の枝への0、1の付加である。前記第1の実施例で述べたように、ハフマン木は、記号の種類だけの葉ノードを有する木であり、枝の個数は、（記号の種類-1）×2である。記号の種類は高々300程度であるので、（1）の処理は、高々600程度の枝について、0または1を付加する処理であり、数キロから数メガバイトの圧縮データに複雑な暗号化を加える従来処理に比べて、無視できるほど小さい時間で実行できる。同様に、（2）の処理時間も無視できるほど小さい。（3）は、圧縮データの先頭部分、例えば100記号分を暗号化する処理であり、圧縮データ全体（通常数万から数十万記号）を暗号化する従来処理に比べて、無視できるほど小さい時間で実行できる。

一方、従来の復号化復元方法における復号化処理では、圧縮暗号データ全体に対して復号化処理を適用していた。これに対し、本実施例では、（1）関数値に基づく記号一ビット列対応情報の生成、（2）記号一ビット列対応情報の変更、（3）圧縮暗号データの先頭部分に対する復号化処理の適用により、圧縮暗号データを復号化する。これら（1）～（3）の処理時間は、暗号化の場合と同様に、無視できるほど小さい。以上から、本実施例によれば、復号化復元の時間を短縮することができる。

産業上の利用可能性

以上のように、本発明によれば、データを圧縮し、かつ、暗号化する処理、

および、圧縮、暗号化されたデータを復号化し、かつ、復元する処理を当該情報処理システムの主記憶装置内で実行できるので、中間結果の二次記憶装置への読み書きを不要化でき、効率向上と電力消費量の削減が達成できる。

また、本発明によれば、データを圧縮し、かつ、暗号化する処理、および、圧縮、暗号化されたデータを復号化し、かつ、復元する処理における計算量を削減できるので、効率向上が達成できる。

請 求 の 範 囲

1. データを入力または受信するデータ入力手段と、当該データを圧縮するデータ圧縮手段と、圧縮されたデータを復元する復元手段と、主記憶装置とを有する情報処理システムにおいて、

データを暗号化する暗号化手段と、暗号化されたデータを復号化する復号化手段とを備え、

前記圧縮手段と前記暗号化手段とが前記データの一部について圧縮と暗号化を行う圧縮暗号化ステップと、前記復元手段と前記復号化手段とが圧縮かつ暗号化されたデータの一部について復元と復号化を行う復元復号化ステップの実行において、

前記圧縮暗号化ステップおよび前記復元復号化ステップが処理する一連のデータの量を、当該処理に必要なメモリの量が前記情報処理システムの前記主記憶装置の容量を越えないように設定し、前記圧縮暗号化ステップおよび前記復元復号化ステップを1回以上繰り返すことを特徴とするデータの圧縮・暗号化方法。

2. 前記情報処理システムが、前記圧縮・暗号化ステップおよび前記復元・復号化ステップの用いるメモリの量を測定あるいは推定し、上記データの量の設定を自動的に行うことの特徴とする請求の範囲第1項記載のデータの圧縮・暗号化方法。

3. データを入力または受信するデータ入出力手段と、当該データを圧縮する圧縮手段と、その逆変換となる復元手段を有する情報処理システムにおいて、暗号化手段とその逆変換となる復号化手段を設け、

前記圧縮手段が、データ中の各記号の出現確率に基づいて記号をビット列に変換し、

前記暗号化手段が、前記記号の出現確率または記号とビット列の対応またはそれと等価な情報を暗号化し、

前記復号化手段が、前記暗号化された情報を復号化することを特徴とするデ

ータの圧縮・暗号化方法。

4. データを入力または受信するデータ入出力手段と、当該データを圧縮する圧縮手段と、その逆変換となる復元手段を有する情報処理システムにおいて、暗号化手段と、その逆変換となる復号化手段を設け、

前記圧縮手段が、データ中の各記号の出現確率の予想に基づいて記号をビット列に変換すると共に、上記出現確率の予想を修正し、

前記暗号化手段が、前記圧縮手段の求めたビット列（圧縮データ）の先頭部分を暗号化し、

前記復号化手段が、上記暗号化された圧縮データを復号化することを特徴とするデータの圧縮・暗号化方法。

5. 圧縮における記号とビット列の対応付けを、偶然あるいは確率に基づく計算を用いて決定することを特徴とする請求の範囲第3項または第4項記載のデータの圧縮・暗号化方法。

6. 圧縮における記号とビット列の対応付けを、記号の頻度、出現確率、それらと等価な情報のいずれとも異なる情報（例えばデータの総ビット数）に基づいて決定することを特徴とする請求の範囲第3項または第4項記載のデータの圧縮・暗号化方法。

7. 圧縮における記号とビット列の対応付けを、データ圧縮の途中で変更することを特徴とする請求の範囲第3項記載のデータの圧縮・暗号化方法。

8. 圧縮における記号とビット列の対応付けを、記号の出現確率の予想以外の情報に基づいて変更することを特徴とする請求の範囲第4項記載のデータの圧縮・暗号化方法。

9. 第2の暗号化手段と、その逆変換となる第2の復号化手段を設け、前記第2の暗号化手段が、前記圧縮データを暗号化し、前記第2の復号化手段が、前記暗号化された圧縮データを復号化することを特徴とする請求の範囲第3項または第4項記載のデータの圧縮・暗号化方法。

10. 第2の圧縮手段と、その逆変換となる第2の復元手段を設け、第2の圧縮手段がデータを圧縮し、圧縮後のデータに対して、前記の圧縮および暗号化

を行い、

前記圧縮および暗号化されたデータに対して、前記の復号化およびデータ復元を行い、その結果のデータを第2の復元手段が復元することを特徴とする請求の範囲第1項または第3項または第4項記載のデータの圧縮・暗号化方法。

1 1. 各データ圧縮手段および各暗号化手段と、各データ復元手段および各復号化手段を、通信回線で接続された二つの情報処理サブシステムに配置し、各データ圧縮手段および各暗号化手段を配置した情報処理サブシステムから、各データ復元手段および各復号化手段を配置した情報処理サブシステムへ、圧縮したデータ、圧縮かつ暗号化したデータ、暗号化した記号とビット列の対応情報のうち少なくとも一つを送信することを特徴とするデータの圧縮・暗号化方法。

1 2. 入力した1組のデータを圧縮し、暗号化する情報処理システムにおいて、前記1組の入力データを前記情報処理システムの主記憶装置の空き容量に応じて区分し、当該区分した単位毎に前記主記憶装置との間で圧縮・暗号化処理を施すことを特徴とするデータの圧縮・暗号化方法。

1 3. 入力した1組のデータを圧縮し、暗号化する情報処理システムにおいて、

前記1組の入力データを前記情報処理システムの主記憶装置の空き容量に応じて区分し、当該区分した単位毎に当該データを圧縮し、当該圧縮結果データを前記主記憶装置に記憶するデータ圧縮手段と、

前記主記憶装置に記憶した圧縮結果データを暗号化処理して出力する暗号化手段と、

前記データ圧縮手段と前記暗号化手段との一連の処理を、前記入力した1組の入力データの全てを圧縮・暗号化するまで繰り返し処理制御する制御手段とを具備して成るデータの圧縮・暗号化装置。

1 4. 主記憶装置と二次記憶装置を有し、データを圧縮し、暗号化する情報処理システムにおいて、

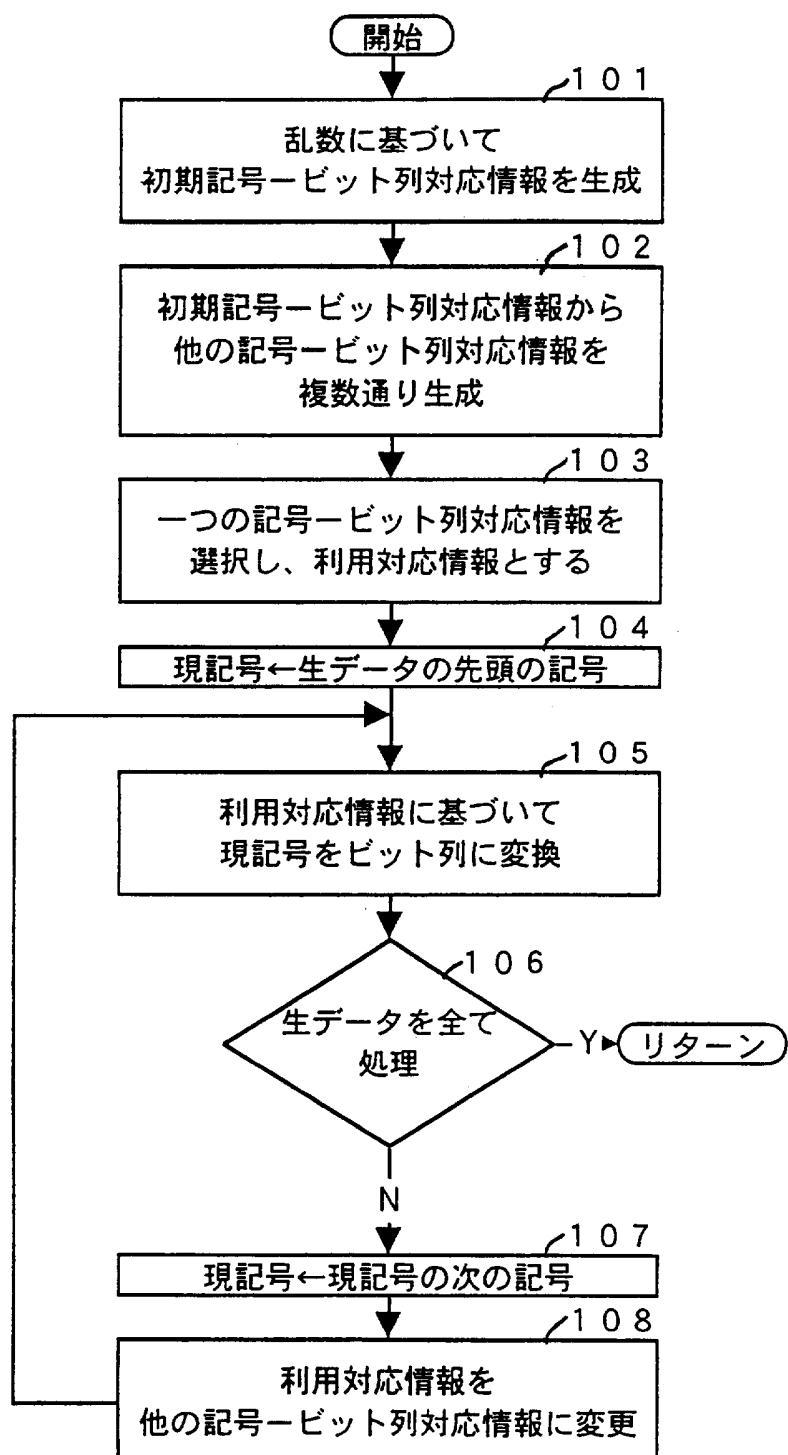
1組のデータを入力し、当該1組のデータを前記二次記憶装置に記憶するデータ入力手段と、

前記1組の入力データを前記情報処理システムの主記憶装置の空き容量に応じて区分し、当該区分した単位毎に当該データを読み出して圧縮し、当該圧縮結果データを前記主記憶装置に記憶するデータ圧縮手段と、

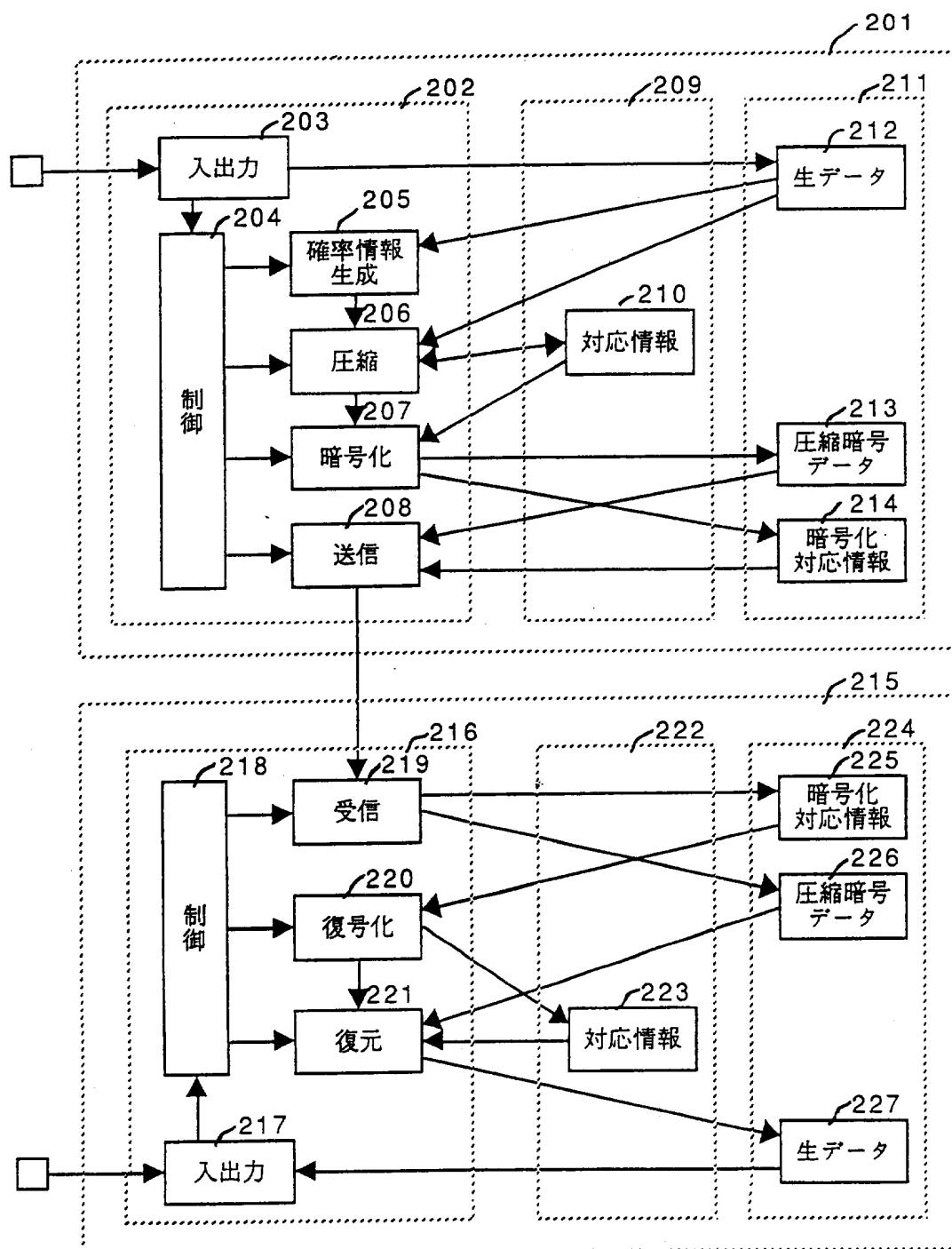
前記主記憶装置に記憶した圧縮結果データを暗号化処理して前記二次記憶装置に記憶する暗号化手段と、

前記データ圧縮手段と前記暗号化手段との一連の処理を、前記入力した1組の入力データの全てを圧縮・暗号化するまで繰り返し処理制御する制御手段とを具備して成るデータの圧縮・暗号化装置。

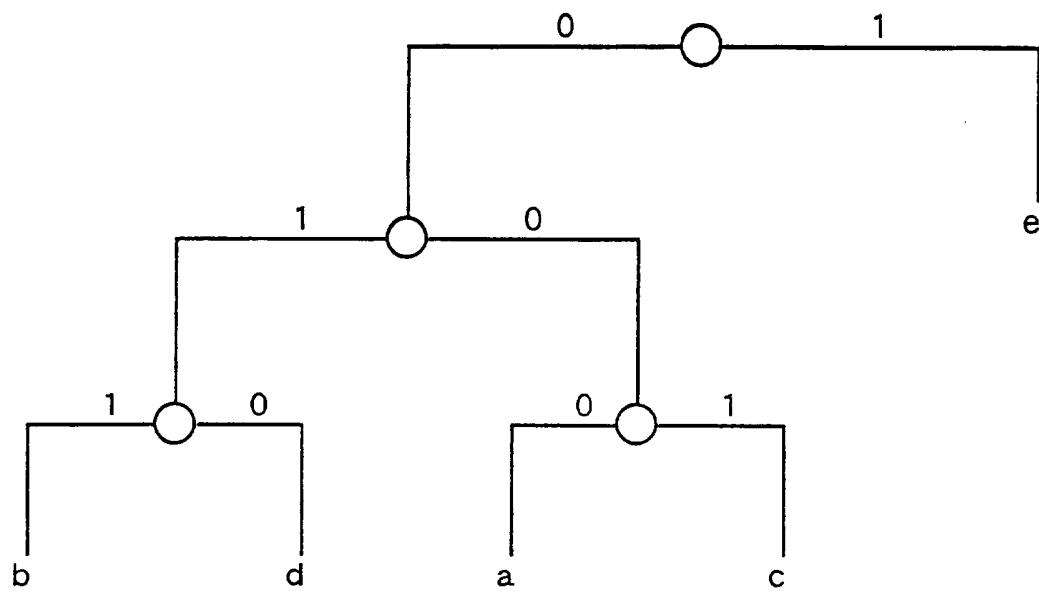
第1図



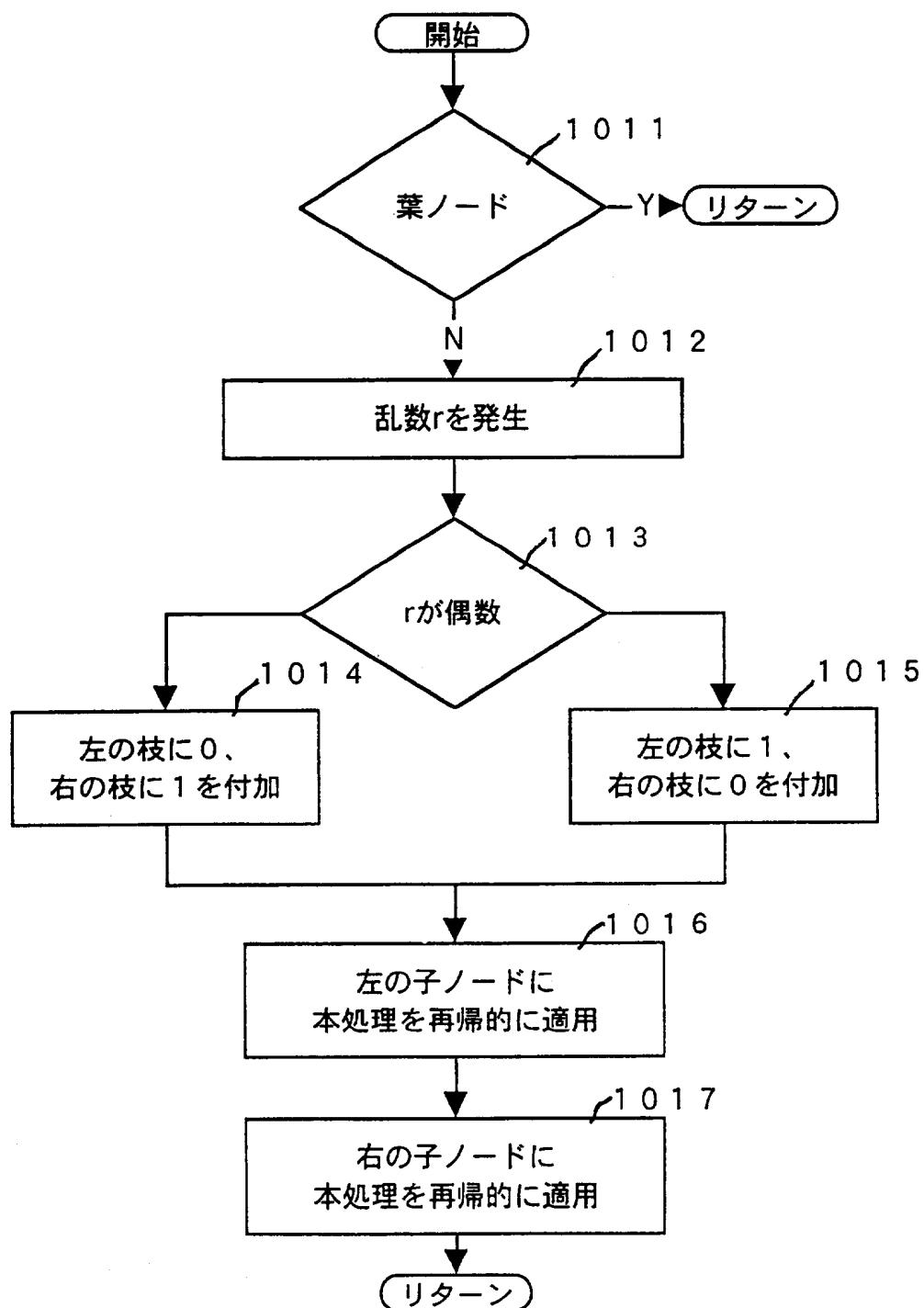
第2図



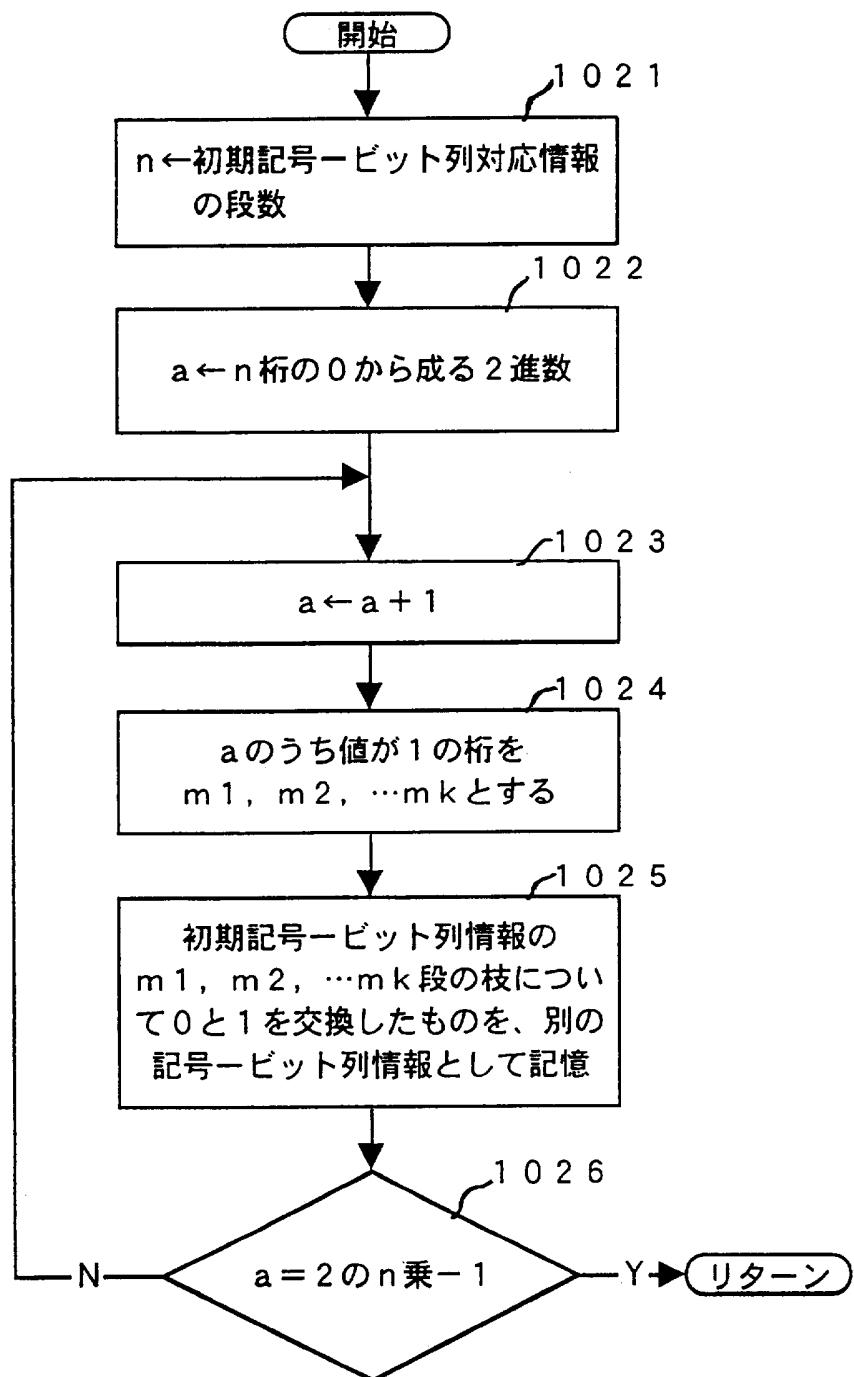
第3図



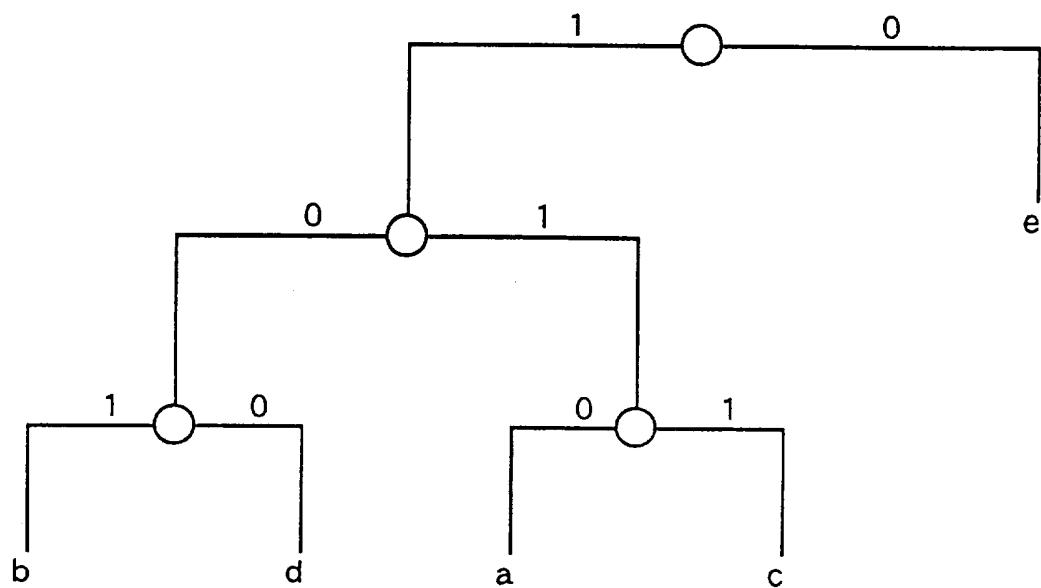
第4図



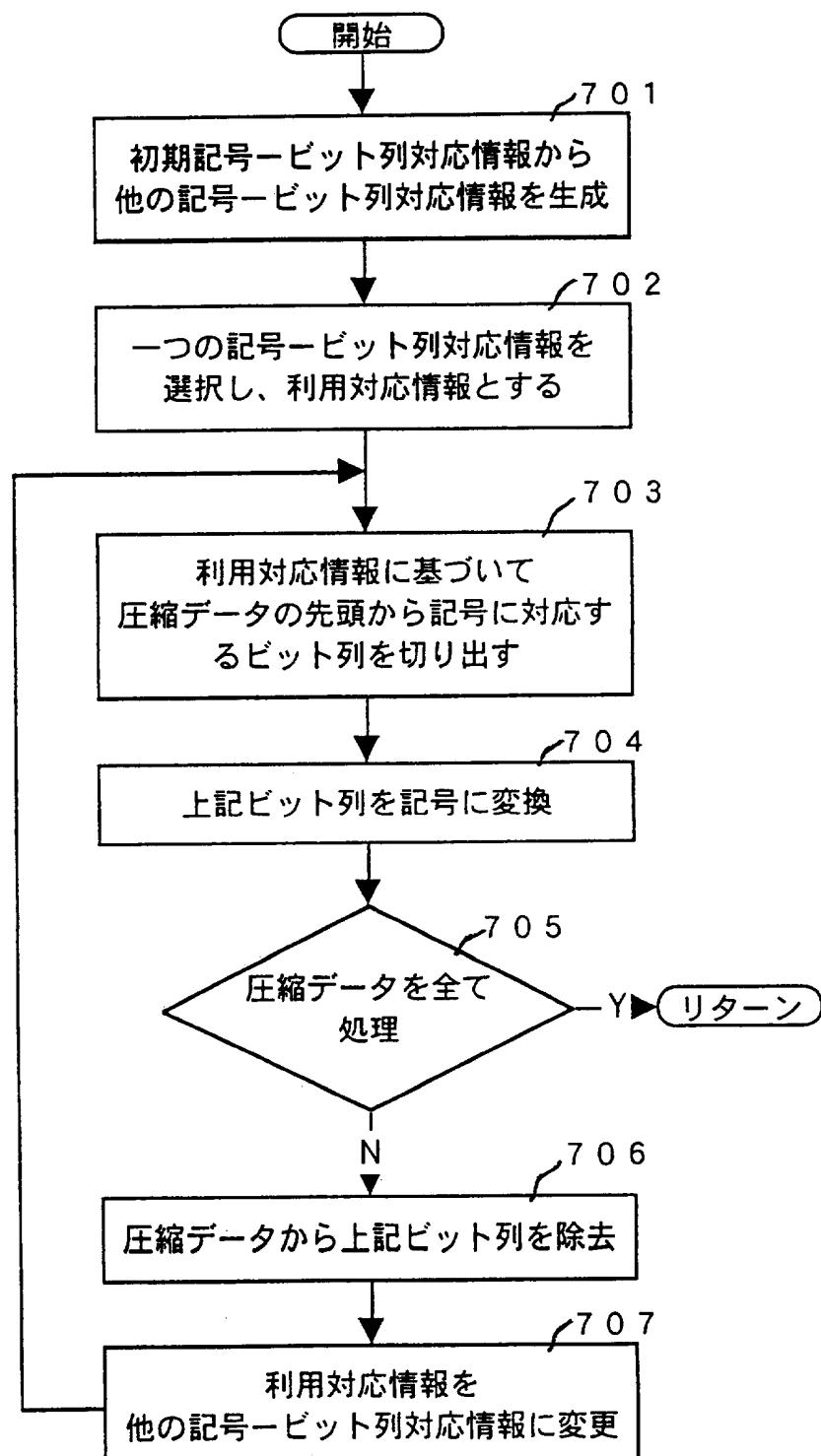
第5図



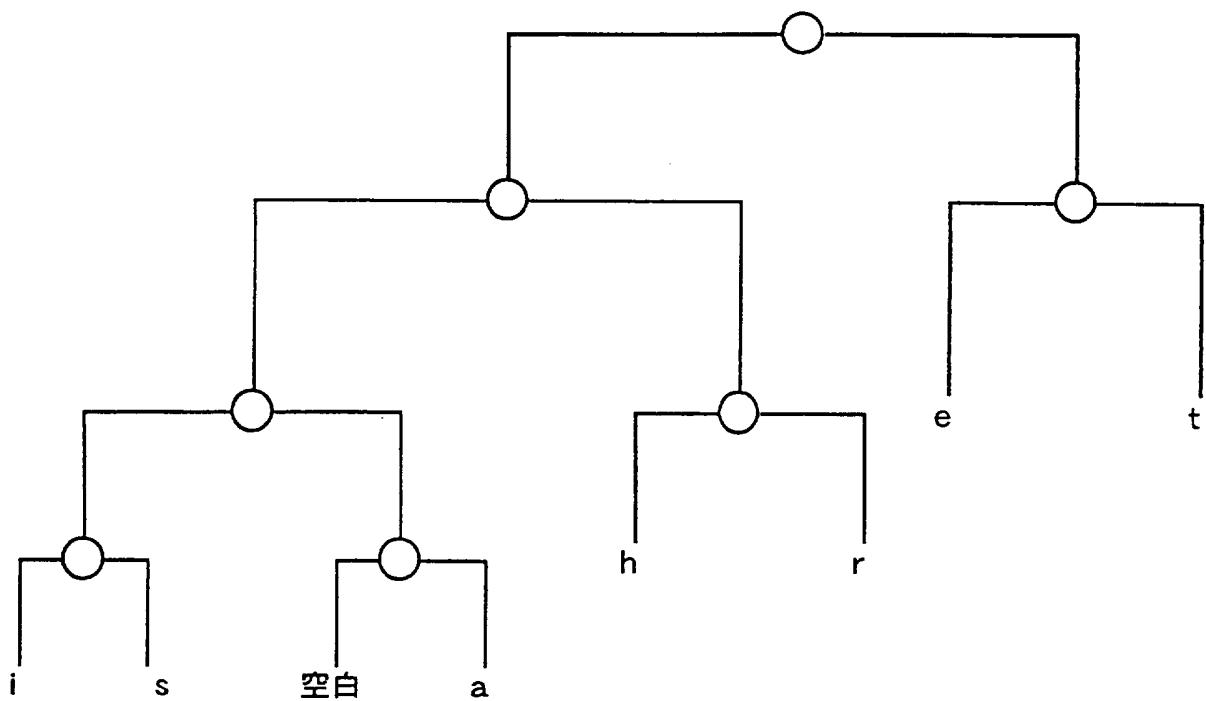
第6図



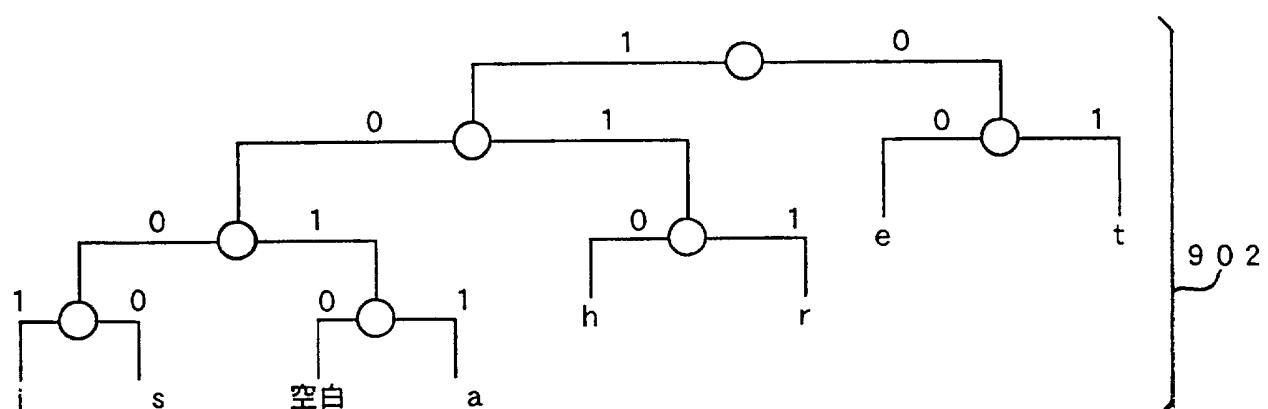
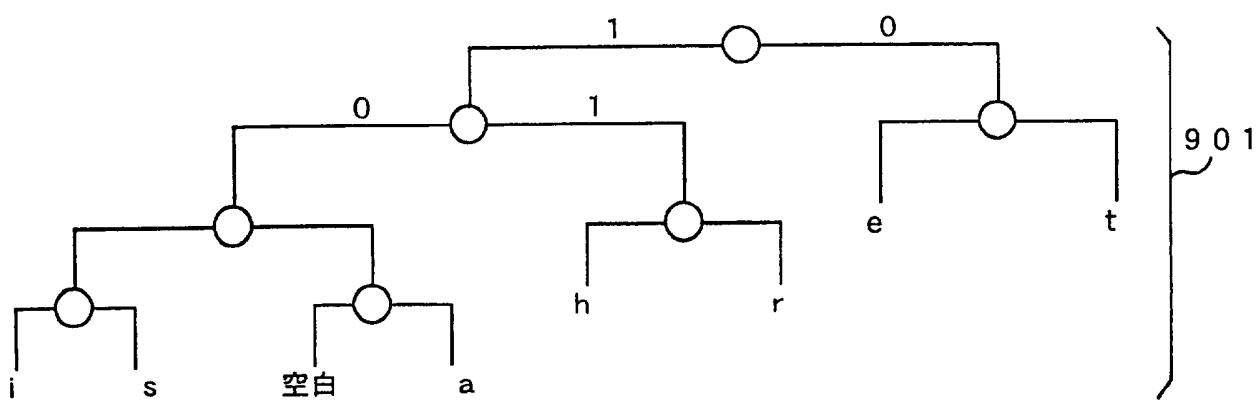
第7図



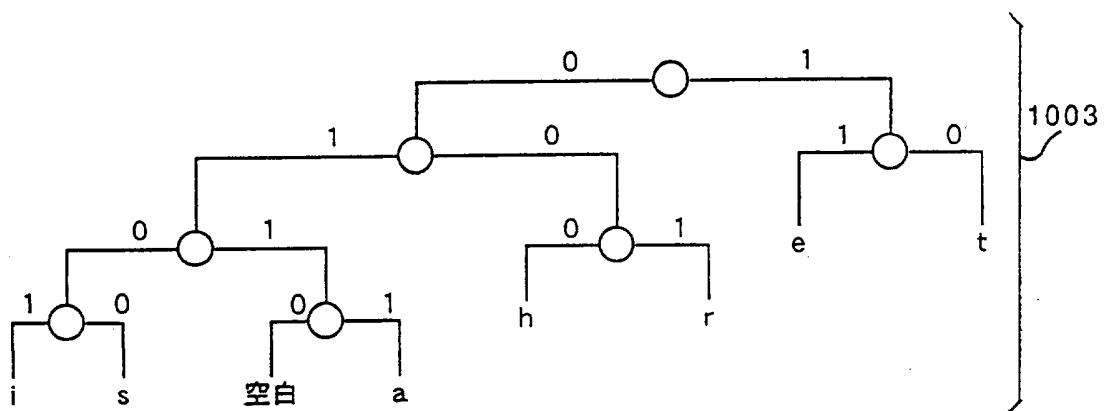
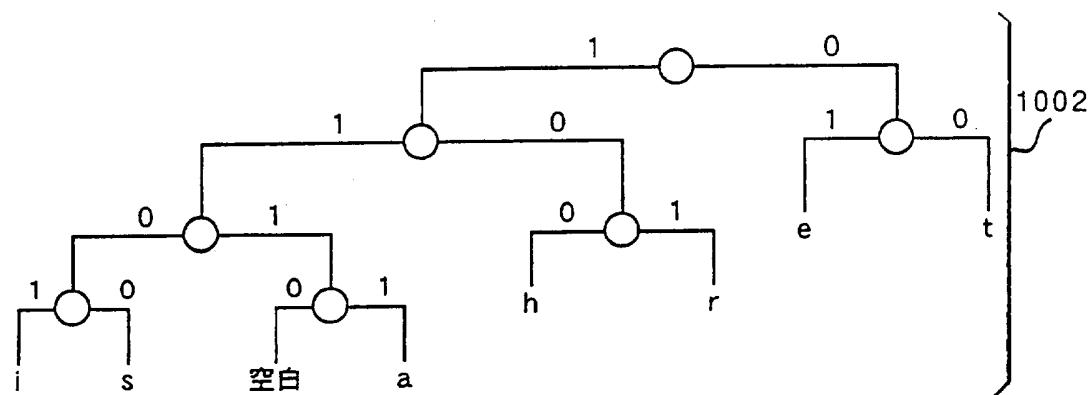
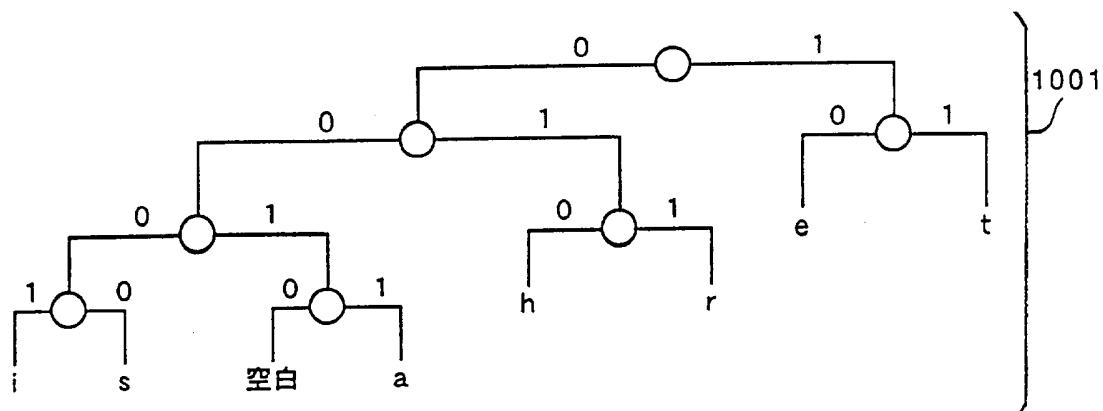
第8図



第9図



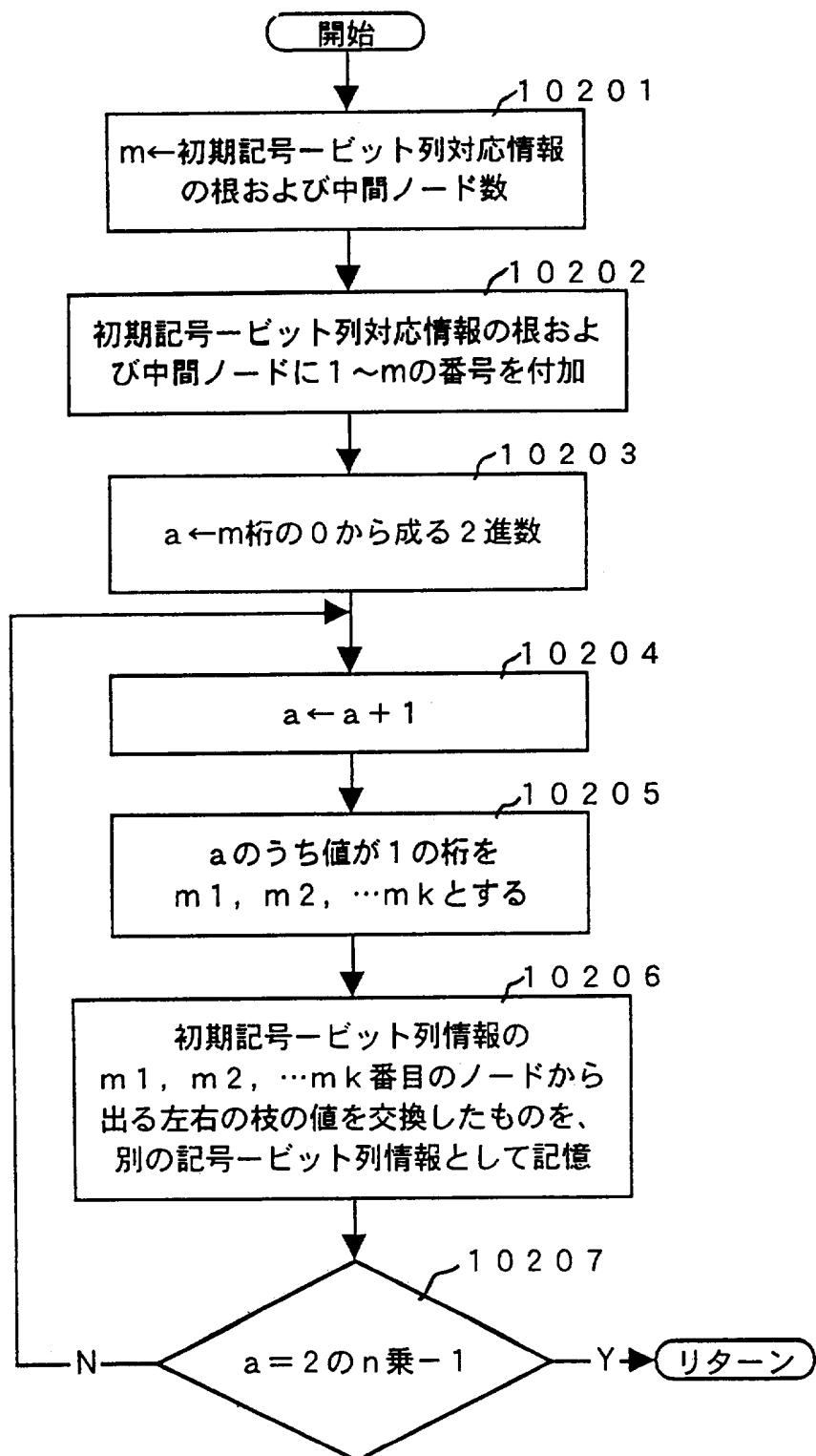
第10図



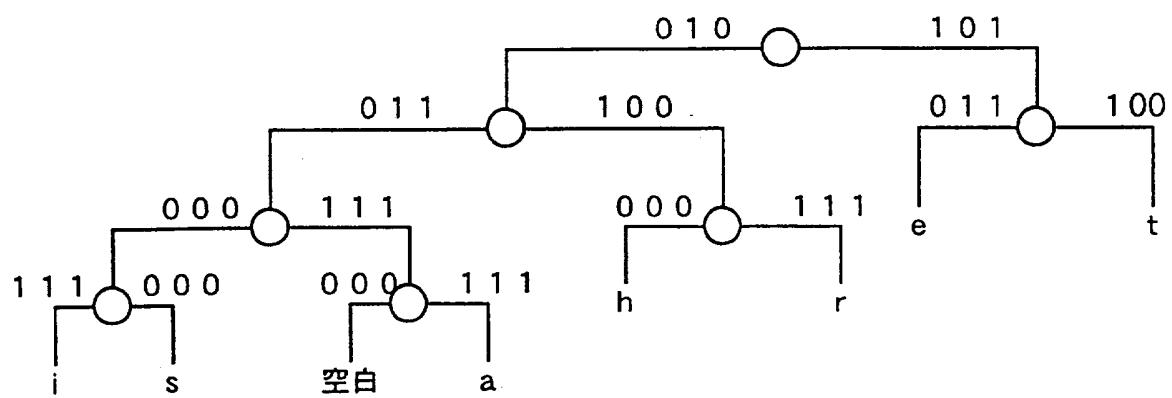
第11図

01010 . . .

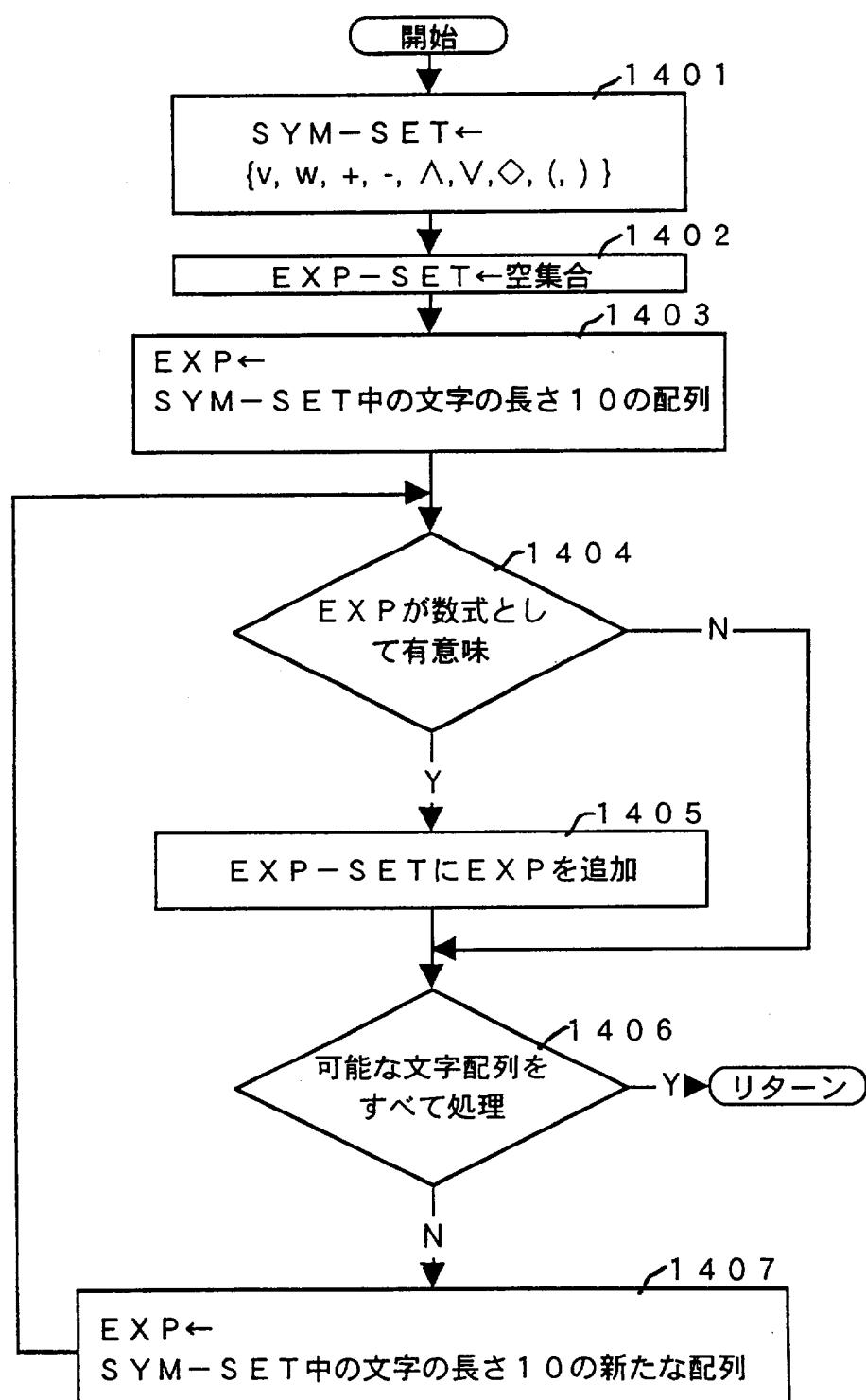
第12図



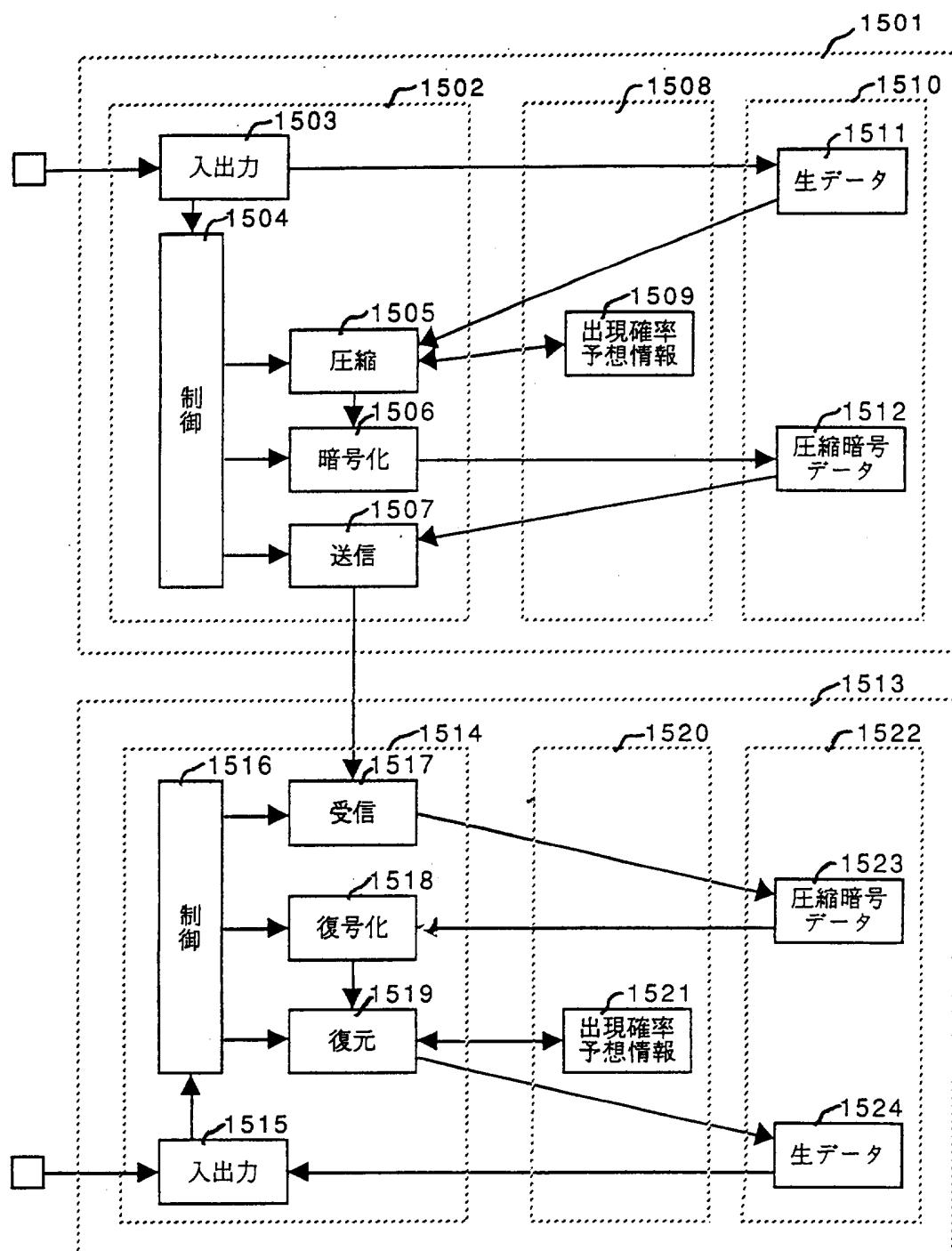
第13図



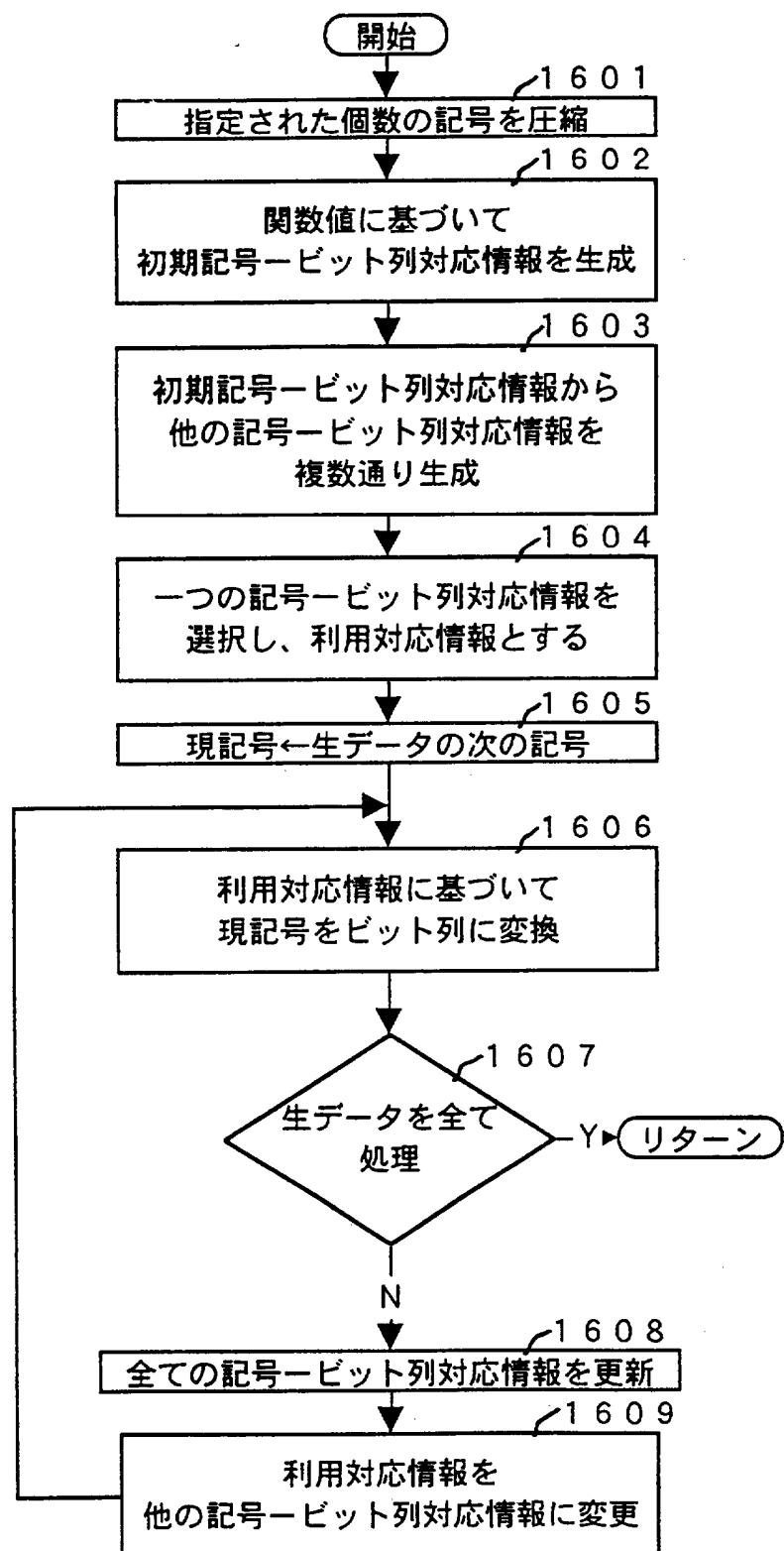
第14図



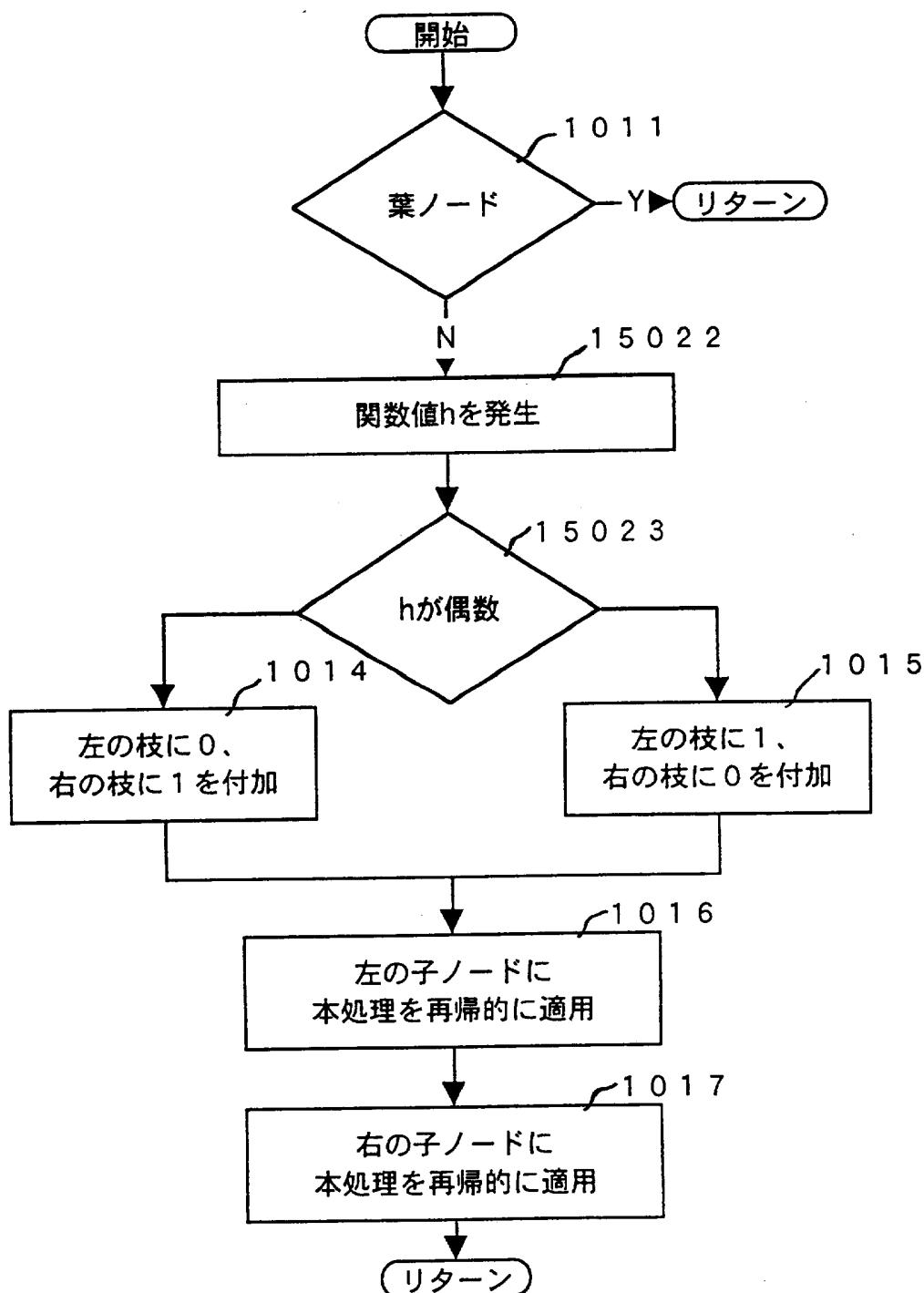
第15図



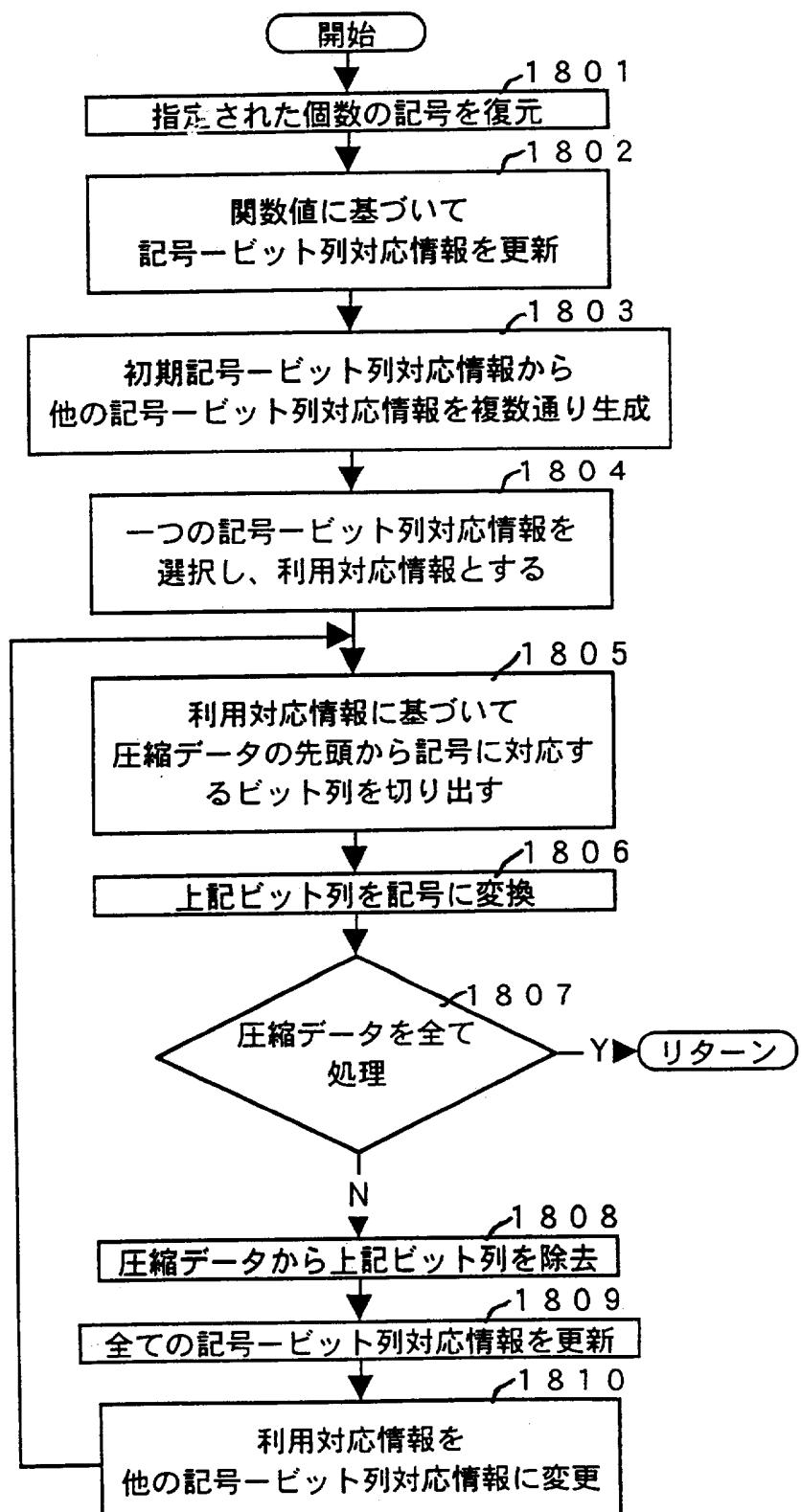
第16図



第17図



第18図



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP95/01815

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl⁶ H04L9/06, H04L9/16

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl⁶ H04L9/06, H04L9/16, G09C1/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1926 - 1995

Kokai Jitsuyo Shinan Koho 1971 - 1995

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP, 6-112840, A (Ricoh Co., Ltd.), April 22, 1994 (22. 04. 94),	1, 11, 3-4, 6, 8
A	Lines 42 to 46, right column, page 2, line 50,	2, 12-14
Y	right column, page 5 to line 18, left column, page 6 (Family: none)	5, 7, 9, 10
Y	JP, 5-333772, A (Toshiba Corp.), December 17, 1993 (17. 12. 93), Claim 6 (Family: none)	5, 7
X	JP, 4-296169, A (Canon Inc.), October 20, 1992 (20. 10. 92), Claims 1 to 5 (Family: none)	3, 4, 6, 7, 8
Y	JP, 4-37367, A (Fujitsu General Ltd.), February 7, 1992 (07. 02. 92), Claim (Family: none)	10
Y	JP, 60-164787, A (Oki Electric Industry Co., Ltd.), August 27, 1985 (27. 08. 85),	9

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

December 8, 1995 (08. 12. 95)

Date of mailing of the international search report

January 16, 1996 (16. 01. 96)

Name and mailing address of the ISA/

Japanese Patent Office

Facsimile No.

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP95/01815

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>Claim (Family: none)</p> <p>JP, 4-119386, A (NEC Corp.), April 20, 1992 (20. 04. 92), Claim (Family: none)</p>	9

国際調査報告

国際出願番号 PCT/JP 95/01815

A. 発明の属する分野の分類(国際特許分類(IPC))

Int. C2 H04L9/06, H04L9/16

B. 調査を行った分野

調査を行った最小限資料(国際特許分類(IPC))

Int. C2 H04L9/06, H04L9/16, G09C1/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国实用新案公報 1926-1995年

日本国公開実用新案公報 1971-1995年

国際調査で使用した電子データベース(データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP, 6-112840, A(株式会社 リコー), 22. 4月. 1994 (22. 04. 94),	1, 11, 3-4, 6, 8
A	第2頁右欄第42-46行, 第5頁右欄第50行-第6頁	2, 12-14
Y	左欄第18行(ファミリーなし)	5, 7, 10, 9
Y	JP, 5-333772, A(株式会社 東芝), 17. 12月. 1993 (17. 12. 93), クレーム6(ファミリーなし)	5, 7

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

- 「A」特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」先行文献ではあるが、国際出願日以後に公表されたもの
 「L」優先権主張に疑義を提起する文献又は他の文献の発行日
 若しくは他の特別な理由を確立するために引用する文献
 (理由を付す)
 「O」口頭による開示、使用、展示等に言及する文献
 「P」国際出願日前で、かつ優先権の主張の基礎となる出願の日
 の後に公表された文献

- 「T」国際出願日又は優先日後に公表された文献であって出願と
 矛盾するものではなく、発明の原理又は理論の理解のため
 に引用するもの
 「X」特に関連のある文献であって、当該文献のみで発明の新規
 性又は進歩性がないと考えられるもの
 「Y」特に関連のある文献であって、当該文献と他の1以上の文
 献との、当業者にとって自明である組合せによって進歩性
 がないと考えられるもの
 「&」同一パテントファミリー文献

国際調査を完了した日 08. 12. 95	国際調査報告の発送日 16.01.96
名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100 東京都千代田区霞が関三丁目4番3号	特許庁審査官(権限のある職員) 村上友幸 ⑤ 5 J 7 2 5 9 電話番号 03-3581-1101 内線 3535

C(続き) 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP, 4-296169, A(キヤノン株式会社), 20. 10月. 1992(20. 10. 92), クレーム1-5(ファミリーなし)	3, 4, 6, 7, 8
Y	JP, 4-37367, A(株式会社 富士通ゼネラル), 7. 2月. 1992(07. 02. 92), クレーム(ファミリーなし)	10
Y	JP, 60-164787, A(沖電気工業株式会社), 27. 8月. 1985(27. 08. 85), クレーム(ファミリーなし)	9
Y	JP, 4-119386, A(日本電気株式会社), 20. 4月. 1992(20. 04. 92), クレーム(ファミリーなし)	9