

氏 名	金 田 佳 郎
	かね だ よし お
学位の種類	工 学 博 士
学位記番号	論 工 博 第 943 号
学位授与の日付	昭 和 52 年 1 月 24 日
学位授与の要件	学 位 規 則 第 5 条 第 2 項 該 当
学位論文題目	漢字情報処理システムの研究

論文調査委員 (主 査) 教 授 矢 島 脩 三 教 授 田 中 哲 郎 教 授 坂 井 利 之

論 文 内 容 の 要 旨

本論文は、「漢字情報処理システムの研究」と題し、一つの実用化につながるシステムの研究をまとめたもので7章よりなっている。

第1章は序論であり、本研究の背景と目的について述べ、漢字を含む日本語の情報処理を行うシステムを構成する場合に生ずる問題点について述べている。

第2章では、当用漢字等の約2000の漢字に対する明朝体の字母の文字パターンについて、それらを30×30の黒または白の画素に分解して分析を試みている。ここでは、画素に分解された文字パターンについて、まず、それらを符号化せずそのまま記憶するとした場合と、黒または白の画素が何画素横方向または縦方向に連続するかに基づきランレングス符号による場合等について比較検討を行っている。さらに、文字全体としての黒の画素の座標位置ごとの出現頻度、文字当りの黒の画素の数の分布等を明らかにしている。

第3章では、この研究において開発された漢字情報処理システムのハードウェアの設計思想と構成について、とくに新しく開発された方式に重点をおいて論じている。本システムにおいては、文字パターンを、前章の分析結果より30を超える2のべき乗ということで $32 \times 32 = 1024$ の画素に分解し、さらに、これを白黒の市松模様の黒画素に対応するもののみに注目して、その他は省き、この1024の画素の半分すなわち1文字当り512ビットに情報を圧縮してフォントメモリを構成することにより記憶容量を半減する手法を示している。この圧縮された情報より元の文字パターンを忠実に復元することは一般にはできないが、実用にたえる程度ではほぼ元のものに近い文字パターンに復元する手順を示し、そのための復元回路をシフトレジスタ等を利用した論理回路で実現している。

さらに、この文字パターンの復元回路による復元がどの程度忠実に行われるかを調べるため復元率などを定義し定量的に検討を行っている。この論文の方式では、縦方向や横方向の要素の多い文字パターンは復元率が高く、斜方向の要素の多い文字パターンは復元率が劣ることを示しており、全体として70～80%

の復元率の文字の多いことを見出している。

これらに対し、本論文ではさらに、この方式で最終結果がよくなるように、元の文字パターンを予め修正してから上記の手法を適用するという方法を考案し、そのための文字パターン復元修正装置を CRT 表示装置などによって開発し、人間—機械系の形で利用し、これによってさらに10%程度良好な復元率を達成できたことを述べている。

つぎに、本漢字情報処理システムの電子式の高速出力装置における光学ファイバ形記録管で問題となっただけや歪の補正をツェナダイオードの非直線性を利用した電子回路で克服したことを述べている。

第4章は研究開発された漢字入力装置について述べてある。鍵盤の設計について述べたあと、鍵盤上にないいわゆる外字の入力方式を説明している。入力した漢字を CRT 表示していく入力装置においては、それらの装置が最高32台まで共用できる漢字パターンのフォントメモリの時分割利用の方式を開発している。

第5章はシステムのソフトウェアについて述べている。種々のものより成立っているが、一つは、まず、フォントメモリのための文字パターン生成のためのソフトウェアである。漢字データの編集のソフトウェアにおいては、従来、作表の場合の表の枠組など、目的ごとに所定の事前印刷フォームを用いデータのみ印刷していたのを改め、漢字等の文字のみならず同時にこのフォームをも印刷できるようにし、事前印刷の不要な方式を開発している。

さらに本論文では、浮動アドレス方式と名付けている方式によって使用頻度の高い文字はフォントメモリに常駐させ、そうでない文字はディスク記憶装置等に入れたままにして必要に応じてフォントメモリに移す、フォントメモリにおける一種の仮想記憶方式を実現し、フォントメモリの容量の制限を實際上なくし、しかも実フォントメモリの容量の節約をはかっている。

第6章は光学ファイバ形記録管を用いて酸化亜鉛記録紙に電子写真法によって漢字を高速に印刷する場合の問題点を材料物性の面から追究している。記録紙の特性と特性を支配する諸因子を調べ、記録紙の表面電位を測定する装置を試作し、材質、印加電圧、紙送り速度と表面電位との関係を明らかにし、またその減衰特性についても論じている。さらに、レーザーを用いた印刷文字の品質を評価する分析装置の試作を行ない各種の実験を行っている。

第7章は本論文の結論として研究成果の要約と将来に残された研究課題について述べている。

論文審査の結果の要旨

計算機技術の進歩によって漢字を含む邦文の情報処理を行なういわゆる漢字情報処理システムの研究開発が進み急速にそれらシステムが実用されるようになってきている。漢字はその文字の種類も極めて多く、字形も複雑であり、同音異字のものや同音異義のものも多い等、計算機による情報処理に当っては、欧米言語にはない特有の困難さを克服していく必要がある。

本論文は、漢字を含む日本語の情報処理を行うためのシステムにつき研究し、我国で比較的早い時期に種々の特徴をもったシステムの研究開発を行って、その研究結果をまとめたものであり、その成果を要約すると次の通りである。

1. 当用漢字など約2000の漢字を選び、それらの明朝体の字母の文字パターンについて、それらを $n \times n$ の白ないし黒の画素に分解して分析を試みている。 $n=20$ 程度では不十分で、 $n=30$ 程度にするとある程度良好で明朝体文字として実用に供しうるようになることを示し、 $n=32$ をシステムで採用している。

つぎに、文字パターンのフォントメモリへの格納に当って、これを符号化せずそのままの場合と、ランレングス符号化した場合の比較検討を行ない、このレベルでは必ずしもこの符号化が有利とは限らないことを示した。

2. $32 \times 32 = 1024$ ビットの文字パターンに対して、これを白黒の市松模様の黒の画素に対応するものに注目して、他は省き、1024の画素の半分すなわち一文字当り512ビットに情報を圧縮してフォントメモリに格納することによりフォントメモリの必要容量を半減した。これより元の文字パターンは近似的にしか再現できないが、そのための復元回路を示し、復元率を定義し、多くの文字は復元率70~80%であることを示した。

さらに、この方式で最終結果がよくなるように、元の文字パターンに予め修正をほどこしてからこの手法をほどこすことを考え、このための人間—機械系の装置を作り各文字に修正をほどこし、さらに復元率を10%程度高めうることを示した。

3. 論文において浮動アドレス方式と名付けたフォントメモリの構成法に関する方式を開発し、使用頻度の高い文字をフォントメモリに常駐させ、他は必要に応じてマスタファイルより取出す一種の仮想記憶方式をフォントメモリで実現し、フォントメモリの容量の制限を実際上無くし、実フォントメモリの容量の節約をはかっている。

4. 漢字の高速出力装置を構成するため、光学ファイバ形記録管(OFT)によって酸化亜鉛記録紙に出力する湿式直接法の電子写真方式を採用し、OFTを高速で駆動する際に問題となったダイナミックフォーカシングや歪の補正のための電子回路を新たに工夫した。さらに記録紙の紙質、チャージャ線の配置、印加電圧、表面電位、OFTの輝度等との関連を詳細に明らかにし、加えてこれらの特性を定量的に計測するための測定装置を開発した。

以上要するに本論文は、漢字を含む日本語の情報処理を行うシステムを実現するために研究を行い一つの実用化につながった方式を実現したものであり、学術上および工業上寄与するところが少なくない。

よって、本論文は工学博士の学位論文として価値あるものと認める。