

## 自動プレゼンテーションシステム(1)

1 Q-1

～プレゼンテーションシステムの概要～

長谷部 浩一 堤竹 秀行 鈴木 薫 堀 修  
山口 浩司 小野口 一則 金澤 博史 恒川 尚

(株)東芝 総合研究所

## 1. はじめに

我々は文書を理解しその内容をプレゼンテーションするシステム(APS)の試作を行った[1]。本システムは文書画像を入力し、その内容を視覚と聴覚に訴える形で聴衆に呈示するマルチメディア変換システムである。

我々が日常行うプレゼンテーションでは、話の要点や口頭では説明しにくい表やグラフなどを視覚に訴えるために、OHPシートやスライドを呈示しながら説明を行う。プレゼンテーションを行う前には、その準備段階で呈示資料と説明する言葉の十分な吟味が必要である。

APSでは、入力文書から文章部分の要点を抽出し、それを用いて箇条書き形式の呈示資料を作成するとともに、説明原稿を生成する。プレゼンテーションの時には、呈示資料を画面に表示しながら、その説明原稿をナレーションとしてスピーカより流す。聴衆の理解を助けるために必要に応じて入力文書中の図表を表示する。さらに、文章の内容や図表データの表示形式について、聴衆がシステムに対して音声によって質問/要求を出すことができる(図1)。

本稿では、APSのシステム構成について報告する。

## 2. システムの概要

## 2.1 文書画像の入力とコード化・ベクトル化

一般に、文書は文章・図表・写真等の様々な表現形態を持つ領域から構成される。当然、これらの領域を計算機で扱うためには、各々の表現形態に適した手段によってコード化する必要がある。そこで、まず入力された文書画像を解析し、文章・図・表・写真等の領域を抽出する。次に、抽出した各領域について、文字認識、図形認識、画像処理技術を利用し、コード化・ベクトル化が行われる。さらにプレゼンテーション時の呈示資料やナレーションの生成のための加工が可能なデータ構造に変換され、保存される。

## 2.2 プレゼンテーション資料の生成

文字コードに変換された文章は、機械翻訳で用いられている構文解析技術と知識処理による箇条書き変換技術により、呈示資料の中で用いられる箇条書きされた文字列へと変換が行われる。同時に、音声によるナレーション用の説明原稿も生成される。

表データとして入力されたデータは、プレゼンテーション時に表として再整形して呈示するほかに、もっと見易いグラフに変換して呈示することが可能である。そのため、表データについても文章と同様、文字列部分の解析を行っている。

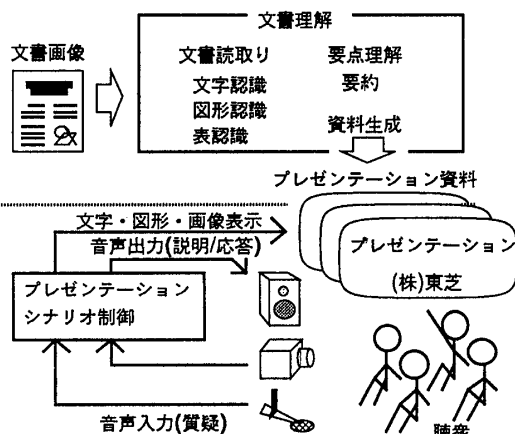


図1 システムの概念図

## 2.3 プレゼンテーションの進行

プレゼンテーションは、基本的には入力文書の内容に従い、聴衆に対して視覚的に資料を呈示しながら、同時にその説明文を音声として発声しながら進められる。このとき、ナレーションと同期させて、呈示資料を段落あるいはフレーズ単位にきめ細かく変化させることができる。表データは必要に応じて各種グラフに変換して呈示する。

## 2.4 会場モニタと質問応答

プレゼンテーションの進行において、会場の状況をモニタすることは欠かせない。聴衆の動作を知ることにより、その進行をダイナミックに制御できる。たとえば次のような処理が考えられる。

- (1) 聴衆がある程度揃ったところでプレゼンテーションを開始する
- (2) 聴衆の表情や動きを観察し、その状態によってプレゼンテーションの進め方を変える
- (3) 質問応答の際、質問者を特定したり多数決の人数カウントを行う

本システムでは、色と動きと位置を組み合わせたカラー画像の認識技術を利用し、着席している人数、および、挙手をしている人数をカウントする。また、音声の認識技術を利用して質問を受ける。

質問応答を実現するために、質問者の音声の認識、文脈の理解、解答の生成、音声の合成の4つの技術を用いている。一般には、質問の文脈を理解し、その答えを生成することは困難な技術である。しかし、対象とする入力文書の内容で、しかもプレゼンテーション時に呈示資料として出力された事項について繰り返し問う、あるいはより詳細なデータを要求するというように限定することで実験を行った。

## Automatic Presentation System (1)

Koichi Hasebe, Hideyuki Tsutsumitake, Kaoru Suzuki,  
Osamu Hori, Koji Yamaguchi, Kazunori Onoguchi,  
Hiroshi Kanazawa, and Shou Tsunekawa  
Toshiba Corporation, R & D Center

### 3. プロトタイプシステム

実験を行うためにプロトタイプシステムを作成した。

前章で述べた処理は、次のように分けられる。

(1)文書入力・呈示資料作成

(2)プレゼンテーション・質問応答

このうち、(1)は逐次的に行う処理、(2)は様々なことを同時に行う処理である。人が行うプレゼンテーションと比較すると、(1)は発表準備、(2)は実際の発表に当たる。このうち、(2)の処理を自動的に実行させるためにはリアルタイム性が要求される。

今回はプロトタイプということとを考慮し、開発環境にも重点をおくことにした。また、機能ごとに分散開発することができ、開発されたものを簡単に接続できる構成を目標の一つとした。さらに、システムのハードウェア構成も柔軟に変更できることとした。

そこで、これらを満たすシステムとして、UNIX システムであるエンジニアリングワークステーション AS4260 を利用することにした。

図2にプロトタイプシステムの構成を示す。まず、(1)は画像処理が中心の文書入力・コード化部と、文字列処理が中心の資料作成部の2つとした。一方、(2)のプレゼンテーション・質問応答プロセスは4つの基本機能(会場モニタ、画面処理、音声認識、音声合成)を持つ。各処理部の詳細は別稿[2][3]で述べる。

前述の通り、本システムでは開発環境とシステムの柔軟性を考慮した。そこで、オブジェクト指向の考え方を導入し、各機能をプロセスとして独立させることにした。各プロセスはお互いに干渉することなく、必要なデータはメッセージとしてプロセス間で交換する。このようにすることで、各機能プロセスを独立して開発できるとともに、プログラムを変更したときも簡単に入替えることができる。さらに、ネットワーク制御プロセスによりネットワークを介してもメッセージ交換ができるようにしたため、各

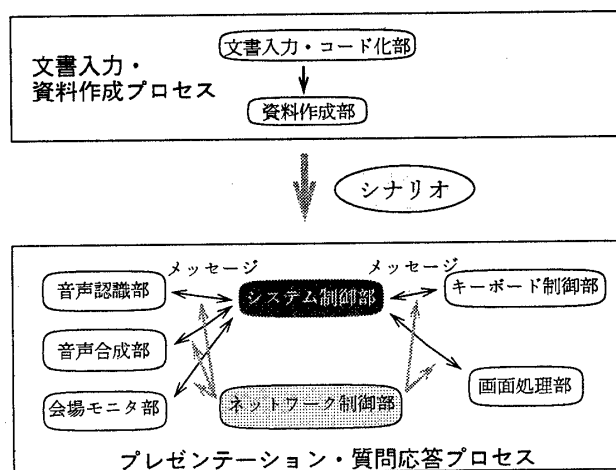


図2 プロトタイプシステムの構成

々のプロセスは別々の CPU で動作させることが可能である。しかも、特別なハードウェアを利用しないプロセスは、ネットワーク上のどの CPU で実行させてもよい。これにより、ハードウェア構成が変化しても、柔軟に対応できるようにになった。

図3は入力文書の例で、この文書に関して生成されたプレゼンテーション時の呈示画面の一部を図4に示す。

### 4. おわりに

実験の結果、ぎこちないが一通りのプレゼンテーションを行うことができた。難しいのは「間」のような時間軸の制御である。単にシナリオだけではなくシステムの遅れなどが問題となる。今後、各要素技術の開発の他に、自然なプレゼンテーションについての研究が必要である。

#### 参考文献

- [1] 恒川他: マルチメディア情報の理解と相互インタフェース, 情報処理学会「マルチメディア情報と分散協調」シンポジウム論文集, pp.67-76 (1989).
- [2] 堤竹, 鈴木, 堀, 恒川: 自動プレゼンテーションシステム(2), 本大会予稿 (1990).
- [3] 山口, 長谷部, 小野口, 金澤, 恒川: 自動プレゼンテーションシステム(3), 本大会予稿 (1990).

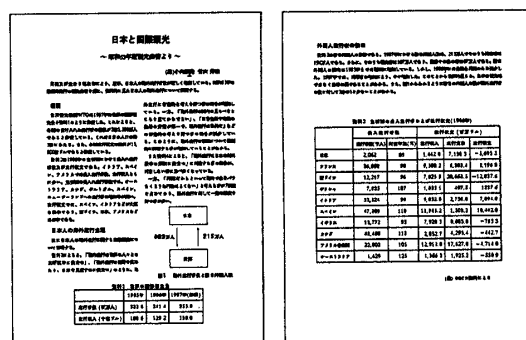


図3 入力文書

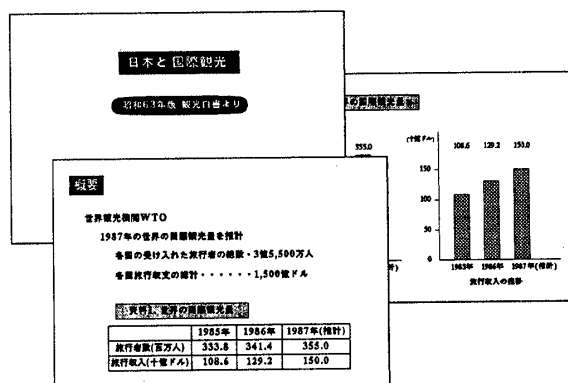


図4 呈示画面