

マルチモーダル情報記録再生システム

山口 浩司 大盛 善啓

(株)東芝 研究開発センター

情報・通信システム研究所

1L-10

1 はじめに

筆者らは、手書きや音声、画像などによるマルチモーダル情報記録再生システムの開発を行っている^[1]。手書きメモや音声などの非定型的情報の蓄積と、それらの効果的な整理・再利用を支援することを目的とする。本発表では、携帯型パソコン上で実現した試作システムについて報告する。

2 アプローチ

2.1 背景

日常生活において、さまざまな情報を入手し、またそれを整理し再利用するための道具として、パソコンやワープロ、電子手帳などが利用されてきた。これらにより、多くの情報が電子的に記録され活用されるようになったが、一方で従来の「紙と鉛筆」に代表される道具による記録も依然行われている。

近年、マルチメディアやペン入力機能により、従来のテキスト中心のシステムでは扱いにくかった情報を、容易にかつ効果的に表現できるようになった。例えば音や画像を含む文書の作成や手書きによる情報入力などである。しかし、このようなシステムでは、その機能ごとに新たな操作が必要であり、実際に会議の場で得た情報を即座に「紙と鉛筆」のような手軽さで記録することは容易ではない。さらに、記録された情報の中から所望の情報を探し出すためには、新たな操作が必要とされる。Whittakerらは、ペンと音声による情報の記録と再現が、従来の「紙と鉛筆」やテープレコーダによるものに比べ効果的であることを示した^[2]。

2.2 非定型的情報とリンク情報

これまで、いわゆる情報機器が扱ってきた情報の多くは、スケジュールや文書という、ある意味で定型化された完成した情報であった。しかし、あくまでもメモ書きとして記述されるような未完成の、

定型化に至らない情報も日常生活には多く存在する。ここでは、前者を定型的情報、後者を非定型的情報と呼ぶ。つまり非定型的情報とは、手書きでメモした、あるいは口で喋った情報などのように、曖昧さやニュアンスを含むような未整理の情報も含むものである。これらは、相互に関連する情報であったり、既存の情報への補足情報であったりするので、そのリンク情報も有用な情報であると言える。

これらのリンク情報は、ハイパーメディアシステムなどにおいては、個々の情報の収集後に、専用のツールによって行われている。また、アイデアの創出から文書作成まで支援する発想支援エディタなども利用されている。非定型的情報の記録を行う場合には、個々の情報間のリンク情報の収集が、その後の整理・検索に当たり重要である。

2.3 マルチモーダル情報記録

個々の非定型的情報ならびにリンク情報を効率よく記録するため、筆者らは、手書きや音声、画像による情報記録を行うと共に、それぞれの情報が有する時刻情報に着目した。つまり、目で見、耳で聞いた、あるいは手書きでメモしたという情報入手の順序関係やその時刻をリンク情報として記録する。図1に示すように、個々のデバイスから入力された情報を、従来のようにその種類毎に記録するのではなく、時刻情報により同期された情報として一括して記録する。

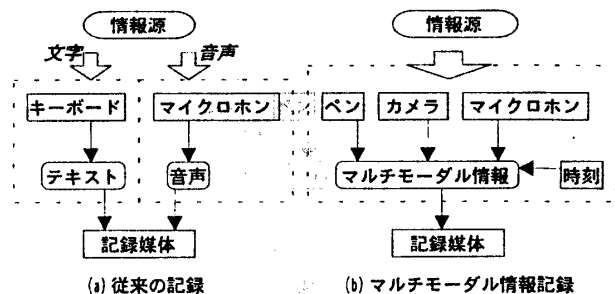


図1 マルチモーダル情報記録

3 実現

3.1 概要

前章での考察に基づき、非定型的情報の記録ならびに記録後の整理・検索等を支援するため、ペンと音声、画像の入力デバイスを備えたシステムの実現を図った。マイクロホンやカメラからの入力情報に加え、ペン入力による手書きストロークの情報を時刻情報とともに記録する。音声や画像は、利用者の選択により任意の時刻あるいは連続した時間での記録を行う。ペンによる入力では、文字や図形の認識処理などは行わず、ペンストロークの軌跡情報をそのまま記録する[図2]。

実現に当たり、利用場面として会議や講演会などを想定し、従来の情報機器ではうまく扱えなかった、口頭での会話や説明に基づくメモ(筆記メモ)や既存情報への付加情報(追記メモ、マーク)のような情報の記録が容易に行えることを特に考慮した。

3.2 記録

会議や講演の進行に合わせ、時刻を基準にして、ペン、マイクロホン、カメラからの情報を一定の入力単位ごとに同期記録する。記録は、画面上の"Record"ボタンを押下することにより開始される。それぞれのデバイスは、必要に応じて"Enable"または"Disable"状態にすることが可能である。手書きストロークについては、前述のように座標系列のみが記録される。カメラを利用した配布文書のイメージ入力を行って、ペンや音声による情報の追記書きができるようにしている。

3.3 再生

"Stop"ボタンによる記録終了後、例えばユーザが特定のペンストロークをペンで指示すると、そのストロークが有する時刻情報を基に、それとほぼ同時刻に記録された音声を探索してそれを再生する。記録された情報は、時刻順や情報の種類毎に選択的に

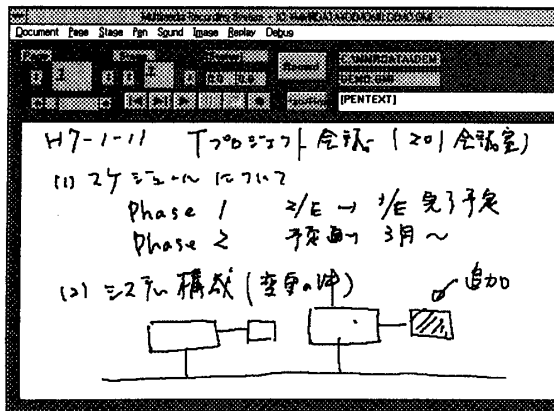


図2 画面例

表1 試作システムの概要

CPU	i486DX2-40MHz (主記憶:12MB)
基本ソフトウェア	Windows 3.1 ならびに Pen Extension 1.0
ペン入力	電磁誘導方式タブレット (STN カラーLCDと一体化)
音声・画像入力	サウンドカード(8bit,11kHz) 及びマイクロホン ビデオキャプチャカード 及びビデオカメラ

呈示することも可能である。

このように本システムでは、利用者が無意識に見たり、聞いたり、あるいは書いたりした情報でも時刻情報によりリンクされ記録されている。勿論、記録後にそのリンク情報の修正・追加を行うことで、採取された非定型情報を基により定型化された情報を構成してゆくことができる。

4 試作システム

本システムの有効性を評価するため、携帯型ペン入力パソコンを用いた試作を行った[表1]。ペンと音声ならびに静止画像による情報入力が可能である。Windows3.1による他のアプリケーションとのデータ交換も可能である。例えば、ワードプロセッサ等で作成された文書の印刷イメージ上に、補足情報をペンや音声で追記することができる。記録できる情報は、HDDの容量に依存するが、PCカード型HDD等の利用により2時間以上の連続記録も可能である。記録したファイルは、デスクトップパソコン上で検索・再生することができる。

現在、本試作システムは、主に会議のレコーダとして使用され機能評価中である。通常"Record"と"Stop"の指示だけで記録が可能であり、非常に実用的なインターフェースであるとの評価を得ている。

5 おわりに

本稿では、ペンと音声の入力機能を利用したマルチモーダル入力による情報記録再生システムについて述べた。今後は、収集したマルチモーダル情報の構造化と検索を支援する機能の開発を主に行ってゆく予定である。

参考文献

- [1] Imai, T., Yamaguchi, K. and Muranaga, T.: Hypermedia Conversation Recording to Preserve Informal Artifacts in Real-time Collaboration, In Proceedings of ACM Multimedia'94 (1994).
- [2] Whittaker, S., Hyland, P. and Wiley, M.: Filochat: Handwritten Notes Provide Access to Recorded Conversations, In Proceedings of ACM CHI'94 (1994).