

モバイル電子会議システム Dynamite における 4 F - 3 音声処理機能の実装

奈良岡将英

山田善大

太田賢

水野忠則

静岡大学

1 はじめに

携帯電話、PHS などのワイヤレス通信機器と、ノートパソコン、PDA などの携帯端末が普及したことにより、ユーザがいつでもどこでもコンピュータを利用できるモバイルコンピューティング環境が整いつつある [1]。しかし、モバイル環境においては帯域幅の制限や高い通信コスト、転送の途切れが生じる恐れなどから、安定した通信環境を必要とする従来の電子会議システムのようなマルチメディアコミュニケーションソフトウェアを利用するのは困難だった。そのため我々は、モバイル環境でもマルチメディア情報を利用できるモバイル電子会議システム Dynamite (DYNAMIC Adaptive Mobile Teleconference system) を開発している (図 1) [2]。

Dynamite サーバは、会議に途中参加したときに不在期間の会議情報を提供するための機能を提供している。モバイルユーザは、高い通信コストやバッテリーの問題などで常にネットワークに接続された状態にいることは難しい。こうした事情で会議を抜け、途中から復帰する場合、抜けた間の会議内容を把握してスムーズに会議に復帰できることが望ましい。しかし、会議を抜けていた間サーバに蓄積されていた発言を単純にそのまま参照していたのでは、その間にも会議は進んでしまうために、いつまでも会議に復帰できないことになる。

本稿では、抜けていた間の会議内容をすばやく理解し、会議へスムーズに復帰できるようにする方法として、会議音声を利用する方法を考察・実装する。ここで音声に着目するのは、マルチメディア会議においても会議の要は発言 (音声) のやりとりであり、会議の内容を把握するには音声を利用するのが

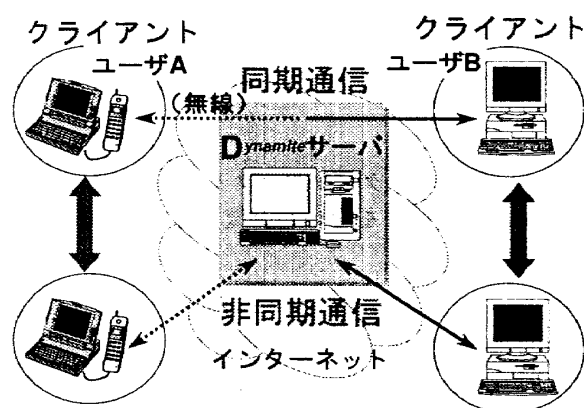


図 1: Dynamite システム構成

理にかなっていると考えからである。

2 音声基本機能の実装

基本機能として、音声をキャプチャ・再生する機能を実装する。

Dynamite のクライアント部は、Windows95 上に実装されている。このクライアント部において、発言時には音声をキャプチャし、順次パケットに分けて送信する。また、音声パケット到着時には、ある程度スプールを行った上で再生する (ストリーム再生)。データ形式は、Windows 標準の WAV フォーマットを用いる [3]。

3 音声処理機能の実装

音声処理機能として次のような機能を実装する (図 2)。

1. 音声再生時間短縮機能
2. 無音部を飛ばして再生する機能

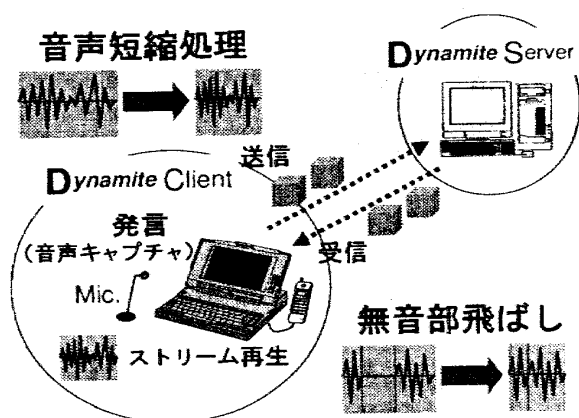


図 2: Dynamite の音声処理機能

音声再生時間短縮機能： 音声の再生時間を短縮する方法は、テープの早回しが一般的である。実際 Dynamite 上で扱われる音声は、映像と同時にキャプチャ（サンプリング）されるデジタルデータであるから、テープでいうところの早回しは再生時に再生周波数を上げることに相当する。これは比較的簡単に実現できるが、アナログテープと同様ピッチ（声の高さ）も上がってしまっていて聞き取りづらくなる。音声データを一定時間で区切り、部分部分を間引いて短縮する方法もあるが、会議音声の場合、ある一つの言葉が欠落ただけで会議の文脈がなくなってしまう可能性があるため、この方法は最適とは言えない。そこで、Dynamite の音声の短縮の方法としては、声の高さは変えずに時間だけを短縮する処理（タイムコンプレッション）[3]を採用する。この方法ではピッチを変化させないため、ある程度までの短縮であれば、聞き取りやすさを損なわずに時間の短縮が可能である。

処理はクライアントが行う。データ量を減らすという観点からすればサーバ側で処理を行った方がよいという考え方もあるが、サーバは同期通信・非同期通信両方をサポートするため、リアルタイムに参加者の発言を転送する処理や、発言の蓄積、蓄積された発言の転送処理などを行っていて [2] 負荷が高い。また実装も複雑になるため、今回はクライアントに実装する。

クライアント・アプリケーションにおいて短縮の比率は簡単に設定でき、また短縮するだけに限らず、速くて聞き取りづらいものについては逆に伸長することもできる。

無音部を飛ばして再生する機能： 会議音声に限らず通常の会話では、話の間や息継ぎ、考えている時間など無音部分（実際には多種多様なノイズを含む）が存在する。この無音部分を飛ばして再生することによって、ある程度時間を短縮することが可能である。

無音部は、サンプルの波形の振幅がない（あるいはある基準より小さい）部分と考えることができる。これらを再生せずに破棄することで無音部を取り除くことができる。

この機能についても、音声短縮機能と同様、クライアント・アプリケーションに実装を行う。

その他： 純粋に会議の内容を参照する時間の他に、そのデータを取り寄せる時間がかかる。音声データ自体を減らすことができれば、取り寄せ時間も短くなり、無線環境の狭い帯域幅にも対応できる。ITU 勧告 G.723.1 など、マルチメディアコミュニケーションのための低ビットレートの音声符号化方式の利用も考えられる [4]。

4 まとめ

本稿では、モバイル電子会議システム Dynamite において、途中から参加するユーザがそれまでの会議内容を素早く把握し、スムーズに会議に参加するための音声処理機能の実装について述べた。

現在、音声キャプチャ・送信機能／受信・再生機能の実装がほぼ完了し、音声短縮処理・無音部を飛ばす機能を実装中である。今後、プロトタイプ実装が終了次第、様々な実験を行ってこの音声処理機能の有効性を評価したい。

参考文献

- [1] 水野, 太田: モバイルコンピューティングの現状と将来, 電子情報通信学会誌 Vol. 80, No. 4, pp. 318-323 (1997.4).
- [2] 太田, 渡辺, 水野: モバイル電子会議におけるユーザ支援システムの設計, 情報ワークショップ論文集 Vol. 97, No. 2, pp. 215-220 (1997).
- [3] 戸田浩: サウンドエフェクト, C MAGAZINE Vol.8, No.12, pp. 22-50 (1996.12).
- [4] Richard V. Cox, Peter Kroon: "Low Bit-Rate Speech Coders for Multimedia Communication", IEEE Communications Magazine, pp. 34-41 (1996.12).