

ジェスチャを用いた実世界マルチメディアメモシステム

中井川 峻[†] 高橋伸[‡] 田中二郎[‡]

筑波大学第三学群情報学類[†] 筑波大学大学院システム情報工学研究科[‡]

1 はじめに

メモを取る, という行動は, アイディアメモや TODO メモ, 伝言メモなど, 記憶を外化して後の自分や他の人間に情報を伝えるために有用である. その性質上コンピュータとの親和性が高く, 方向に関連付けた音声メモシステム [1] や GPS を利用した場所に対しメモを行うシステム [2] など, 近年コンピュータによりメモ支援を行う研究, 製品が増加している.

しかし, 他人のためのメモとしてオフィスや研究室などで広く利用されている「伝言メモ」を支援する研究は少ない. したがって, 現在も紙とペンを利用したメモの利用やメールによる代用がなされているが, その場合比較的簡易な視覚情報しか用いることが出来ないため, 情報を残すユーザ, 読み取るユーザどちらも負担を増加させ, 情報の伝達を制限しかねない. さらに, 情報を残す対象が他人の机であるため, その場にあるペンや紙, コンピュータなどを利用する訳にはいかず, 思いついた時に伝言を残すことが難しい.

本論文では, オフィスにおける伝言メモの支援のために, 上記の問題を解決するための手法を提案する. また, それを実装した, 実世界に残した音声・映像メモをハンドジェスチャにより操作するシステムについて述べる.

2 提案手法

オフィスにおいて他人の机に伝言メモを残すことを考えた場合, その場の道具を利用出来ないため, 道具を使用することなく全ての操作を行えることが好ましい. これに対し我々は, メモをプロジェクトにより実世界上へ投影し, ハンドジェスチャを用いて操作を行う手法を提案

する. これにより, 道具を使用せずに操作することが可能なだけでなく, デジタル化された既存のメモでは利用できなかった, 実空間上の物体との近接性や机上の位置といった情報を利用したメモの記録が可能となる. また, ハンドジェスチャを用いることにより, 実世界上に投影されたメモをより直感的に操作することが可能であると考えられる.

さらに, 本研究ではメモとして残す情報の種類として, 音声や映像を用いることを提案する. 音声は筆記に比べ即時記録性に優れ, 考えたことをよりそのままに近い形で記録することが可能である. また, 映像による視覚情報を付加することにより, 身振り手振りなどを用いた素早く正確な情報伝達を実現する.

3 システム概要

3.1 メモの記録

紙のメモを書く際の動作に着目し, メモを行うジェスチャを考案した. まず, メモを押さえるように手を机に置くことにより, その手の位置にメモ用紙オブジェクトを表示する. そして, 反対の手でメモを書くように, 表示されたメモ用紙オブジェクトを指差すことにより, 音声と映像の記録を開始する.

音声は机に設置したマイクから, 映像は天井に設置したカメラから取得する. 天井からの映像を用いることにより, 物体や文字, 図などを指示した情報の記録, 伝達を可能とする.

もう一度机に手を置くことでメモの記録を停止し, 置いた手の周辺にメモの種類を選択するメニューを表示する. これを指差して選択することにより, 音声のみを記録するか, 音声と映像を共に記録するかを選ぶことができる. これにより, メモの即時記録性を保持し, また, メモの最中に映像を残したくなった時にも柔軟に対応することができる.

3.2 メモの閲覧

記録したメモは, 図 1 のように縮小されたアイコンとして投影される. このサイズはメモを記録した日時から時間が経つにつれて大きくなる

The real world multimedia memo system using hand gesture

[†]Shun Nakaigawa [‡]Shin Takahashi [‡]Jiro Tanaka

[†]College of Information Sciences, the Third Cluster of Colleges, University of Tsukuba

[‡]Graduate School of Systems and Information Engineering, University of Tsukuba

ことにより、伝言対象者へメモの存在を気付かせる。メモの再生は、メモアイコンを一定時間指差すことにより行う。メモの再生時には実際の大きさと同様に映像を拡大する。

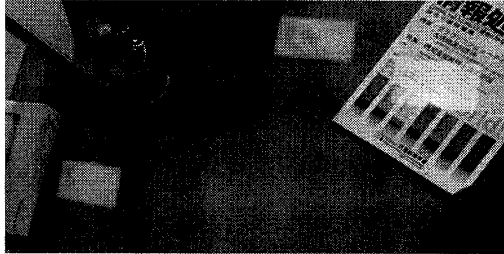


図 1 机の上に投影されたメモを示すアイコン

4 システムの実装

システムは C++ によって開発し、画像処理用のライブラリとして OpenCV を用いた。ハードウェアは机撮影用の Web カメラ、音声取得用のマイク、情報投影用のプロジェクタ、音声出力用のスピーカ、それらの入出力や画像処理、音声処理を行う計算機によって構成される。

4.1 手と指の認識

Web カメラから取得した RGB 画像を、修正彩度 [4] を用いた HSV に変換し、各チャンネルの閾値によって二値化することにより、肌色領域を検出する。肌色領域の大きさを求め、大きさがある 2 値の間に含まれる場合のみ、それを手領域として認識する。

次に、手領域を基にして指先の認識を行う。手領域を包含する最小の矩形を求め、図 2 に示すように、各辺 a, b, c, d から指の第一関節分だけ矩形の内側に入った線 a', b', c', d' を探索する。一本の指を伸ばした場合、指が存在する線 x' では指の太さ分の物体のみが検出される。この時、検出した物体を指であると認識し、指先の座標を辺 x と手領域との接点とする。また、座標が複数検出された場合、手領域の重心位置からより距離の離れたものを指先座標として採用する。これにより、テンプレートマッチングなどを用いた手法よりも、高速な検出を実現する。

4.2 ジェスチャの認識

システムは起動から終了まで、机に設置したマイクから音声を取得し続ける。音声の音量が閾値を超えた時に手領域が検出された場合、手領域の重心位置を基にメモ用紙オブジェクトを表示する。メモ用紙オブジェクトと、置いた手

とは別の手領域における指先座標の当たり判定を行い、その動作をメモの開始ジェスチャとして認識する。また、メモが行われている状態において、音量が閾値を超えた時に手領域が検出された場合、その動作を、メモを終了するジェスチャとして認識する。

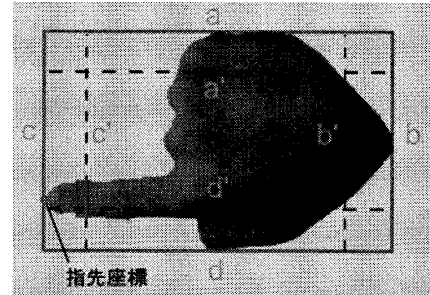


図 2 指の認識

5 まとめと今後の課題

オフィスにおける伝言メモにおける問題解決の手法を提案し、ハンドジェスチャを用いた実世界マルチメディアメモシステムとして実装を行った。これにより、道具を使用したり、身につけたりすることなく、素早い伝言メモの記録や閲覧を行うことが可能となる。また、音声や映像を利用した多くの情報を、短時間の内に記録することが可能となる。

今後は伝言メモのコミュニケーション性を高めるため、ネットワークを利用したメモの送受信や共有を行う機能を実装する。さらに、その際必要であると思われる個人の認証について検討していきたい。

参考文献

- [1] Tomoko Yonezawa, Hirotake Yamazoe, and Hiroko Terasawa, "Portable Recording/Browsing System of Voice Memos Allocated to User-relative Directions," Pervasive2009 Adjunct Proceedings, pp.241-244, 2009.
- [2] 白井旬, 小林敦信, 華井毅, 前島隆行, 辻正人, 奥出直人, コミュニケーションメモ "Putto" の開発, インタラクシオン 2004 論文集, pp.69-70, 2004.
- [3] 松橋聡, 藤本研司, 中村納, 南敏, 顔領域抽出に有効な修正 hsv 表色系の提案. テレビジョン学会誌, Vol.49, No.6, pp.787-797, 1995.