

# インタラクション状況のリアルタイムな判別と利用

伊藤禎宣<sup>\*1</sup>, 高橋昌史<sup>\*1 \*2</sup>, 中原淳<sup>\*1</sup>, 角康之<sup>\*1 \*2</sup>, 間瀬健二<sup>\*1 \*3</sup>

<sup>\*1</sup>(株)ATRメディア情報科学研究所 <sup>\*2</sup>京都大学 <sup>\*3</sup>名古屋大学

## 1. はじめに

近年、実世界環境に情報通信機器を遍在させ、いつでも、どこでも、ユーザに適切なサービスを行うことを目的とした研究が盛んに行われている。適切な情報サービスの実現には、システムがユーザや周辺の状況を遅滞無く判別し、状況に即したタイミングと方法で、適切な内容の情報提示を行うことが求められる。

しかし、判別すべき"状況"は、ユーザの存在位置などの単純な行為から、ユーザ間の社会的関係などの行為集合からなる抽象的状況まで、多様である。とりわけ、複数対象が関係する継続的社会的状況の判別には、時間的空間的に広範囲のセンサ入力装置から得た情報を集約する必要があり、即時的に状況を判別し、何らかのサービスを提供することは困難と言える。また、ユーザの実世界活動を阻害するような情報提供手法が煩雑な印象を与えることは、携帯の着信音などで周知であり、ユーザ負荷の少ない提示手法が求められている。

本稿では、我々がユビキタス環境実現の試行対象としている展示会場を焦点に、

- ・適切なタイミングで情報サービスを行うために必要な即時的に判別すべき状況の検討
- ・即時的状況判別に必要なセンサ入力及び処理プロセス
- ・即時的情報サービスに適切な提示手法

について検討し、展示会場閲覧過程のリアルタイム支援サービスを試行実験用アプリケーションとして実装したので報告する。

## 2. システムの検討

我々は、昨年度、展示会参加ユーザの状況を判別し、体験として記録する装置[角 2003]を提案している。同装置では、赤外線 ID トラッカなどを使ったセンサ入力から、体験の要約ビデオ生成を行っている。入力ノイズの排除などが必要であったため、要約のための状況判別は、閲覧体験終了後に統計的に分析するという方法であった。同様にセンサデータの分析的利用を主眼とした研究には、[Choudhury 2003]などがある。しかし、リアルタイムな状況依存型のサービスをユーザからの要求を待たずに提供するためには、体験の要約に重要な社会的状況 (e.g. グループディスカッション) を分析的に判別するのは別に、より即時的にユーザがそれらのサービスを受け入れられる状況か否かを判定する必要がある。

### 2.1. "状況"

閲覧過程での情報提供には、主に正対する対象物に付

加的な注視情報や個人化情報が想定される。特定対象物への情報付加では、即時的な表示がユーザ理解を助けるため、何が注視や傾聴の対象物か、といった状況の判別は、即時的に行う必要がある。また、例えば、白熱した議論の最中に閲覧対象の推薦などを行うことは避けるべき、というように、ユーザが情報提供を受け入れる状況にあるか否か、発話や傾聴といったユーザ個人の行為の判別は、即時的に行われる必要がある。そのため、これらの状況に関わるセンサ入力は、後処理なく即時利用可能な特性を持つ必要がある。

### 2.2. センサ装置の検討

前述のセンサ特性を満たすため、実際の利用を想定し、展示会場を模擬した環境と外部参加者による実験を行い、センサ入力の検討と調整を行った。

ユーザの位置や注視対象の判別に、本研究ではビジョンによる絶対測位(absolute positioning)センサである赤外線 ID トラッカ[角 2003]を使っている。しかし、ID 認識が移動や蛍光灯光源などの外乱に弱く、単体では運用上 3 割程度の認識エラーがあり、一定時間収集したセンサデータから、統計的にユーザ位置や注視対象の判定を行うことで対処していた。しかし、統計的判定ではリアルタイム性の確保が困難なので、ID のデコード手法にセンサの電気的特性を考慮したヒューリスティクスを導入する、エラー検出手法を改善する、といった方法を行った。模擬実験の結果、エラー率を 0.38%にまで下げることができ、ほぼリアルタイムにユーザ位置や注視対象の判別が可能な精度を得ることができた。

また、ユーザの発話状況のセンサ装置について、接話マイクによる検出を行っていたが、対話者の音声など外乱要因が多く、ユーザ本人の音声をリアルタイムに切り出すことは困難であった。そこで、声帯振動を咽喉部から直接ピエゾ素子で検出し、振動の強弱で発話の有無を検出する手法を採用した。検出する振動レベルの閾値は、展示会場を模擬した参加者及び発表者で構成される実験を行い、呼吸による振動を排除し、発声された相槌などは検出できる程度を目安にして経験的に決定した。同様に、咳払いなどを排除し、有意義な発話を検出する目安として、発話の最小継続時間を導入し、これを 500ms とした。これにより、高い S/N 比でほぼリアルタイムに発話を検出することができた。

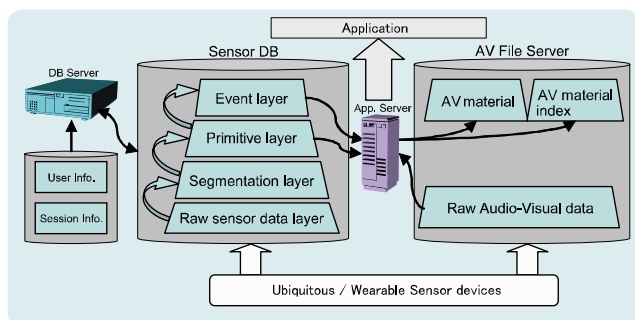


図1 システム構成図

### 2.3. 処理プロセスの検討

状況の判別に抽象度毎の階層的処理モデル[高橋2003]を導入した。発話や注視、傾聴といった各種インタラクションのプリミティブに相当し、抽象度が低く比較的高精度なセンサ入力が見られる状況は、その開始時点から即時的に判別、記録する。グループディスカッションのような抽象度の高い社会的状況は、インタラクションイベントと呼び、判別に必要な情報量が集まり次第、オンザフライで逐次処理する。このように、状況判別に必要なセンサ情報の量（解釈の抽象度）に応じた階層的データベースを作成し、アプリケーションは、必要な即時性や状況の抽象度に応じて、適切なDBへアクセスできるようにした（図1）。これにより、情報内容のリソースとなるユーザに関わる社会的状況の判別と、情報提供タイミングの決定に必要なユーザ状況の即時判別を統合的に扱うことができるようになった。

### 2.4. 提示手法の検討

脇見型の単眼式HMDを情報提示デバイスとして採用した。対話交流が主な展示会場で、音を利用した情報提示は適切ではなく、短時間で伝達可能な情報量も限定されてしまう。情報提示時に画面のブリンクなどでユーザの受動的注意を喚起し、サムネイル画像などを用いた一覧性の高い画面で、ユーザが情報の要不要を即座に判断できるようにした。各情報提示は、展示物注視、対面関係形成といったインタラクションのプリミティブに応じて即座に発行することができる。



図2 アプリケーション画面例

## 3. 試作アプリケーション

以上のシステム構成を元に、即時的な情報サービスのプロトタイプを構築した。（図2）

### 3.1. 興味類似度の表示

展示会場において来訪者が出会ったときお互いの興味類似度を知ることは、互いの体験や知識に関する対話を始める一つのきっかけになると考えられる。来訪者同士が対面関係プリミティブを検出後、蓄積された興味類似度情報を参照し、これを表示した。類似度は、各ブースの滞在時間をもとに、ブース数次元の興味ベクトルから、ベクトルの平均を引いたものに対して内積をとり、大きさを正規化したものである。

### 3.2. ブース人気度の表示

特定の展示物が過去にどれだけ人気があったかを知ることが、閲覧を決める上で一つの指標になると考えられる。ブース前への滞在というプリミティブを検出後、そのブースの訪問者数を10分毎の平均で1時間までさかのぼって表示した。

### 3.3. ガイド情報(推薦展示)の表示

展示会場において、ブースの推薦度合いが表示されることは、効率良く会場を閲覧するために有効な情報であると考えられる。この情報は、プリミティブの検出が無く、ユーザが何もしていないと考えられる待機状態に表示した。推薦情報はユーザ間興味類似度をもとに協調フィルタリングを行い生成した。

## 4. おわりに

展示会場を想定したインタラクション状況にもとづくリアルタイムな情報サービスについて検討し、プロトタイプを実装した。今後は、展示会場以外の場面で想定されるインタラクションへの適用、アプリケーションの検討などを行う予定である。

### 謝辞

本研究にご協力頂いたATR研究発表会2003参加者と関係者の皆様に感謝します。本研究は通信・放送機構の研究委託により実施した。

### 参考文献

- [角 2003] 角康之, 伊藤禎宣, 松口哲也, Sidney Fels, 内海章, 鈴木紀子, 中原淳, 岩澤昭一郎, 小暮潔, 間瀬健二, 萩田紀博, 複数センサ群による協調的なインタラクションの記録, インタラクション 2003, 情報処理学会, 2003.
- [高橋 2003] 高橋昌史, 伊藤禎宣, 角康之, 間瀬健二, 複数センサを利用したインタラクション・パターンの自動抽出, 情処研報, ユビキタス, 2003.
- [Choudhury 2002] Tanzeem Choudhury and Alex Pentland, The Sociometer: A Wearable Device for Understanding Human Networks, MIT Media Lab TR# 554. (2002)