

J-043

## 時空間画像データ構造を用いた情報保障ロボットビジョンシステム

Information Assurance Robot Vision System  
with Spatio-Temporal Image Data Structure

鍵山 泰尋† 塩瀬 隆之† 川上 浩司† 片井 修†  
Yasuhiro KAGIYAMA Takayuki SHIOSE Hiroshi KAWAKAMI Osamu KATAI

## 1 はじめに

これからの超高齢化社会において生き甲斐を創出する1つとして、生涯学習の環境整備が求められている。これに対して視覚障害を持つ参加者への情報保障などが、生涯学習のコンテンツと期待される市民講座や講演会が抱える課題となる。従来の情報保障の手段としては、事前に講演資料を点字変換したり、テキスト情報をデータで送付し読み上げソフトを利用する手法などが一般的に知られる。

しかし、たとえば、講演者が「このように...」「こちらが...」といった指示語を多用し、スクリーンに投影される情報の説明を十分に行わない場合や、スライドや図表が無音で切り変わる場合など、従来の保障手段では適切な保障が不可能な状況が存在している。

情報保障として重要なことは、発表者の指示行為などにより遷移する講演中の焦点を傍らで「代読」してくれる存在である [1]。この「代読者」と呼ばれる方法は、視覚障害者が講演内容に興味を持った箇所を的確に説明し、また、事前に点字や触図などの資料が配られる場合には、現在話題になっている事柄に対応した点字や触図上の箇所へと視覚障害者の手先を補助することも可能となる。以上のように、視覚障害者の注意を読んだ上での情報保障の可能な「代読者」は市民講座や講演会という場面において有効である。だが、「代読者」制度はまだ十分は整備がなされておらず、また一定のトレーニングが必要となるため「代読者」の数も現状としてはまだ少ない。

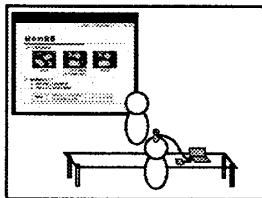


図1 Information Assurance Robot by Reading for a Person

よって我々はこれまで、この「代読者」の立場をロボットによって実現することを目的とし、その視覚情報処理部分にあたる情報保障ロボットビジョンシステムの開発を行ってきた。

本稿では、開発してきたロボットビジョンシステムの概要と、触図上での視覚障害者の探索行為を時間軸・空間軸両方向にて分析するために導入した、時空間データ構造に基づく情報提示手法について説明する。

視力や体力の衰えが学習に及ぼす影響に関して、たとえば Tatumi ら [2] による実践報告などは多数存在するが、こ

れらの影響を推し量る具体的な研究というのは少なく、上に示した問題は視覚障害者に限ったものではない。この点においても、本研究が提案するようなやり取りを重視した情報保障システムの開発は今後ますます重要となると考えられる。

## 2 情報保障ロボットビジョンシステム

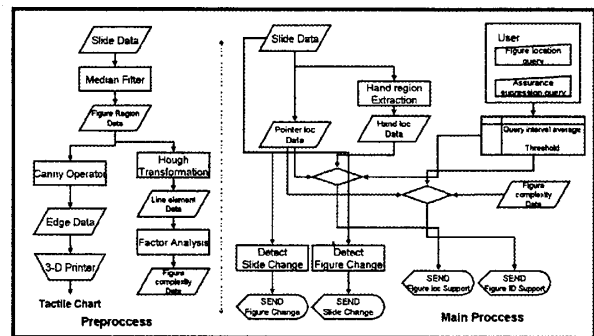


図2 The Overview of Information Assurance Robot Vision System

本研究では視覚障害者の生涯学習支援研究の一環として、情報保障ロボットビジョンシステムを開発した [3]。システムの概略は図2のようになる。またビジョンシステムとユーザとのやり取りの媒介として、本研究では視覚障害者の情報保障手段として一般的な触図を利用している。

本システムの処理は、講演前と講演時の2段階に大別される。まず講演前の段階において、プレゼンテーション中の図表領域の抽出、触図に用いる画像データ（以下「触図画像」）の作成を行う。そして、講演時には触図上でのユーザの探索行為などを基にした音声による情報の提示、その際の提示タイミングの調整を行う。

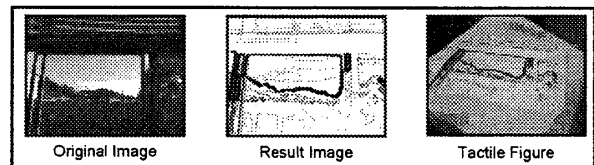


図3 Tactile Figure with Our Proposal Technique

## 2.1 図表領域の抽出

シーン中の図表領域を抽出する手法としては、たとえば Ichino らの方法などが挙げられるが [4]、これは書類内の図表の抽出に特化した前提条件が存在し、本研究で目的とする

† 京都大学, Kyoto University

ようなプレゼンテーション内の図表抽出には適切ではない。

よって本システムでは新たに、高速かつ前提条件の少ない手法として、カーネルサイズを大きく設定した上でメディアンフィルタを施すことによって図表領域を抽出する方法を提案し用いた。上記のフィルタを施すことによって、スライド中の文字領域は外れ値として除去される。そして、図表領域が存在しないとされる最初のスライドなどからの差分を計算し、図表領域を抽出する。

## 2.2 触図画像の作成

触図とは、図3右に示すような、印刷された箇所に熱を加えることでその部分が膨らむ特殊なカプセルペーパーを用いて作成されるもので、印刷面が紙面上の凹凸で知覚可能となる。本システムでは、印刷する画像の作成段階の処理を行う。

抽出された図表領域の画像はコントラストが小さく触図化に不向きであるため、エッジ検出の手法では一般的なCannyオペレータ[5]を施し、その後立体プリンタにて触図化を行う。

なお、Cannyオペレータの性質上、エッジの強弱の程度を分けて抽出することが可能である。よって、弱いエッジの連結度・検出するエッジの強度ともに段階的に増加させつつ繰り返し、強いエッジに対してはその程度に応じて線を太くする処理を加えた。

結果として、たとえば図3中央に示す画像が得られる。

## 2.3 スライド・図表変化タイミングの検出

スライド変化に関しては、取得したフレーム間で画像が大域的に変化した時点を変化タイミングとして検出するものとした。まず連続したフレーム間で差分画像を作成し、さらに差分画像を縦横に均等に分割した小領域を設定する。そして、各小領域ごとに非零領域の探索を行い、これが発見された小領域数の全体に対する割合で判断を行った。

一方、図表の変化検出に関しては、2.1にて説明した図表領域抽出と同じ手法を用い、変化タイミングを検出した。

## 2.4 音声による情報提示

これまでの処理を基に、音声による情報の提示を行う。提示する内容は、

- 触図上の手先位置から見た、講演者の示すポインタの方向・距離の提示
- 話題となっている図表・スライド情報の提示

上の2つである。また、後者においては図表上に講演者がポインタを向けている場合は、その位置も提示している。提示するタイミングは、前者は過去数フレームに渡り

- 手先が動いている
- 手先とポインタの示す先が一定値以上離れている
- この提示が行われなかった

以上を満たす場合、即座に提示している。これに対して後者は

- スライド・図表の変化が見られた場合、即座に提示
- ユーザからの要求があった場合、即座に提示
- 間隔を調整した上、定期的に提示

という3条件の下、「尋ねられたら答える」プル型の対応と、「聞かれなくても定期的に伝える」プッシュ型の対応を平行して実行している。

また、このプッシュ型の対応においては、ユーザの提示要求の頻度が高い場合には提示間隔を密にするような処理

をシグモイド関数を基に実装し[3]、ユーザの探索行為を阻害しないような情報の提示が可能となっている。

## 3 時空間画像データ構造に基づく触知行為の分析

本システムにて、ユーザの理解の深度や興味領域の推定によりユーザの意図を反映したより適応的な保障を行うため、触図上のユーザの探索行為を時空間軸両方向で任意に範囲指定が可能な時空間画像データ構造を提案した[6]。

### 3.1 提案する時空間画像データ構造

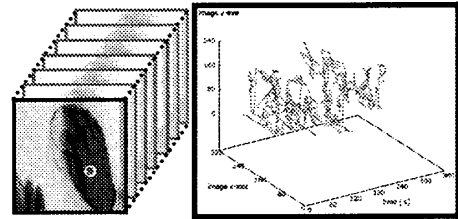


図4 Expression of Hand Orbit by Spatio-Temporal Image Data Structure

提案するデータ構造は、取得してきた手先軌道の遷移を各時点での手先位置の重心と次時点へのフローとの対の集合で構成される。これを用いて触知において重要となる特定の箇所での往復運動の認識や、k-means法[7]を用いた単位時間に変異距離のクラスタリングなどを行い、ユーザの触図に対するアプローチの仕方を手先軌道から推定することが可能となる。

### 3.2 探索行為に対する履歴情報の提示

時空間画像データ構造を用いて、一連の探索行為のなかで、序盤から終盤にかけてどのような箇所をどのような場合に探索するかなど、これまでのユーザが触図に対して採ってきた探索戦略の提示を行うといった、提示する内容に関して適応的な提示処理を実装した[8]。

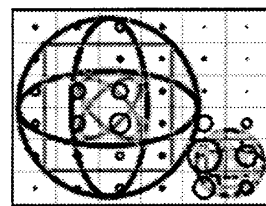


図5 Complexity by Numbers of Line Segments

処理としては、触図画像とカメラから随時得られる手先領域の動画像データを入力とし、まずはじめに触図画像に対して、主要な直線検出アルゴリズムの1つであるHough変換[9]を施す。これにより検出された線分量を用いて、分割された領域ごとの複雑さを定義し、その分布を図5のように求める。次に、カメラから得られた手先位置と過去の履歴から得られる各時点で探索を行った領域の情報、そして現在の触図画像を用いて音声による情報の提示を行う。そして最後に、提示後の触知行為の動向から、各時点での興味領域情報の更新を行う流れとなる。



図6 Set of Test Images (UPPER : Source Images, UNDER : Tactile Images)

### 3.3 実験

以上の提示処理を実際に被験者に対して行った際の結果が、次の図7となる。なお、テスト画像としては図6に示すような、複雑さについて違いを持つ5枚を用意した。これらについて各5分ずつ、被験者として20代の健常者の男性1人に目隠しをしてもらった上で探索してもらい、実験に協力していただいた。

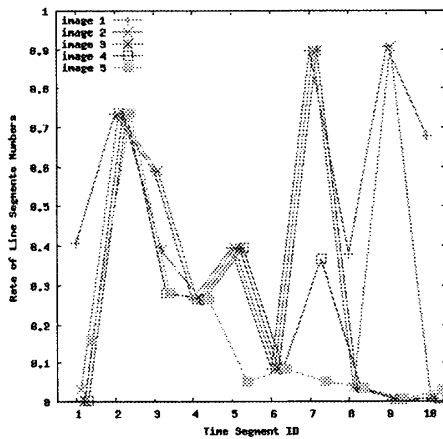


図7 Transition of the Presumed Region of Interest

図7は、分割された各時間範囲ごとに推定された興味領域情報として、対応する時間帯においてこれまで探索し、またその時点において探索するであろう領域を持つ線分量を縦軸にプロットしたものである。試行ごとに繰り返される中で、はじめに全体の単純な構造を持つ領域を探索した後、一度複雑な箇所へと注目して、そこを起点に再度大まかなところへと探索範囲を広める、というようなユーザの持つ戦略が、試行ごとに調整を加えながら推定されている様子が見受けられる。

## 4 結言

本研究では視覚障害者の生涯学習支援研究の一環として、「代読者」の存在を受けた新たな情報保障ロボットビジョンシステムを提案した。また、その枠組の提案に加え、触図上のユーザの探索行為を時空間軸両方向にて任意に範囲指定が可能な時空間画像データ構造を用い、これまでのユーザの触図に対して採ってきた探索戦略の提示を行うといった、提示内容に関して適応的な提示処理を実装した。

今後の課題としては、3節にて提案した提示処理などの情報保障ロボットビジョンシステムへの実装や、シーン中の3

次元構造が記述可能な触図の表現手法の開発など、本システムとユーザとのインタフェースとして重要な位置を占める触図作成の際の処理の高機能化が挙げられる。

注意を推定する仕組みを取り入れたインタラクションを扱っている従来研究としては視覚的なものが主であり、たとえば Nakano ら [10], Toda ら [11] の研究がある。これらは、相手の顔の向きや視線の動きの偏在度によってシステム内の処理に調整を加えるものである。だが、両者とも視線の推定のみで終始し、本来注視方向の先に必ずしも存在するとは限らない、真の注意対象について扱っているものではない。

この問題に対して本研究では、注視方向による両者のすり合わせが上手く得られない視覚障害者とのやり取りを通し、本システムにて重要な位置を占めた触覚など、他の手がかりとなりうるものを加えた、マルチモーダルに新たな注意対象を推定するモデルの提案が可能となる。

## 参考文献

- [1] ワンブックワンライフ. 日本ライトハウス盲人情報文化センター, 2007.
- [2] 辰己佳寿子. 少子・高齢化社会と生涯学習に関する研究(1)-山口大学の公開講座等に関するアンケート調査から-. 大学教育, Vol. 1, pp. 149-160, 2004.
- [3] 鍵山泰尋, 塩瀬隆之, 川上浩司, 片井修. 視覚障害者の生涯学習支援のための情報保障ロボットビジョンシステム. 計測自動制御学会 システム・情報部門学術講演会 2007 講演論文集, pp. 351-356, Nov. 2007.
- [4] 市野順子, 箕牧数成, 山口和泰, 垣智, 東郁雄, 古田重信. 図表検索のための図表情報自動抽出の試み. 情報処理学会研究報告. 情報学基礎研究会報告, Vol. 2002, No. 28, pp. 143-150, 20020315.
- [5] J. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 8(6), pp. 679-698, 1986.
- [6] Y.Kagiyama, T.Shiose, H.Kawakami, and O.Katai. Information assurance robot vision system based on the uneven distribution level in the hand orbit. *ICKS2008*, 2008.
- [7] J.McQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281-297, 1967.
- [8] 鍵山泰尋, 塩瀬隆之, 川上浩司, 片井修. 時空間画像データ構造に基づく触知行為保障システム. 第35回知能システムシンポジウム資料, pp. 97-102, 3月17-18日 2008. 東京工業大学.
- [9] J.Matas, C.Galambos, and J.Kittler. Progressive probabilistic hough transform. *British Machine Vision Conference*, 1998.
- [10] 中野有紀子, 岡兼司, 佐藤洋一, 西田豊明. ユーザの視線に気づく会話エージェント-アテンションの知覚と制御を利用した会話の円滑化-. *JSAI*, Vol. 3B2-08, pp. 1-4, 2005.
- [11] 戸田真人, 鷲見和彦, 松山隆司. 装着型能動視覚センサを用いた注目対象映像の獲得と理解. *MIRU2005*, Vol. IS4-150, pp. 1-8, Aug 2005.