

## 人間の視覚空間探索モデルに基づく視覚障害者 のための感覚代行技術の検討

伊賀 聡一郎<sup>†</sup>, 門田光司<sup>†</sup>, 安村 通晃<sup>†</sup>

慶應義塾大学大学院政策・メディア研究科<sup>†</sup>

〒252 神奈川県藤沢市遠藤 5322

{igaiga,orecchi,yasumura}@sfc.keio.ac.jp

<http://www.sfc.keio.ac.jp/~igaiga/>

### Abstract

人間の視覚による空間探索モデルの提案と本モデルを画像解析に応用したプロトタイプシステムについて報告する。人間の空間における周囲の状況把握について、「注視」と「見回す」という行動特性に着目したモデルを提案する。また、本モデルを実際に画像解析に利用した感覚器代行技術への試みを行なった。プロトタイプシステムでは、周囲の視覚的情報を音声で示すことができ、探索行動に基づいてユーザに提示される情報を連続的に変化させる。さらに、初期の評価実験とその結果について報告する。

## Sensory Substitution Technology for The Blind based on Human Visual Space Searching Model

Soichiro Iga<sup>†</sup>, Mitsuji Kadota<sup>†</sup>, Michiaki Yasumura<sup>†</sup>

Graduate School of Media and Governance, Keio University<sup>†</sup>

5322 Endo, Fujisawa, Kanagawa, 252 JAPAN

### Abstract

This paper proposes a human visual space searching model and a prototype system which applies this model to support visually handicapped. The model design is based on behaviors of sighted people when searching visual information in real space: "look closely" and "look around". We applied this model to sensory substitution technology and developed a prototype system which represents visual information by auditory information. We have conducted an experiment and reports on a preliminary evaluation of the prototype system.

# 1 研究の背景と目的

一般的に健常者の得られる情報の80%は視覚から得られる情報であるといわれている。視覚障害者が視覚情報にあふれた環境で生活し、様々な視覚的情報を把握するには周囲の情報を視覚以外の情報で知覚する必要がある。障害者が生活する上で抱えている困難な点を援助するために、ボランティアや福祉関係者による努力がなされてきた。しかし、常時ボランティアによる支援を得ることは難しく、また、障害者自身の自立を促し、人間としてのプライバシーなどの問題を守る意味においても過度の人的支援は問題である。

近年様々なテクノロジーによる視覚障害者の歩行支援に関する技術として、超音波を利用した盲人用歩行補助器 [5] や万歩計と地磁気センサを利用した歩行誘導システム [4] などが開発されている。

これら歩行支援に加えて、視覚障害者が生活していく上で必要な情報としては周囲の状況把握ということがあげられる。視覚障害者の周囲の視覚的情報を視覚以外の情報で伝達することで、視覚障害者自身による視覚情報に関する部分的な支援が可能になると考える。

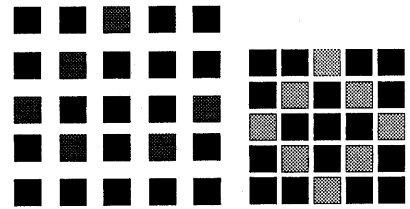
本研究では、人間の視覚による空間探索モデルの提案を行ない、イメージキャプチャによる画像解析に応用し、計算機システムによる視覚障害者のための感覚器代行技術に関する提案を行なう。

## 2 人間の視空間探索モデル

人間の視野は、左右に180度、上下に150度ほどの広がりを持つ。視力は中心視から周辺視に向かうにつれて急速に低下し、視野の中心から1.5度ほど離れただけで解像度は中心視力の30%程度になるといわれている [2]。一般的に我々は視覚による空間探索を行なう際には、サッケードといわれるように急速な眼球運動を行なっている。

計算機のカメラにより視覚障害者の視覚代行を考える際には、次の2点が課題となる。

1. 空間における周囲の状況把握: 視覚的情報を空間内で探索する際には、多くの視覚的情報の中から障害者にとって必要な情報を抽出する必要がある。そのためには、視覚代行技術



(a) 「見回す」 (b) 「注視」

情報 疎 密  
視野 広い 狭い

図1: 人間の視空間探索モデル

により解釈できる視空間を広くとり、その中から障害者が情報をしぼりこめるような仕組みを持たせる必要がある。

2. 空間内のある焦点に関する詳細な情報提示: 空間内から情報を絞り込み、目的の情報を検索できた場合、対象となる情報に関する詳細な情報が必要になる。

通常イメージキャプチャによる画像解析手法では、常に詳細な認識結果を求める傾向がみられるが、視覚障害者が利用する際には、障害者自身が求めている精度で解析を行なうことが重要である。

ここでこれらの課題に取り組むべく、人間の視覚的な空間を探索する際のモデルの提案を行なう。通常、晴眼者が空間の中からある情報を探索する際には、眼球運動や身体運動により、目的の情報にマッチする焦点を探す。探索の結果、目的の情報に近いものが視野に入ると、その焦点に関してのより詳細な情報を得ようとする。

例えば、晴眼者がある部屋である本を探している場合を考えると、まず、対象物の粗い情報(色、大きさ、形状など)に関する部分を空間から探す。対象物の粗い情報に近い物体が見つかったら、その周辺に視点を移動し、その対象物のより詳細な情報(より詳細な色、デザイン、タイトルなど)を探索し、目的の本であるかどうかを判断する。

空間内の視覚的情報の探索を行なう場合の、これら課題で挙げた視覚的な環境での動作をここで「人間の視空間探索モデル」とする。本モデルでは、晴

眼者が空間内で「見回す」動作を起こす際と、「注視」する際の視覚的および身体的運動特性に着目し、双方における情報の粒度と視野の大きさとの関係を元に、視覚による空間内での探索行動をモデル化した。

図1に本モデルの仕組みを説明する。晴眼者が「見回す」ことで空間内での探索行動を行なう場合、視覚による情報は疎になり、反面その視野は広くとられる。ある対象物の絞り込みを行ない「注視」する場合、その視点に対する情報は密になり、それにとともに視野は比較して狭くなる。情報量 ( $I$ )、1つの受容器が表わす情報量 ( $g$ )、空間の識別粒度 ( $n$ )、動きによる受容器情報上の変化 ( $v$ ) とすると、式 (1,2) のように定義することができる。なお、 $F$  と  $F'$  は単調増加関数である。

$$I = g \cdot F(n^{-2}) \quad (1)$$

$$n = F'(v) \quad (2)$$

晴眼者はこれらの視覚的および身体的行動を連続的に行なうことで空間内の様々な情報の中から自分に必要な情報を探索していることが考えられる。

### 3 プロトタイプシステム

ここでは試作した、視空間探索モデルに基づく計算機システムによる支援システムについて述べる。

#### 3.1 システム概要

本システムは、視覚による空間探索における「見回す・注視する」という視覚的および身体的行動に着目したモデルを計算機システムにおけるイメージキャプチャによる画像解析部に応用し、解析された視覚的情報を音声により提示するものである。プロトタイプシステムでは、空間内の色情報に注目し、見回す行動と注視する行動の連続的な身体的動作からイメージ情報からの粒度と視野の広さを得る。

#### 3.2 システム構成

図2にシステムのハードウェアの構成を示す。ユーザは小型 CCD カメラを携帯し、カメラを手などで移動させることで空間内の視覚的情報を探索する。

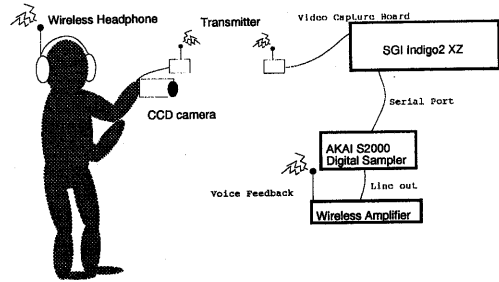


図 2: ハードウェア構成

カメラより得られた画像はビデオボードの付属したワークステーション (SGI 社 Indigo2 XZ) に送られ、画像の解析を行なう。画像の解析結果から、フィードバックすべき音声を選択され、合成音声はサンプリングされたデジタルサンプラ (AKAI S2000) 内の音声で再生される。再生された音声はワイヤレスのヘッドフォンなどでユーザに提示される。

#### 3.3 画像解析

ビデオフレーム内のピクセルを RGB 色空間から以下のクロマテックベクタへ変換する。

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B} \quad (3)$$

2つの連続したフレーム ( $f_n, f_{n-1}$ ) 間の各 ( $r, g$ ) 値に対して閾値に基づいて差分計算を行なう。差分計算の値から、モザイク処理を行ない情報の粒度が決定される。

$$p = \sum_j \sum_i |f_n(r_i, g_j) - f_{n-1}(r_i, g_j)| \quad (4)$$

フレーム間の画像の変化量が大きいほど情報が粗くなり (粒度が大きくなる)、変化量が小さいほど細かい情報となる。さらに、情報の粒度に従って中心から  $5 \times 5$  の配列からなる「視野」に相当するポイントが決定される。このポイント内に関する情報がユーザに音声として提示される。現在のシステムでは、それぞれの粒内の色に関する情報 (約 5 色程度) を検索できるようにしている。

図3に解析例を示す。図3(a)は「見回す」動作での解析結果を表している。見回す動作では、画像

の変化量が大きく、情報の粒度は粗くなっている。それにともなって、視野に相当するポイントは広くとられている。図3(b)は「注視」に相当する。注視するとは、ユーザが認識対象とする焦点のある程度絞り込み、カメラを固定的な位置に向けている状態を指す。注視動作により、情報の粒度は密になり、視野は狭く中心部分に集まる。これにより、ある焦点に関する詳細な情報を得ることができる。これらの解析は、フレーム内の変化率に従って連続的に変化する。現状では、対象は静止しているものを仮定している。また、カメラ自身のズームなどは固定されており、画像の変化はユーザによるカメラの物理的な移動からのみおこる。

### 3.4 音声情報提示

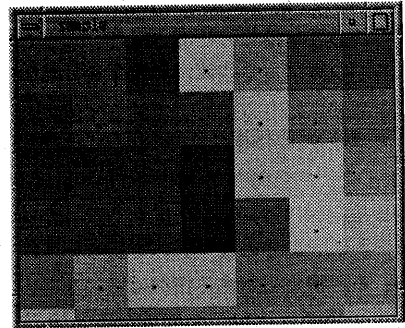
本プロトタイプでは、空間内の色情報の検索支援を行なっている。デジタルサンプラーにあらかじめ色の名称に関する合成音声をサンプルとして記録しておき、画像解析部での視野配列の各ポイントの情報を元にサンプルのピッチとパンが決定され、それらに対応した合成音声を提示している。左右情報はパン、上下はピッチで表現されている。

## 4 評価

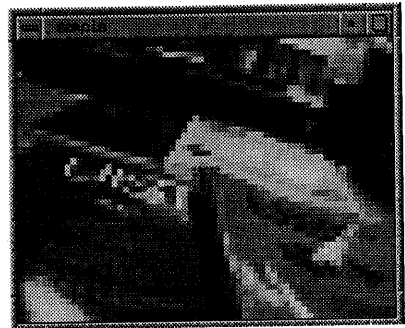
### 4.1 実験

本システムの有効性を示すべく、表1に挙げたシステムで比較実験を行なった。実験1では、本論文で提案する視覚代行支援の仕様によるもので、 $5 \times 5$ のすべてのマトリクスにおいて色情報の音声化を行なう。実験2では、カメラで捉えられた中心にあたる部分の $1 \times 1$ のみで色情報の音声化を行ない、視野は固定されている。実験3では、中心から左右3つの $3 \times 1$ のみで色情報の音声化を行ない、視野は固定されている。

実験では、計算機画面に5色(白, 赤, 緑, 青, 黄色)のうちランダムに塗り分けられた $3 \times 3$ の正方形のマトリクスを提示し、目隠しをした被験者にそれぞれのシステムを利用してもらい、発話および筆記により結果を記録した。それぞれの実験での正答率と時間を計測した。また、画面上に提示するマ



(a) 「見回す」



(b) 「注視」

図3: 画像解析結果: フレーム間の差分計算からモザイク処理による情報の粒度が決定される。情報の粒度に従って視野に相当する配列が決定される。

トリクスの位置別の正答率を計測した。被験者は5名で、それぞれの仕様について実験を行なった。

### 4.2 結果

表2に実験の結果を示す。平均正答率は、それぞれの試行で提示された画面上のマトリクスを正しく認識できた割合を表している。また、平均操作時間は、それぞれの実験について被験者が実験開始から提示された画面上のマトリクスを認識できたと判断したところまでの時間を平均したものである。

図4は各実験において画面上に提示されたマトリクスの位置別による正答率の割合を表している。 $x$ は横方向、 $y$ は縦方向を表しており、添字はそれぞ

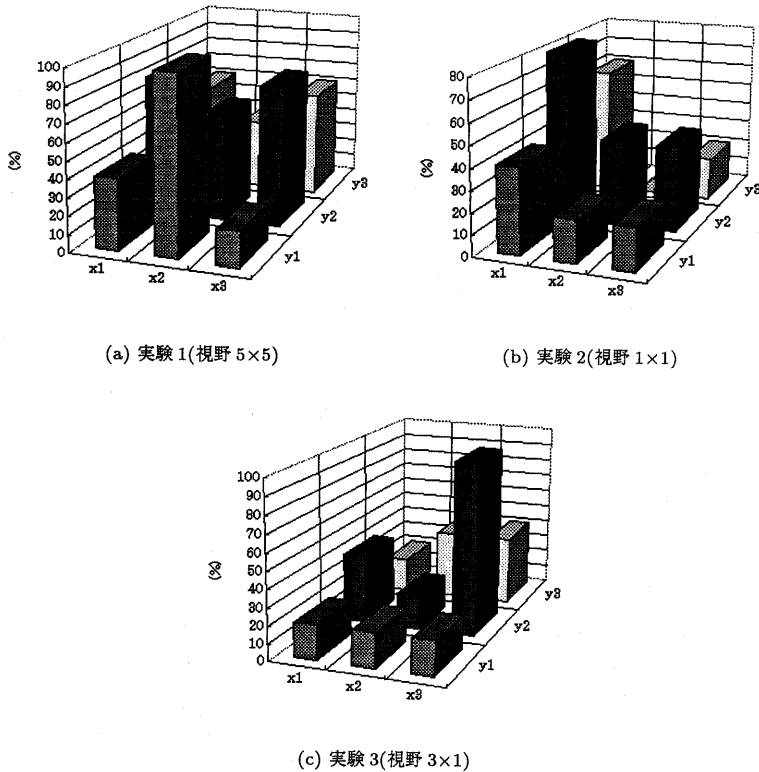


図 4: マトリクスの位置別の正答率

表 1: 比較実験

	視野の設定	仕様
実験 1	5×5	合計 25 の視野マトリクスで音声化
実験 2	1×1	中心の 1 つのみの視野で音声化
実験 3	3×1	中心から左右 3 つのみの視野で音声化

表 2: 平均正答率と平均操作時間

	平均正答率	平均操作時間
実験 1(視野 5×5)	60.0%	15.1 sec
実験 2(視野 1×1)	35.6%	14.3 sec
実験 3(視野 3×1)	35.6%	10.0 sec

れ 3×3 のマトリクスの座標位置を表している。

#### 4.3 考察

表 2 より、本研究で提案する手法(実験 1)での正答率が他の実験 2,3 と比較して高いことがわかる。これは、本手法が複数の視覚的な情報の構造を把握するのに有効であることを示している。

一方、操作時間においては、実験 1 が他の実験 2,3 に比較して長かった。これは、本手法では、マトリクス表現された複数の視野の音声化を行なっているため、提示されている音声情報が複雑になっていることが考えられる。

図 4(a)(b)(c) では、各実験における提示される情報の位置と正答率の関係を表している。この結果、実験 1(図 4(a)) では、すべての位置において 50% 以上の正答率が得られている。特に、中心付近( $x_2, y_2$ )での正答率が高くなっている。実験 2(図

4(b))では、特に $x$ 方向における位置と正答率の傾向が中心から偏っている。実験3(図4(c))でも同様に $x$ 方向に偏りが見られた。これは、視野マトリクスが少ない分、手の移動範囲が広くとられるため、空間情報を認知する際の基準となる認識ポイントが判断できなくなることが原因と考えられる。

また、実験において、システムの認識における遅延が被験者の認知の誤りをひきおこしていることが観察された。これは、被験者がカメラで捉えている個所と実際に認識を行なってフィードバックを行なうまでの遅延時間によるものである。

## 5 関連研究

計算機の3Dグラフィックスの中のある定点の位置を把握するための手法が提案されている[3]。オーケストラの楽音を利用して、3D空間の左右と奥行きに関する実験を行なっている。本システムでは、計算機のグラフィカルな情報のみならず実空間の情報を視覚障害者に伝えることを目的としている。本研究のアプローチにおいては、オーケストラで利用する楽器の音よりも、合成音声による音の方が、音と認識された情報との関係に関する記憶の負荷を軽減できるものとする。

磁気センサを利用して視覚障害者の歩行を支援するシステムが提案されている[4]。万歩計と地磁気センサをもとにした歩行軌跡を予め記録し、教師データと比較しながら音声とブザーで視覚障害者を目的地まで誘導するものである。本システムでは、このように従来の歩行を支援する形態に加えて、従来認知することが困難であった実世界のものの色や配置といった情報を音声化して伝えることで、新しい計算機による障害者支援が可能になると考える。

## 6 本研究の今後の展開

利用評価実験から、本研究の提案する人間の視空間探索モデルに基づいたシステムの有効性が確認できた。実験の結果、音声化するための視野をマトリクス状に配置することで、空間認知を行なう際の基準となるポイントを容易にできるようになり、視覚障害者のための視覚代行の手がかりとなることがわかった。しかし、同時に音声による情報提示

が複雑になるため、操作時間が長くなる傾向が見られた。今後はマトリクスの配置と音声化手法の検討を行ない、より詳細な実空間把握を目指す。

実験において、システムの認識する部分と実際のユーザの動作との遅延時間が問題となった。遅延時間が長くなるに従って、ユーザがカメラでとらえているフレームと音声により提示されている情報の食い違いが大きくなり、視覚障害者がシステムを利用する際には問題となる。システムの遅延とユーザの行動との関係に関しても実験を行ない、適切な情報提示を行なう必要がある。

現状のシステムでは、対象物が静止しておりユーザが移動することを想定しており、対象物が移動する場合には対応していない。加速度センサーなどを利用して、画像からの動きベクトルを計算する必要がある。

## 7 まとめ

人間の視空間探索モデルを元にした視覚障害者のための感覚器代行の支援システムに関しての提案を行なった。また、初期実験を通じて、本システムの利用により、視覚障害者による実空間における色情報を元にしたものの選択の可能性を示し、本研究で提案している視空間探索モデルの有効性を示した。今後は、音声情報やその他の提示手法に関する検討を行ない、実際に視覚障害者による実空間での評価実験を予定している。

## 参考文献

- [1] アクセス研究会編, 障害者のためのコンピューターアクセス'96, 慶應義塾大学湘南藤沢学会 SFC-RM97-002, ISBN4-906483-95-X, 1997.
- [2] 川人光男, 行場次郎, 藤田一郎, 乾敏郎, 力丸裕, 岩波講座 認知科学 3 視覚と聴覚, 岩波書店, 1996.
- [3] S.W. Mereu, R. Kazman, Audio Enhanced 3D Interfaces for Visually Impaired Users, In *Proc. of CHI'96*, pp.72-78, 1996.
- [4] 中村和弘, 青野嘉幸, 田所嘉昭, 視覚障害者誘導型歩行支援システム, 信学論, D-II Vol.J79-D-II No.9, pp.1610-1618, 1996.
- [5] 佐々木忠之, 超音波を利用した盲人用歩行補助器, 音響学会誌, vol.43, no.5, pp.344-348, 1987.