

NAIST-IS-DD0361016

博士論文

実空間における注視推定とインタフェースへの応用

竹村 憲太郎

2006年 3月 24日

奈良先端科学技術大学院大学  
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に  
博士(工学) 授与の要件として提出した博士論文である。

竹村 憲太郎

審査委員：

小笠原 司 教授 (主指導教員)

木戸出 正繼 教授 (副指導教員)

松本 吉央 助教授 (副指導教員)

# 実空間における注視推定とインタフェースへの応用\*

竹村 憲太郎

## 内容梗概

本論文は、実空間における注視推定手法及びインタフェースへの応用を提案するものである。「見る」という行為は人の興味・関心を示す代表的な動作であり、視線計測技術は自動車工学やヒューマンコンピュータインタラクション等の多くの分野で期待されている。本研究では人の興味・関心を抽出する手掛かりとして特に注視対象及び注視位置に注目する。従来、ディスプレイ上の注視点の推定とその応用に関して盛んに研究が行われてきたが、更に多くの分野への応用を実現するために、実空間における注視推定手法の確立が必要である。

まずはじめに静的対象に対する注視判定手法の提案を行った。非装着型の視線計測装置を用いて静的対象に対する注視判定を行い、ドライバモニタリング及びヒューマンコンピュータインタラクションへの応用を実現した。ドライバモニタリングではドライバの注視対象及び頭部の運動に注目し、車線変更及びカーブ走行の実験にて、先行研究によって報告されているドライバの特徴的な注視行動の計測を実現した。また、ヒューマンコンピュータインタラクションとしては、「見る」という行為による直感的な操作対象の切り替え手法の提案を行い、シームレスなコンピュータ及び家電の操作切り替えを実現した。一般に用いられている操作対象の切り替え手法と比較し、有効性を示した。

次に複数人の注視情報を用いて注意の集中する箇所(場の注意)の推定手法を提案した。異なる対象を注視する人が存在する場合にも安定して推定できることや、複数個の注意点が存在する場合にも本提案手法が有効であることをシミュ

---

\*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 博士論文, NAIST-IS-DD0361016, 2006年3月24日.

レーション及び実際の対話シーンで確認した。また、応用として推定した「場の注意」を用いた映像の自動編集の提案を行った。複数個の注意点が推定される場合には、注意点それぞれに対応した映像の自動編集を実現した。非装着型の計測システムとして複数カメラを用いたカメラシステムを提案し、実際の対話におけるオンライン計測及び推定を行い実現可能性を示した。

本研究では、実空間における注視推定及び注視情報に基いたアプリケーションを実現した。これらの成果は視線計測技術の大きな可能性を示すものである。

## キーワード

視線計測，注視位置・対象，ドライバモニタリング，ヒューマンインタフェース，場の注意，映像自動編集

# Estimation of Focus of Attention in Real World and Its Applications to Human Interfaces\*

Kentaro Takemura

## Abstract

In this paper, methods to estimate the focus of attention and applications are proposed. Behavior of looking reflects interests of a person, and the number of researches on driver monitoring and human computer interaction based on the gaze measurement technology has been increasing. The point of regard (POR) and the focused object are estimated as a key to estimate intention of a person. In traditional researches, the POR on the display has been measured. However it is necessary to measure the POR in real world for expanding the application area.

First, a method to estimate the focused object using non-contact measurement system in real world is proposed. Then it is applied to driver monitoring, and recognition of the objects at which the driver is looking is realized. The feasibility of the system is verified through measurement experiments using an in-car measurement system. With regard to human interfaces, a simple and intuitive interface to control multiple devices without explicit switching operation is proposed based on “What you look at is what you control” concept. The practicality of the proposed interface is verified through experiments using an implemented system.

Next, a method to estimate the group attention using attention-related information of multiple people is proposed. Even when there is a person who is

---

\*Doctoral Dissertation, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD0361016, March 24, 2006.

looking at different point(i.e. an outlier), the results of estimation are stable, and multiple group attentions can also be estimated. In our experiments, the focus of attention during conversations are estimated and the feasibility of the method is confirmed. The proposed method is applied to automated video editing. Group attention is compared with speech information, and the practicability of automated camera work is corroborated. Furthermore group attention is estimated online using multiple camera system.

In this research, estimating the focus of attention and applications based on attention-related information are realized. Achievements of this research present the possibility of gaze measurement technology

**Keywords:**

gaze measurement, focus of attention, driver monitoring, human interfaces, group attention, automated video editing

# 目次

第1章 序論	1
1.1. 研究背景	1
1.1.1 視線計測の現状と可能性	2
1.1.2 注視推定技術の重要性	3
1.1.3 従来の注視計測の問題点	3
1.2. 研究目的	4
1.3. 期待される効果	5
1.4. 本論文の構成	5
第2章 注視推定とヒューマンインタフェース	7
2.1. 本章の趣旨	7
2.2. 注視推定手法の分類	7
2.3. 注視計測の応用に関する従来研究	8
2.3.1 ドライバの視線計測に関する研究	8
2.3.2 ヒューマンコンピュータインタラクションに関する研究	9
2.3.3 複数人の注視計測に関する研究	10
2.4. 本研究の位置づけ	11
2.4.1 ドライバモニタリング	11
2.4.2 ヒューマンコンピュータインタラクション	11
2.4.3 複数人の注視情報を用いた注意推定	12
2.5. 本章の結論	13
第3章 実空間における注視対象の推定	14
3.1. 本章の趣旨	14

3.2.	注視判定手法	15
3.3.	応用1：ドライバモニタリング	20
3.3.1	実験システムの構成	20
3.3.2	車線変更時の注視行動計測	23
3.3.3	カーブ走行時の注視行動計測	26
3.3.4	考察	27
3.4.	応用2：注視による操作対象の切り替え	31
3.4.1	注視による複数コンピュータ操作の切り替え	32
3.4.2	評価実験	34
3.4.3	考察	34
3.4.4	注視による家電操作の切り替え	36
3.4.5	評価実験	37
3.4.6	考察	39
3.5.	本章の結論	39
<b>第4章</b>	<b>複数人の注視情報を利用した注意推定</b>	<b>41</b>
4.1.	本章の趣旨	41
4.2.	複数人の注意計測に基づく注意推定	42
4.2.1	頭部情報の利用	42
4.2.2	頭方向ベクトルの定義	43
4.2.3	場の注意の推定	45
4.2.4	ロバスト推定の利用	46
4.2.5	複数注意点の推定	49
4.2.6	楕円体を用いた注意推定	50
4.3.	対話中における「場の注意」の推定実験	53
4.3.1	しりとり中の場の注意	53
4.3.2	考察	57
4.3.3	ディベートの場の注意	59
4.3.4	考察	59
4.4.	応用：映像自動編集	61



4.4.1	場の注意に基づく映像自動編集 . . . . .	61
4.4.2	複数注意点に対応した映像自動編集 . . . . .	62
4.5.	カメラシステムによる複数人の同時計測 . . . . .	65
4.5.1	カメラキャリブレーション . . . . .	65
4.5.2	カメラシステムによる対話シーンの計測 . . . . .	69
4.6.	本章の結論 . . . . .	69
<b>第5章</b>	<b>結論</b>	<b>72</b>
5.1.	まとめ . . . . .	72
5.2.	課題と今後の展望 . . . . .	74
	謝辞	76
	参考文献	79
	付録	87
A.	顔情報計測システム . . . . .	87
B.	半直線と三角形の交差判定 . . . . .	90
C.	注視頻度マップ . . . . .	91
D.	M 推定 . . . . .	92
E.	平面パターンによるカメラキャリブレーション . . . . .	93
	著者研究業績	96

# 目次

3.1	注視判定手法 (球によるモデル化)	15
3.2	注視判定手法 (三角ポリゴンによるモデル化)	16
3.3	机上の注視対象	17
3.4	机上の対象モデル	17
3.5	視線を用いた注視判定結果	18
3.6	頭部方向を用いた注視判定結果	18
3.7	注視対象モデルと処理時間	19
3.8	実験車両: (a) 外観, (b) ステレオカメラ, (c) 全方位カメラ, (d) 前方カメラ	21
3.9	記録画像: (1) 前方カメラ画像, (2) 全方位画像, (3) カーナビ画面, (4) 処理画面	21
3.10	トラッキング結果: (a) 昼間, (b) 夜間	22
3.11	車内の注視対象の位置関係	22
3.12	車内での注視対象の推定結果: (左) サイドミラー, (中) スピードメータ, (右) ルームミラー	23
3.13	車線変更中のドライバの様子と注視点	24
3.14	車線変更時の頭部運動	25
3.15	サイドミラー, 及びルームミラーへの注視頻度	25
3.16	カーブ走行中のドライバの様子	27
3.17	頭部方向と車両のヨーレイトの関係	28
3.18	視線の水平方向の動きと車両のヨーレイトの関係	28
3.19	左カーブ走行中の注視頻度マップ	29
3.20	右カーブ走行中の注視頻度マップ	29

3.21	トラッキングが失敗したケース	30
3.22	複数コンピュータの操作切り替えシステム	32
3.23	注視によるコンピュータ切り替えの様子	33
3.24	タスク完了平均時間 (PC 操作)	35
3.25	家電操作切り替えシステム	36
3.26	ライトの点灯の様子：(左) 点灯前，(右) 点灯後	37
3.27	タスク完了平均時間 (家電操作)	38
4.1	高再帰性反射マーカ	43
4.2	マーカと頭部方向ベクトルの関係	44
4.3	頭部方向のキャリブレーション	44
4.4	注視点の推定手法	45
4.5	場の注意の推定手法	45
4.6	注意点と頭部方向の距離	46
4.7	注意の共有が成立しない状況	46
4.8	注意点の軌跡	47
4.9	M 推定のシミュレーション結果	49
4.10	前フレームの推定結果を反映しない場合	51
4.11	前フレームの推定結果を反映した場合	51
4.12	楕円体を用いた注意の推定手法	52
4.13	楕円注意領域の推定結果	52
4.14	しりとりの様子	54
4.15	しりとり中の注意点の軌跡	54
4.16	推定された注意方向	55
4.17	「しりとり」のターン，注意対象 (人)，発話者のタイムチャート	55
4.18	設定した注意対象 (人) の範囲	55
4.19	注意点の共有	58
4.20	注意点の分離	58
4.21	注意点の分離 (全方位画像上にオーバーレイ表示)	58
4.22	複数注意点の推定結果	60

4.23	複数注意点の推定エラー	61
4.24	映像の自動編集 (注意点が1つの場合): (左)「場の注意」の推定結果, (中) 注意方向の提示した全方位画像, (右) 切り出された注意画像	63
4.25	映像の自動編集 (注意点が複数の場合): (左)「場の注意」の推定結果, (中) 注意方向の提示した全方位画像, (右) 切り出された注意画像	64
4.26	ミーティングキャプチャシステム (プロトタイプ)	66
4.27	平面パターンによる特徴点の抽出, 左:オリジナル画像 右:特徴点抽出結果	66
4.28	カメラキャリブレーション	68
4.29	キャリブレーション結果	68
4.30	複数人の顔情報計測結果	70
4.31	カメラシステムによる場の注意の推定結果	70
1	顔情報計測システムの外観	88
2	顔全体のテンプレート画像	88
3	特徴領域テンプレート画像	88
4	視線ベクトル計算	89
5	顔情報計測結果	89
6	注視頻度マップ	92

# 表 目 次

3.1	入力インタフェースと制御対象 . . . . .	37
4.1	頭部方向ベクトルの角度誤差 . . . . .	48
1	代表的な M 推定法 . . . . .	93

# 第1章 序論

## 1.1. 研究背景

近年，コンピュータの処理能力の向上や画像処理技術の発展に伴って，リアルタイム処理や非装着型での視線計測が実現されている．従来，視線計測は研究者によって視覚特性や認知プロセスの解明のために多く利用されてきたが，技術の発展に伴い，一般ユーザの利用が期待されている．

自動車工学の分野ではドライバが運転に使用する情報の90%以上が視覚情報であると報告されており，ドライバモニタリング等において非常に重要な要素と考えられている．視覚特性には多くの重要な知見が報告されていることから，事故防止や運転支援につながると期待されている．また，「見る」という行為は直感的で分かり易いことや日常的な動作であることから，ヒューマンインタフェースへの応用にも適していると考えられている．最近では，心理学の研究成果から「見る」という行為が発達過程やコミュニケーションにおいて重要な役割を担っていることが明らかになってきた [1]．注視情報を共有することで円滑な協同作業を実現できると Computer-Mediated Communication(以下 CMC) や Computer Supported Cooperative Work(以下 CSCW) の分野への応用も注目を集めている．

このように視線計測技術は，多くの分野の研究者・技術者が注目する技術に発展し，応用が期待されている．しかしながら，視線計測技術の利用は一般ユーザへの普及には至っていない．計測技術やアプリケーション等様々な問題点が考えられるが，普及には画期的なアプリケーションが必要不可欠である．

### 1.1.1 視線計測の現状と可能性

#### 視線計測技術

視線計測の歴史は非常に古く，非侵襲な計測技術は1901年にDodgeとClineによって提案されたのが最初とされている [2]．その後約100年間に渡り，多くの研究者によって視線計測技術の改良が行われてきた．計測手法は大きく以下のよう  
に5つに分類される．

- 1: 眼球の電位差を測定する方法 (EOG 法)
- 2: サーチ・コイルを内蔵したコンタクトレンズを装着する方法 (サーチ・コイル法)
- 3: 強膜反射率の違いを検出する方法 (強膜トラッカー法・強膜反射法)
- 4: 角膜反射光を検出する方法 (角膜反射光法)
- 5: 角膜反射光と瞳孔中心を同時計測する方法 (瞳孔-角膜反射法)

2000年に入ってからには画像処理やアクチュエータの発展に伴って，非装着型の視線計測技術の研究が盛んに行われ，大きな成果を残している．海老澤は明瞳孔及び暗瞳孔の検出を併用した頑強な瞳孔検出手法 [3] を提案した．また，Morimotoらはこの技術を応用した視線計測システム BlueEyes [4] を提案している．松本らはステレオカメラを用いて顔向き及び位置と虹彩を同時計測するシステムを提案し [5]，非装着型の計測システムにおける頭部計測の重要性を示した．現在では，大野らによって頭部の位置計測と視線計測を組み合わせたシステムが提案されている [6]．この計測システムはユーザの頭部位置に応じて，目領域を大きく撮影するカメラがパンチルトを行い，非装着型のシステムにおいて高精度な計測が実現されている．現在，このような計測装置は大学等からスピンオフした企業によって販売が行われている．

#### 応用分野

視線計測技術は視覚特性の研究に多く利用され，リーディングのメカニズムの解明やパイロット視線特性の研究に用いられてきた．1970年代からドライバの視

線計測等の研究も盛んに行われている。また、ユーザビリティ等の検証にも多く用いられてきた。

従来、視線計測技術は視覚特性の研究のため研究者によって利用されてきたが、1981年にJ.L.Levineによってコンピュータの入力方式として利用することが提案された[7]。その後、一般のユーザが使用することを目的としたアプリケーションが多く提案されている。今後、近いうちにコンピュータに内蔵されたカメラによる視線計測が実現され、多くのユーザに利用されることが期待できる。

### 1.1.2 注視推定技術の重要性

視線計測技術の利用は視覚特性の研究から、ヒューマンコンピュータインタラクション(以下 HCI)等へと応用範囲が広がってきている。視覚特性を研究する分野では固視微動(fixation)や跳躍運動(saccade)等の眼球運動の計測が重要となるが、HCI等の研究ではユーザの興味・関心を抽出する手掛かりとして注視対象・位置の推定が重要である。

一般にアイカメラによって計測される情報から注視点を求めるためには測定した視線情報を投影する必要がある。装着型の計測装置の場合は、頭部に装着しているカメラの画像上に投影するのが一般的である。これに対して、非装着型の計測装置の場合には既知である平面(ディスプレイ等)上に投影し、注視点を求める手法が一般的である。

視線計測を HCI やその他の分野に応用する場合、注視対象の形状や位置は非常に複雑になる。そこで実空間における注視対象・位置の推定手法を確立する必要がある。

### 1.1.3 従来の注視計測の問題点

装着型のアイカメラを用いた場合、カメラ画像上に注視点を投影することは可能であるが、頭部運動に拘束はないため撮像画像は、頭部運動と同様に自由に並進・回転が可能である。これによって、頭部位置と環境の対応関係を明らかにすることが難しい。現在では、環境との対応関係を明らかにするために、環境に設



置したマーカを頭部に装着したカメラでトラッキングする手法やステレオカメラや磁気センサによって頭部位置を計測する手法が用いられている。

注視対象の推定を行う場合のアプローチとしては、装着したカメラの画像から物体認識を行い注視判定を行う手法も考えられる。しかしながら、頭部運動の影響がカメラ画像に影響するため困難である。非装着型による視線計測では、アイカメラと環境の位置関係が明らかであり、幾何学的に注視対象の推定が可能である。しかしながら、ディスプレイ等の平面上への投影が一般的であり、応用範囲が限られている。

また、従来の注視判定手法は環境の情報を用いたものが殆どであり、予め設定した対象のみに対しての注視判定が一般に行われてきた。より高度なインタラクションを実現するには、注視対象と注視位置の情報を別けて考えることも重要な場合がある。

「見る」という行為は非常に重要度が高く、その行為はドライバ計測及び、HCI、CSCW 等において注目されている。応用範囲を広げるためにも注視対象・位置の推定技術の確立が求められる。注視対象の推定技術を確立することで、その応用範囲は更に広がると考えられる。

## 1.2. 研究目的

近年、コンピュータや家電、自動車等の機器はより高い機能を求められる傾向にある。ユーザインタフェースにおいても同様に、グラフィカルユーザインタフェース(以下 GUI) から Perceptual User Interfaces(以下 PUI) や Attentive User Interfaces(以下 AUIs)[8] へと高機能が求められる傾向がある。ユーザにインタラクティブな機能を提供する際には、現在の興味・関心の情報が非常に重要である。「見る」という行為は人の興味・関心等を表す代表的な動作であり、注視対象・位置の推定はユーザの意図を抽出する手掛かりとして重要である。

本研究ではユーザの興味・関心が反映されている注視行動に注目し、注視推定手法の確立を目指す。注視対象・位置の推定は計測状況や装置に依存するものであるが、本研究ではユーザに負担の少ない非装着型の計測装置における注視対象・

位置の推定手法を提案する．非装着型の視線計測システムの利用は一般にディスプレイ上における注視点推定に用いられてきたが，実空間における注視対象・位置の推定手法を確立し，応用範囲の拡大を目指す．

また，ドライバモニタリングやヒューマンインタフェース，CSCW 等の分野において，実空間へと利用を拡張し，新しいアプリケーションの提案と評価を行う．

### 1.3. 期待される効果

本研究では実空間における注視推定手法を確立し，その成果は今後の視線計測技術の応用に貢献するものと思われる．従来，視線計測技術はディスプレイ上の注視点推定等に利用されてきたが，本研究ではコンピュータ上からドライバモニタリング，HCI，CSCW 等への応用を提案し，視線計測技術利用の有効性・可能性を示す．視線計測技術の新しい応用を提案することで，一般ユーザへと普及させることが可能であり，計測技術の向上にも大きな影響を与えることが期待できる．また，従来は視線計測の応用分野はユーザ 1 人を対象としたものが主流であったが，実空間への拡張及び CSCW 等への応用を実現することで，複数人の視線情報利用の有効性及び重要性を示し，今後の視線計測技術の発展に大きな影響を与えられる．

### 1.4. 本論文の構成

本論文は以下の 5 章から構成される．

第 1 章：序論 研究背景として視線計測の現状と可能性について述べ，注視推定の問題点を明確にし，本研究の目的及び期待される効果を明確にした．

第 2 章： 従来行われてきた注視計測に関する研究を，注視推定手法及び応用分野の観点から整理/分類する．応用分野としてはドライバの視線計測，HCI 及び複数人の注視計測という枠組みで分類する．各々の研究の位置づけを推定手法及び応用分野から明確にし，本研究の位置付けを明確にする．

第3章： 実空間における注視計測に注目し，注視判定手法の提案を行う．実世界における注視判定を実現する際は対象のモデル化が重要であり，本研究では球及び三角ポリゴンによってモデル化された対象に対する注視判定手法を提案する．また，提案した推定手法を用いて，ドライバモニタリングやHCIの分野に応用を行い，評価実験を通して有効性を示す．ドライバモニタリングでは実車において，先行研究によって報告されている重要な知見と同様の結果が，提案するシステムを用いて得られるか検証を行う．また，HCIの分野では「見る」という動作による操作対象の切り替えを提案し，コンピュータ及び家電の操作切り替えを提案する．

第4章： 複数人の注視情報を利用した注意推定手法の提案を行う．モーションキャプチャを用いて顔向きを推定し，それらの情報から注意点を推定する．人の注意が異なる対象に向けられた場合や注意点が複数存在する場合を考慮したアルゴリズムを提案し，有効性を示す．応用例として映像の自動編集を行い，音声等の比較から有効性を示す．

また，複数の非装着型計測システムを用いたミーティングキャプチャシステムの提案を行う．統一した座標系における複数人の注視情報の同時計測を実現し，実際のミーティングにおける実現可能性を示す．

第5章：結論 本論文を結ぶ．本研究を通して得られた成果を総括し，今後の展開/可能性を述べる．

# 第2章 注視推定とヒューマンインタフェース

## 2.1. 本章の趣旨

本章では，従来の注視計測に関する研究を注視推定手法という観点から整理/分類し，各々の研究成果と位置づけを明確にする．また，注視計測技術の応用分野に関しても，ドライバの注視計測，ヒューマンインタフェース，及び複数人の注視計測と整理/分類を行い，本研究で提案する手法及びシステムの意義を明確にする．

## 2.2. 注視推定手法の分類

注視対象・位置の推定手法は，状況に応じて様々な手法が提案されている．一般に良く利用されているコンピュータ上の注視点は，アイカメラとディスプレイの位置関係から画面上の注視点を推定している．

また，近年はコンピュータ以外の機器に対する注視に関しても注視判定が行われている．JefferyらはeyeCONTACTセンサをホームアプライアンス等のデジタル機器に接続したeyePLIANCEを提案している．これらは注視によるインタラクティブな機器の動作を実現している[9]．また同様のeyeCONTACTセンサをサングラスに装着し人とのアイコンタクトを推定する手法も提案されている[10]．

これらの手法と異なり，注視対象の情報を使用せずに注視点を求める試みも行われている．満上らはアイカメラによって両眼の計測を行い，輻輳角から注視点を推定した[11]．また，Sumiらはアイカメラとステレオカメラを併用することで注視点を3次元情報として推定している[12]．

## 2.3. 注視計測の応用に関する従来研究

注視計測は、HCI の分野をはじめとして様々な分野に応用されている。本節では各分野において行われている応用を整理/分類する。本研究で注目している分野は、注視利用の代表的な分野である HCI とドライバの視線計測、また複数人の注視情報を利用した応用である。以下、これらの分野における代表的な研究について述べる。

### 2.3.1 ドライバの視線計測に関する研究

ドライバの視線計測の研究は 1970 年代から盛んに行われていて、ドライバが運転に必要としている情報のうち 90%以上は視覚情報であると報告されている。運転中の視覚特性の解明が重要課題として、カーブや交差点等の様々な状況において計測が行われてきた。以下、代表的なドライバの視覚特性に関してまとめる。

カーブ走行中の注視特性としては、Land らはステアリングと視線の動きには相関があることを示している [13]。また、ドライバの視線はカーブ内側の tangent point に集中する傾向があること [14] やステアリングが道路のどの部分に影響を受けているかを明らかにした [15]。交差点付近の注視特性については Salgian らが停止線と標識の切り替えの周期が早まる特徴があることを報告している [16]。車線変更においては Salvucci らによってミラーやレーンに対する注視時間に特徴ある変化があることが報告されている [17]。視線の遠近運動特性に関しては Miura らによって明らかにされた [18]。これによってカーナビゲーションや各メータに対する注視の影響が示され、各メーカーによってセンタメータ等が提案されている。藤森らはドライバの注視行動における分担比に注目し計測を行い、その特性を明らかにした [19]。

近年ではこのような注視特性をドライバの運転支援に用いることが注目を集めている。運転支援の代表的な例としてはドライバの意図推定があり、A.Liu らが注視エリアから Hidden Markov Dynamic Models(以下 HMDMs) を用いて行った [20]。また、運転支援を目的としたドライバの計測システムが提案されており、ドライバの非装着型計測システムとして代表的なものは Volvo によって提案された

システムがある [21] . これは松本らによって提案された顔情報計測システム [5] を用いてドライバの計測を実現している . また , Luke らは , 実車において交通標識の注視判定を実現し , 運転支援を行っている [22] .

### 2.3.2 ヒューマンコンピュータインタラクションに関する研究

入力インタフェースとしての歴史は , J.L.Levine によって提案された 1981 年に提案された [7][23] のがはじめて , その後 Ware らによってターゲットの選択などポインティングデバイスとして視線入力の効果的であることが示された [24] . 視線入力インタフェースは見たものがすべて入力として扱われてしまう”Midas Touch Problem” があったが , この問題に対して Jacob らはメニューの選択には一定時間の注視を必要とする手法を提案した [25] . また , 同様の手法において Hansen らは選択時間をユーザにフィードバックする Eyecon を提案した [26] . これによってユーザは視線による GUI の選択が自由に行えるようになった . 大野らはメニューの情報表示領域と選択領域を分離することで “Midas Touch Problem” を解決している [27] . また , Zhai らによって視線の入力とマウスによる入力を融合することでマウスのみ利用よりも良いパフォーマンスが得られることが報告されている [28] .

視線計測はディスプレイの表示技術に対しても応用が行われている [29][30] . 大型のディスプレイや 3D 表示ではリアルタイム表示を実現するため , 注視しているエリアの解像度を高めることやそのエリアのモデルを詳細に表示するなどの工夫が , 視線情報を利用し行われてきた . また , 視線の履歴情報を利用した写真や絵画の表示手法も検討されている [31][32] . 履歴情報を反映させることで , 重要な情報をユーザに分かりやすく伝えることも可能である .

実世界における視線利用に注目すると , R.Bolt らによってビデオシステムにおける映像のズームやクロージングに視線情報の利用がいち早く提案された [33][34][35] . また , 機器とのインタラクションの代表的な研究としては , J.Shell らによって提案された eyePLIANCE がある [10] . これらは AUIs としてユーザの注視に応じたインタラクションを実現している .

### 2.3.3 複数人の注視計測に関する研究

コミュニケーションにおける視線の重要性が注目され、複数人の注視計測が行われるようになってきた。R.Vertegaalらは複数人の会話シーンに注目し、人は話者に特に注目する傾向があることを示した [36]。

複数人のシチュエーションとして代表的なものが会議であり、通常の会議や遠隔会議など状況に応じて注視情報の利用法が提案されている。遠隔コミュニケーションではR.VertegaalらがGAZE Groupware Systemを提案している [37]。Gaze Groupware Systemでは、ユーザの顔をVRML空間上に注視対象の方向を向けて配置することで、誰に注意が向けられているかを提示した。また、Jenkinらによって最も注視されている人物を拡大して提示するeveViewというビデオ会議システムも提案されている [38]。

また、Stiefelhagenらは実際の会議に注目し、全方位カメラによって撮像した画像から顔向きを推定し、ベイズ推定を用いてミーティング中のFocus of Attention(以下FOA)を推定している [39][40]。また、竹前らは、会議中の注意状況を映像の自動編集へ応用することを提案している [41][42]。また、複数人の注視情報を用いたカメラワークの研究は、会議以外にもスポーツ [43] や講義 [44] において行われている。

1対1のビデオ会議では、アイコンタクトの重要性も注目されている。通常、ディスプレイ上にカメラを設置するが、ユーザはディスプレイ上の相手の顔を注視するためアイコンタクトを実現することは難しい。この問題に対して、Gaze-2と呼ばれるハーフミラーを用いたアイコンタクト可能なシステム [45] や2台のカメラの中間画像を生成する手法等が提案されている [46]。これらのシステムを用いることで、ビデオ会議においてもアイコンタクトを実現することが可能である。

オフィスにおいてはコミュニケーションを取りたい時のみ透過するAttentive Cubicle Wallが提案されている [47]。これは頭部の向きに注目し、互いに注視をした場合に液晶の原理を応用した瞬間調光ガラスによって透過への切り替えを実現している。

以上のように、複数人の注視計測を用いた研究ではコミュニケーションを円滑に行うための支援としてCMCやCSCWの研究が盛んに行われている。

## 2.4. 本研究の位置づけ

本研究では、実空間における注視判定手法を確立し、ドライバモニタリングやHCIへと応用を行う。また、複数人の注視情報から、注意の集中する箇所の推定を実現する。各応用分野における本研究の位置付けの詳細を以下に述べる。

### 2.4.1 ドライバモニタリング

ドライバの注視計測では、先行研究によって多くの重要な知見が報告されており、中でも注視対象や注視時間はドライバの意図推定につながる要素であると考えられている。ドライバの特徴的な動作の計測を実現することで、高機能な運転支援の実現を期待することが出来る。また、ドライビングレコーダ等が登場し、運転状態の記録が行われている。このような背景からもドライバの定量的な行動記録は社会的に意義のあることであると考えられる。

従来研究において、非装着型のドライバモニタリングシステムの提案は行われているが、ドライバの特徴的な動作の計測は行われていない。そこで本研究ではドライバモニタリングとして、ドライバの頭部運動、視線方向の定量的な計測及び注視対象の推定を実現する。ドライバの注視計測では、装着型のシステムでの計測が一般的であったが、実車へ搭載する場合には非装着型で拘束感のない必要がある。非装着型の計測装置を用いることで、ドライブレコーダ等で常時記録可能な計測システムへと発展をさせることが可能である。実車に非装着型の顔情報計測システムを搭載し、実際に走行中の注視対象の推定を行い得られたデータと重要な知見を比較し、提案する非装着型のシステムによる計測の実現可能性を検証する。

### 2.4.2 ヒューマンコンピュータインタラクション

HCIの分野では従来行われてきた入力方式としての利用やAUIsのようなインタラクションとは異なり、複数の機器の操作切り替えに視線情報を用いる。近年、コンピュータや家電の高機能化が進み、利便性が非常に向上しているが、同様の



機能を有するインタフェースが多数存在することが問題として指摘されている。一般的な対策として、コンピュータの切り替えにはKVMスイッチ、家電の切り替えにはユニバーサルリモコンが利用されている。

本研究では、「見る」という行為と操作対象の切り替えを対応させ、ユーザの注視という自然な動作でシームレスに操作対象の切り替えを実現する。ユーザは操作する対象を注視する傾向があることから、その動作を用いて直感的な操作対象の切り替えを実現し、ユーザビリティの向上を目指す。複数機器のインタフェースの切り替えにユーザの視線を用いた例は少なく、本研究では注視による操作対象の切り替え手法を提案すると共に評価を行う。複数コンピュータ及び家電における操作切り替えにおいて、本提案手法の実装を行い、従来手法と比較し有効性を示す。

### 2.4.3 複数人の注視情報を用いた注意推定

インターネットや記録媒体の発展に伴い、映像記録の用途が非常に増加している。映像は分かり易く記録することが重要であり、カメラワークや映像の自動編集という用途において注意推定は非常に注目されている。また会議等においては、ホワイトボードやテーブルの上の対象など様々なものに対して注意が注がれる。そのような情報を計測し、遠隔コミュニケーションを支援することは勿論であるが、自動要約への応用など会議システムにおける注意計測は非常に応用範囲が広いと考えられる。

複数人の視線計測を用いた注意点の推定では、注意対象を人に限定し推定を行う手法が一般的である。しかしながら、実際のコミュニケーションでは注視対象を予め設定しておくことは非常に難しい。このような理由からより柔軟に空間中の注意を推定することが求められる。そこで本研究では、注視対象の位置・大きさ等の情報は利用せずに、複数人の視線情報を用いて注意の集中する箇所(場の注意)を推定する。これによって事前に注意対象を定義する必要もなく、柔軟な注意推定が可能となる。

従来、注意点はある1点に集中しているものと仮定し推定が行われ、異なる対象を注視するユーザの情報を考慮した研究はない。しかしながら、グループの人

数が増えると注意点の数は1つに限定することが出来ない。そこで本研究では、複数の注意点や異なる対象へ注視するユーザの情報への対処等の問題に取り組む。これらを考慮した注意点の推定手法を提案し、実際の対話シーンにおける推定実験を通して、有効性を検証する。

## 2.5. 本章の結論

本章では、従来より数多く行われてきた視線計測技術の応用研究を、ドライバ計測、HCI 及び複数人の注視計測と分類、整理を行った。また、各研究の成果と位置づけを明確にし、現在の問題点を指摘した。そして、各応用分野における本研究の位置づけを従来研究との比較から明確にした。

# 第3章 実空間における注視対象の 推定

## 3.1. 本章の趣旨

本章では，実空間における注視対象の推定手法及び各分野への応用の提案を行う．初めに，提案する注視対象の推定手法についての詳細を述べる．提案手法は，予め登録した機器のモデル(球・三角ポリゴン)と頭部(位置・姿勢)及び視線情報から，幾何学的に注視判定を行う．

次に注視対象の情報をを用いた応用例を示す．本研究では応用例として，ドライバモニタリング及びヒューマンインタフェースを提案する．ドライバモニタリングでは注視対象の情報が重要であり，先行研究ではこの情報を用いてドライバの意図推定等が行われている [20]．しかしながら，非装着型の計測装置を用いた実車での計測が行われている例は少なく，計測はシミュレーションや装着型の計測装置を用いたものがほとんどである．そこで本研究では，非装着型の計測装置である顔情報計測システム及び提案する注視判定手法が，ドライバのモニタリングに有効であるか公道にて実験を行い，そのデータを過去の知見と比較し有効性を示す．

また，ヒューマンインタフェースの分野では「注視による操作対象の切り替え」を提案し，複数コンピュータ及びネットワーク家電に対して，注視情報を用いた直感的な操作切り替えを実現する．評価実験では従来手法と比較し，操作対象の切り替えに提案手法の有効性を示す．

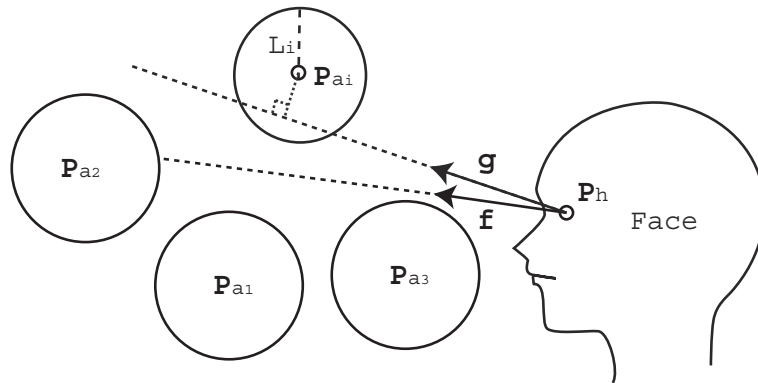


図 3.1 注視判定手法 (球によるモデル化)

### 3.2. 注視判定手法

従来，注視推定はコンピュータ画面を主な対象として研究が行われてきた．応用範囲の拡大を考えた場合，コンピュータ画面上に限定せず，実空間における様々な対象への注視判定を実現する必要がある．本節では，実空間の対象に対する注視判定手法の提案を行う．本研究では非装着型のアイカメラを用いて頭部位置，姿勢及び視線の計測を行うが，装着型のアイカメラを用いた場合も頭部位置計測をカメラや磁気センサを用いて行うことが出来る場合，本手法を適用することが可能である．

まずはじめに図 3.1 に示すように，3次元空間中に物体が配置されている環境において，視線及び頭部方向を用いた注視判定手法を提案する．視線・頭部方向ベクトルと物体の交差を判定することでユーザの注視対象を推定する．頭部位置，姿勢及び視線計測には顔情報計測システム (付録 A) を用いる．対象  $i$  を注視しているとすると，頭部位置及び視線ベクトル，対象  $i$  の 3次元位置は次式の条件を満たす．

$$\left\| \frac{(\mathbf{p}_{a_i} - \mathbf{p}_h) \cdot \mathbf{g}}{\|\mathbf{g}\|^2} \mathbf{g} + \mathbf{p}_h - \mathbf{p}_{a_i} \right\| \leq L_i \quad (3.1)$$

$\mathbf{p}_h$ ， $\mathbf{p}_{a_i}$  はそれぞれ頭部位置と  $i$  番目のオブジェクトの位置である．また， $\mathbf{g}$  は視線ベクトルであり， $L_i$  は  $i$  番目のオブジェクトの半径である．すべてのオブ

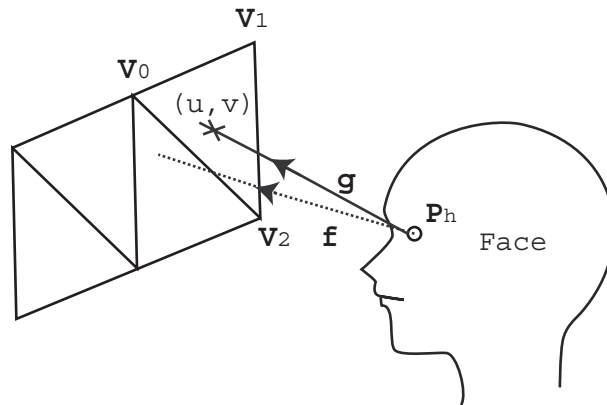


図 3.2 注視判定手法 (三角ポリゴンによるモデル化)

ジェクトは取り扱いが容易であるため球として扱い、オブジェクトの位置及びサイズは事前に定義する．左辺は点  $p_{a_i}$  と視線ベクトル  $g$  の距離であり、それと対象の半径  $L_i$  の比較を行う．視線ベクトルが球を通過するか否かで注視判定を行う．視線方向を  $g$  の代わりに頭部方向ベクトル  $f$  を用いても注視判定を行うことが可能であり、頭部方向を用いた場合は次式を用いる．式 3.1 の視線ベクトル  $g$  を頭部方向ベクトル  $f$  に置き換えることで、注視判定を行うことが可能である．

$$\left\| \frac{(p_{a_i} - p_h) \cdot f}{\|f\|^2} f + p_h - p_{a_i} \right\| \leq L_i \quad (3.2)$$

対象のモデル化は図 3.1 に示したように、球を用いて行うことで容易にモデルを作成することが可能である．しかしながら、対象に対してより正確な注視判定を考えた場合、詳細なモデル化が必要となる．そこで、本研究では球によるモデル化が適していない場合には、三角ポリゴンを用いてモデル化を行い注視判定を行う．現在のところモデル化はマニュアルで行っているため、複雑な形状に対応することはできないが、平面で構成されているものに関しては三角ポリゴンの集合としてマニュアルでも容易にモデル化することが可能である．

コンピュータグラフィックス (以下 CG) の分野では Ray と Triangle に対する交差判定手法が提案されている [48]．本研究では、三角ポリゴンで構成したモデルに対する注視判定に、Tomas Möller の交差判定手法 (付録 B) を用いる．初めに

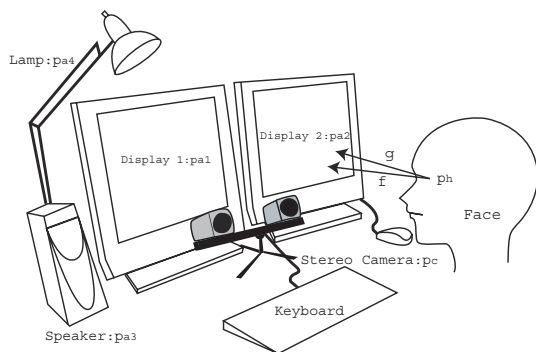


図 3.3 机上の注視対象

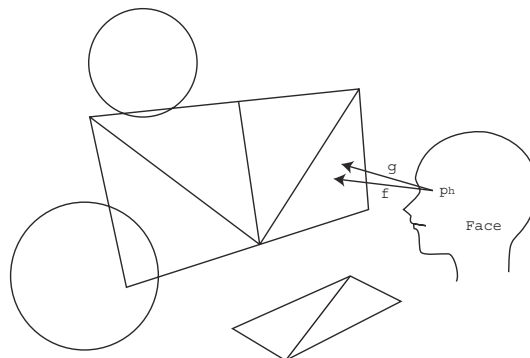


図 3.4 机上の対象モデル

図 3.2 に示すように対象を三角ポリゴンで表す．注視方向ベクトルは視線及び頭部方向ベクトルを用いて以下のようになる．

$$\mathbf{R}_g(t) = \mathbf{p}_h + t\mathbf{g} \quad (3.3)$$

$$\mathbf{R}_f(t) = \mathbf{p}_h + t\mathbf{f} \quad (3.4)$$

また，三角形は重心座標系  $(u, v)$  を用いると

$$\mathbf{T}(u, v) = (1 - u - v)\mathbf{V}_0 + u\mathbf{V}_1 + v\mathbf{V}_2 \quad (3.5)$$

となる．式 3.4 と式 3.5 または式 3.3 と式 3.5 より Cramer の公式を用いて求めた交点  $(u, v)$  が  $u \geq 0$  ,  $v \geq 0$  及び  $u + v \leq 1$  の条件を満たすとき，三角形を注視していると判定する．三角ポリゴンを用いることでより，対象に関してより詳細なモデル化及び注視判定性能を向上させることが可能である．

以上に述べた手法を用いて，実際に図 3.3 に示すような机の上の環境において，注視判定を行った．ツインディスプレイ，卓上ライト及びスピーカーを配置し，視線ベクトル及び頭部方向ベクトルを用いて注視判定を行った．図 3.4 に示すように，注視対象のモデル化を行う．ツインディスプレイのように隣接している対象に関しては，三角ポリゴンによるモデル化が有効である．卓上ライトやスピーカーについては球でモデル化を行った．視線を用いた場合の注視判定結果を図 3.5 に示す．図の下部に表示している結果が推定した対象の名称である．虹彩から伸

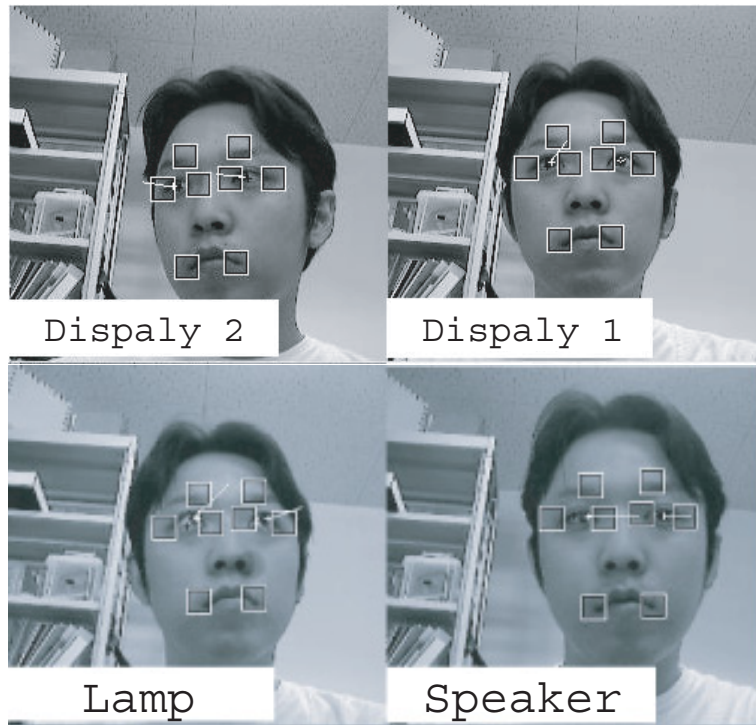


図 3.5 視線を用いた注視判定結果

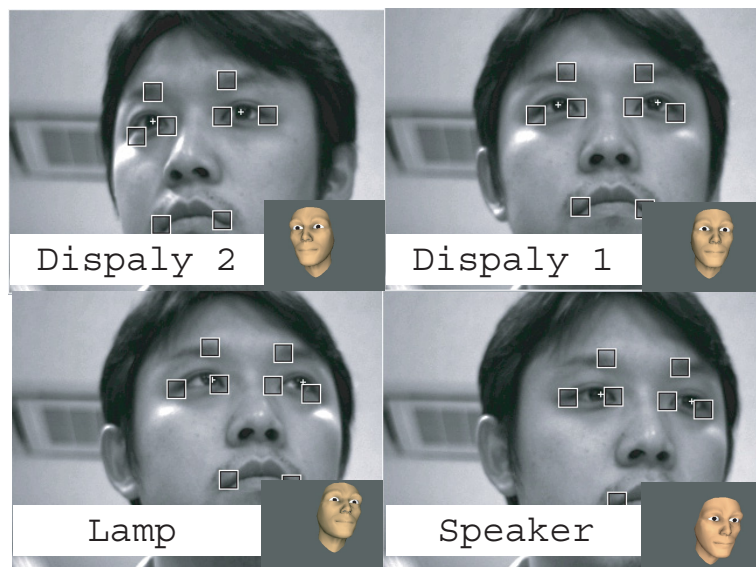


図 3.6 頭部方向を用いた注視判定結果

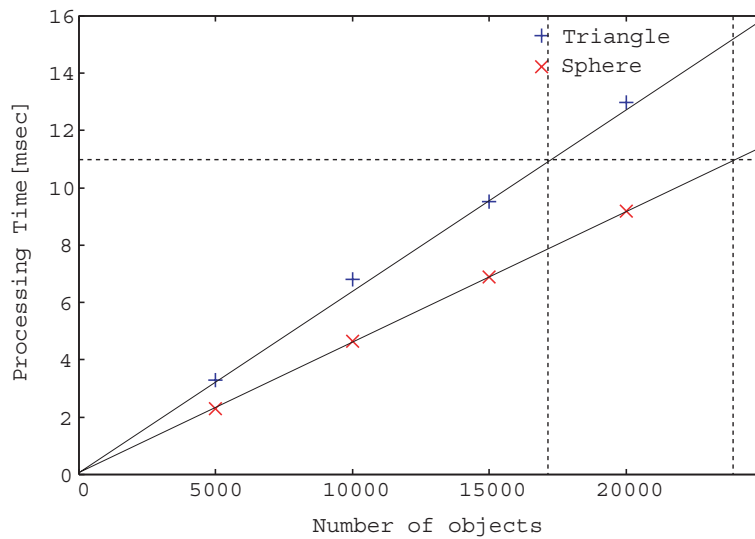


図 3.7 注視対象モデルと処理時間

びている白い線は計測された視線方向となっている．また頭部方向を用いた場合の注視判定結果を図 3.6 に示す．図の右下には推定した頭部姿勢を示す．視線及び頭部方向を用いた注視判定手法が有効であることを，実験を通して確認した．

HCI への応用を考えた場合，リアルタイム (ビデオレート) での処理が重要である．そこで処理速度に関して考察を行う．本研究で提案した，球を用いた場合と三角ポリゴンを用いた場合の注視判定処理時間を計測し，ビデオレート内での注視判定可能な対象数を算出する．今回用いた PC は Pentium4 3[GHz] でカメラのフレームレートは 30[fps] となっている．図 3.7 に示すように，球を用いた場合は約 17000 個の球に関して注視判定を行うことが可能である．これに対して，三角ポリゴンを用いた場合は約 24000 個のポリゴンに対して注視判定を行うことが可能である．図は注視判定のみに必要な時間となっている．

本研究では詳細なモデルを作成する際には三角ポリゴンの利用を提案したが，球を用いた場合にも，複数の球の集合として詳細なモデルを作成することが可能である [49][50]．また，詳細なモデルを用いた場合に，処理速度を向上させる手法としては multi-resolution を用いた手法が有効であると思われる．P.Hubbard らは球を用いた階層的なモデル提案し，高速な衝突判定を行っている [51]．本研究



においても、より複雑なモデルに対する注視判定を実現する場合は、Hubbard らが提案したような階層的なモデルが有効であると思われる。

注視対象の情報は、ユーザの意図を知る手がかりとして様々な分野で非常に重要とされている。そこで次節では、提案した注視判定手法を用いて応用システムを提案する。

### 3.3. 応用1：ドライバモニタリング

交通事故総合分析センター (ITARDA) が報告する交通事故発生件数・死者数・負傷者数の統計データによると、平成5年以降、交通事故における死者数は減少傾向にある。これは車両の衝突安全性に対する取り組みの成果であると思われる。しかしながら、その一方で交通事故の発生件数・負傷者数は共に増加傾向にあることが問題であると指摘されている。このような背景から国土交通省の主導によって、高度道路運転支援 (以下 ITS) や先進安全自動車 (以下 ASV) のプロジェクトが進められている。

事故原因のほとんどは、ヒューマンエラーであるとの報告もあり、ASV の取り組みには、事故予防としてカメラによるドライバ計測も盛り込まれている。ドライバのモニタリングには様々な制約があるが、負担の少ない非接触な計測が重要な課題となっている。ドライブレコーダなどの実用化も行われており、ドライバの様子を常時記録可能なシステムの需要は非常に高いと考えられる。本節では非接触なドライバモニタリングシステムを提案し、有効性の検証を行う。

まずはじめに3.3.1 節で、実験車両のシステム構成について述べる。次に3.3.2 節では車線変更中、3.3.3 節ではカーブ走行中の注視計測について述べる。

#### 3.3.1 実験システムの構成

ドライバのモニタリングを行うため、図 3.8 に示すように実験車両を構築した。車両のダッシュボード上には (b) 顔情報計測用のステレオカメラを設置し、ドライバの頭部位置、姿勢及び視線が計測可能となっている。また計測した注視点を重畳表示するために、環境計測用のカメラとしてボンネット上に (c) 全方位カメ

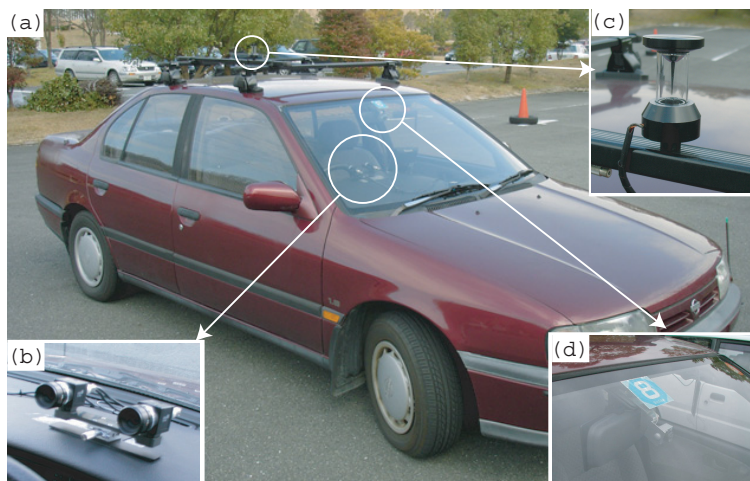


図 3.8 実験車両：(a) 外観，(b) ステレオカメラ，(c) 全方位カメラ，(d) 前方カメラ

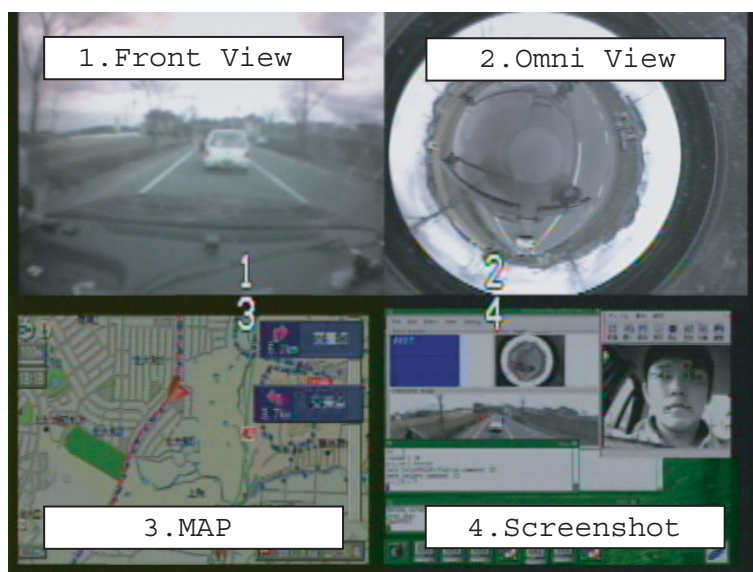


図 3.9 記録画像：(1) 前方カメラ画像，(2) 全方位画像，(3) カーナビ画面，(4) 処理画面

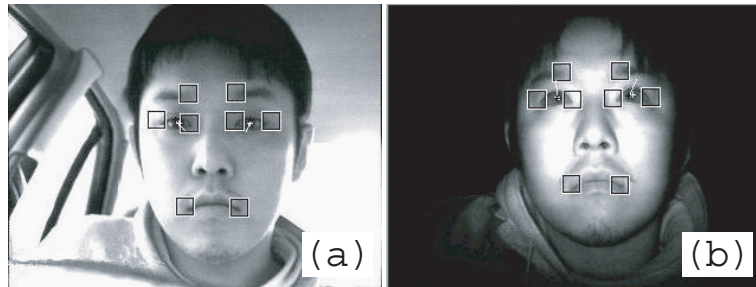


図 3.10 トラッキング結果：(a) 昼間，(b) 夜間

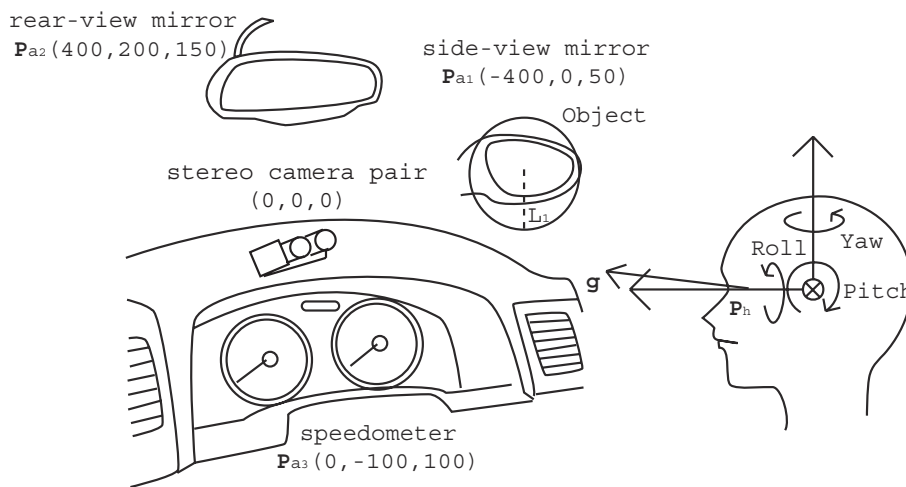


図 3.11 車内の注視対象の位置関係

ラ，ルームミラーの裏側には (d) 前方カメラを設置した．ドライバの注視行動と車両運動の関係を計測するために，ジャイロ스코プ，加速度計及び GPS を搭載する．これらの各センサはコンピュータに接続され記録されている．また，オフラインでの解析も可能にするためハードディスク及び DV レコーダを用いて映像を記録する．記録される映像は図 3.9 に示す通りであり，(1) 前方カメラ画像，(2) 全方位画像，(3) カーナビ画面，(4) 処理画面となっている．

ドライバモニタリングでは昼夜を問わず注視状況を計測する必要がある．そこで本研究では，近赤外をダッシュボード上からドライバの顔方向に照射することで夜間の計測を実現した．昼間と夜間の計測の様子は図 3.10 に示す．

車内の対象 (ミラー，スピードメータ等) の注視計測を行うにあたって，3.2 節

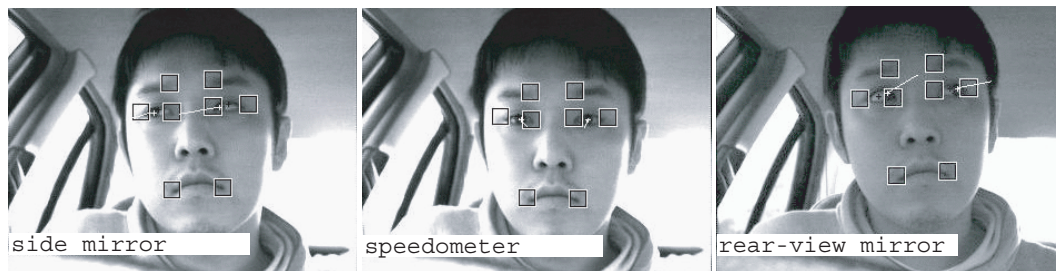


図 3.12 車内での注視対象の推定結果：(左) サイドミラー，(中) スピードメータ，(右) ルームミラー

で述べた注視判定手法を用いる．車内の対象はすべて球でモデル化を行った．本研究で設定した注視対象は図 3.11 に示すとおりであり，ルームミラー，サイドミラー及びスピードメータへの注視を判定することが可能である．ステレオカメラの設置場所をワールド座標の原点として，各注視対象の位置は設定されている．各ドライバーの身長(頭部位置)に応じてステレオカメラの角度を変更する必要があるが，スイベルステージを用いて調整を行うことで，対象物の位置はワールド座標系からカメラの角度に応じてカメラ座標系に変換し，注視判定を行う．実際に注視判定を行った結果を図 3.12 に示す．図はルームミラー，サイドミラー，及びスピードメータを注視している様子であり，推定結果は画像の左下に表示されている．

### 3.3.2 車線変更時の注視行動計測

車線変更中の注視行動の特徴として，Salvucci らは車線変更前及び車線変更中においてミラーや変更先のレーンへの注視頻度が変化することをシミュレーションにおける実験から報告している [17]．

本実験では，実車において車線変更中のドライバーの頭部ヨー方向の動き及び注視対象の測定を行った．図 3.13 は車線変更中のドライバーの様子と注視点を表示したものである．前方にバスが停車していたため，車線変更を行った様子であり，図 3.13(6) ではバスを注視している様子を確認することができる．また，図 3.14 は左から右への車線変更中の頭部のヨー方向の角度であり，計測結果にローパス

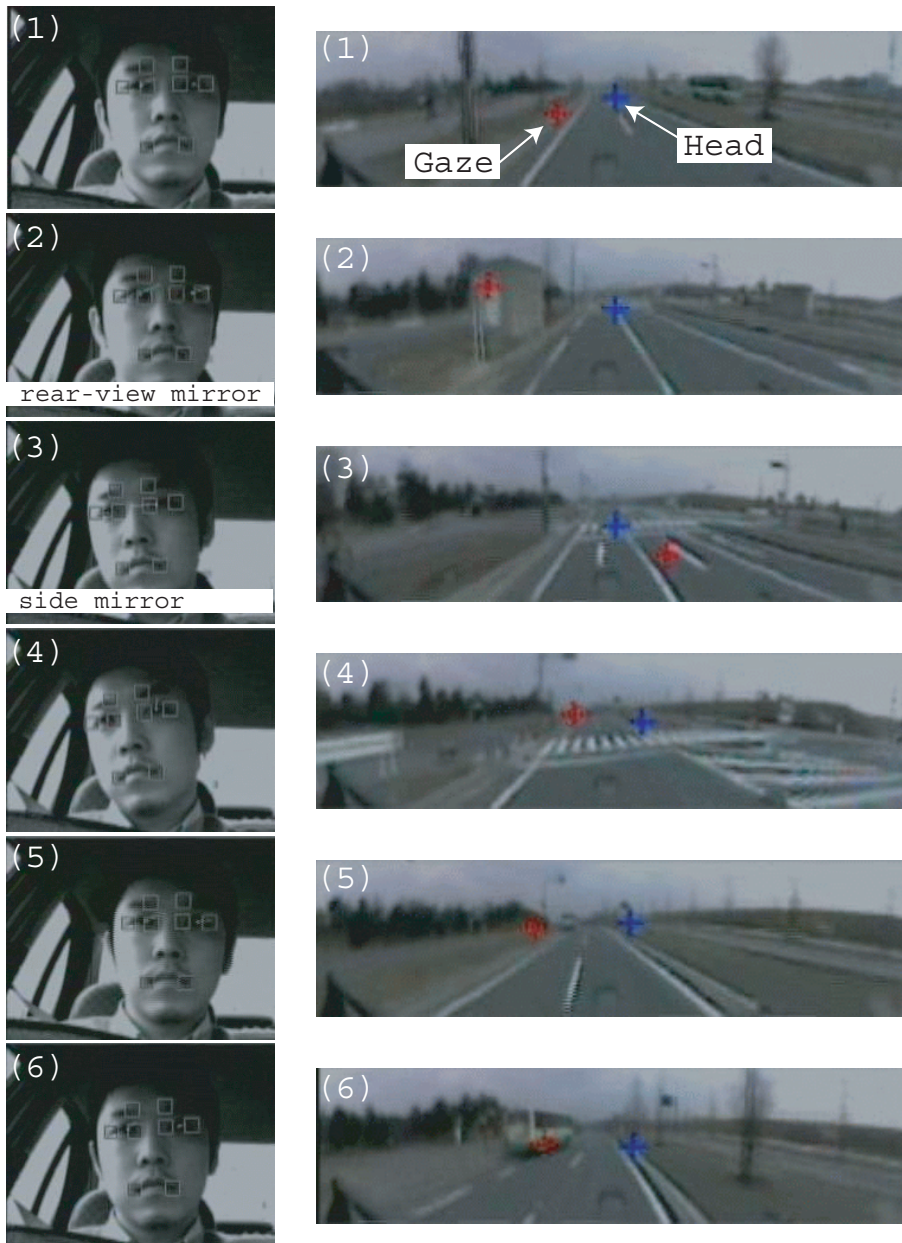


図 3.13 車線変更中のドライバーの様子と注視点

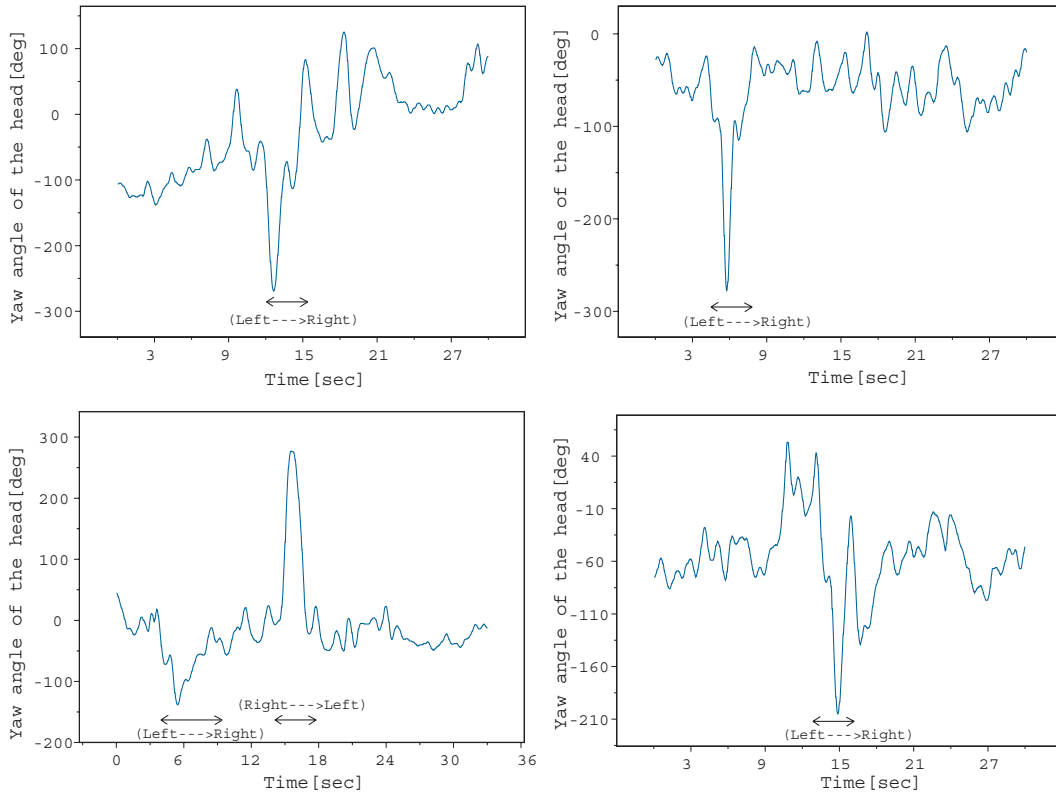


図 3.14 車線変更時の頭部運動

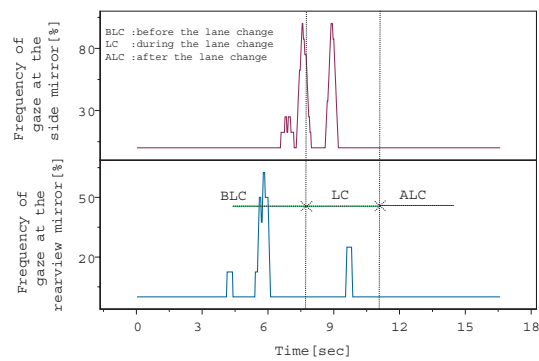


図 3.15 サイドミラー，及びルームミラーへの注視頻度

フィルタを適用したものである。図は4回の計測結果であり、図中の矢印は車両のヨーレイトから推定された車線変更の開始から終了まで区間である。本実験においても車線変更時には変更車線の方向へ大きな頭部運動が計測されており、Salvucciらの報告と同様の結果が得られた。

また、ルームミラー及びサイドミラーに対する注視頻度にも注目し、算出した結果を図3.15に示す。注視頻度は過去8フレーム(約264[msec])中の注視フレーム数の割合で算出している。Sperlingは部分報告法によってアイコニックメモリーの存在を確認し、250[msec]残存することを示した[52]。本研究では視覚情報が一時的に貯蔵されるアイコニックメモリーの残存時間に合わせて設定を行った。Salvucciの研究報告では、ミラーへの注視は車線変更直前から頻度が上がることが報告されている。本研究では同様の結果を実車で確認したことに加えて、サイドミラーとルームミラーを区別し注視頻度の変化を示した。

### 3.3.3 カーブ走行時の注視行動計測

カーブ走行中の注視行動はLandらによって、ステアリング角度と視線方向の相関関係やtangent point付近を注視する特性が報告されている[14]。本実験では大学構内の周回コースを走行し、車両のヨーレイト及び頭部・視線運動の計測を行った。実験は普段から運転を行っている被験者4名(A,B,C,D)に対して行われた。図3.16は運転中のドライバの様子である。頭部のヨー方向の角度及び車両のヨーレイトの関係は図3.17に示すような結果であり、頭部運動と車両のヨーレイトの関係は非常に相関が高いことが確認できる(相関値はそれぞれ0.91, 0.86, 0.85, 0.92)。また、視線と車両のヨーレイトの関係についても、図3.18に示すように同様に相関が高い傾向にあることが確認できる。Landらによって頭部及び視線運動は共に車両より若干先行して運動が生じることが報告されているが、これに関しても本実験の結果から同様に確認することができた。これらの結果はLandらによって報告されているステアリング操作の際の注視特性と同様の結果であると言える。また、注視頻度マップ(付録C)をカーブ走行の際に生成した結果を図3.19, 3.20に示す。図のように注視頻度マップを用いることで、複雑に動く注視情報を分かりやすく可視化することが可能であり、tangent point付近の注



図 3.16 カーブ走行中のドライバの様子

視頻度が高くなっていることから，Land らの報告する知見と一致していることが確認できた．

### 3.3.4 考察

本研究では 2 種類の走行実験（車線変更及びカーブ走行）を行った．車線変更実験の結果は Salvucci らがシミュレーションにより得ている知見と，またカーブ走行における実験結果は Land らが報告している知見と一致したものが得られた．本研究によって得られた具体的な成果は以下の通りである．

- 1: 実車における非接触な計測
- 2: ドライバの注視対象の推定
- 3: 先行研究で報告されている重要な知見の計測

走行実験を通して，先行研究によって報告されているドライバの特徴的な注視行動の計測が可能であることが確認できた．ドライバのモニタリングシステムの研究では，非接触で計測することに焦点が当てられてきたが，本研究ではドライ



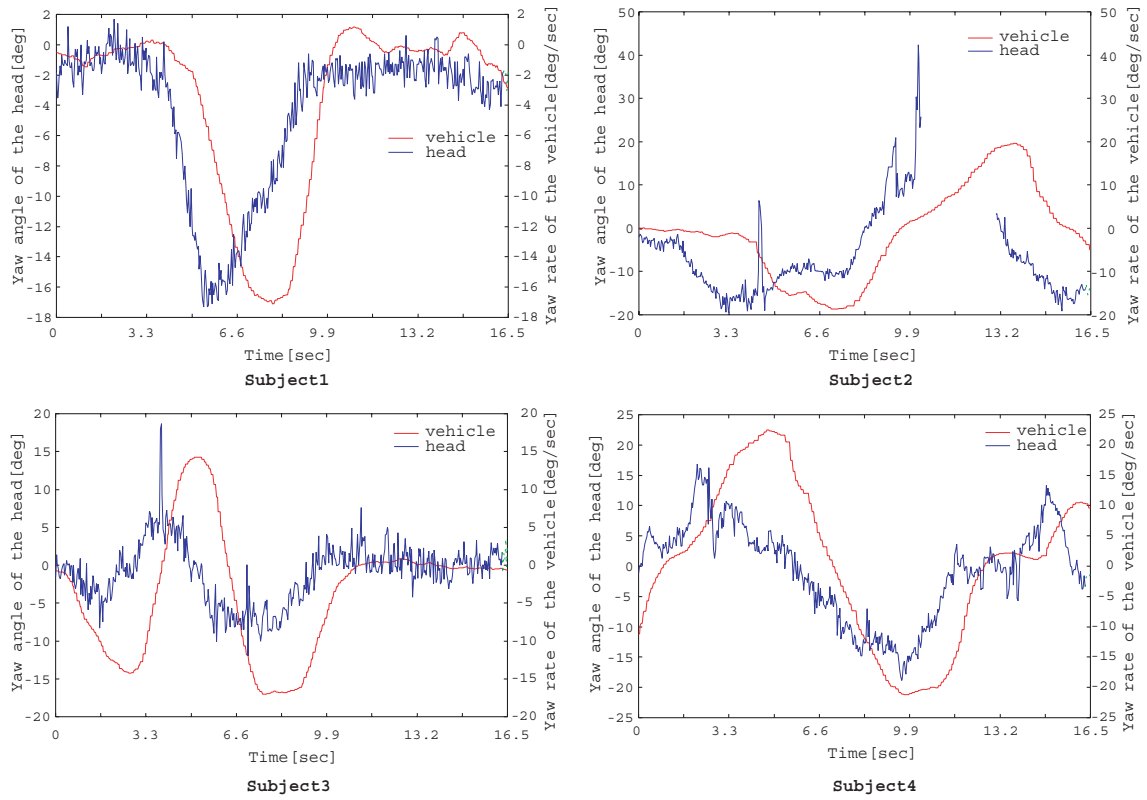


図 3.17 頭部方向と車両のヨーレイトの関係

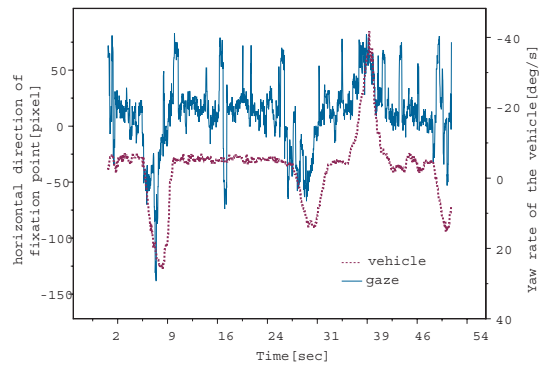


図 3.18 視線の水平方向の動きと車両のヨーレイトの関係

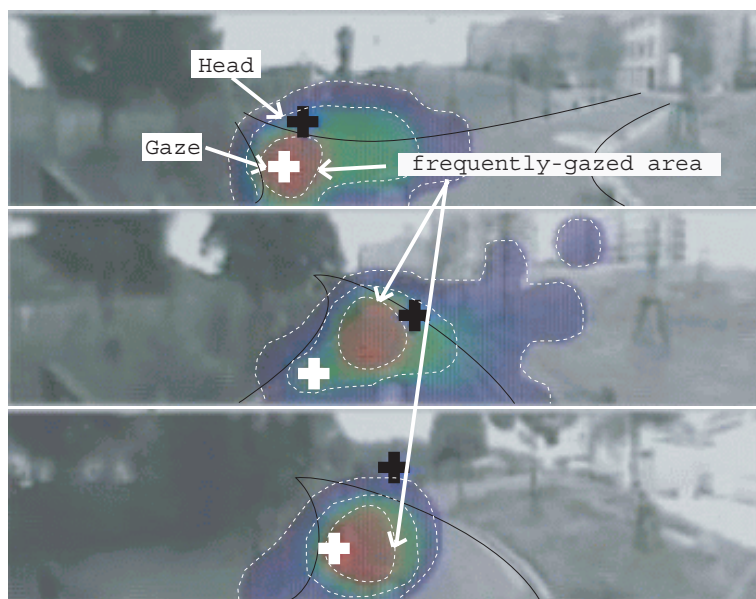


図 3.19 左カーブ走行中の注視頻度マップ

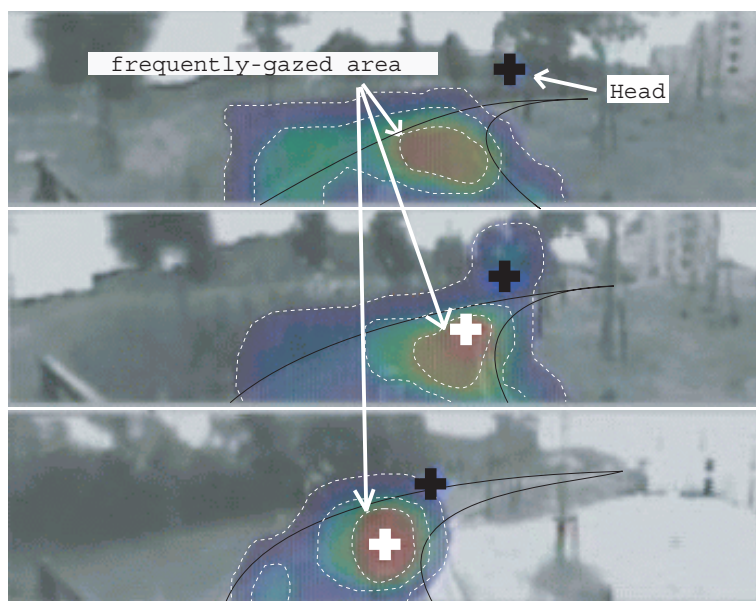


図 3.20 右カーブ走行中の注視頻度マップ

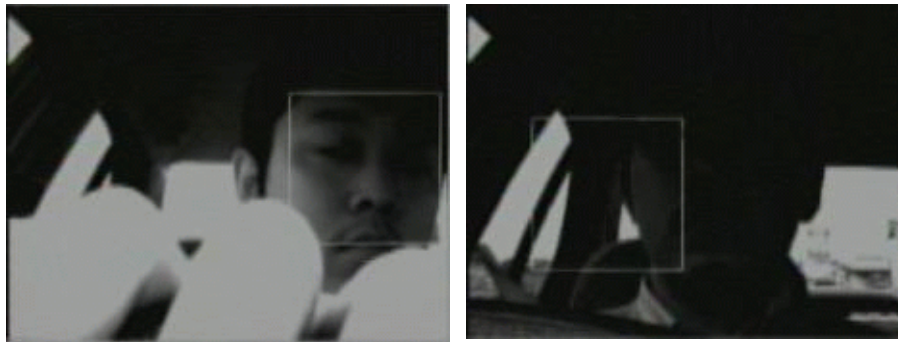


図 3.21 トラッキングが失敗したケース

バの特徴的な注視行動を計測することが重要であることを示した。重要な知見と同様の結果を計測可能なシステムの開発によって、運転支援への応用が可能であると考えられる。モニタリングシステムとしては計測成功率が重要であるが、曇天の15時から16時にかけて行った約1時間の走行実験では、走行時間の約92%で計測が可能であった。この成功率はシステムとして十分に有効な数値である。

走行実験においてトラッキングが外れることが確認されたが、外れやすい状況を特定することが出来たのでここでまとめる。代表的な例の1つ目は交差点である。図3.21の左図のように交差点での頭部運動が大きいことは予想されていたが、本実験で特にはずれる傾向があったのは交差点での左折の場合である。一般に右折より左折の方が頭部運動の動作が大きく、トラッキングの可能範囲からはずれやすいことが確認された。また左折の場合、ハンドル操作が大きく、手が画像中に写ることがしばしば確認された。カメラの設置場所はダッシュボード上であるが、検討する必要があることが分かった。2つ目は図3.21の右図に示すような日陰の問題である。トラッキングは一度はずれても復帰が可能であるが、日陰や逆光に入った際に本システムが対応できる範囲からはずれてしまうことが何度か確認された。そこで、日陰や逆光等の影響に対応するため輝度値に応じて、シャッタースピードの変更や近赤外線等を自動で照射するようなシステムも必要となることが分かった。

### 3.4. 応用2：注視による操作対象の切り替え

注視情報を用いた応用システムの2つ目として、本節では注視による操作対象の切り替えを提案する。近年ユーザを取り巻くコンピュータ環境は変化し、1人で複数台のコンピュータを用途に応じて、利用するケースが増えてきている。ノートPCとデスクトップの利用や、用途に応じて異なるOSを利用するユーザが増えてきていることが背景にある。また、家電もネットワークに接続し、ユーザに利益のある情報等を提供している。このような環境において問題となるのがそれぞれの機器に専用のインタフェースが存在することである。これにより机の上等は”Mouse jungle”と呼ばれるようなケーブルの絡まった状況となってしまう。家電に対しても複数のリモコンが存在することで、どのリモコンがどの機器のものであるのか瞬時に判断するのが困難な場合がある。

同様の機能を有するインタフェースが複数存在することは好ましくないため、一般的にPCの切り替えにはキーボード切り替え器、AV機器では学習リモコンなどにより、この問題への対応が行われている。また、研究段階ではあるがこのような問題に対していくつか新しい手法が提案されている。操作切り替えに関する代表的な研究としては綾塚らによって提案されたGaze-Link[53]がある。これはコンピュータに搭載されたカメラによってCybercodeを認識し、コンピュータが操作対象に接続するインタフェースを提案している。また、HITACHIのAir-realはリモコンをに向けた対象に接続するインタフェースであり、テレビやビデオ間のやり取りなどで直感的で分かりやすいインタフェースを実現している[54]。

本研究では、注視情報に基づいた操作対象を切り替え手法の提案を行う。1セットのキーボードやマウスで複数のコンピュータの操作を行うことやマウスのクリックやスクロールでテレビ、オーディオ等の機器の電源や音量等を制御することが可能となる。ユーザは操作する機器を注視することで、直感的かつシームレスに操作対象を切り替えることが可能である。

3.4.1節では、コンピュータの操作切り替えについて、3.4.4節では家電の操作切り替えについての実装及び評価の詳細を述べる。

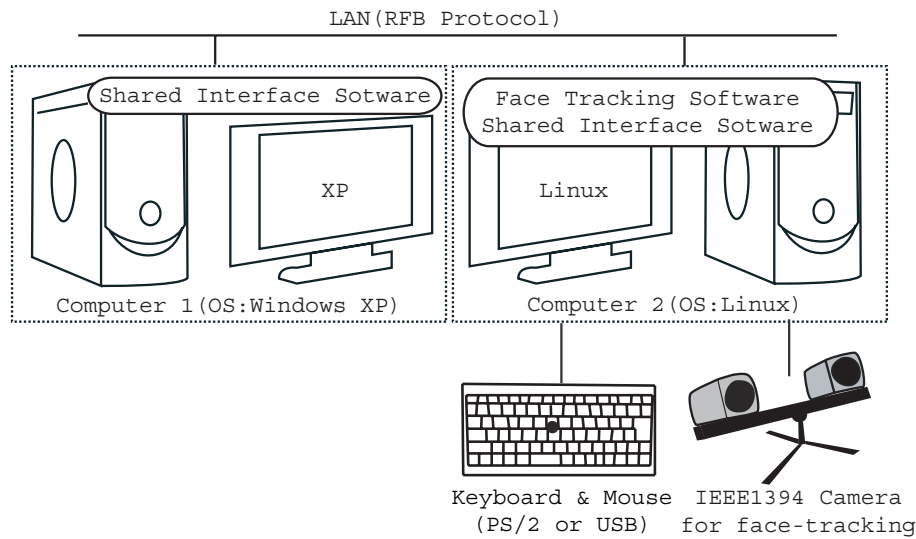


図 3.22 複数コンピュータの操作切り替えシステム

### 3.4.1 注視による複数コンピュータ操作の切り替え

複数のコンピュータがあるような環境において、注視に基づく操作対象の切り替えを実現する。本研究では複数の OS をサポートするために、Virtual Computing に用いられる Remote Frame Buffer(以下 RFB)/Virtual Network Computing(以下 VNC) Protocol を用いてインタフェースの情報を転送し、共有及び切り替えを実現する。システム構成は図 3.22 に示す通りであり、Linux にインタフェース(キーボード・マウス等)及び顔情報計測システムのカメラを接続する。Linux に接続されたキーボードやマウスの情報を RFB 経由で転送することで Windows や Mac 等のコンピュータを制御することが可能となっている。顔情報計測システム(付録 A)によって頭部方向の計測を行い、3.2 節で述べた注視判定手法を用いて、注視対象の推定を行った。これによって図 3.23 のように異なるコンピュータ間において、シームレスな操作切り替えを実現することが出来た。図は 2 つのコンピュータに対して交互にマウスで文字を描いている様子である。提案手法はマウスのみでなく、コンピュータの入力デバイスとして認識可能なものであれば、同様に共有を実現することができる。

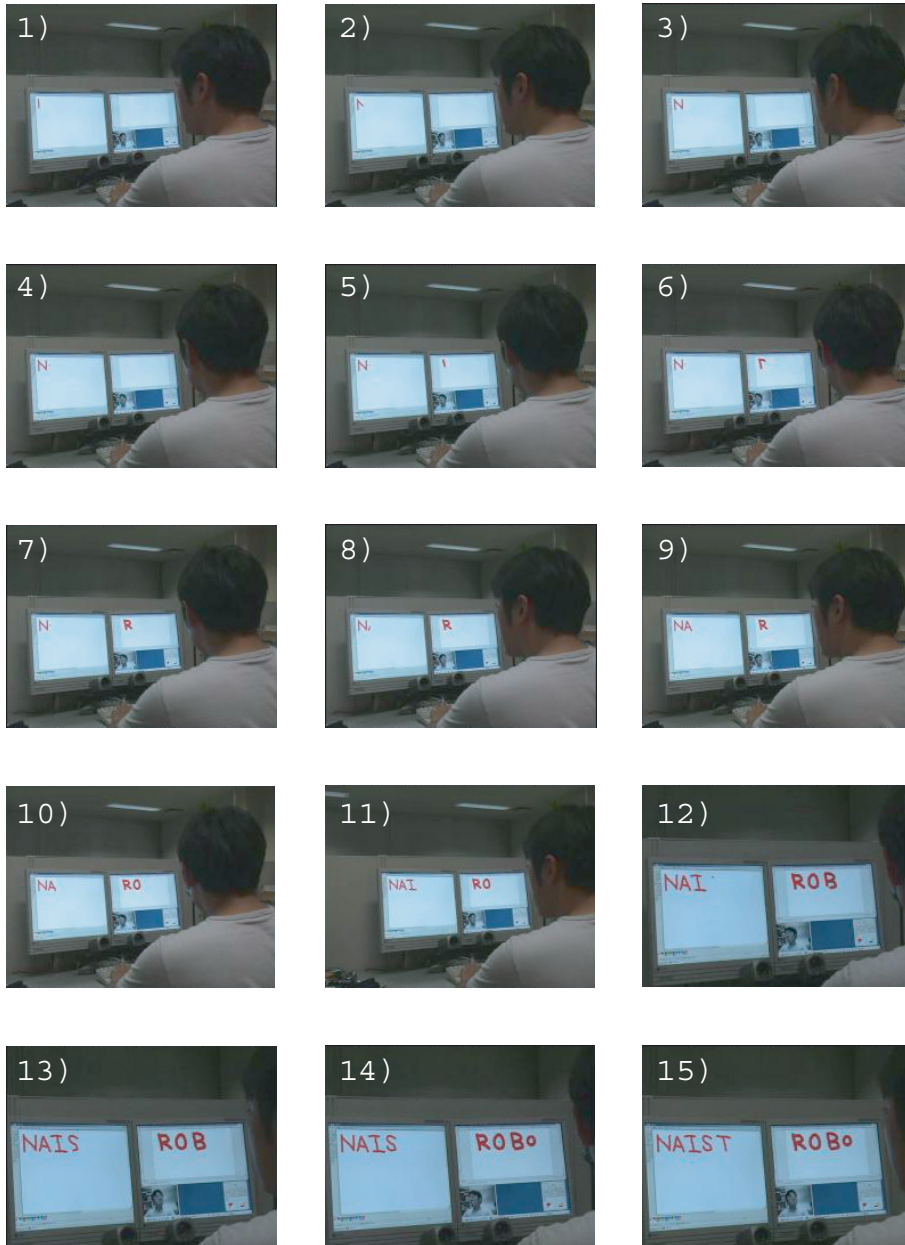


図 3.23 注視によるコンピュータ切り替えの様子

### 3.4.2 評価実験

提案する注視によるコンピュータの操作切り替えについて評価を行う。評価手法は、2台のコンピュータに対して被験者が交互に一定回数の単語入力を行い、その際の所要時間を計測する。被験者は2つの方法で操作(入力)対象の切り替えを行う。操作対象の切り替え方法は以下の通りである。

- 1: 注視情報を利用したPCの切り替え
- 2: マウスの移動によるPCの切り替え

タスク1は提案手法である注視による操作対象切り替えを利用した場合であり、タスク2はマウスの移動によって操作対象を切り替えるもので、マウスカーソルのあるコンピュータを操作対象とするフリーのツールを用いた場合である。ディスプレイの範囲を超えてマウスカーソルを移動させた場合に他のコンピュータ上にカーソルが移り、操作を切り替えることができる。

実験は、23歳から30歳までの男性4名(被験者A, B, C, D)に対して行われた。被験者は日頃からコンピュータを使用していて、タッチタイピングを行うことが可能であるものとする。被験者には実験前にタスク1及びタスク2の操作方法について数分間の習熟する時間を与え、実験の順番により入力速度に影響がないよう配慮した。1セットのキーボード及びマウスを用いて、指定された単語を2台のコンピュータに対して、交互に10ワードずつ入力を行う。各被験者はタスク1及びタスク2に対してそれぞれ5回ずつ試行を行い、所要時間の計測を行った。各被験者のタスク1及びタスク2の所要時間から平均値及び標準偏差を求めたところ、その結果は図3.24のようになった。

### 3.4.3 考察

実験結果によると被験者ごとのタスクの達成時間には個人差があるが、タスク1及びタスク2の所要時間はすべての被験者において注視によるコンピュータの切り替え(タスク1)の方が、マウスの移動によってコンピュータを切り替える場合(タスク2)より所要時間が短いことが確認できる。そこで提案手法(タスク1)

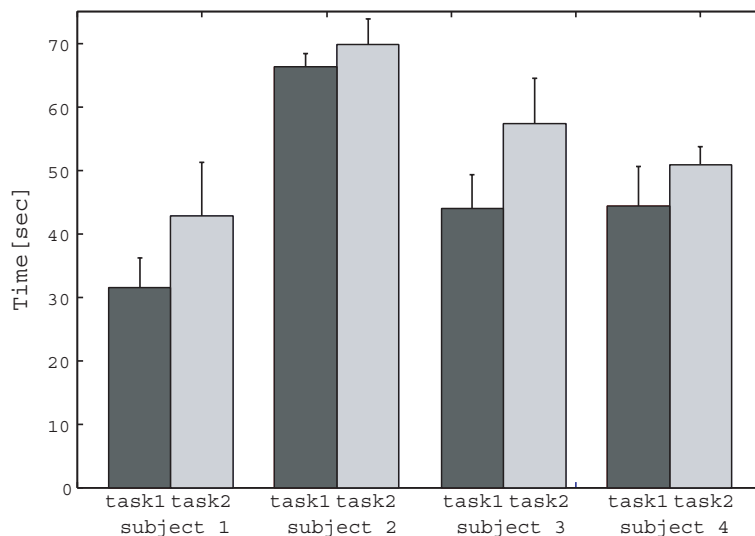


図 3.24 タスク完了平均時間 (PC 操作)

と従来手法 (タスク 2) に有意な差があるか t 検定を行ったところ, 被験者 1~3 は有意であった ( $p < 0.05$ )。また, 被験者 4 においても有意傾向であることが確認できた ( $p < 0.1$ )。検定の結果, 提案手法は従来手法よりもタスク完了時間を短縮することが可能であり, その差は有意であることが確認できた。これより注視を用いた操作対象の切り替えが有効であり, 作業効率を向上させることが確認できた。

実験を通して問題点を確認されたので以下にまとめる。まず本提案手法を利用することが出来ないユーザがいることが確認された。有効に働かなかった理由を, 提案手法を利用後にインタビューによって確認したところ, タッチタイピングが出来ないことが理由であった。注視をしているときのみ入力が有効であるように設定されていたため, ディスプレイを注視した状態で入力の出来ないユーザに対する配慮が行われていなかった。タッチタイピングできないユーザにも利用が可能なように, 最後に注視した対象を入力対象として保持する機能を付け加えた。これによって, 多数のユーザに利用してもらうことが可能となった。

また, 入力対象とは異なるものを注視しながら入力する場面があることが確認された。例を挙げると, 2つのディスプレイ使用時には入力する画面ではなく他



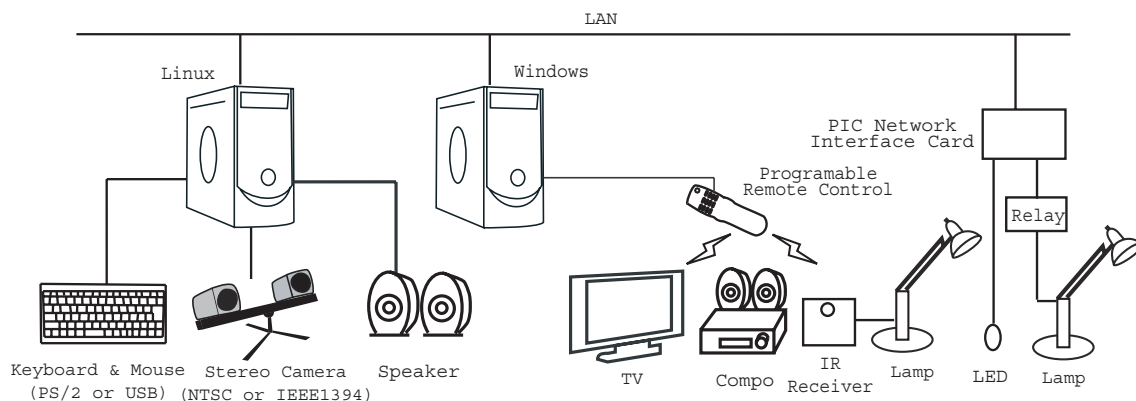


図 3.25 家電操作切り替えシステム

方の画面の情報を見ながら作業を行う場合がある．このような場合は，提案手法のみでは対処することが困難であると考えられる．この問題への対応は今後の大きな課題である．

### 3.4.4 注視による家電操作の切り替え

3.4.1 節では，注視情報を用いたコンピュータの操作切り替えを提案し，評価実験を通してその有効性を示した．本節では，同様の手法を家電に応用し注視を用いた円滑な操作を実現する．

近年，ネットワークインタフェースを有する家電が登場し，ネットワークを介したコントロールやサービスが期待されている．しかしながら，その制御用 API 等の提供は行われてはいないため，本研究では仮想的にネットワークコントロールが可能な家電システムを構築する．本システムの構成は図 3.25 に示す通りであり，ライトなど制御が単純なものに関しては，ネットワークインタフェースを有するマイコンボードを用いて外部からの制御を可能にした．また TV やオーディオシステムに関しては，プログラマブルな学習リモコンを用いてネットワークからの制御を可能にしている．

本研究では，機器の操作インタフェースとしてマウスを用いる．注視対象がコンピュータ (ディスプレイ) の場合は通常のインタフェースとして働くが，その他

表 3.1 入力インタフェースと制御対象

入力インタフェース	操作対象	操作機器
マウスホイール	音量	テレビ, オーディオ
マウスボタン	電源 On/Off	PC 以外の機器

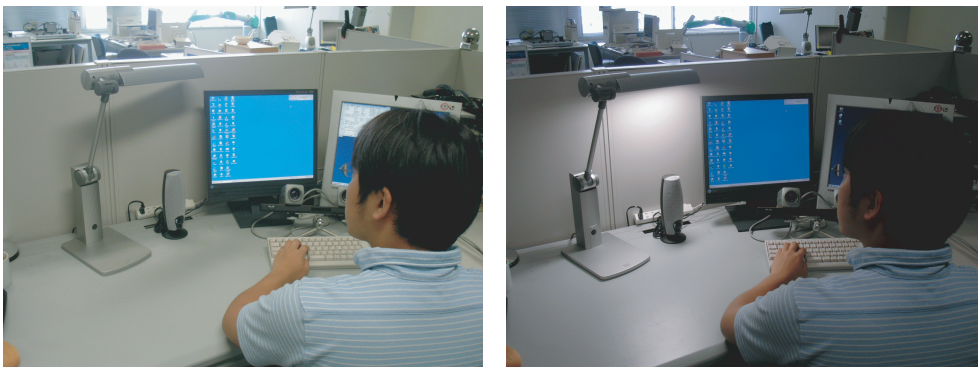


図 3.26 ライトの点灯の様子：(左) 点灯前，(右) 点灯後

の機器に関しては ON/OFF やボリューム等の制御がマウスによって可能である。本研究における機器の動作とインタフェースの対応は表 3.1 に示すようになっている。図 3.26 のようにライトを注視し，マウスをクリックすることで点灯することが可能である。

### 3.4.5 評価実験

提案する注視による家電の操作切り替えについて評価を行う。評価実験は，ユーザの前方に配置された 4 台のライトを，指示通りに点灯/消灯させるというものである。注視対象であるライトは球によってモデル化を行った。各タスクにおいて，ライト点灯/消灯の操作を 10 回行う。タスクの内容は以下に示す通りである。

- 1: 各リモコンによる操作
- 2: 学習リモコンによる操作
- 3: 視線による操作対象の切り替え

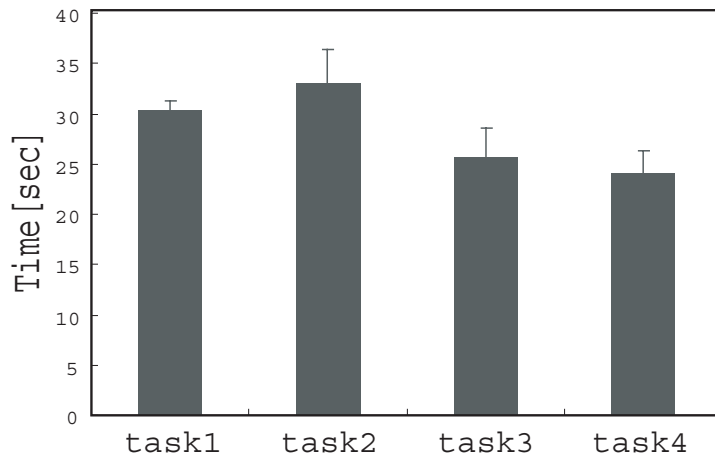


図 3.27 タスク完了平均時間 (家電操作)

#### 4: 視線による操作対象の切り替え (LED 点灯有り)

タスク 1 はそれぞれのライトに割り当てられたリモコンを使い分け、モニターに表示された指定のライトの点灯/消灯を行う。タスク 2 は学習リモコンを用いて操作対象を切り替え、ライトの点灯/消灯を行う。指定したライトを操作するためにはダイヤルを回す必要がある。タスク 3 は視線によって操作対象を切り替え、マウスの入力によって点灯/消灯を行う。タスク 4 は手法に関してはタスク 3 と同様であるが、注視対象として認識された機器に関しては、機器に取り付けられた LED が点灯する仕組みになっている。実験は、23 歳から 30 歳までの男性 5 名 (被験者 A, B, C, D, E) によって行われた。被験者は各タスクにおいて、モニターに出されるランダムな操作指令に応じてライトの操作を行う。この動作をそれぞれのタスクで 10 回ずつ行い所要時間を計測する。

タスク達成の平均時間及び標準偏差を算出したところ、その結果は図 3.27 に示すようになった。提案手法である視線と LED を併用する手法 (タスク 4) が最も操作時間が短く、続いて視線による操作切り替え (タスク 3)、各リモコンによる操作 (タスク 1)、学習リモコンによる操作 (タスク 2) という結果になった。

### 3.4.6 考察

提案手法(タスク3, 4)が従来手法(タスク1, 2)に比べて有意な差があるか検定を行った。タスク1とタスク3及びタスク2とタスク3においてt検定を行ったところ、有意であることが確認できた( $p < 0.01$ )。また、LEDによる注視判定の表示有(タスク3)と表示無(タスク4)でt検定を行ったところ、平均は等しいという帰無仮説を棄却することはできなかった。

これによって提案手法である注視による家電操作の切り替えは、各リモコンや学習リモコンを用いる場合に比べて操作の所要時間が短く、家電操作の切り替えに対しても有効であることが確認できた。システムの認識状況をユーザに表示することが重要であると考え、注視による操作切り替えにおいて、LED点灯の有無(2通り)の実験を行った。認識ミスや操作ミスを減少させることにつながると考えたが、検定結果ではタスク3とタスク4には有意な差はなかった。

## 3.5. 本章の結論

本章では、注視判定手法及び注視情報に基づく応用システムの提案を行った。実空間における対象を球及び三角ポリゴンでモデル化し、視線及び頭部方向を用いて注視判定を実現した。また、これらの提案手法を「ドライバモニタリング」及び「注視による操作対象の切り替え」に応用し、システムの有効性の評価を行った。ドライバモニタリングでは車線変更及びカーブ走行における重要な知見を非接触なシステムで計測を実現した。従来のモニタリングシステムでは、先行研究によって報告されている重要な知見の計測は行われていなかったが、本研究ではモニタリングシステムとして、ドライバの特徴的な注視行動を計測することが重要であることを示した。

また、コンピュータ及び家電の注視による操作切り替えでは、従来用いられてきた切り替え手法よりも操作時間を短縮することが可能であり、有効性を確認することができた。

本研究では、注視対象の情報に基づいたユーザ支援システムの提案を行った。ユーザの意図を把握し支援する手法が様々研究されているが、手がかりとして注

視対象の情報が重要であることを示すに十分な結果であったと思われる。注視対象の情報を用いてドライバモニタリング及び注視による操作対象の切り替えを提案したが、その他にも様々な応用が期待できる。

提案した注視判定手法において、一番の大きな課題は対象のモデル化である。球及び三角ポリゴンで対象のモデルを作成したが、詳細なモデルを作成する必要がある他、配置等に変更があった際にも柔軟に対応する必要がある。この問題を改善し、より高度な HCI を実現したいと考えている。

# 第4章 複数人の注視情報を利用した 注意推定

## 4.1. 本章の趣旨

本章では、複数人の注視情報を用いた「場の注意」の推定手法について述べる。初めに「場の注意」の推定アルゴリズムに関して詳細を述べる。提案する注意推定手法は、複数人の頭部方向を用いて注意の集中する箇所(場の注意)の推定を行うものである。モーションキャプチャを用いて計測された頭部方向のデータを用いて場の注意の推定を行い、動的な対象を注意した場合や注意の共有が成立していない場合に対する有効性を示す。

次に実際の対話シーンにおいて「場の注意」の計測を行いシステムの有効性を検証する。評価に適したシーンとして発話の順番が限定される「しりとり」を行い、その際の発話状況と場の注意を比較検討する。またディベート中の注意推定を行い、複数の注意点が存在する状況で推定を行う。

また、提案手法の応用として映像の自動編集を行う。対話中のシーンを記録した全方位画像を「場の注意」に基づき展開を行い、映像の自動編集への応用を検証する。

最後に複数の顔情報計測システムを用いたミーティングキャプチャシステムを提案する。非接触な計測システムを用いて「場の注意」の推定を行い、実現可能性及び問題点を示す。

## 4.2. 複数人の注意計測に基づく注意推定

「見る」という動作はコミュニケーションにおいて、非常に重要な役割を担っていることが心理学等の分野で報告されている [1]。最近では CMC, CSCW の分野への注視情報の利用が注目され, Gaze Groupware System [37] のように注視情報を利用し, コミュニケーションを円滑に行えるようにサポートするシステムの研究が行われている。また, 注意情報はカメラワークや映像の自動編集等に 응용が行われており, 会議映像の自動編集 [42] や球技のカメラワーク [43], 及び講義の映像選択 [44] への利用が提案されている。

注意推定は映像の自動編集やカメラワークだけでなく, 会議記録等においては自動要約など, 情報を自動的に整理分類する際に重要な要素となると思われる。今後, 映像記録の需要はますます増え, 様々な作業の自動化が求められると予想される。本研究はその手がかりの一つとして注意情報が非常に重要であると考えている。そのような理由から注意点の推定は社会的に意義が非常に高いといえる。

このような背景から複数人の注意が集中する箇所として「場の注意」の推定手法を提案する。注意推定の従来研究では人や対象の位置情報を利用し, 注意の最も集中する箇所 1 つを推定する手法がほとんどである。しかしながら, 注意対象や数を限定してしまうと, 新しい注意点に柔軟に対応することができない。そこで, 本研究ではこれらの問題点に注目し, 注意推定アルゴリズムを提案する。

### 4.2.1 頭部情報の利用

画像処理技術の向上により, 頭部や視線方向の計測技術も実用レベルに近づいてきている。本研究でも画像処理を利用したオンラインでの頭部, 視線方向の計測を目標としているが, 各人に視線計測システムを用意し, 座標系を統一して扱うことは容易ではない。そこで, 複数人の動きがまとめて取り扱いやすい頭部方向のデータを使用し「場の注意」を推定する。なお, 頭部方向を注視方向として取り扱うことの妥当性に関しては R.Stiefelhagen が頭部と視線の水平方向の向きに関しては 87% 一致するというデータを示している [39]。また, 垂直方向に関しては水平方向と同様に大きな視線移動 (眼球運動 + 頭部運動) に対しては, 頭部



図 4.1 高再帰性反射マーカ

運動が移動の 60～70%を占めることが報告されている [55] ことから、同様に妥当であると考えられる。

#### 4.2.2 頭方向ベクトルの定義

本研究では頭部方向の推定にモーションキャプチャシステム (VICON 及び MAC3D System) を用いる。図 4.1 に示すように、高再帰性反射マーカを装着し頭部の位置及び姿勢を計測する。頭部位置や方向の算出方法は計測システムに大きく依存するものである。本研究では図 4.2 に示すように、マーカ 3 点によって定義される平面上のベクトルを基準ベクトルとして定義した。

しかしながら、この計測方法ではマーカは頭部表面に装着しているため、実際の頭部方向とピッチ (上下) 方向の角度が異なる。そこで本研究ではピッチ方向の角度調整を行い、頭部方向ベクトルを定義する。ピッチ方向の調整は一定時間キャリブレーション用のマーカを注視した際の頭部位置・方向及びキャリブレーション用マーカの位置データを用いて行う。図 4.3 に示すようにキャリブレーション用のマーカを任意に動かし、被験者はそれを注視する。ここではマーカを注視するようにと指示をただけで、その他に細かな指示は行わない。この一定時間の頭部方向ベクトルとキャリブレーション用マーカのユークリッド距離に注目す



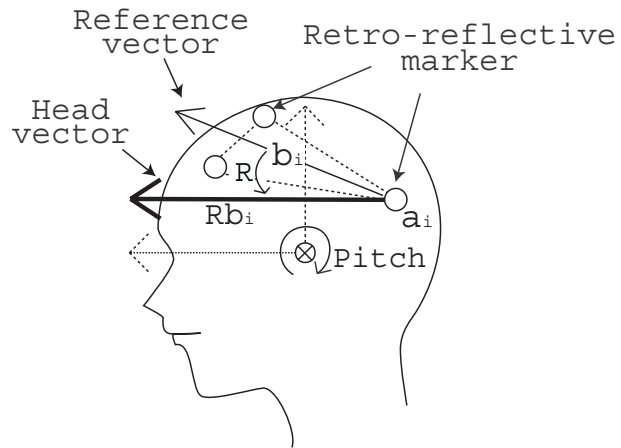


図 4.2 マーカと頭部方向ベクトルの関係

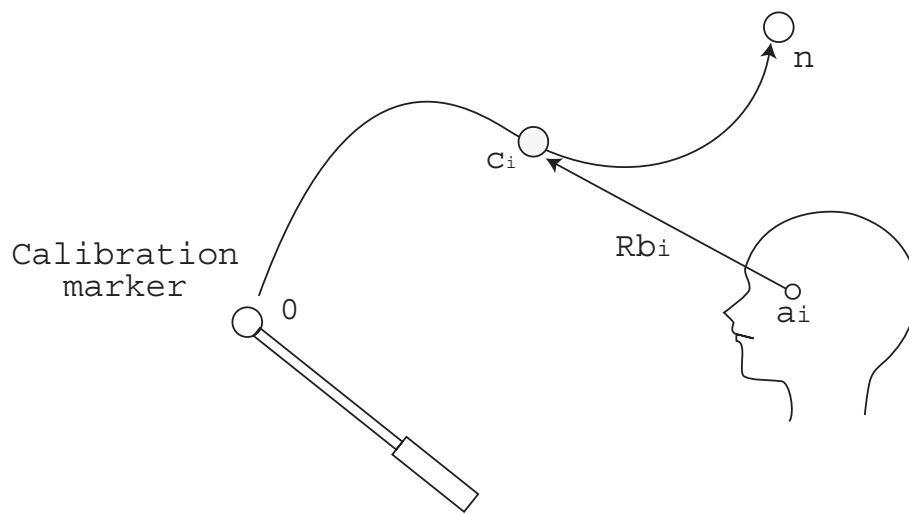


図 4.3 頭部方向のキャリブレーション

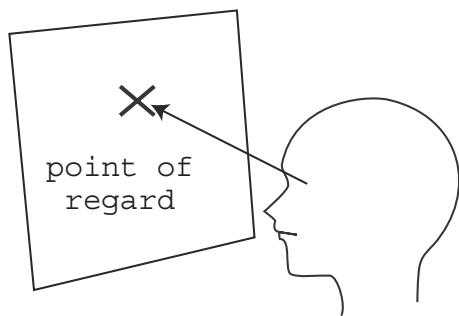


図 4.4 注視点の推定手法

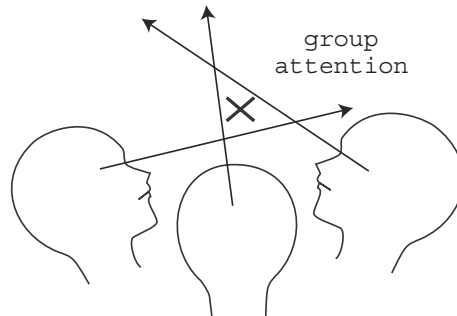


図 4.5 場の注意の推定手法

る．式 4.1 はピッチ方向の回転行列を求めるための誤差関数であり  $D$  を最小にするような回転行列  $R$  を求める．被験者すべてのキャリブレーションは同時に行うことが可能であるが，被験者ごとに評価が必要となる．

$$D = \sum_{i=0}^n \left\| \frac{(c_i - a_i) \cdot Rb_i}{\|Rb_i\|^2} Rb_i + a_i - c_i \right\| \quad (4.1)$$

誤差関数  $D$  はキャリブレーション中の各フレーム  $i$  の頭部方向ベクトルとキャリブレーション用マーカのユークリッド距離の和である． $c$  はキャリブレーション用マーカの位置， $Rb$  はピッチ方向の調整を行った頭部方向ベクトル，また  $a$  は頭部位置である．

### 4.2.3 場の注意の推定

従来の注視位置の推定は図 4.4 に示すように，個人の視線ベクトルとディスプレイ等の平面の交差を考えるものがほとんどである．本研究では図 4.5 に示すように視線ベクトルの交差を考えることで，空間中の注意点を推定するというものである．複数人の注視情報を用いることで，従来必要とされていた平面等の情報を利用せずに推定することが可能である．しかしながら，一般に複数の 3 次元ベクトルは交差しない．そこで，式 4.2 のように各人の頭部方向ベクトルから注意点までのユークリッド距離  $\varepsilon_i$  について考える．図 4.6 は  $\varepsilon_i$  について図示したもの

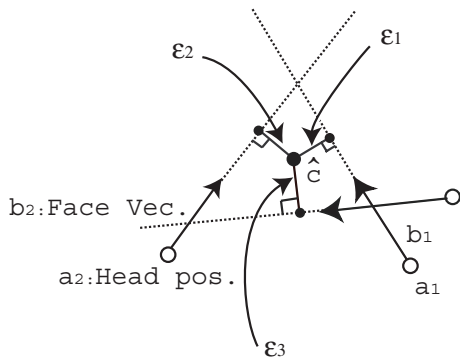


図 4.6 注意点と頭部方向の距離

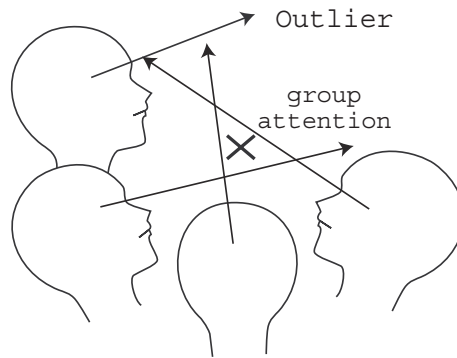


図 4.7 注意の共有が成立しない状況

である． $\varepsilon_i$  の和を最小にするような点  $c$  を注意点  $\hat{c}$  とする．注意点  $\hat{c}$  は式 4.3 に示すように，最小自乗法を用いて求めることが可能である．

$$\varepsilon_i = \left\| \frac{(c - a_i) \cdot b_i}{\|b_i\|^2} b_i + a_i - c \right\| \quad (4.2)$$

$$\hat{c} = \operatorname{argmin}_c \sum_i^n (\varepsilon_i)^2 \quad (4.3)$$

$a$  は頭部位置， $b$  は 4.2.2 節で述べたピッチ方向の調整を行った頭部方向ベクトルである．また， $\varepsilon_i$  は  $i$  番目の人の頭部方向ベクトルと注意点のユークリッド距離であり，最小自乗法により求めた点  $\hat{c}$  を注意点として定義する．

#### 4.2.4 ロバスト推定の利用

複数人が同一対象へ注視を行っている理想的な場合は 4.2.3 節で提案したように最小自乗法を用いて注意点を求めることができる．しかしながら提案手法で問題となるのが，図 4.7 のように注意の共有が成立しない第三者(異なる対象へ注視を行っている人)への対応である．実際のコミュニケーションシーンにおいては，全員が同一対象を常に注視するとは限らないため，この問題に対して考慮することは非常に重要である．4.2.3 節において述べた手法で推定を行うと，異なる対象を注視している第三者の注視方向ベクトルの影響が大きく，推定した注意点と実

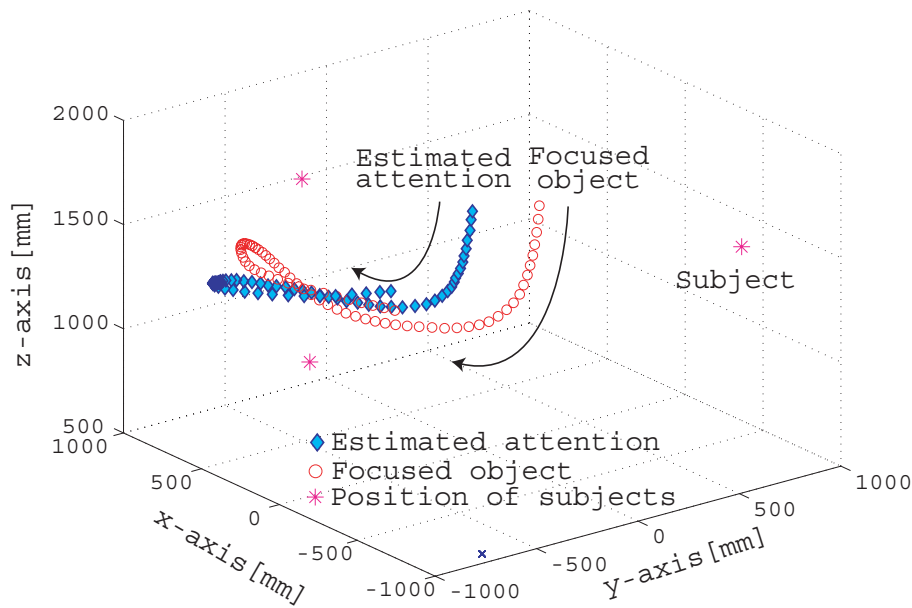


図 4.8 注意点の軌跡

際の注視対象物の位置誤差が大きくなってしまふ．そこで本研究では異なる対象を注視している第三者の影響を少なくするため M 推定 (付録 D) を用いる．最小自乗法はデータを均等な重みで取り扱うため，例外値に大きな影響を受けるが，M 推定では例外値は小さな重みになるように収束計算を行う．M 推定を用いると式 4.4 に示すように，重み付き最小自乗問題として注意点を求めることができる [56][57] ．

$$\hat{c} = \underset{c}{\operatorname{argmin}} \sum_i^n \left\{ \sqrt{\omega(\varepsilon_i^{(k-1)})} \cdot \varepsilon_i \right\}^2 \quad (4.4)$$

ここで  $\omega$  は重み関数であり，本研究では Welsch の重み関数を使用する． $\varepsilon_i$  は式 4.2 にて示したように  $i$  番目の人の頭部方向ベクトルと注意点のユークリッド距離である．また  $k$  は繰り返し回数であり， $k = 0$  の時， $\varepsilon_i^{(0)} = 0$  を初期値として，重み  $\omega(\varepsilon_i^{(k-1)})$  を更新し，注意点  $\hat{c}$  を求める．推定に必要な人数は，頭部方向ベクトルの交点を求めることから最低 2 人となるが，2 人の位置の直線上及びその付近は安定して推定することが出来ない．

表 4.1 頭部方向ベクトルの角度誤差

	水平誤差 [deg]	垂直誤差 [deg]
被験者 A	12.5	11.1
被験者 B	7.8	10.8
被験者 C	7.2	13.8
平均	9.2	11.9

3人の被験者に指定した対象を意識的に注視する実験を行い，対象の位置データと推定した注意点を比較することで精度評価を行った．3人の被験者は直径200[cm]程度の輪を描くように向かい合い，ほぼ等間隔で座っている．注視対象は高再帰性反射マーカを取り付け，人によってその輪の中を自由に動かし，位置を記録した．図 4.8 は3人の被験者の位置，推定した注意点及び注意対象の軌跡の一部を表示したものである．

頭部方向ベクトルの注視対象に対する角度誤差は表 4.1 に示すように，水平方向には平均約 9.2 度，垂直方向には平均約 11.9 度となった．この結果から水平及び垂直方向に大差はなく，同等に扱うことが妥当であることが確認できた．また，注意点と注視対象の位置誤差の平均は約 30[cm] であり，水平方向には約 25[cm]，垂直方向には約 15[cm] の誤差が確認された．本実験は意識的に対象を注視した実験であることから，誤差は視線の代わりに頭部方向を用いたことが要因で生じた誤差と考えられる．本研究では，カメラワーク等への応用を考えているが，30[cm]程度の誤差であれば，カメラの視野内に対象を捉えることが可能であり，頭部方向のみの使用でも，注意点は注視対象の運動の軌跡を十分に表現することが可能であることが確認できた．また，従来研究では定義された場所のみへの注意の推定が可能であったが，本手法によって注意対象の定義なしに，注意点の連続的な推定を実現した．

また，異なる対象を注視している人が存在する場合に対する頑健性を評価する．異なる対象を注視する第三者が存在する場合を想定してシミュレーションを行った．精度評価の実験データ（意識的に対象を注視する3人のデータ）に任意に設定したベクトルを加えて注意点を算出した．今回は任意に設定したベクトルが第三

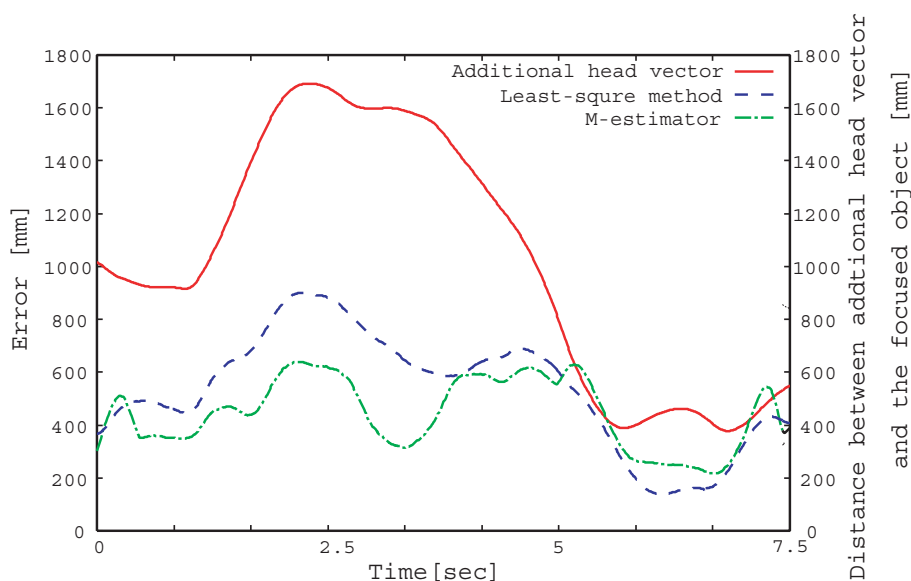


図 4.9 M 推定のシミュレーション結果

者 (Outlier) の役割をする。シミュレーションを行い、最小自乗法を用いた場合と M 推定を用いた場合の推定誤差は図 4.9 に示すようになった。最小自乗法の場合は任意のベクトルの影響を受け、誤差が大きくなっているが、M 推定を用いている場合は、任意のベクトルによる影響を軽減し、安定した結果を出していることがグラフから確認できる。

#### 4.2.5 複数注意点の推定

4.2.4 節で M 推定を用いて安定に、注意点を推定することが可能であることを示したが、ミーティング等の状況を想定した場合、参加人数が増えると「場の注意」が複数存在することが予想される。そこで、前節まで述べた場の注意推定アルゴリズムを用いて複数注意点の推定を行う。具体的には以下のような手順を用いて複数注意点の推定を行う。

##### 複数注意点推定アルゴリズム

- 1: 頭部位置・方向から最小自乗法を用いて初期位置を算出。

- 2: 初期位置から M 推定による収束計算で注意点を推定．
- 3: M 推定で算出された各人の重み関数の値から注意点への関与を判定．
- 4: 推定された注意点を次フレームの初期値とする．
- 5: 関与の小さいデータから新たな注意点を求める．  
前フレームより注意点が多い場合は1 から 4 の手順を繰り返す．  
前フレームより注意点が少ない場合は2 から 4 の手順を繰り返す．
- 6: 注意点の候補がなくなった時点で推定を終了．

以上のようなアルゴリズムによって複数注意点の推定シミュレーションを行う．シミュレーションは人数は6人，2つの注意点は円軌道を移動するように設定を行い，3人ずつのグループでそれぞれの注意点を注視するものとする．各人の頭部方向は注意点を注視しているものとして算出するが，顔情報計測システムを用いた場合の計測誤差を想定してノイズを頭部位置，及び頭部方向に加える．

前フレームの推定結果を考慮しない場合及び考慮する場合(提案アルゴリズム)の結果はそれぞれ図 4.10, 4.11 となる．前フレームの推定結果を考慮しない場合は，注意点が本来の軌跡とは異なる箇所を推定するケースが多く観測された．これは複数の注意が存在することの影響で，推定結果がローカルミニマムに陥ったことが原因であると考えられる．M 推定の場合，初期値の設定が非常に重要であり，初期値によってはこのようにローカルミニマムに陥りやすくなる．一方，初期値に前フレームの推定結果を反映した提案アルゴリズムでは，ローカルミニマムに陥ることなく，安定した複数注意点の推定を実現した．

#### 4.2.6 楕円体を用いた注意推定

空間中の注意は注視対象や状況において，集中の度合いが異なるものである．そこで本研究では注意を楕円体として扱うことで，空間中の注意の集中度を大きさで表現する．図 4.12 に示すように，M 推定を用いて求めた注意点  $\hat{c}$  を各人の頭部方向ベクトルに沿う直線上に射影する．各直線上に射影された点  $d_i$  に注目し，主成分分析を用いて楕円体を構成する．M 推定を用いた時に算出される重み  $\omega(\varepsilon_i^{(k-1)})$  が低い場合には，対応する  $d_i$  を除くことで，異なる対象を注視している

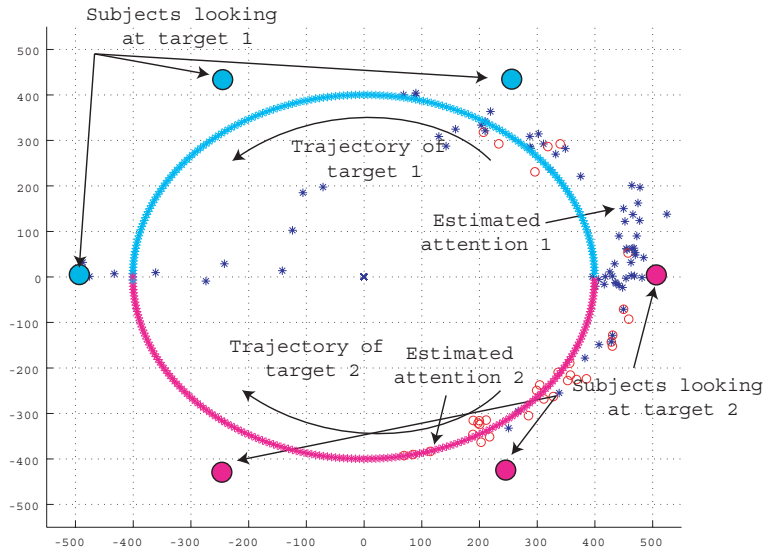


図 4.10 前フレームの推定結果を反映しない場合

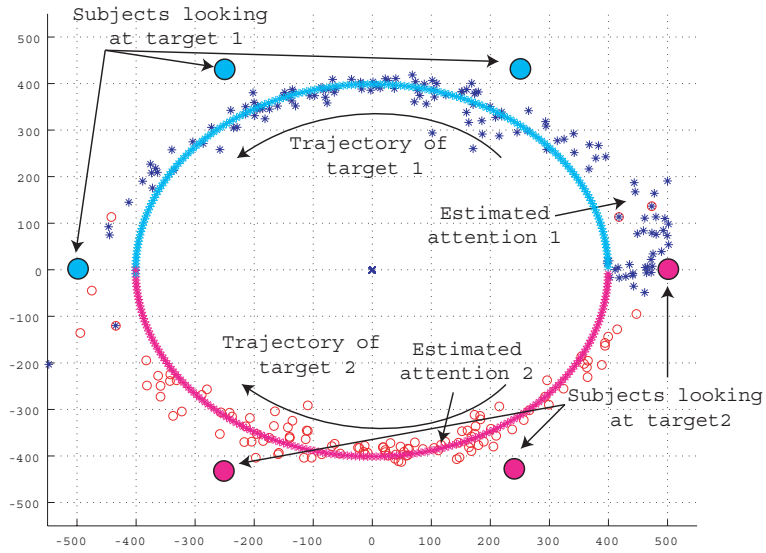


図 4.11 前フレームの推定結果を反映した場合



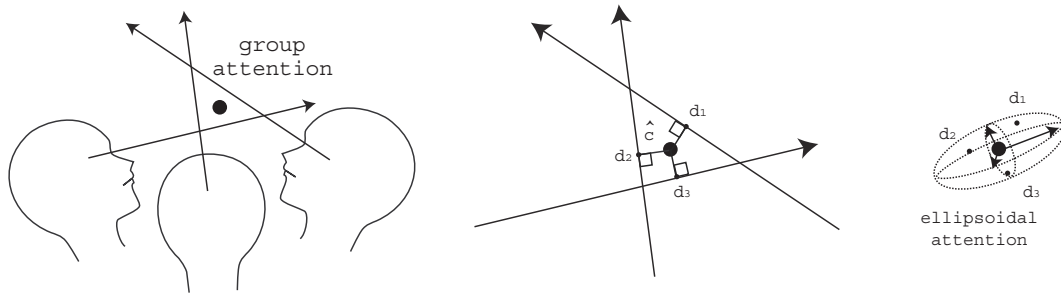


図 4.12 楕円体を用いた注意の推定手法

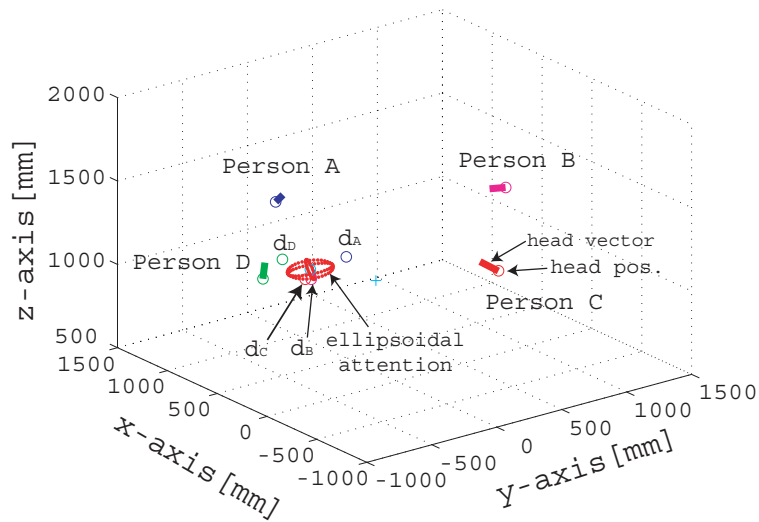


図 4.13 楕円注意領域の推定結果

第三者の影響を除いた楕円体を構成することができる。

分散共分散行列  $V$  は点  $d_i$  を注意点  $\hat{c}$  を原点とした座標系で考えると、式 4.5 のように求められる。

$$V = \frac{1}{N} \sum_{i=0}^n d_i d_i^T \quad (4.5)$$

行列  $V$  の固有値，固有ベクトルを  $\lambda_i, v_i$  とすると，楕円体は式 4.6 と定義される。

$$\frac{x^2}{\sqrt{\lambda_1} v_1} + \frac{y^2}{\sqrt{\lambda_2} v_2} + \frac{z^2}{\sqrt{\lambda_3} v_3} = 1 \quad (4.6)$$

注意が 1 点に集中しているときは楕円体が小さくなり，注意が分散しているとき

は楕円体が大きくなる。図 4.13 に示すように楕円体を用いることで空間中の注意の集中度を表現することが可能である。

楕円体で注意を表現することは、注意の分散を可視化する他にも大きな意味がある。「場の注意」をカメラワーク等へ応用した場合、注意の大きさをカメラのズーム等をコントロールするパラメータとして応用することが可能であると考えられる。

### 4.3. 対話中における「場の注意」の推定実験

本研究では対話中の「場の注意」の推定実験を行った。まずはじめに「しりとり」中の「場の注意」の推定を行った。自然な対話やミーティングなどの状況において実験は行うべきであるが、注目されるべき対象や位置を定義することは困難であり、推定した注意点の評価を行うことは非常に難しい。そこで、本研究では注意されるべき箇所の定義が比較的容易である「しりとり」中の注意点を推定し評価を行う。「しりとり」は発話順番が決まっているので、注意されるべき対象(人)を定義することができる。また対話をベースとしているため、比較的实际のミーティング等の状況と類似度が高いとも考えられる。

次に自然な対話状況として、ディベート中の注意推定を行った。「しりとり」では発話者が限定されているため複数注意点は存在しないが、ディベートでは複数の注意点が存在すると予想できる。複数注意点推定実験として、計測及び推定を行う。以下、実験の詳細及び推定結果について述べる。

#### 4.3.1 しりとり中の場の注意

図 4.14 に示すように、被験者 4 人で「しりとり」を行った際の「場の注意」を推定した。頭部方向に関しては 4.2 節と同様にモーションキャプチャを用いて算出した。サンプリングレートは一般的なビデオのフレームレートと合わせて 30[Hz] で実験を行った。原点はモーションキャプチャのキャリブレーションによって設定された位置となっている。図 4.15 は最初の 15 秒間の推定結果を表示したものである(図は分かりやすいようにデータを間引いている)。注意点は「しりとり」



図 4.14 しりとりの様子

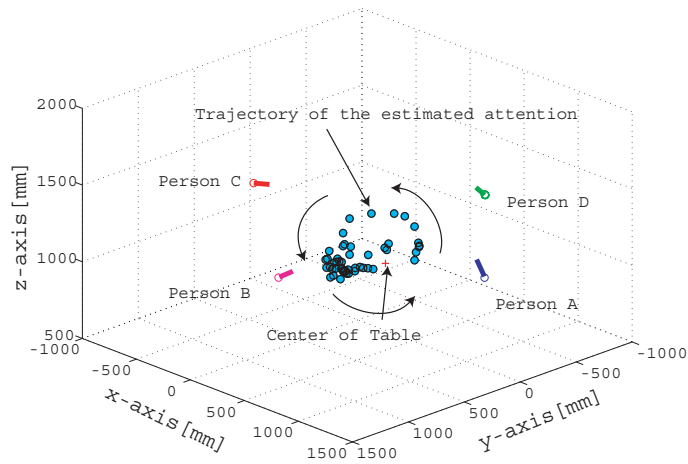


図 4.15 しりとり中の注意点の軌跡

の順番を示すような円形の軌道を描いた。また、被験者 B 付近に注意点がたくさんプロットされていることがわかるが、これは注意が被験者 B に停留したことを表している。

注意点を X-Y 平面上に射影した 2 次元のデータとし、テーブル中央を原点とした座標系において注意点の方向は図 4.16 に示すように変化した。また、図 4.16 に対応した以下のような 3 種類のタイムテーブルを図 4.17 に示す。

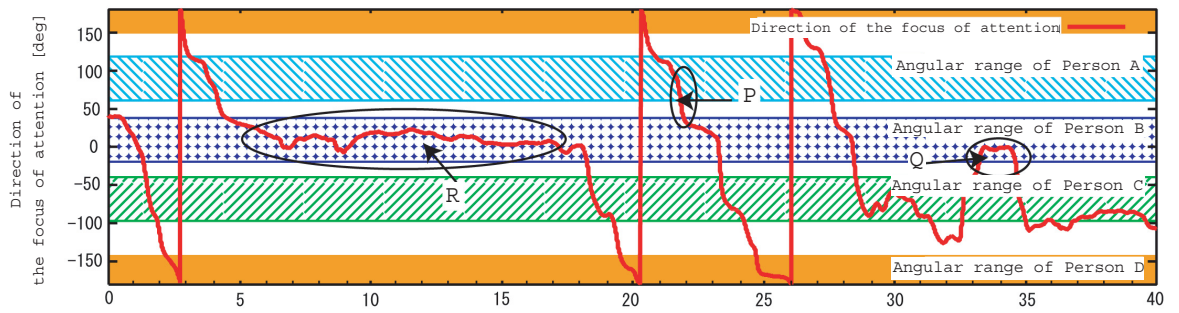


図 4.16 推定された注意方向

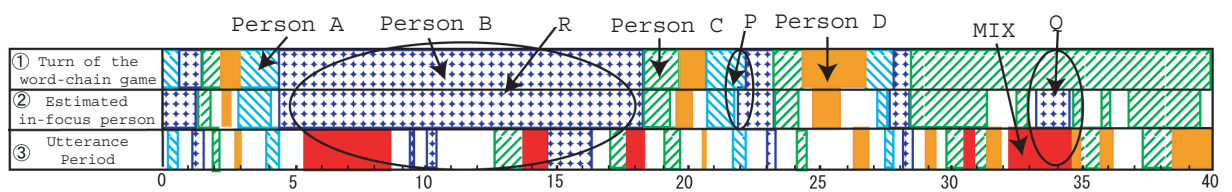


図 4.17 「しりとり」のターン，注意対象(人)，発話者のタイムチャート

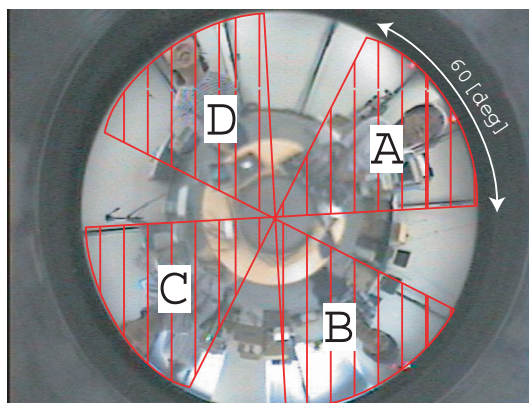


図 4.18 設定した注意対象(人)の範囲

- ① 「しりとり」のターン
- ② 「場の注意」から推定した注意対象(人)
- ③ 発話区間

このうち①, ③は音声信号処理ツール wavesurfer[58] を用いてラベリングを行った。①「しりとり」のターンは, 前の人発話直後から発話終了までを区間としてラベリングを行った。また, ③発話区間は各被検者の発話した区間を A, B, C, D, 複数人が同時に発話を行った区間を”MIX”としてラベリングした。②注意対象(人)は図 4.18 のように各被検者の方向を基準に  $\pm 30$  度以内に注意点が含まれている場合を注意対象とした。 $\pm 30$  度という範囲は人の幅や, 実験中に生じる上体のぶれを考慮した範囲になっている。

今回の実験では, 注意が集中すべき箇所として「しりとり」のターンに注目する。「しりとり」の全時間を  $T[\text{sec}]$ , ターンとの一致時間を  $H[\text{sec}]$  とすると, 一致率  $\kappa[\%]$  は式 4.7 のように算出する。

$$\kappa = H/T \times 100 \quad (4.7)$$

その結果, ①ターンと③発話区間の一致率は 25% である。一方, ①ターンと②注意対象(人)の一致率は 72% となった。また各人の発話区間と注意方向の関係に注目すると, 停留している区間の注意方向は話者の方向とほぼ一致していることが分かる。例えば B の発話区間 (4 ~ 18 秒, 22 ~ 23 秒, 33 ~ 34 秒) に注目するとどの区間も 30 度付近に停留している。その他の人の場合も同様に一定の角度 (A は 100 度, C は 80 度, D は -160 度付近) で停留していることが分かった。

また特徴ある区間として, 注目したい区間は P, Q, R である。区間 P は注意が次の発話者へ移動している区間であるが, ①ターンと②注意対象に注目すると注意対象の方が先行して変化していることが確認できる。これは, 発話より先に注意方向に変化が現れていることを示している。区間 Q では「しりとり」の順番どおりに階段状に変化するのではなく, 順番と逆方向への変化が観測された。また, 区間 R では③発話区間が示すように, ターンとは関係なく発話が行われたが, ②注意対象が示すように発話に影響されずに注意対象は A に注意が停留していた。

### 4.3.2 考察

「しりとり」のターンに注目したところ、注意対象(人)との一致率は72%となり、発話区間との一致率に比べて非常に高いことが確認できた。複数人の発話が同時に起こっている場合、音声だけでは注意の集中している箇所を特定することは困難であるが、複数人の注視情報を用いることから実現することができた。また、発話は無音区間が頻繁にあるため、注意の推定を行うのには注視情報がより適していると考えられる。

発話と注意方向の関係に注目すると図4.16、図4.17のP、Q、Rのように、非常に興味深い動きをする区間がある。区間Pのように発話より先に注意方向に変化が生じるケースは「しりとり」のように発話の順番が決まっている場合、次に注意すべき対象を予測していることから生じると考えられる。また区間Qでは、③発話区間が示すように複数人の発話の区間と一致していることから、それによって一時的に注意が移ったことを表している。またこれとは対照的に区間Rでは、他者の発話が行われても影響されず、注意方向はターンと一致していた。発話内容を確認したところ、影響を与えた区間Qでは「次の文字はどの文字からか」という「しりとり」の進行に影響を与える内容となっていた。複数人の注視情報を用いることで発話により影響を受けている区間とそうでない区間を区別できることが確認された。

また、図4.19、4.20はある被験者の注意が注意対象から外れる瞬間の推定の様子を示したものである。図4.19では全ての注視情報を用いて注意点が推定されている。これに対して図4.20はその33[msec]後の状態であるが、M推定によって算出されたDの視線の重みが閾値( $\omega(\varepsilon_D) < 0.5$ )よりも低く、楕円体を構成する際にはその注視情報は使用されなくなり、残りの3つから注意点が推定されているため、楕円体は小さくなる。このように同一対象への注意が成立しないケースは、本実験(167秒)の中で13回発生した。図4.21は注意の分離が観測された瞬間の全方位画像に頭部方向と推定した注意点を重畳表示したものであるが、この画像からも確認できるように、Dは注意が異なる方向に向けられている。同様に記録したビデオから13件の分離についても同様の状況が確認できた。

このように、提案手法では共通の対象への注意が成立していない人が混在して

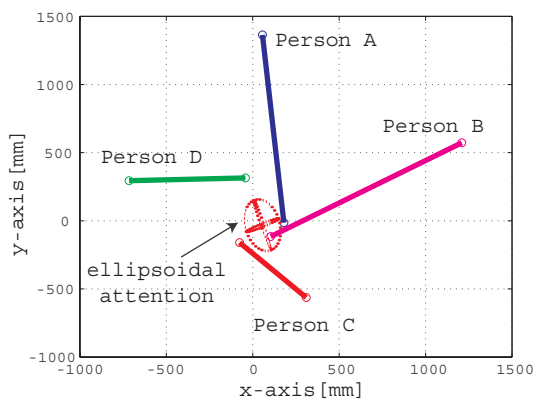


図 4.19 注意点の共有

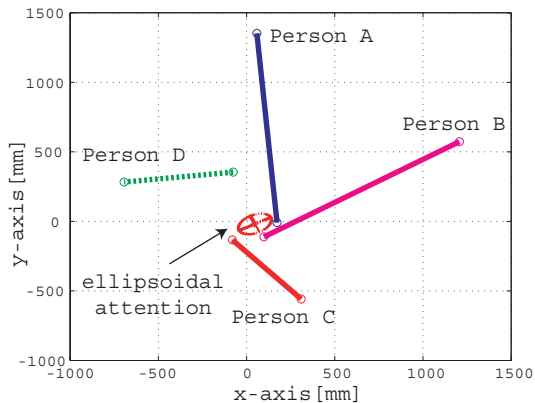


図 4.20 注意点の分離

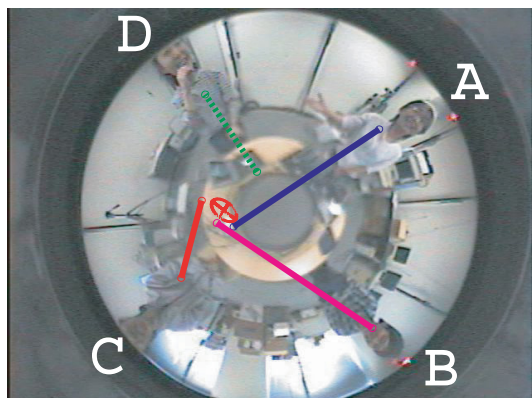


図 4.21 注意点の分離 (全方位画像上にオーバーレイ表示)

いる中でも、注意推定を行うことが可能であることが確認できた。注意を向けられている人の重みも閾値より大きく推定に関与しているが、頭部位置に近いことから推定に問題はないと考えられる。この重みを利用することで、注意を向ける人及び向けられる人をグループとして認識することが可能である。

### 4.3.3 ディベートの場の注意

複数注意点の推定を行うため、4.3.1 節の実験よりも人数を増やし、6 人の対話シーンにて「場の注意」を推定した。対話は以下のような条件のもとに行われた。

- 1: 議論のテーマを設定 (例：ペットを飼うなら犬?)
- 1: テーマに関して、賛成及び反対のグループに分かれる。
- 2: グループごとに、意見をまとめる。
- 3: 各グループの代表者が中心にグループの意見を主張する。

実験は 4.3.1 節と同様にモーションキャプチャによって頭部方向の計測を行った。サンプリングレートは 200[Hz] で計測を行ったが、カメラ画像と対応をとるため、注意点の推定に実際に用いたデータは 30[Hz] となるようにデータを間引いた。

図 4.22 は対話中の「場の注意」の推定結果である。対話は約 3 分間行われたが、図に示したのはグループごとのディベートから全体のディベートへと移る区間である。提案する複数注意点の推定アルゴリズムを適用することによって、グループに分かれている区間 (115[sec] まで) は注意点は二つに推定され、またその後全体の対話に移った区間 (120[sec] から) は注意が 1 つにまとまる傾向が確認できた。

### 4.3.4 考察

グループごとのディベートと全体のディベートに自動的に対応し、注意点の推定数を変化することができた。この成果によって、自動的に対話のグループをラベル付けすることが可能になると考えている。グループの情報を用いて、ミーティングの自動要約をはじめとして、様々な応用が可能であると考えられる。

実験においていくつか問題点が確認されたのでここでまとめる。人及び注意点の位置によっては、異なる対象を注視している人の情報が推定に大きな重みとして加わることがあることが確認された。これによって、図 4.23 に示すように本来、注意が集中している箇所とはまったく異なる場所を推定結果となることが何



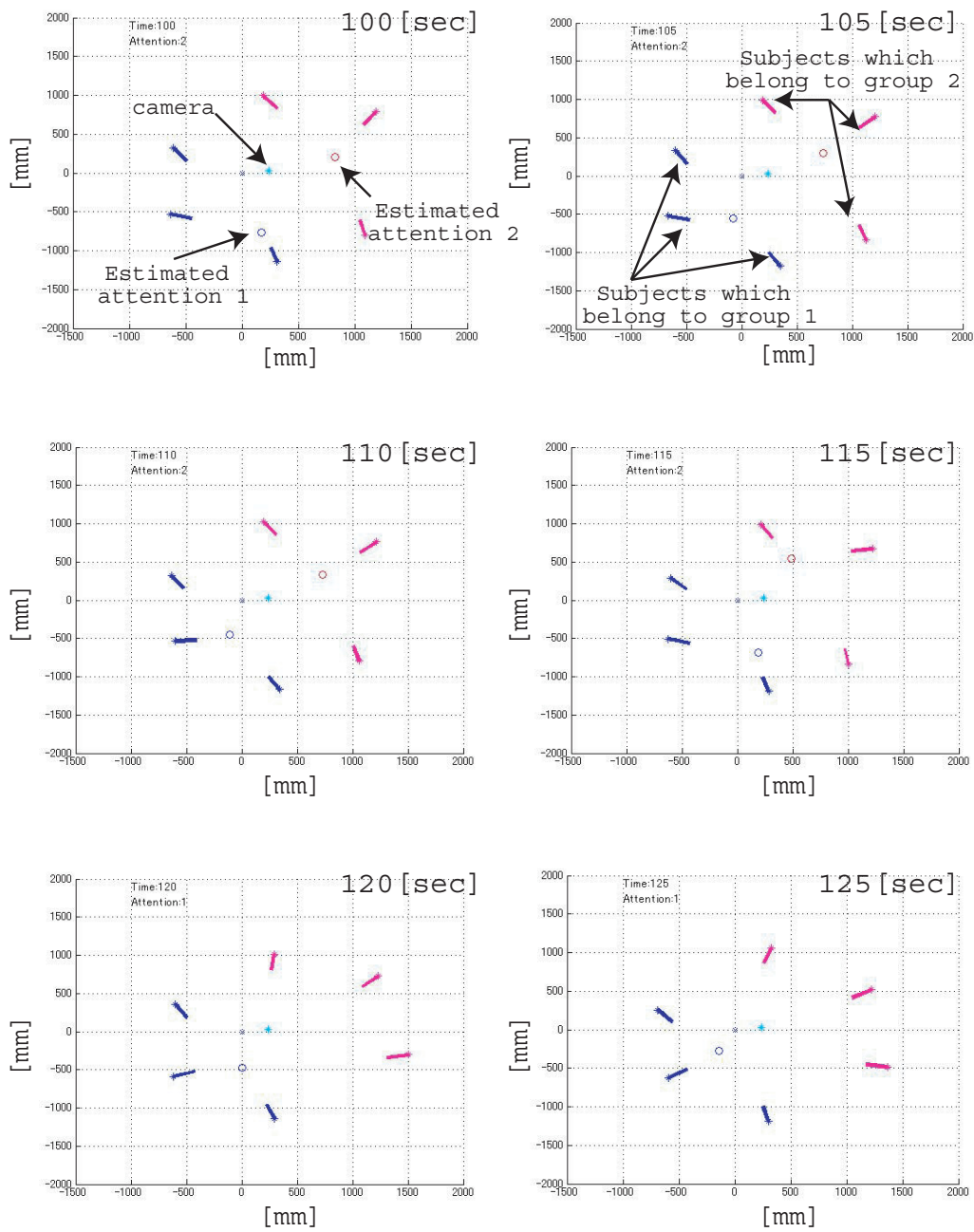


図 4.22 複数注意点の推定結果

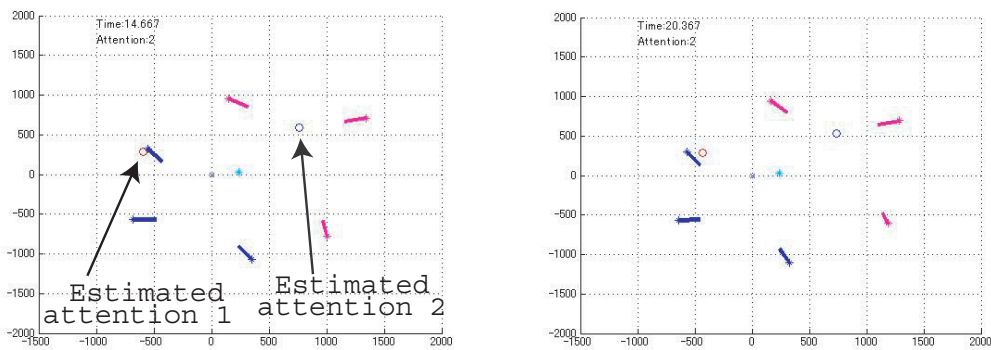


図 4.23 複数注意点の推定エラー

度か確認された．これを防止するために注意点に対して距離が離れるほど，影響力(重み)が小さくなるような工夫が必要である考えられる．

## 4.4. 応用：映像自動編集

「場の注意」は，複数人の注意の集中する箇所を推定するものであり，その応用分野として映像の自動編集やカメラワーク，または会議の自動要約等が挙げられる．本研究では，応用例の一つとして映像の自動編集に注目した．注意点の推定実験のデータを基に「しりとりに」及び「ディベート」に関してそれぞれ映像の自動編集を行った．以下詳細について述べる．

### 4.4.1 場の注意に基づく映像自動編集

場の注意を捉えることで，状況把握し易い映像への編集が可能である．従来研究においてシーンを自動撮影する場合には，注目喚起行動 [59] や音声情報 [60] 等を利用し撮影が行われている．しかし，提案手法は注意点を捉えるように映像を編集するため，注意対象の設定の必要がなく，また無音状態でも注意点が推定できるという利点がある．対面式会議等の状況において使用可能なカメラシステムは，複数のカメラを用いる方法，アクティブカメラを使用する方法など様々なシ

ステムがあるが，本研究では容易に複数カメラの切り替えやコントロールと同様の効果を得るために全方位カメラを用いて記録を行った．全方位カメラはテーブルの中央付近に設置した．注意点の位置情報は3次元の情報を持っているが，本研究では水平面の注意方向という1次元の情報を映像に反映させる．注意の方向を画面中央に表示するように映像の自動編集を行った．推定した注意点の結果，及び注意情報を映像に反映させた結果を図 4.24 に示す．図は 5[sec] 間隔の画像であり，それぞれの推定結果に全方位画像と展開画像が対応している．全方位画像上の矢印は推定した注意点の方向である．また，注意方向を基準として，全方位画像の展開を行った [61]．展開画像はターン（発話順番）の人を中央付近に提示することができた．

#### 4.4.2 複数注意点に対応した映像自動編集

4.4.1 節では，注意点が1つの場合における映像の自動編集を行った．本節では注意点が複数ある場合において映像自動編集を行う．ディベートにおける実験においても「しりとり」と同様にテーブルの中央に設置した全方位カメラを用いて映像の記録を行い，これを基に映像編集を行う．

推定した注意点の結果，及び注意情報を映像に反映させた結果を図 4.25 に示す．図は 5[sec] 間隔の画像であり，それぞれの推定結果に全方位画像と展開画像が対応している．全方位画像上の矢印は推定した注意点の方向であり，2つの矢印は注意方向を示している．矢印が1つしか表示されていない区間は，注意点が1つしか推定されなかった区間である．また，注意方向を基準として，全方位画像の展開を行った．展開画像は2つまたは1つの注意方向を基準として展開する．図に示すように，対話中は注意点の数を基にして展開画像の提示も動的に変更する．グループでの対話中は注意点が2つ推定されたので，提示画像も2つとなっている．その後，全体での対話に推移したことによって，推定される注意点の数も1つとなり，展開画像も1つとなった．複数注意点を映像の自動編集に用いることで，対話のグループに基づく編集を実現することができた．

提示画像が複数存在すると提示手法に関しても工夫が必要となり，多チャンネルの映像の提示方法に関しても研究が盛んに行われている [38][62]．本研究では，

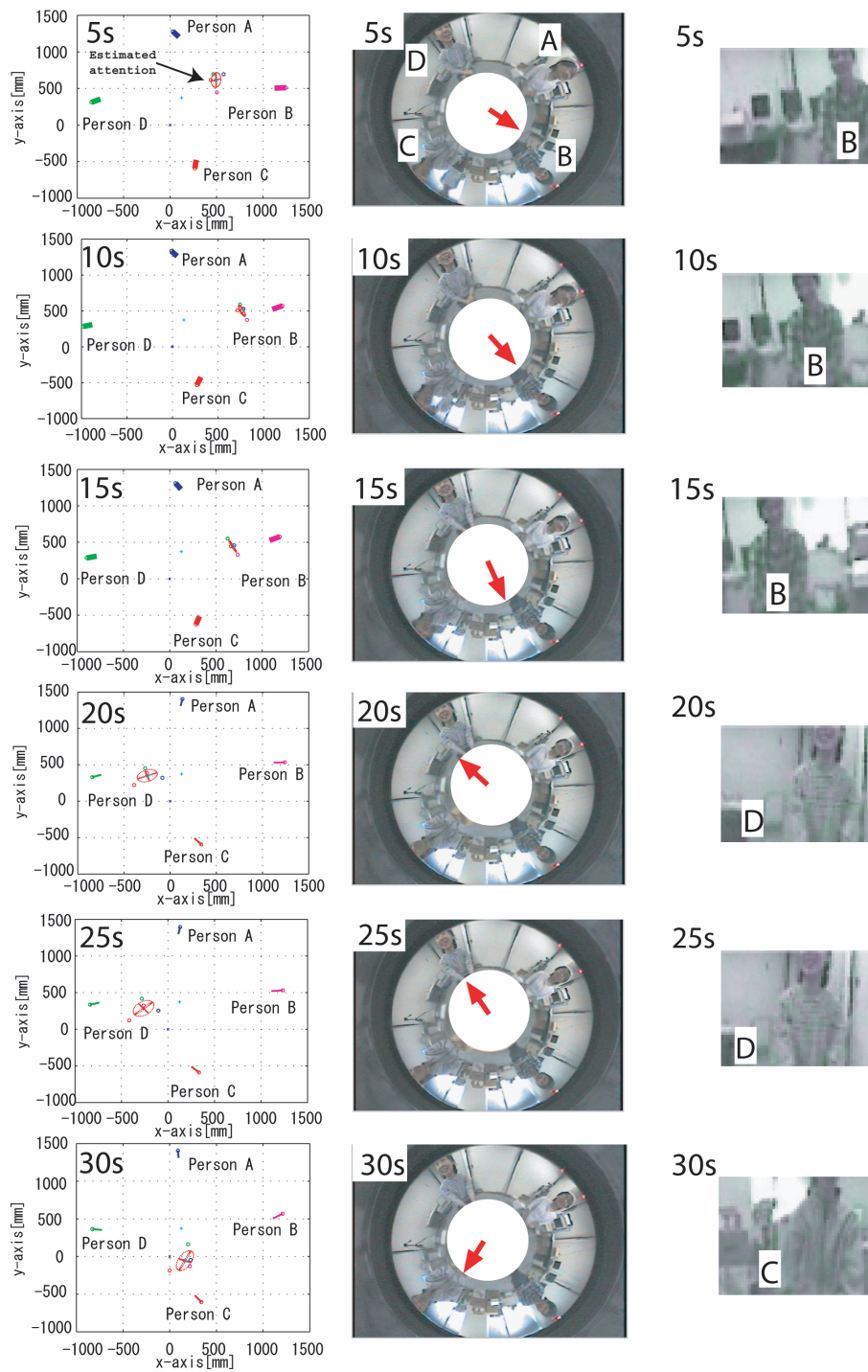


図 4.24 映像の自動編集 (注意点が1つの場合) : (左)「場の注意」の推定結果, (中) 注意方向の提示した全方位画像, (右) 切り出された注意画像

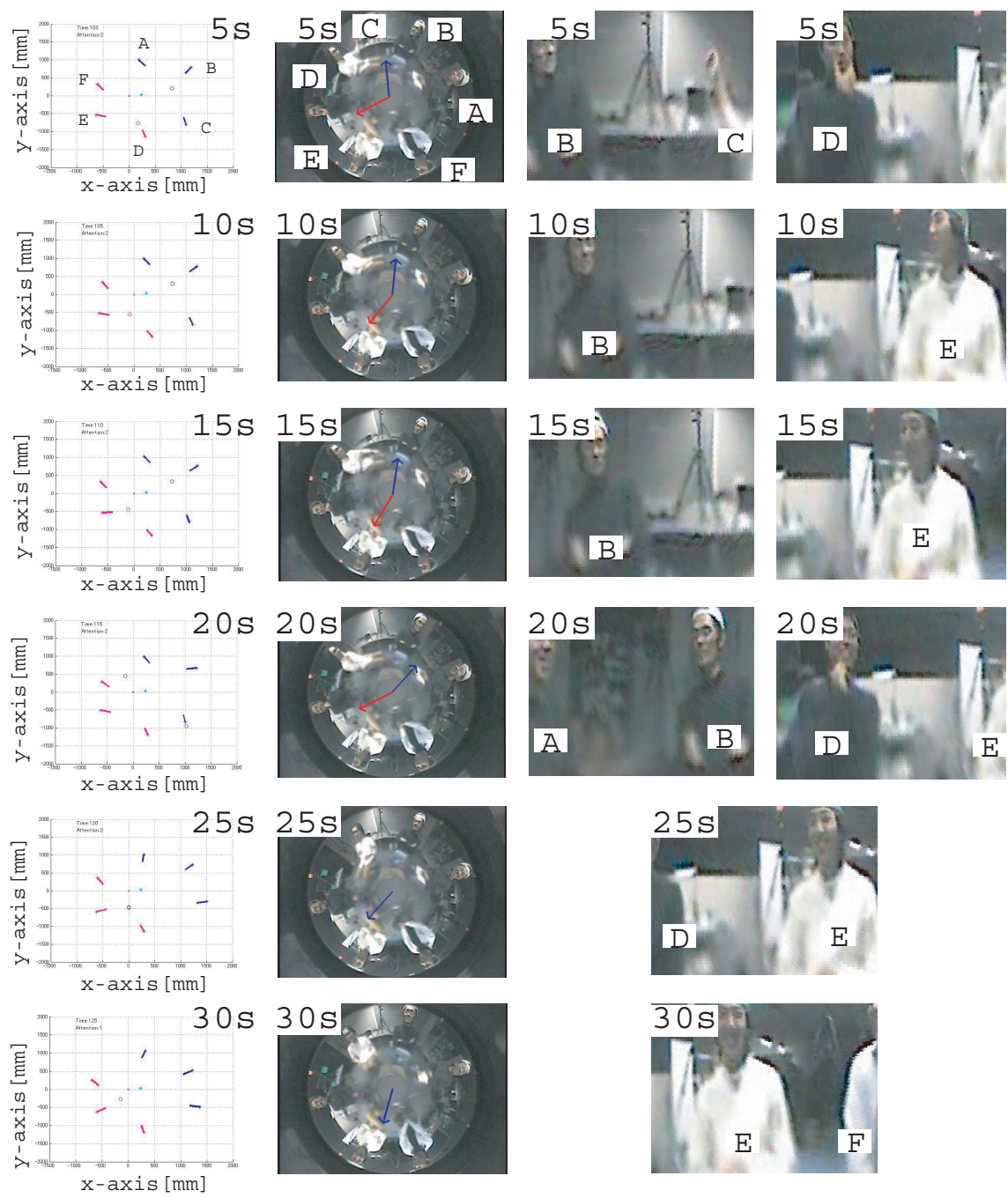


図 4.25 映像の自動編集 (注意点が複数の場合) : (左) 「場の注意」の推定結果 , (中) 注意方向の提示した全方位画像 , (右) 切り出された注意画像

複数の注意点に基づく映像の提示手法まで議論が及ばなかった．今後の課題として検討の必要がある．また，本研究では楕円体として注意推定を行い分散を考慮している．楕円体のパラメータをカメラのズーム等のコントロールへも応用することが可能である．「場の注意」を用いて，より高機能なカメラコントロールを実現したいと考えている．

## 4.5. カメラシステムによる複数人の同時計測

前節まで，本研究で提案する「場の注意」の推定手法の評価を行い，実際の対話において有効であることを確認した．しかしながら，モーションキャプチャを利用していため，頭部に再帰反射マーカを装着する必要がある，実際の利用を想定したものとはなっていない．そこで本節ではモーションキャプチャシステムを利用せず，マーカレスな計測の実現手法について検討し，実際のミーティングにおける実現可能性について述べる．

本節では，複数人の頭部位置・方向(6DOF)の同時計測を実現するために図 4.26 に示すような，ステレオカメラベースである多人数同時計測可能なミーティングキャプチャシステムを構築する．従来のミーティングキャプチャシステムはテーブルの中央から外側を撮影するものが一般的であり，マルチカメラによるミーティングキャプチャシステムが報告されている [63][64]．しかしながら，本研究では頭部方向の定量的計測を実現する必要があるため，図 4.26 のようなステレオ計測が可能なカメラシステムとなっている．ステレオとすることで，顔情報計測システム ((付録 A)) のアルゴリズムを利用することが可能となる．構築したシステムはプロトタイプであり，4 人用のカメラシステムとなってる．従来のミーティングキャプチャシステムと同様にテーブル中央から外側を撮像するようにカメラを配置している．

### 4.5.1 カメラキャリブレーション

一般に複数カメラのキャリブレーションにおいて，撮像空間が共有されている場合には直線によるキャリブレーションが一般的に利用されている [65]．2 次元



図 4.26 ミーティングキャプチャシステム(プロトタイプ)

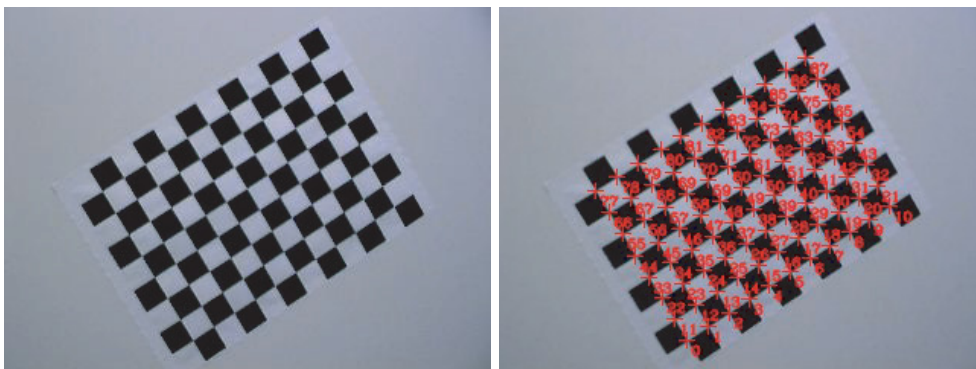


図 4.27 平面パターンによる特徴点の抽出，左：オリジナル画像 右：特徴点抽出結果

平面や3次元の参照物体を利用した場合には，参照点を観測することができる方向が限定されるが，直線状に配置したマーカを用いると，どの方向からも観測できる利点がある．

しかしながら，ミーティング用のカメラシステムはテーブルの中心から外側を撮影するものが一般的である．このように配置すると撮像空間が共有できないため，通常のキャリブレーション手法を用いることができない．そこで本研究では

以下のような手順でキャリブレーションを行い，複数人の頭部位置・方向を統一した座標系で計測する．

- 1: 平面パターンを用いてカメラパラメータの推定
- 2: トータルステーションを用いて世界座標系の統一

まず，Zhang らによって提案された平面パターンを用いたキャリブレーション手法 [66](付録 E) を用いてカメラパラメータを推定する．まずはじめに，図 4.27 に示すように，平面パターンから特徴点を抽出を行う．この作業を 3 回以上繰り返し，内部パラメータの推定を行い，その後外部パラメータを求める．これによって各ステレオカメラに関しては平面パターンを用いてキャリブレーションを行うことができる．しかしながら，複数人の頭部位置及び方向を統一座標系で取り扱うためには，すべてのステレオカメラの座標系を統一する必要がある．

そこで次に，トータルステーションを用いて図 4.28 に示すように，全てのカメラにおいて統一された座標系で外部パラメータの推定を行う．通常，ワールド座標系の原点は平面パターン上に設けるが，本研究では池田らの全方位マルチカメラシステムのキャリブレーション手法 [67] と同様にトータルステーションを用いて，外部パラメータの推定を行う．外部パラメータの推定手法の手順は以下のようになっている．

- 1: 内部パラメータを推定した後，カメラで平面パターンを撮像する．
- 2: 撮像した平面パターンの 4 隅の座標をトータルステーションで計測し，平面の 3 次元位置を算出する．
- 3: 4 隅の座標から，各格子の座標系を線形補間によって算出する．
- 4: 算出された座標を用いて外部パラメータの算出を行う．
- 5: すべてのカメラで 1~4 の手順を繰り返す．

トータルステーションの位置を途中で変更する必要があるが，その際は後方交会を用いて，統一された座標系で平面パターンを計測する．

以上のような手順によって統一された世界座標系による外部パラメータを推定することが可能である．推定された結果を図 4.29 に示す．提案するミーティング



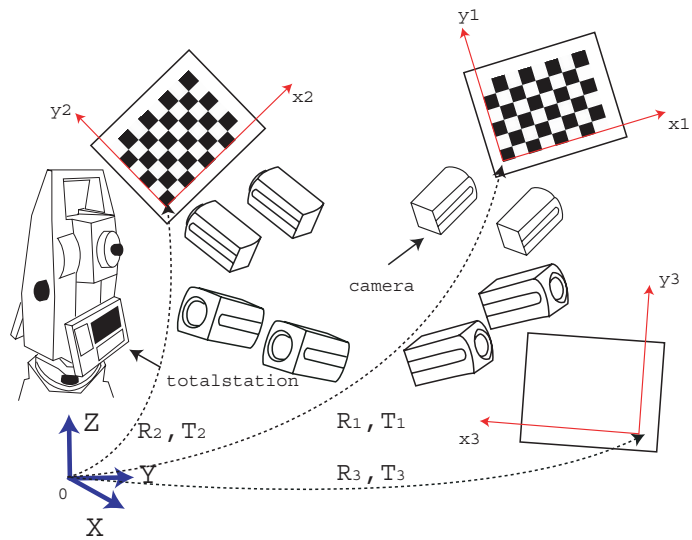


図 4.28 カメラキャリブレーション

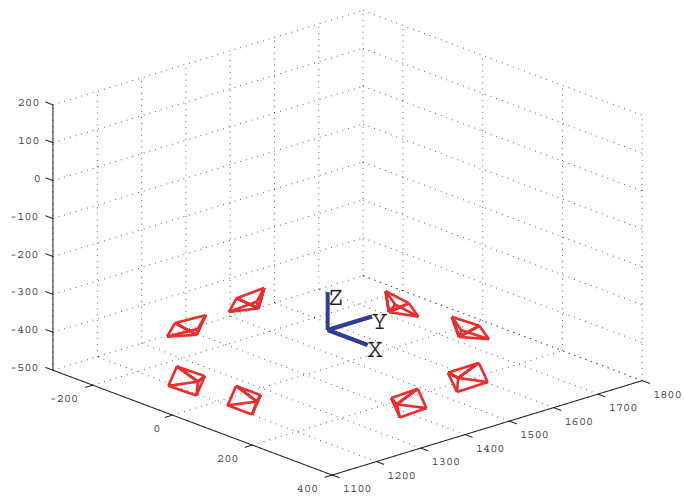


図 4.29 キャリブレーション結果

キャプチャシステムによって、統一した座標系における頭部位置及び頭部方向の計測を実現することが可能となる。

#### 4.5.2 カメラシステムによる対話シーンの計測

4.3.1 節と同様に「しりとり」中の場の注意の推定を行う。複数の顔情報計測システムを用いて図 4.30 に示すように、複数人の同時顔情報計測を実現した。その結果、頭部位置・方向を用いて図 4.31 に示すように「場の注意」が推定される。中央付近に配置されているのがカメラであり、各人の頭部位置と方向は点とベクトルで表示されている。2 番の人の前方に表示されている楕円体が「場の注意」であり、2 番へ注意が集中していることが分かる。また 3 番は他の人と異なり、4 番の方向を向いていて、注意推定に情報が反映されていない。

以上のように、顔情報計測システムを複数用いることによって非接触かつリアルタイムに「場の注意」の計測を実現した。アプリケーションへの応用を考えた場合にこのような非接触かつリアルタイムでの計測技術は重要であり、この実現によって様々な応用が可能となる。問題点としては、頭部方向の計測範囲が挙げられる。顔情報計測システムを用いた場合、頭部方向の計測範囲は $\pm 30$ 度程度であることから、トラッキングを行うことができない場合がしばしば確認された。今後の課題として頭部方向計測及び画像処理技術の改善が必要事項として挙げられる。

### 4.6. 本章の結論

本章では、複数人の注視情報を利用した注意点の推定手法について述べた。注意推定に M 推定を用いて、注意の共有が成立しない第三者の情報に対する考慮を行った。シミュレーションによってその効果を確認し、評価実験を行ったところ注意点と注視対象の平均誤差は約 30[cm] 程度となっていた。従来研究と異なり人や物体に対する注意を推定するのではなく、複数人の注視情報を用いることで注意の集中する箇所(場の注意)を推定した。これによって従来難しかった空間上を移動する物体に対する注意や新しい対象に対する注意にも柔軟に対応すること

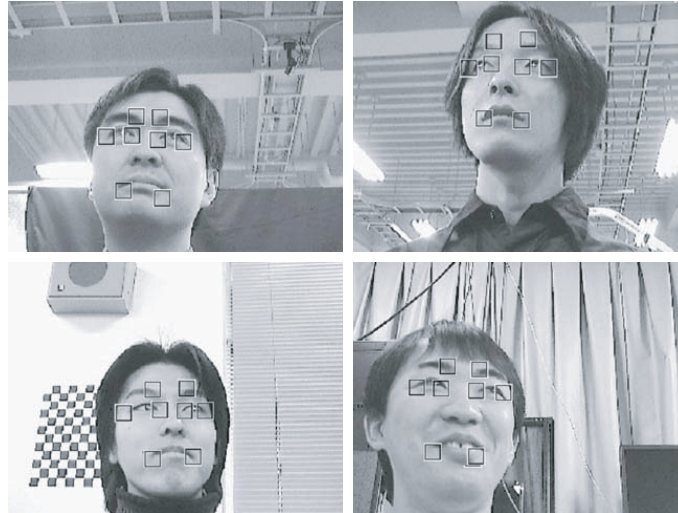


図 4.30 複数人の顔情報計測結果

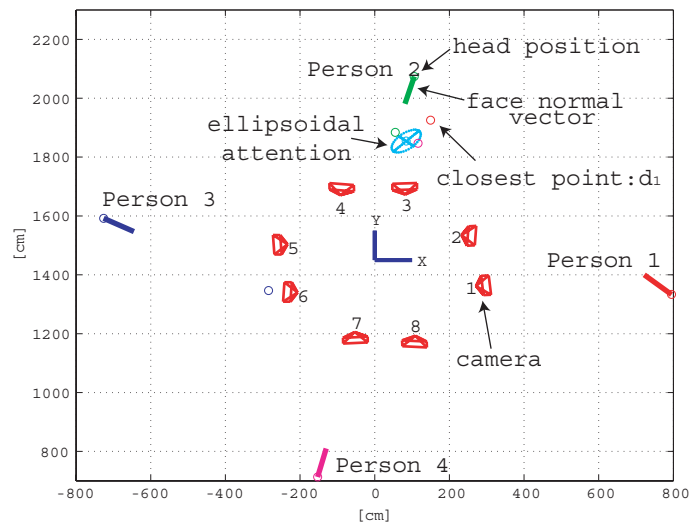


図 4.31 カメラシステムによる場の注意の推定結果

が可能となった。実際に「しりとり」中の4人の頭部運動から注意点の推定を行い、発話と注意点の方向に注目したところ、注意点の方向は「しりとり」のターンとの一致度が高いことが確認できた。また、「ディベート」によって複数注意点の推定が可能であることを確認した。

応用例としては、「場の注意」に基づく映像の自動編集を行った。自動編集された映像はターンを表現するようにスクロールされたことから本手法の有効性を確認できた。複数注意点が存在する場合も同様に、注意点の方向を全方位画像から切り出し提示を行った。

複数の顔情報計測システムを用いたカメラシステムを提案し、統一された座標系でのオンライン計測を実現した。撮像空間が共通でないため、トータルステーションを用いて座標系の統一を行った。カメラシステムによって実際に対話の計測を行い、本提案手法の実現可能性を確認した。

今後の課題としては、まず「場の注意」の応用があげられる。本研究では対話状況に適用したが、人の注意はいかなる状況においても行われており、複数人が存在する状況に本提案手法が適用可能である。応用分野としてはミーティングのみでなく、授業やTV番組、スポーツなど多くの応用が考えられる。今後は各分野へ応用を広げていきたいと考えている。また、カメラシステムではオンライン計測の実現可能性を確認したが、今後の課題として頭部方向計測の頑健性が求められる。現在の顔情報計測システムでは $\pm 30$ 度程度の範囲が計測可能範囲となっているが、計測範囲を広げることで頑健なミーティングキャプチャシステムとなると考えている。

## 第5章 結論

### 5.1. まとめ

本研究では、実空間における注視推定技術の確立及び、応用範囲の拡大を課題とした。本研究では日常活動の支援を行うため、ユーザの負担の少ない非装着型の計測システムにおける注視判定手法及び応用に注目した。

「見る」という行為は人の興味・関心を表す代表的な動作であり、視線計測技術はドライバモニタリングやHCI等の多くの分野で期待されている。本研究では人の興味・関心を抽出する手掛かりとして特に注視対象に注目した。従来の視線計測に関する研究は注視特性の解明をはじめとし、コンピュータ上の注視点に基づくインタフェースが中心であった。本研究では実空間の注視点に注目し、ドライバモニタリングやコンピュータ及び家電の操作、また複数人の注視計測に基づく注意推定と応用範囲の拡張を行った。

第3章では、静的な物体に対する注視判定手法の提案を行った。注視対象を球及び三角ポリゴンでモデル化を行い、頭部方向及び視線方向によって注視判定を実現した。注視判定手法をドライバモニタリングへ応用し、メータやミラー等に対する注視判定を実現した。ドライバモニタリングシステムとして車線変更及びカーブ走行中の注視行動の記録を行った。車線変更では、Salvucciらがシミュレーションにより得ている注視対象の変化が代表的な知見であるが、本研究ではこれと同様の結果を実車かつ非装着型の計測システムを用いて、実車で計測することを実現した。またカーブ走行時では、Landらが報告している頭部・眼球運動の知見と一致した結果が得られた。これらの結果から提案するモニタリングシステム及び注視判定手法の有効性が確認できた。本研究では、ドライバモニタリングシステムにおいて、先行研究によって報告されている特徴的な注視行動の計

測が重要であることを示した。これによって運転支援システム等への応用が可能である。また、注視と操作対象の切り替えを対応付け、直感的でシームレスなコンピュータ及び家電の操作切り替えを実現した。評価実験にて従来手法と比較し、コンピュータ及び家電の操作において注視による操作対象の切り替えが有効であることを示した。

第4章では、複数人の注視情報を利用した注意推定手法について提案を行った。従来研究では、人や対象の位置情報を利用して注意の最も集中する個所の推定が行われている。しかしながら注意対象や数が限定され、柔軟な推定が推定を行うことが出来ない問題点があった。そこで本研究ではこれらの問題に対して、複数人の視線ベクトルの交差を考えることで注視対象を限定することなく、柔軟な注意推定を実現した。複数人の注視を取り扱う場合、異なる対象に注視を行う人を考慮する必要があるが、M推定を用いることで異なる対象に注視を行う人が存在する場合にも安定した推定結果が得られた。また、場の注意の推定時に算出された重みに応じて、グループの推定を行うことが可能であり、複数の注意点を推定することが出来ることもシミュレーション及びディベート等の対話実験によって確認した。また、注意を楕円体として取り扱うことで、注意の集中度の可視化を実現した。推定した「場の注意」を映像の自動編集へ応用し、実際の対話シーンのデータから評価を行った。評価実験として「しりとり」中の「場の注意」の計測を行った。注意が集中すべき箇所としてターンに注目し、ターンと音声の一致率及びターンと「場の注意」の一致率を比較した。ターンと「場の注意」の一致率は72%であり、音声とターンの一致率の25%と比べて非常に高い一致率を示した。このことから本研究で提案する「場の注意」は注意箇所を的確に示し、カメラワークや映像編集への利用が有効であることが確認された。また、複数の顔情報計測システムを用いたミーティングキャプチャシステムを構築し、オンライン推定を行い実現可能性を確認した。

以上のように本研究では、非装着型の計測システムを用いた注視判定手法の提案を行った。人の興味・関心を表す情報として注視対象に注目し、ドライバモニタリング、HCI及びCMC、CSCWの分野への応用を行い、提案したシステムの有効性を確認した。

## 5.2. 課題と今後の展望

本研究では非装着型の計測システムにおける注視判定手法を提案し，評価を行った．静的な対象に対する注視は，視線と対象の位置の情報を基に幾何学的に判定を行った．しかしながら注視対象のモデルに関して議論が及ばず，現状では簡易なモデルをマニュアルで作成している．視線計測の精度が向上すると共にモデルは詳細でかつ容易に作成することが求められる．また，配置等に変更があった場合には動的に対応することが出来ないため，位置計測が今後の課題である．

また，提案手法では自由に移動するユーザの注視対象を計測することは困難である．より柔軟に人の注視行動の計測を実現するためには，計測技術の改良が求められる．装着型の場合に頭部の位置計測は，磁気センサやステレオカメラを用いて行うのが一般的である．しかしながら，これらの手法では，ユーザが自由に動き回ることができない．非装着型の場合には，カメラが固定であることから同様に計測範囲が制限されてしまう．今後，注視を用いたより高度なインタラクションを実現するために，計測範囲の制限の問題に装着型，非装着型の両方からアプローチを行う必要がある．

ドライバモニタリングへの応用としては，ドライバの頭部運動や注視対象を自動車の運動特性と同時に計測を行った．これによって先行研究の重要な知見と同様の結果を得られたが，注視対象の判定は車内の対象であるスピードメータやルームミラー等に限定されている．本研究ではパノラマ展開画像上に視線情報を投影し注視点の推定を行ったが，屋外の対象に対する注視の自動判定は実現できていない．屋外の対象である道路領域や自動車，歩行者等の自動検出と視線計測技術を用いて注視を自動的に判定する手法が求められる．

「場の注意」に関しては，今後は頭部情報でなく視線情報も考慮し，注意点の精度向上を行う必要がある．頭部運動は視線移動と比べて動きが小さく，推定において誤差の要因になるが視線計測を併用して安定した結果が得られるように改善する必要がある．

本研究では，ドライバモニタリング，HCI，及び映像の自動編集へと注視情報の応用を行ったが，自動カメラワークや遠隔ミーティング支援，ロボットインタラクションなど応用可能な分野が多数存在する．今後，新しい分野において注視

情報の利用の提案を行い，応用範囲の拡大を更に進める必要がある．最後に本研究の成果が，今後の視線計測技術及び応用分野において今後の発展を促すことができれば，筆者にとって大きな喜びである．



# 謝辞

本論文に書き記した一連の研究は、筆者が奈良先端科学技術大学院大学 情報科学研究科在学中に、ロボティクス講座(小笠原研究室)において執り行われたものです。この研究を行うにおいて多くの方々のお力添えを頂きました。ここで御礼を申し上げます。

本学 情報科学研究科 小笠原 司 教授には、修士及び博士の5年間にわたりご指導を頂きました。研究テーマの大きな転換を決断した際にも、筆者の希望を尊重して頂き、研究をここまで遂行することができました。また、海外への研修の機会を与えて頂き、今後の研究・教育活動の大きな糧を得ることができました。先生のお力添えがなかったら本研究をまとめる事ができませんでした。深く御礼申し上げます。

本学 情報科学研究科 木戸出 正繼 教授には、本研究をまとめるにあたり、数々の御助言、御指導を頂きました。特に中間発表、公聴会等では筆者の研究の問題点等を的確に指摘して頂きました。深く御礼申し上げます。

本学 情報科学研究科 松本 吉央 助教授には、研究をいちからご指導を頂きました。ドライバ計測では先生所有の自動車までご提供頂きました。また、論文執筆の際には、何度も原稿の添削を行って頂き、発表に関しても適切なご指導、ご助言を頂きました。有難うございました。

本学 情報科学研究科 上田 淳 助手には、研究会等で多くのご指導を頂きました。特に、複数人の注意を楕円体として推定する際には、手法をご教示頂き実現することができました。また、研究活動に対する姿勢は学ぶことが多く、筆者も研究成果を残すことができました。有難うございました。

本学 情報科学研究科 上田 悦子 助手には、研究の取り込み方に対してご指導を頂きました。研究テーマ転換後に短期間で論文をまとめあげることができたの

も、御助言頂いたからであります。有難うございました。

和歌山大学 中村 恭之 助教授には、ロボカップ等を通じて貴重な経験を沢山させて頂きました。これらの経験を今後の研究・教育活動に生かしていきたいと思っております。有難うございました。

株式会社 国際電気通信技術研究所 メディア情報科学研究所 片桐 恭弘 室長(現 函館未来大学 教授)には研修研究員として、プロジェクトに参加する機会を与えて頂きました。有難うございました。

株式会社 国際電気通信技術研究所 メディア情報科学研究所 認知メディア情報学研究室 馬田 一郎 氏, 鈴木 紀子 氏, Nik Swoboda 氏(現 Universidad Politecnica de Madrid), 深谷 拓吾 氏には研究, 論文に対して沢山の御助言を頂きました。深く御礼申し上げます。

豪州連邦科学産業研究機構 Alex Zelinsky 博士には国際学会で多くのコメントを頂いた他, オーストラリア国立大学への研修の際に多くのご協力を頂きました。深く御礼を申し上げます。

オーストラリア国立大学の Luke Fletcher 氏には, 実験やオーストラリアでの生活の際に多くの配慮を頂きました。有難うございました。また, Smart Cars のメンバーにはミーティング等において, 多くの御助言を頂きました。有難うございました。

本学 情報科学研究科 佐藤 智和 助手, 池田 聖 氏にはトータルステーションの使用方法をご教示頂きました。有難うございました。

本学 情報科学研究科 柴田 智広 助教授, 為井 智也 氏, 野村 知広 氏にはモーションキャプチャの使用方法をご教示頂きました。有難うございました。

東京理科大学 竹村 裕 助手, 広島大学 栗田 雄一 特任教員, 立命館大学小枝 正直 講師, 本学情報科学研究科 怡土 順一 氏には先輩として多くのアドバイスを頂きました。有難うございました。

秘書 金岡 恵 様には多くの支援を頂き, 5年間の研究生生活を円滑に進めることができました。有難うございました。

21世紀 COE プログラム情報科学研究拠点推進室 桑野 美紀 様, 足立 敏美 様には, COE 研究発表会の準備や物品等の購入の際に大変お世話になりました。有

難うございました。

本学 情報科学研究科 南出 隼人 君（現 三洋電機株式会社）には、注視による家電操作の切り替えに関する評価を担当してもらいました。筆者の指導の至らない点があり、研究を遂行する上で多くの苦勞をかけてしまいました。そんな中、最後まで実装、評価を粘り強く担当していただきました。有難うございました。

また、筆者が在学中、研究生活を様々な面で支えて頂いた本学 情報科学研究科 ロボティクス講座の卒業生、在校生の皆様感謝いたします。

最後に、私の研究活動を理解し暖かく見守り続けてくれた両親、応援してくれた姉に深く感謝をし、論文の結びとしたいと思います。

## 参考文献

- [1] 福井康之. まなざしの心理学-視線と人間関係. 創元社, 1984.
- [2] 古賀一男. 眼球運動実験 ミニ・ハンドブック. (財)労働科学研究所出版部, 1998.
- [3] Y.Ebisawa. Improved Video-Based Eye-Gaze Detection Method. *IEEE Transactions on Instrumentation and Measurement*, Vol. 47, No. 4, pp. 948–955, 1998.
- [4] C.H.Morimoto, D.Koons, A.Amir, and M.Flickner. Pupil Detection and Tracking Using Multiple Light Sources. *Image and Vision Computing*, Vol. 18, No. 4, pp. 331–334, 2000.
- [5] Y.Matsumoto and A.Zelinsky. An Algorithm for Real-time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement. In *Proceeding of the 4th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 499–505, 2000.
- [6] T.Ohno and N.Mukawa. A Free-head, Simple Calibration, Gaze Tracking System That Enables Gaze-Based Interaction. In *Proceedings of the Eye tracking research & applications*, pp. 115–122, 2004.
- [7] J.L.Levine. An Eye-Controlled Computer. Technical Report RC-8857, IBM Thomas J. Watson Research Center, 1981.
- [8] R.Vertegaal. Attentive User Interfaces. *Communications of the ACM*, Vol. 46, No. 3, pp. 30–33, 2003.

- [9] J.S.Shell, T.Selker, and R.Vertegaal. Interacting with Groups of Computers. *Communications of the ACM*, Vol. 46, No. 3, pp. 40–46, 2003.
- [10] J.Shell, R.Vertegaal, D.Cheng, A.W.Skaburskis, C.Sohn A.J.Stewart, O.Aoudeh, and C.Dickie. ECSGlasses and EyePliances: Using Attention to Open Sociable Windows of Interaction. In *Proceedings of the symposium on Eye tracking research & applications*, pp. 93–100, 2004.
- [11] I.Mitsugami, N.Ukita, and M.Kidode. Estimation of 3D Gazed Position Using View Lines. In *Proceedings of 12th International Conference on Image Analysis and Processing*, pp. 466–471, 2003.
- [12] K.Sumii, A.Sugimoto, and T.Matsuyama. Active Wearable Vision Sensor: Recognition of Human Activities and Environments. In *Proceedings of international conference on informatics research for development of knowledge society infrastructure*, pp. 15–22, 2004.
- [13] M.F.Land. Predictable Eye Head Coordination During Driving. *Nature*, Vol. 359, pp. 318–320, 1992.
- [14] M.F.Land and D.N.Lee. Where we look when we steer. *Nature*, Vol. 369, pp. 742–744, 1994.
- [15] M.F.Land and J.M.Horwood. Which Parts of the Road Guide Steering. *Nature*, Vol. 377, pp. 339–340, 1995.
- [16] G.Salgian. *Tactical Driving Using Visual Routines*. PhD thesis, University of Rochester, 1998.
- [17] D.Salvucci and A.Liu. The Time Course of a Lane Change: Driver Control and Eye-Movement Behavior. *Transportation Research Part F*, Vol. 5, No. 2, pp. 123–132, 2002.
- [18] 三浦利章. 行動と視覚の注意. 風間書房, 2002.

- [19] 藤森充, 上迫宏計, 川村幹也. 高速道路における頭部運動と視線挙動. 計測自動制御学会論文集, Vol. 37, No. 1, pp. 36–43, 2001.
- [20] A.Liu. What the Driver’s Eye Tell the Car’s Brain. In *Eye Guidance During Reading and Scene Perception*, pp. 431–452. Elsevier Science Ltd, 1998.
- [21] T.Victor. A Technical Platform for Driver Inattention Research. In *Proceedings of internet forum The Safety Impact of Driver Distraction When Using In-Vehicle Technologies*, NHTSA, 2000.
- [22] L.Fletcher, G.Loy, N.Barnes, and A.Zelinsky. Correlating driver gaze with the road scene for driver assistance systems. *Robotics and Autonomous Systems*, Vol. 52, No. 1, pp. 71–84, 2005.
- [23] J.L.Levine. Performance of an Eyetracker for Office Use. *Computational Biology and Medicine*, Vol. 14, No. 1, pp. 77–89, 1984.
- [24] C.Ware and H.T.Mikaelian. An Evaluation of an Eye Tracker as a Device for Computer Input. In *Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*, pp. 183–188, 1987.
- [25] R.Jacob. The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get. *ACM Transactions on Information Systems*, Vol. 9, No. 2, pp. 152–1169, 1991.
- [26] J.P.Hansen, A.W.Andersen, and P.Roed. Eye-Gaze Control of Multimedia Systems. In *Symbiosis of Human and Artifact: Proceedings of the Sixth International Conference on Human-Computer Interaction*. Elsevier Science, 1995.
- [27] 大野健彦. 視線を用いたインタフェース. 情報処理, Vol. 44, No. 7, pp. 726–732, 2003.

- [28] S.Zhai, C.Morimoto, and S.Ihde. Manual And Gaze Input Cascaded Pointing. In *Proceeding of the SIGCHI conference on Human factors in computing systems*, pp. 246–253, 1999.
- [29] P.Baudisch, D.DeCarlo, A.T.Duchowski, and W.S.Geisler. Focusing on the Essential: Considering Attention in Display Design. *Communications of the ACM*, Vol. 46, No. 3, pp. 60–66, 2003.
- [30] C.O’Sullivan, J.Dingliana, and S.Howlett. Eye-movements and Interactive Graphics. In *The Mind’s Eye: Cognitive and Applied Aspects of Eye Movement Research*, pp. 555–571. North-Holland, 2003.
- [31] D.Wooding. Fixation maps: Quantifying Eye-movement Traces. In *Proceedings of the symposium on Eye tracking research & applications*, pp. 31–36, 2002.
- [32] D.DeCarlo and A.Santella. Stylization and Abstraction of Photographs. *ACM Transactions on Graphics*, Vol. 21, No. 3, pp. 769–776, 2002.
- [33] R.A.Bolt. Gaze-Orchestrated Dynamic Windows. In *Proceedings of the 8th annual conference on Computer graphics and interactive techniques*, pp. 109–119, 1981.
- [34] R.A.Bolt. Eyes at the interface. In *Proceedings of the 1982 conference on Human factors in computing systems*, pp. 360–362, 1982.
- [35] I.Starker and R.Bolt. A Gaze-responsive Self-disclosing Display. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 3–10, 1990.
- [36] R.Vertegaal, R.Slagter, Gerrit C. van der Veer, and A.Nijholt. Eye Gaze patterns in Conversations: There is More to Conversational Agents than Meets the Eyes. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 301–308, 2001.

- [37] R.Vertegaal. The GAZE Groupware System: Mediating Joint Attention in Multiparty Communication and Collaboration. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 294–301, 1999.
- [38] T.Jenkin, J.McGeachie, D.Fono, and R.Vertegaal. EyeView: Focus+Context Views for Large Group Video Conferences. In *CHI '05 extended abstracts on Human factors in computing systems*, pp. 1497–1500, 2005.
- [39] R.Stiefelhagen. Tracking Focus of Attention in Meetings. In *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, pp. 273–280, 2002.
- [40] R.Stiefelhagen, J.Yang, and A.Waibel. Modeling Focus of Attention for Meeting Indexing based on Multiple Cues. *IEEE Transactions on Neural Networks*, Vol. 13, No. 4, pp. 928–938, 2002.
- [41] Y.Takemae, K.Otsuka, and N.Mukawa. Video Cut Editing Rule based on Participants' Gaze in Multiparty Conversation. In *Proceedings of the eleventh ACM international conference on Multimedia*, pp. 303–306, 2003.
- [42] 竹前嘉修, 大塚和弘, 武川直樹. 対面の複数人対話を撮影対象とした対話参加者の視線に基づく映像切り替え方法とその効果. *情報処理学会論文誌*, Vol. 46, No. 7, pp. 1752–1767, 2005.
- [43] 醍醐真治, 小沢慎治. 顔方向追跡に基づいた球技放映カメラのデジタルパン制御～バスケットボールを例として～. *映像情報メディア学会誌*, Vol. 58, No. 9, pp. 1299–1306, 2004.
- [44] 先山卓朗, 大野直樹, 椋木雅之, 池田克夫. 遠隔講義における講義状況に応じた送信映像選択. *電子情報通信学会論文誌 (D-II)*, Vol. J84, No. 2, pp. 248–257, 2001.



- [45] R.Vertegaal, I.Weever, C.Sohn, and C.Cheung. GAZE-2: Conveying Eye Contact in Group Video Conferencing Using Eye-Controlled Camera Direction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 521–528, 2003.
- [46] R.Yang and Z.Zhang. Eye Gaze Correction with Stereovision for Video-Teleconferencing. Technical Report MSR-TR-01-119, Microsoft Research, 2001.
- [47] M.Danninger, R.Vertegaal D.P.Siewiorek, and A.Mamuji. Using Social Geometry to Manage Interruptions and Coworker Attention in Office Environments. In *Proceedings of the 2005 conference on Graphics interface*, pp. 211 – 218, 2005.
- [48] T.Möller and B.Trumbore. Fast, Minimum Storage Ray/Triangle Intersection. *Journal of Graphics Tools*, Vol. 2, No. 1, pp. 21–28, 1997.
- [49] M.Levoy and T.Whitted. The Use of Points as a Display Primitive. Technical Report 85-022, University of North Carolina, 1985.
- [50] S.Rusinkiewicz and M.Levoy. QSplat: A Multiresolution Point Rendering System for Large Meshes. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 343 – 352, 2000.
- [51] P.Hubard. Approximating Polyhedra with Spheres for Time-Critical Collision Detection. *ACM Transactions on Graphics*, Vol. 15, No. 3, pp. 179–210, 1996.
- [52] 乾敏郎 (編) . 知覚と運動 認知心理学. 東京大学出版会, 1995.
- [53] 綾塚祐二, 松下信行, 暦本純一. 実世界指向ユーザインタフェースにおける「見ているものに接続する」というメタファ. *情報処理学会論文誌*, Vol. 42, No. 6, pp. 1330–1337, 2001.

- [54] 星野剛史, 堀井洋一, 丸山幸伸, 片山淳詞, 柴田吉隆, 吉丸卓志. Air-real:ホームネットワークのユーザーインタフェース. *インタラクティブシステムとソフトウェア IX*, pp. 113–118. 近代科学社, 2001.
- [55] 山田光穂. 2次元平面上の指標を注視させた時の頭部運動と眼球運動の協調関係の分析. *電子情報通信学会論文誌*, Vol. J75-D-II, No. 5, pp. 971–981, 1992.
- [56] 徐剛, 辻三郎. 3次元ビジョン. 共立出版, 1998.
- [57] Z.Zhang. Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting. *International Journal of Image and Vision Computing*, Vol. 15, No. 1, pp. 59–76, 1997.
- [58] S.Kare and B.Jonas. WaveSurfer - an Open Source Speech Tool. In *Processing of International Conference on Spoken Language Processing*, Vol. 4, pp. 464–467, 2000.
- [59] 尾関基行, 中村裕一, 大田友一. 注目喚起行動に基づいた机上作業映像の編集. *電子情報通信学会論文誌*, Vol. 88, No. 5, pp. 844–853, 2005.
- [60] M.Onishi, T.Kagebayashi, and K.Fukunaga. Production of Video Images by Computer Controlled Camera. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 131–137, 2001.
- [61] K.Yamazawa, Y.Yagi, and M.Yachida. Obstacle Detection with Omnidirectional Image Sensor HyperOmni Vision. In *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 1062–1067, 1995.
- [62] 池田大介, 苗村健. 多チャンネル映像の同時閲覧インタフェースの検討. *電子情報通信学会技術研究報告 HIP*, Vol. 104, No. 526, pp. 49–54, 2004.
- [63] R.Cutler, S.Silverberg, Y.Rui, A.Gupta, JJ Cadiz, I.Tashev, L.He, A.Colburn, Z.Zhang, and Z.LiuRoss. Distributed meetings: A meeting capture and broadcasting system. In *Proceedings of the tenth ACM international conference on Multimedia*, pp. 503–512, 2002.

- [64] P.Rybski, F. De la Torre Frade, R.Patil, C.Vallespi-Gonzalez, M.Veloso, and B.Browning. CAMEO: The Camera Assisted Meeting Event Observer. Technical Report CMU-RI-TR-04-07, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 2004.
- [65] Z.Zhang. Camera Calibration with One-Dimensional Objects. *IEEE Transactions Pattern Analysis Machine Intelligence*, Vol. 26, No. 7, pp. 892–899, 2004.
- [66] Z.Zhang. A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, pp. 1330–1334, 2000.
- [67] 池田聖, 佐藤智和, 横矢直和. 全方位型マルチカメラシステムを用いた高解像度な全天球パノラマ動画の生成とテレプレゼンスへの応用. *日本バーチャルリアリティ学会論文誌*, Vol. 8, No. 4, pp. 443–452, 2003.
- [68] D.Salvucci and J.Gorlberg. Identifying Fixations and Saccades in Eye-Tracking Protocols. In *Eye Tracking Research & Applications Symposium*, pp. 71–78, 2000.
- [69] C.M.Privitera and L.W.Stark. Algorithms for Defining Visual Regions-of-Interest : Comparison with Eye Fixations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 9, pp. 970–982, 2000.
- [70] A.Santella and D.DeCarlo. Robust Clustering of Eye Movement Recordings for Quantitation of Visual Interest. In *Proceedings of the symposium on Eye tracking research & applications*, pp. 27–34, 2004.

# 付録

## A. 顔情報計測システム

顔情報計測システム [5] は、図 1 に示すようなステレオカメラのシステムを用いて眉毛、目、口等の部位の 3 次元位置計測を行う。その結果に対して、3 次元顔モデルのフィッティングを行うことで、頭部の 3 次元位置・姿勢を推定する。以下に顔情報計測アルゴリズムの詳細について述べる。

顔情報計測を行う場合、まず各人の顔モデルの登録を行う。顔モデルの詳細は以下のようにになっている。

- 1: 顔領域画像 (図 2)
- 2: 特徴領域画像及び各特徴点の 3 次元位置 (図 3)

まず、入力画像から顔領域を発見するために図 2 のような顔領域の画像を作成する。画像は  $60 \times 30$  [pixel] であり、処理時間を短縮するため低解像度となっている。また、照明変動の影響を押さえるために微分画像を用いる。

顔領域を発見した後、図 3 に示すように両目と口の両端、そして眉の内側の端等の特徴領域について、正規化相関演算を行い対応領域を求め、3 次元位置を求める。各特徴の 3 次元位置の観測結果から頭部の位置・視線を求めるのに仮想バネモデルを用いたフィッティングを行う。このフィッティング手法は、モデルの観測値の間の変位が小さいと仮定できる場合には収束が速く、リアルタイムのトラッキングに適している。各特徴領域の座標は、剛体としての一貫性を保つように求めた位置・姿勢パラメータとモデル中の 3 次元座標を用いて修正される。

視線は虹彩の中心と眼球中心を通るベクトルとして推定される。眼の特徴領域から眼球中心までのベクトルを設定することで、図 4 に示すように視線ベクトルを推定することが可能である。以上のような手順で、計測した結果を図 5 に示す。



図 1 顔情報計測システムの外観

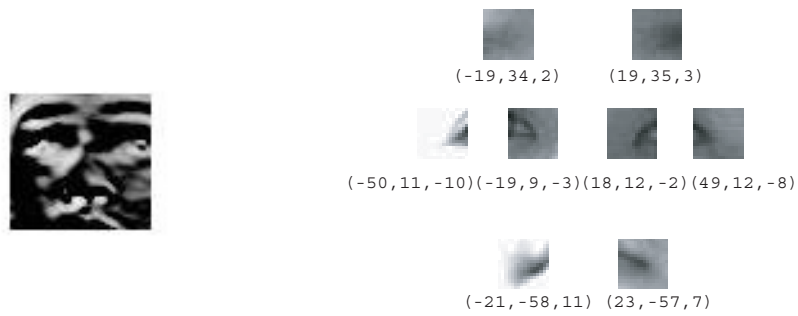


図 2 顔全体のテンプレート画像 図 3 特徴領域テンプレート画像

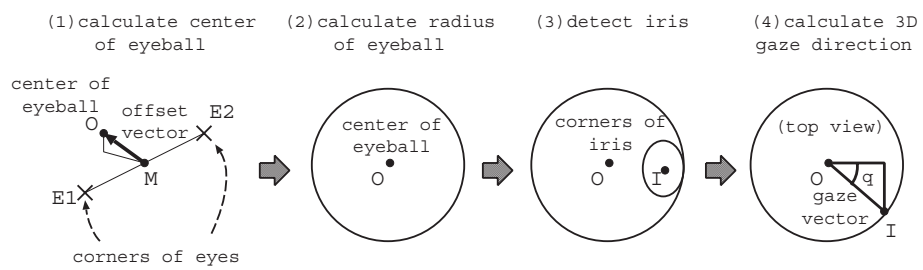


図 4 視線ベクトル計算

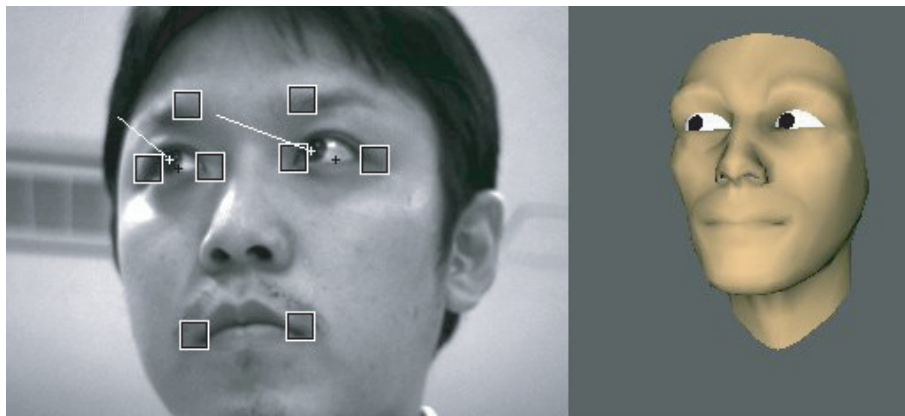


図 5 顔情報計測結果

四角形で囲まれた部分が検出した各部位であり，また2つの線が推定した視線方向である．

ステレオカメラで取得したそれぞれの画像サイズは  $320 \times 240$ [pixel] であり，計測精度は頭部位置は約  $2$ [mm]，姿勢は約  $2$ [deg]，視線方向の精度は約  $5$ [deg] となっている．以上の画像処理時間は Pentium4 2.53[GHz] で約  $10$ [msec] であり，NTSC のビデオレートで処理が可能である．

## B. 半直線と三角形の交差判定

レイトレーシングの分野で交差判定に関しては多くの研究成果が報告されている．半直線と三角形の交差判定で有名な Tomas Möllerr の手法について述べる [48].

半直線  $R$  は基点を  $O$ ，方向ベクトルを  $D$  とすると，以下のように定義される．

$$R(t) = O + tD \quad (1)$$

また，三角形上の1点  $T(u, v)$  は

$$T(u, v) = (1 - u - v)V_0 + uV_1 + vV_2 \quad (2)$$

のように定義され，これらは  $u \geq 0$ ， $v \geq 0$  及び  $u + v \leq 1$  の条件をみたく．半直線と三角形が交差するので

$$O + tD = (1 - u - v)V_0 + uV_1 + vV_2 \quad (3)$$

となり，これを変形すると

$$[-D, V_1 - V_0, V_2 - V_0] \begin{bmatrix} t \\ u \\ v \end{bmatrix} = O - V_0 \quad (4)$$

が得られる． $E_1 = V_1 - V_0$ ， $E_2 = V_2 - V_0$ ， $T = O - V_0$  とし， $t, u, v$  について整理する．Cramer の公式を用いると

$$\begin{bmatrix} t \\ u \\ v \end{bmatrix} = \frac{1}{|-D, E_1, E_2|} \begin{bmatrix} |T, E_1, E_2| \\ |-D, T, E_1| \\ |-D, E_1, T| \end{bmatrix} \quad (5)$$

となり,  $|A, B, C| = -(A \times C) \cdot B = -(C \times B) \cdot A$  であるから,

$$\begin{bmatrix} t \\ u \\ v \end{bmatrix} = \frac{1}{(\mathbf{D} \times \mathbf{E}_2) \cdot \mathbf{E}_1} \begin{bmatrix} (\mathbf{T} \times \mathbf{E}_1) \cdot \mathbf{E}_2 \\ (\mathbf{D} \times \mathbf{E}_2) \cdot \mathbf{T} \\ (\mathbf{T} \times \mathbf{E}_1) \cdot \mathbf{D} \end{bmatrix} = \frac{1}{\mathbf{P} \cdot \mathbf{E}_1} \begin{bmatrix} \mathbf{Q} \cdot \mathbf{E}_2 \\ \mathbf{P} \cdot \mathbf{T} \\ \mathbf{Q} \cdot \mathbf{D} \end{bmatrix} \quad (6)$$

となる. ここで  $\mathbf{P} = \mathbf{D} \times \mathbf{E}_2$ ,  $\mathbf{Q} = \mathbf{T} \times \mathbf{E}_1$  である. 求められた  $(u, v)$  に関して, 条件判定を行うことで容易に交差判定を行うことが可能である.

## C. 注視頻度マップ

ユーザの注視状況を定量的に計測・解析を行うため, 視線情報から注視領域を特徴づける研究は非常に重要な課題として取り組まれている. 代表的な研究例として, Salvucci らは注視点の動きを Hidden Markov Models(以下 HMMs) を用いて自動的に fixation と Saccade を推定した [68]. また, 注目点を自動的に抽出するため, 注視点の軌跡から Regions-of-Interest(以下 ROI) の抽出 [69] や注視点のクラスタリング [70] 等の研究が行われている. また, 注視点の移動は複雑で分かりにくいいため, 視線の履歴情報を画像表示に反映させたものがある. Wooding らは視線情報の履歴から頻度の高い箇所を明るく表示させる fixation map を提案した [31]. DeCarlo らは視線の履歴を写真の抽象化に利用している [32].

そこで, 本研究においても注視点の傾向をつかむため, 定量的な可視化を行う. 本研究では絵画や写真等の静的なものが対象ではなく, 刻々と変化するシーンを対象としている. そこで動的なシーンに対応する可視化手法を提案する. 以下, 本研究での可視化技術である注視頻度マップの生成手法について述べる.

画像上に投影した注視点の情報を基準に 2 次元正規分布  $g(x, y)$  を投影する. 前フレームまでの影響を考慮するため, 式 (7) に示すように前フレームの投票結果  $f_{n-1}(x, y)$  に減衰係数  $\xi$  を乗算したものを現フレームの分布に加算する.

$$\begin{aligned} f_n(x, y) &= g(x, y) + \xi f_{n-1}(x, y) \quad 0 < \xi < 1 \\ f_1(x, y) &= 0 \quad (n = 1, 2, 3 \dots) \end{aligned} \quad (7)$$





図 6 注視頻度マップ

このように可視化を行うことで，図 6 に示すように，動的なシーンに対する注視頻度のマップの生成が可能である．

## D. M 推定

点の集合から直線を推定する際に用いられる手法として最小自乗法が最も良く知られている．しかしながら，外れ値がある場合，求める直線の方程式は大きくずれが生じてしまう．実際のセンサによる入力はこのようなノイズの影響を考慮しなくてはならない．そのような際に用いられる推定法として，ロバスト推定がある．ロバスト推定には代表的なものとして M 推定 [57] や LMedS 推定がある．本研究では，注意点を求めるために M 推定を用いる．

$i$  番目のデータから推定値までの残差を  $\varepsilon_i$  とすると，最小自乗法は  $\sum_i \varepsilon_i^2$  を最小化するものであるが，M 推定では外れ値の影響を減らすため  $\sum_i \rho(\varepsilon_i)$  の最小化を行う． $\sum_i \rho(\varepsilon_i)$  の最小化は重み付き最小自乗法の繰り返し計算で行う．推定するパラメータを  $\mathbf{p} = [p_1, \dots, p_m]^T$  とすると，M 推定における  $\mathbf{p}$  の解は

$$\sum_i \Psi(\varepsilon_i) \frac{\partial \varepsilon_i}{\partial \mathbf{p}} = \mathbf{0} \quad (8)$$

となる．ここで  $\Psi(x) = d\rho(x)/dx$  は影響力関数であり，重み関数を  $\omega(x) = \Psi(x)/x$  と定義すると式 8 は

$$\sum_i \omega(\varepsilon_i) \varepsilon_i \frac{\partial \varepsilon_i}{\partial \mathbf{p}} = \mathbf{0} \quad (9)$$

となり，これは

$$\min_{\mathbf{p}} \sum_i^n \left\{ \sqrt{\omega(\varepsilon_i^{(k-1)})} \cdot \varepsilon_i \right\}^2 \quad (10)$$

表 1 代表的な M 推定法

型	$\rho(x)$	$\phi(x)$	$\omega(x)$
$L_2$	$x^2/2$	$x$	1
$L_1$	$ x $	$\text{sgn}(x)$	$\frac{1}{ x }$
$L_1 - L_2$	$2(\sqrt{1+x^2/2} - 1)$	$\frac{x}{\sqrt{1+x^2/2}}$	$\frac{1}{\sqrt{1+x^2/2}}$
$L_p$	$\frac{ x ^\nu}{\nu}$	$\text{sgn}(x)  x ^{\nu-2}$	$ x ^{\nu-2}$
Fair	$c^2 \left[ \frac{ x }{c} - \log \left( 1 + \frac{ x }{c} \right) \right]$	$\frac{x}{1+ x /c}$	$\frac{1}{1+ x /c}$
コーシー	$\frac{c^2}{2} \log(1 + (x/c)^2)$	$\frac{x}{1+(x/c)^2}$	$\frac{1}{1+(x/c)^2}$
Geman-McClure	$\frac{x^2/2}{1+x^2}$	$\frac{x}{(1+x^2)^2}$	$\frac{x}{(1+x^2)^2}$
Welsch	$\frac{c^2}{2} [1 - \exp(-x/c)^2]$	$x \exp(-(x/c)^2)$	$\exp(-(x/c)^2)$

と等価である． $k$  は繰り返し回数であり， $\varepsilon_i = 0, i = 1$  を初期値として，計算を行う．重み関数である  $\omega(\varepsilon_i^{(k-1)})$  は繰り返し更新される．このアルゴリズムは， $\rho$  関数によって必ずしも最適解に収束することが保障されないため，初期値の設定に注意が必要である．表 1 に代表的な M 推定法を示す．

## E. 平面パターンによるカメラキャリブレーション

画像上の点を  $\mathbf{m} = [u, v]^T$  3次元点を  $\mathbf{M} = [X, Y, Z]^T$  との間に式 11 が成り立つ．

$$s\tilde{\mathbf{m}} \cong \mathbf{A} [\mathbf{R} \ \mathbf{t}] \tilde{\mathbf{M}} \quad (11)$$

ここで  $\mathbf{R}$  と  $\mathbf{t}$  はそれぞれ，カメラ座標系から世界座標系への回転及び並進ベクトルであり，カメラの内部行列は以下のように設定する．

$$\mathbf{A} = \begin{bmatrix} \alpha_u & b & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

世界座標系は平面パターン上が $Z=0$ となるようにおくと，式 11 は式 12 のようになる．

$$s\tilde{\mathbf{m}} \cong \mathbf{A} [\mathbf{R} \mathbf{t}] [X, Y, 0, 1]^T = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} [X, Y, 1]^T \quad (12)$$

ここで  $\mathbf{r}_i$  は回転行列  $\mathbf{R}$  の  $i$  番目の列ベクトルである．射影変換行列を

$$\mathbf{H} \cong \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$$

と定義すると式 12 は

$$s[u, v, 1]^T \cong \mathbf{H} [X, Y, 1]^T \quad (13)$$

となり，世界座標平面からカメラ座標平面への変換を表す． $\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix}$  と表すと式 14 が得られる．

$$\begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix} \cong \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \quad (14)$$

$\mathbf{r}_1$  と  $\mathbf{r}_2$  が直行する単位ベクトルであることから，次式が得られる．

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 = 0 \quad (15)$$

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 \quad (16)$$

1 枚の画像からは式 15, 16 の内部変数に関する拘束式が得られる．行列  $\mathbf{B} = \mathbf{A}^{-T} \mathbf{A}^{-1}$  を定義し，その要素を並べたベクトルを  $\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T$  を定義すると， $\mathbf{h}_i^T \mathbf{B} \mathbf{h}_j$  は

$$\mathbf{h}_i^T \mathbf{B} \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b} \quad (17)$$

と表現することができる．ここで

$$\mathbf{h}_i = [h_{i1}, h_{i2}, h_{i3}]^T$$

$$\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i1}h_{j3} + h_{i3}h_{j1}, h_{i2}h_{j3} + h_{i3}h_{j2}, h_{i3}h_{j3}]^T$$

さらに，式 15,16 は

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0} \quad (18)$$

となる． $n$  枚の画像について，上記の式を積み重ねれば，

$$\mathbf{V}\mathbf{b} = \mathbf{0} \quad (19)$$

が得られ,  $\mathbf{b}$  は  $\mathbf{V}^T\mathbf{V}$  の最小固有値に対応する固有ベクトルとして求められる.  
 $\mathbf{B}$  が求まれば,  $\mathbf{B} = \lambda\mathbf{A}^{-T}\mathbf{A}^{-1}$  から内部変数は次のように計算される.

$$v_0 = (B_{12}B_{13} - B_{11}B_{23}) / (B_{11}B_{22} - B_{12}^2) \quad (20)$$

$$\lambda = 1 - [B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})] / B_{11} \quad (21)$$

$$\alpha_u = \sqrt{\lambda/B_{11}} \quad (22)$$

$$\alpha_v = \sqrt{\lambda B_{11} / (B_{11}B_{22} - B_{12}^2)} \quad (23)$$

$$u_0 = -B_{13}\alpha_u^2/\lambda \quad (24)$$

$$b = -B_{12}\alpha_u^2\alpha_v/\lambda \quad (25)$$

$$u_0 = bv_0/\alpha_u - B_{13}\alpha_u^2/\lambda \quad (26)$$

さらに,  $r_1, r_2$  と  $t$  も式 14 から求められる.

# 著者研究業績

## 学術雑誌論文

1. 竹村 憲太郎, 松本 吉央, 小笠原 司, “ドライバ行動解析のための非侵襲的な注視計測システムの開発”, 日本機械学会論文集, C 編, Vol.71, No.702, pp.165-170, 2005.
2. 竹村 憲太郎, 松本 吉央, 小笠原 司, “複数人の視線計測に基づく「場の注意」の推定”, ヒューマンインタフェース学会, Vol.8, No.1, pp.185-194, 2006.

## 国際会議議事録 (査読あり)

1. K.Takemura, Y.Matsumoto, and T.Ogasawara, “Estimation of Focus of Attention of Multiple People for Video Conferencing”, In *Proceedings of CHI'05 extended abstract on Human factors in computing systems*, 2005.
2. K.Takemura, H.Minamide, Y.Matsumoto, and T.Ogasawara, “What You Look at Is What You Control: A Universal Remote Control Based on Gaze Measurement Technology”, In *Proceedings of 1st IEEE Technical Exhibition Based Conference on Robotics and Automation*, 2004.
3. Y.Matsumoto, J.Ido, K.Takemura, M.Koeda, and T.Ogasawara, “Portable Facial Information Measurement System and Its Application to Human Modeling and Human Interfaces”, In *Proceedings of IEEE 6th International Conference on Face and Gesture Recognition*, pp.475-480, 2004.

4. K.Takemura , J.Ido , Y.Matsumoto , and T.Ogasawara , “Development of Non-Contact Drive Monitoring System for Advanced Safety Vehicle” , In *Proceedings of IEEE/ASME International Conference on Advanced Intelligent Mechatronics* , pp.1119-1122 , 2003.
5. K.Takemura , J.Ido , Y.Matsumoto , and T.Ogasawara , “Drive Monitoring System Based on Non-Contact Measurement System of Driver’s Focus of Visual Attention” , In *Proceedings of IEEE Intelligent Vehicles Symposium* , pp.581-586 , 2003.

## 研究会発表

1. 松本 吉央, 竹村 憲太郎, 上田 悦子, 小笠原 司, “顔・注視点計測を用いた知的コンピュータインタフェース” , 情報処理学会研究報告 , 2005-CVIM-150-5 , pp.33-38 , 2005.
2. 竹村 憲太郎, 松本 吉央, 小笠原 司, “複数人の視線計測に基づく「場の注意」の推定” , 情報処理学会研究報告 , 2004-HI-110-5 , pp.25-30 , 2004.

## 国内発表

1. 竹村 憲太郎, 松本 吉央, 小笠原 司, “ミーティングにおける注意計測システムの構築” , 第4回情報科学技術フォーラム (FIT2005) , pp.517-518 , 2005.
2. 竹村 憲太郎, 南出 隼人, 松本 吉央, 小笠原 司, “複数コンピュータ環境における注視行動に基づく操作対象の切り替え” , ヒューマンインタフェースシンポジウム 2004 , pp.1149-1152 , 2004.
3. 竹村 憲太郎, 松本 吉央, 小笠原 司, “ミーティングにおける複数人の視線計測に基づく共同注意の撮影” , 画像認識・理解のシンポジウム (MIRU2004) , pp.I-445-450 , 2004.

4. 竹村 憲太郎，松本 吉央，小笠原 司，“非侵襲ドライバ行動計測に関する研究-注視点の計測注視対象の判別-”，ロボティクス・メカトロニクス講演会 2003 講演論文集，1A1-2F-E4，2003.
5. 竹村 憲太郎，怡土 順一，松本 吉央，小笠原 司，“ドライバ注視点計測システム”，日本ロボット学会創立 20 周年記念学術講演会予稿集，3C17，2002.