

修士論文

生体情報に基づいた
題 目 グラフ描画における探索行動の
分析に関する研究

指導教員

小山田 耕二 教授

京都大学大学院工学研究科 電気工学専攻

氏名

上野 裕貴

平成31年2月1日

Analysis of Exploration Behavior in Graph Drawing Based on Physiological Information

Yuki Ueno

Abstract

To improve the visualization, it is needed to optimize the design by analyzing the users' behavior as well as to evaluate users' performance using the conventional evaluation index (e.g., correct answer rate and completion time). Although some studies have been conducted to investigate users' behavior in evaluation of visualization, most of them were directed to simple visualization tasks. A simple task does not mean a task for simple visualization consisting of a few visualization elements, but a task in which information obtained from visualization is the only clue to accomplish the task. On the other hand, there are few studies that targeted the complicated task, in which regardless of the number of elements that compose the visualization, multiple information obtained from the visualization is a clue to accomplish the task. Therefore, in this study, we investigated the behavior of participants who perform the complicated task.

We selected group-in-a-box (GIB) layouts as a target of task. GIB layout is an efficient graph-drawing method designed to visualize the group structure of graphs. GIB is composed of multiple elements. Therefore, we can obtain a lot of information from GIB. In the experiment, participants were asked to perform an exploration task specific to GIB layouts: which group has the maximum number of intra-edges? In addition to the task performance, we collected eye-tracking data.

As a result of gaze trajectory analysis, we found that participants focus on the intra-edges regardless of the size of boxes. However, the correct answer rate is influenced not only by the intra-edges but also by the size of box and the size of circle enclosing all nodes belong to the group. From this, we revealed that the participants focus on the visualization element, which is important clue of the task, but the other elements affect the task performance. Our result suggest that visualization elements that are not intended by the visualization designer can influence the task of extracting information from the data. Therefore, designers need to configure the visualization considering the character of visual cognition of the users.

目 次

第 1 章 序論	1
第 2 章 関連研究	4
2.1 可視化技術利用時のユーザの行動分析	4
2.2 生体情報を用いた評価	5
2.3 GIB レイアウト	6
2.4 生体計測手法	8
第 3 章 予備実験 1:GIB レイアウトに対する複雑なタスクの選択に関する検討	9
3.1 実験の目的	9
3.2 タスク	9
3.3 データ・レイアウトの生成	10
3.4 実験設計	11
3.5 実験環境	12
3.6 実験手順	12
3.7 被験者	12
3.8 結果	13
第 4 章 予備実験 2:被験者の内部状態を捉える生体指標の選択に関する検討	14
4.1 実験の目的	14
4.2 データ・レイアウトの生成	14
4.3 実験設計	15
4.4 実験環境	15
4.5 実験手順	16
4.6 被験者	17
4.7 分析	17
4.7.1 信号処理	17

4.7.2 特徴量抽出	18
4.8 結果	19
4.8.1 タスクパフォーマンスに影響を与える可視化要素	19
4.8.2 生体信号の分析	21
4.8.3 まとめ	23
第 5 章 実験：視線追跡データを用いた被験者の探索行動分析に関する検討	24
5.1 実験の目的	24
5.2 タスク	24
5.3 データ・レイアウトの生成	25
5.4 実験設計	28
5.5 実験環境	29
5.6 実験手順	29
5.7 被験者	30
第 6 章 実験結果と考察	31
6.1 結果	31
6.1.1 ボックスの大きさと正答率の関係	31
6.1.2 グループ内エッジの密度と正答率の関係	31
6.1.3 ボックスの大きさ・グループ内エッジの密度と正答率の関係	33
6.2 考察	33
6.2.1 ボックスの大きさが正答率へ与える影響	33
6.2.2 ボックスの有無が正答率に与える影響	33
6.2.3 円の大きさが正答率に与える影響	35
6.2.4 GIB レイアウトの違いによる正答率の比較	36
6.2.5 視線追跡データの分析	37
第 7 章 結論	39
謝 辞	41
参考文献	42

付 錄 A 生体計測の原理	47
A.1 脈波	47
A.2 脳波 (EEG)	47
A.3 瞬目	47

第1章 序論

可視化技術とは、ヒトがデータを認識できるように画像化し、その画像を通して、データの理解や、新しい発見を可能にするものである。したがって、ヒトと可視化技術との間のインタラクションが基本原理であり、可視化技術の設計および評価のためにヒトを可視化システムの一部として組み込む必要がある。すなわち、可視化技術を有効に設計及び利用するためには、計算実験から求められる評価指標やユーザ実験における正答率などのタスクパフォーマンスを向上させるだけではなく、ユーザ実験におけるユーザの行動を分析し、その結果を考慮してデザインを最適化することが必要である。そうすることでより良い可視化技術の利用と科学的発見が促進される。

ユーザの行動を記録するための方法としては、従来から用いられているプロトコル分析や、インタラクションログの分析などが存在する。近年では、視線計測技術が発展しており、専門的な知識がなくても簡単にユーザの行動を定量的に記録することが可能になった。このような背景から、可視化の分野では、可視化技術の評価にこれらの方法が取り入れられてきた。例えば、多属性意思決定を支援する可視化技術の評価¹⁾や、ノードリンクダイアグラムの評価^{2, 3)}のために用いられてきた。また、視線追跡データやプロトコル分析、インタラクションログなど、従来別々に用いられてきた方法を組み合わせて新しい評価方法も提案されている⁴⁾。

可視化技術評価の一環として、ユーザの行動を分析している研究は多いが、対象とするタスクが単純なものが多い。ここで、単純なタスクとは、少ない可視化要素から構成される可視化に対するタスクではなく、構成要素の多い、少ないに関わらず、可視化から得られる1つの情報がタスクパフォーマンスに影響を与える唯一の要因であるタスクである。一方で複雑なタスクとは、可視化要素の数に関わらず、可視化から得られる複数の情報がタスクパフォーマンスに影響を与えるタスクである。我々の知る限り、単純なタスクに対するユーザの行動分析を行った研究は存在する^{2, 3, 5)}が、複雑なタスクから情報を抽出し、タスクを遂行する際のユーザの行動について調べた研究はない。より一般化した可視化技術の設計改善を行うためにも、複雑なタスクに対する行動分析を行う必要がある。

このような背景から、本研究では複雑なタスクを行うユーザが可視化から情報を抽出する際の行動の分析を行う。複雑なタスクを設計するためには、複数の要素から構成される可視化を対象とした方が取得できる情報量が多く、複雑なタスクとなる可能性が高いため、group-in-a-box (GIB) レイアウトを本研究の対象とした。GIB レイアウトはグラフのグループ構造を可視化するために設計された効率的なグラフ描画方法である^{6, 7, 8)}。実際の応用例としては、ツイッターデータにおけるユーザの関係性や生命科学分野における生物の表現型特徴の関係性の可視化などが考えられる。様々なタイプの GIB レイアウトが提案されており、これらのレイアウトを比較、評価するために計算実験が行われてきた^{7, 8)}。GIB レイアウトは、基本的にグループを構成するノード、エッジ、それらを囲むボックス、そしてグループ間の関係を示すエッジから構成されており、そこから得られる情報が多い。また、複数の要素から構成されることに加えて、ボックスの大きさがグループに属するノードの数に比例して変化し、それに伴ってグループ内のノードをつなぐエッジの見え方が変わることから、同じレイアウトでもデータの性質によってタスクパフォーマンスに影響を与える可視化要素が変化する可能性が高い。

本研究では、GIB レイアウトからグラフ特徴を評価するタスクを行うように求められた 27 名の被験者から、視線情報を取得し、タスク実行中の探索行動について調べた。予備実験の結果を踏まえ、force-directed GIB (FD-GIB) と tree-reordered GIB (TR-GIB) をユーザ実験の対象とし、これらの GIB レイアウトに対して、最大のグループ内エッジ数を持つグループを探すという評価タスクを選択した。さらに、予備実験から本タスクにおいてユーザの探索行動を捉えるために有効な生体指標も検討した。その結果、視線追跡データがユーザの探索行動の特徴を捉えるうえで有効な生体指標であることが明らかとなった。我々は、被験者がタスクを行ったときの正答率から定義されるタスクパフォーマンスに加え、タスク遂行中の探索行動を調べるために視線追跡データを取得した。

実験の結果、被験者はグループ内エッジ数が最大のグループを探すというタスクにおいて、グループ内エッジの数あるいは密度に注目していることが視線追跡データから分かった。それにも関わらず、ボックスの大きさやグループに属するすべてのノードを囲う円といった GIB を構成する他の要素が正答率に影響を与えることが分かった。このことから、被験者は注目すべき可視化の特徴に焦点を当てていたが、それ以外の設計者が意図していない他の可視化要素がタスクパフォーマンスに影響を与えることが明らかとなった。

本稿の以降の構成は次のとおりである。まず第2章では本研究の背景と関連研究について説明する。続いて第3章で複雑なタスクを選択するための予備実験について述べる。第4章で被験者の探索行動を捉えるのに最適な生体指標を検討するための予備実験について述べる。第5章で視線追跡データを用いたユーザの探索行動を分析するための実験について述べ、第6章でその結果・考察について述べる。最後に第7章で結論を述べる。

第2章 関連研究

本研究では、可視化技術の評価に注目している。近年、計測技術や計算技術の発展により、データ量が爆発的に増加した。このようなビッグデータを工夫せずにそのまま可視化すると、大量の線や点が画面を埋め尽くし、可読性が低下して、データから情報を読み取れなくなってしまう。このような問題に対処するために、可読性が高く、理解のしやすい可視化が求められてきた。例えば、エッジの交差であるエッジクロッシングは可読性を下げてしまうため、エッジクロッシングを減らすために様々な手法が研究されてきた⁹⁾。このように、可読性を下げる可視化図の指標を向上するために、可視化図の生成アルゴリズムを工夫するという観点も重要であるが、ヒトの認知特性及び行動を対象として、ヒトを可視化システムの一部として考えることも可視化技術の向上に資する。最近では、可視化分野においてそのような研究が増えている。本章では、ヒトの行動や認知に焦点を当てた可視化技術の評価という観点から関連研究を紹介し、その後に本研究で用いるGIBレイアウト及び生体計測手法について紹介する。

2.1 可視化技術利用時のユーザの行動分析

可視化技術利用時のユーザの行動分析は、多くの研究で行われてきた^{1), 2), 3), 5), 10)}。これらの研究において、視線追跡システムは、ユーザの眼球運動を記録するためには使用してきた。視線追跡データを使用すると、ユーザが可視化技術をどのように使用したかを理解したり、ユーザの推論や問題解決の方法についての洞察を得ることさえ可能になる¹¹⁾。したがって、視線追跡データを使用することで、可視化技術の使いやすさや可読性を評価することができ、視覚認知や行動特性に焦点を当てた可視化技術の改善が可能となる。

例えば、Netzelらは地図における4つのアノテーション手法（画像内アノテーション、グリッド参照アノテーション、方向アノテーション、縮小版アノテーション）を評価した¹⁰⁾。被験者はできるだけ早く、そして正確に地図の中で指定されたラベルを見つけるよう指示された。彼らがタスクを遂行している間、視線追跡データとタスクの完了時間が記録された。その結果、画像内アノテーションは他

のどの方法よりも劣ることが分かった。そして、縮小版アノテーションにおいて完了時間が一番早かった。さらに、視線追跡データは、被験者が各アノテーション手法に対して異なるタスク戦略を実行したことを明らかにした。Burch らは、ツリーダイアグラムの従来法、直交法、放射状法の3種類を比較、評価した²⁾。被験者は典型的な階層探索タスクである、マークされた葉ノードの最下位の共通親ノードの探索を行うように求められた。その際、タスクの正答率と完了時間に加え、アイトラッカーを用いて視線追跡データが記録された。その結果、視線追跡データから、タスク実行時の探索戦略が可視化手法ごとに異なることが明らかになった。放射状法では他の手法に比べてタスク解決策の検討が頻繁に起こることが分かり、完了時間が他の手法よりも長くなった原因を明らかにした。

可視化技術利用時のユーザの行動分析は多くの研究で行われているが、単純なタスクを対象としているものが多い¹²⁾。本研究では、ユーザ実験の対象として、より複雑なタスクを設計し、GIB レイアウトをタスクの対象として選択した。GIB レイアウトでは、ノードやリンク、ボックスといった複数の可視化要素がそれぞれ何かしらの量を表している。そのため、GIB レイアウトから複数の情報を取得することができ、ユーザがそこから得た情報を用いてタスクを遂行するとき、複数の情報がタスクパフォーマンスに影響を与える。したがって、同じタスク、同じレイアウトにおいてタスクパフォーマンスが変化する可能性が高い。複雑なタスクにおいて、可視化から情報を抽出してタスクを遂行する際のユーザの行動分析を行った研究はないため、本研究の意義は大きい。

2.2 生体情報を用いた評価

本研究では、可視化評価タスクを遂行する被験者の行動を生体情報から捉えようとしている。視覚メディア処理の分野では、生体情報に基づいた QoE の評価がよく行われている¹³⁾。例えば、Moon らは、High Dynamic Range (HDR) 動画と Low Dynamic Range (LDR) 動画による QoE の比較を行った¹⁴⁾。被験者は HDR 動画と LDR 動画を視聴し、その動画についての主観的なアンケートに答えるよう求められた。また、動画視聴中の被験者から脳波、皮膚電気活動、呼吸、脈波、皮膚温度を測定し、それらの生体信号からどちらの動画を視聴していたかを識別した。その結果、脳波の γ 波が識別において重要な役割を果たすことが分かった。また、Abadi らは、感情価のある映画に対する没頭の度合いが高いか低いかを生

体信号によって識別した¹⁵⁾. 被験者は複数の短い映像を視聴し, 各映像に対して没頭の度合いを答えるよう求められた. また, 映像視聴中の被験者から脳波, 皮膚電気活動, 顔の動きを測定し, それらの生体信号から没入の度合いを識別した. その結果, それぞれの信号が識別において重要であることが分かった.

このように, 視覚メディア処理の分野では, 映像視聴中の被験者の生体信号を用いて視覚刺激に対する評価が行われている. 一方, 可視化の分野では, 利用可能な様々なタイプのモダリティのうち, 視線追跡データのみが評価に使用されている. 本研究では, 視線追跡データ以外の生体情報が可視化の評価タスクに対しても応用できると考え, 複数のモダリティの中から本研究のタスクに有効な生体指標を調べた.

2.3 GIB レイアウト

本研究でユーザ実験の対象として選んだ GIB レイアウトは, グラフ内のグループ構造を可視化するために設計されたグラフ描画手法である^{6, 7, 8)}. GIB は, グループ内のすべてのノードを, ノードの数に比例する大きさのボックス内に配置する. そのため, GIB を使用すると, グループ構造, グループ間の関係, グラフ内のグループの大きさなどを同時に可視化することができ, グラフ内のグループ構造が理解しやすい. 例えば, GIB レイアウトをツイッターデータに応用すると, 複数のユーザがノードで表され, ユーザ間のメッセージのやり取りをエッジで表すことができる. また, よくやり取りをするユーザでコミュニティが形成されると考えられ, これらのコミュニティはグループに相当するため, ツイッターにおけるユーザネットワークを可視化することができる. さらに生命科学分野において, 生物の表現型特徴のデータに応用すると, 複数の表現型特徴がノードで表され, 相関の強い表現型特徴がエッジで結ばれる. そして各表現型特徴が現れる細胞期をグループとして考えることができ, 表現型ネットワークの可視化が可能となる. 本研究では, 次の二つの理由から GIB レイアウトを評価対象とした.

- GIB レイアウトは複数の情報を可視化しているため複雑であり, 同じタスクを実行する際に, 複数の可視化要素がタスクパフォーマンスに影響を与える. そのような可視化から情報を取得する際のユーザの行動を分析した研究がないため, 複雑な可視化の 1 つとして GIB レイアウトを選択した.

- GIB レイアウトは、様々な分野に応用可能であり、より良い GIB レイアウトを作るためにも、評価することが求められる。GIB レイアウトをユーザ実験によって評価した研究がないため、本研究では GIB レイアウトを対象として選択した。

GIB レイアウトの手法はいくつかあり、以下では、本研究でユーザ実験の候補とした GIB レイアウトである ST-GIB, CD-GIB, FD-GIB, TR-GIB について説明する。各レイアウトの例は図 2.1 に示されている。

ST-GIB Squarified-Treemap GIB (ST-GIB) (図 2.1(a)) は Bruls らによる四角形のツリーマップレイアウト¹⁶⁾に基づいており、Rodrigues らによって提案された⁶⁾。ST-GIB では、各グループはそのグループに所属するノードを持つボックスとして表される。この手法では、アスペクト比の低いボックスによって空間を埋めることができると容易にできる。しかし、ボックスが配置された時のノードの関係性を考慮していないため、ネットワークの理解を妨げるエッジクロッシングが多くなりやすいという特徴を持つ^{17, 18, 19, 20)}。

CD-GIB Chaturvedi らは Croissant-and-Doughnut GIB (CD-GIB) (図 2.1(b)) を開発した⁷⁾。彼らはあるノードと他のグループのノードとのつながりに関するネットワーク情報を理解できるようにするために ST-GIB を改良した。ボックスは、G-degree と G-skewness に応じて配置される。グループの G-degree は、あるグループと繋がりのある他のグループの数として定義され、G-skewness は、G-degree が最も大きい 2 つのグループのノード数の比として定義される。G-degree と G-skewness に基づいて、レイアウトは Croissant-GIB, Doughnut-GIB、および ST-GIB から選択される。我々は Chaturvedi らによって定義された基準を使用して、3 つのレイアウトから 1 つを選んだ。Croissant-GIB, Doughnut-GIB のどちらのレイアウトも他のグループとの繋がりが最も多いボックスが中心付近に配置ため、エッジクロッシングの数は少なくなり、可読性は ST-GIB よりも良好であると予想される。しかし、ボックスのアスペクト比が悪化する傾向があり、可読性に影響を与える可能性がある¹⁶⁾。

FD-GIB Force-directed GIB (FD-GIB) (図 2.1(c)) も Chaturvedi らによって提案された⁷⁾。このレイアウトは各ボックスを中心からの重力とボックス同士の反発力によって配置するレイアウトである。このレイアウトを用いると、ボックスが重なってしまう可能性があるので、PRISM 手法²¹⁾を用いてこれら

の重なりを除去した。このレイアウトはネットワーク全体のトポロジーを示すことに長けているが、それぞれのボックスが使える面積が少ないため、グループ内の関係を読み取ることが難しいと考えられる。しかし、このレイアウトでは各ボックスのアスペクト比を一定にすることができるので、ボックスの大きさの比較は容易になる。

TR-GIB Onoue らは ST-GIB のすべてのエッジの長さが最小になるように最適化した Tree-Reordered GIB (TR-GIB) (図 2.1(d)) を提案した⁸⁾。このレイアウトの見た目は ST-GIB に似ているが、グループ間の距離が最小になるようにボックスの配置が最適化されているため、ST-GIB よりもエッジクロッシングが少ない。このレイアウトは良いアスペクト比と画面の効果的な使用という ST-GIB の利点とエッジクロッシングが少ないという利点を持つと期待されている。

このように、種々の GIB レイアウトを生成する過程で、エッジクロッシングや G-degree, G-skewness を指標として設計しているが、ユーザ実験によって、ヒトの視覚認知や行動特性といった観点から GIB を評価した研究はない。

2.4 生体計測手法

本研究では、ユーザの内部状態を測るために様々な生体情報を記録した。以下に本研究で使用した眼球運動の計測原理について述べる。脳波、脈波、瞬目の計測原理については付録 A に示す。眼球運動計測の方法は複数ある。代表例を挙げると、目の周りに電極を配置し、眼球運動に伴う皮膚表面の電位変化を計測する眼電図法や、コイルのついたコンタクトレンズを装着し、磁場中のコイルに眼球運動から誘導される電位を計測するサーチコイル法、角膜（黒目）と強膜（白目）の反射率の違いを利用し、それぞれに照射した赤外光を複数の受光素子で計測する強膜反射法、そして角膜上に光の反射点を生じさせ、その画像をカメラで撮影し、光の反射点やその他の幾何学的特徴を基に画像処理によって眼球の方向を求める角膜反射法がある²²⁾。本研究では、非接触で手軽に使用することのできる角膜反射法によって眼球運動を計測した。今回使用した視線追跡システム (Tobii Pro X3-120) では、従来の角膜反射法を改良し、画像処理に加え、眼球の生理学的 3D モデルを使用することで空間中の目の位置と視点を高精度で推定する。

第3章 予備実験1：GIB レイアウトに対する複雑なタスクの選択に関する検討

3.1 実験の目的

この実験の目的は、5章の実験で対象とするGIBレイアウトと評価タスクを選択することである。5章の実験では、複雑なタスクを設計する必要がある。そのため、さまざまなユーザ実験を予備実験として行った。4種類の評価タスクを使用して2.3節で紹介した4つのGIBレイアウトを評価した。そして、これらのGIBレイアウトをタスクパフォーマンスの観点から評価した。

3.2 タスク

我々は、Vehlowら²³⁾とSaketら²⁴⁾が用いた方法を基に評価タスクを作成した。GIBレイアウトはノード、エッジ、ボックスから構成され、それらの可視化要素から、グループの数や、グループの大きさ、グループ内外の関係を読み取ることができる。彼らはそういったGIBレイアウトの特徴を考慮して次のようにタスクを分類した。Vehlowらはグループのみに関連するタスク(GOT: Group Only Tasks)，グループと頂点(=ノード)に関連するタスク(GVT: Group Vertex Tasks)，グループとエッジに関連するタスク(GET: Group Edge Tasks)，グループ、ノード、エッジといったネットワークを構成するすべてに関連するタスク(GNT: Group Network Tasks)に分類した。GOT, GVT, GETは本研究で使用することができるが、GETはエッジがグループ化されているネットワークでしか使えないため、本研究では選ばなかった。SaketらもGOT、グループとノードに関するタスク(group node task)，グループとエッジに関するタスク(group link task)，GNTに分類した。どちらの研究でも、各タスクのいくつかの例を示しており、これらの例の中から、以下に説明する4つのタスクを選択した。

タスク1 (GOT) グラフの中にグループがいくつあるか？

タスク2 (GVT) どのグループが最も多く、あるいは最も少ないノードを持つか？

タスク 3 (Intra-GNT) どのグループが最も多い, あるいは最も少ないグループ内エッジを持つか?

タスク 4 (Inter-GNT) どのグループが最も多くのグループ間エッジを持つか?

これらのタスクをツイッターデータを可視化した GIB レイアウトにあてはめて考えると, それぞれ次のようなタスクに対応する.

タスク 1 (GOT) ツイッターネットワークの中にコミュニティがいくつあるか?

タスク 2 (GVT) どのコミュニティのユーザ数が最も多いか, あるいは最も少ないか?

タスク 3 (Intra-GNT) どのコミュニティにおいてユーザ間の交流が最も多い, あるいは最も少ないか?

タスク 4 (Inter-GNT) どのコミュニティが他のコミュニティとの交流が最も多い, あるいは少ないか?

GIB レイアウトは, グループ間およびグループ内の関係を可視化するための可視化手法であるため, グループ内エッジおよびグループ間エッジの両方を示すのに優れていなければならない. この特徴から, Intra-GNT と Inter-GNT という 2 種類の GNT タスクを行った. グループ内エッジは, 同じグループ内のノード同士を接続し, グループ間エッジは, あるノードを別のグループのノードに接続する. さらに, GIB レイアウトはグループの大きさに関わらずグループ内の特徴を可視化する. そのため, 小さいグループと大きいグループの両方の構造を示すのに適しているべきである. したがって, タスク 2, 3 では, 2 種類のタスクを設定した.

3.3 データ・レイアウトの生成

本実験では, データ生成の効率, 実データとの類似性, タスクの難易度を考慮して, Onoue ら⁸⁾ が用いた手法を参考に以下のようにデータを生成した.

- まず, グループ数を m とし, m を m_{min} から m_{max} の範囲で定義される m_{mean} と m_{stdev} というパラメータを使った正規分布に基づいてランダムに決定する. 続いて, 各グループでのノードセットを $\{V_1, \dots, V_m\}$ とする. ノードセット V_i のノード数 $|V_i|$ は, $v_{mean}, v_{stdev}, v_{min}$ というパラメータを使った正規分布に基づいてランダムに決定する. グループ V_i に属するノードのペア $(u, v) \quad \forall u, v \in V_i, u \neq v$ 間のエッジは確率 p_{in} で生成される.

2. グループ V_i と V_j 間のグループ間エッジは p_{group} の確率で生成される。そしてグループ V_i に属するノード u とグループ V_j に属するノード v の間に p_{bridge} の確率でエッジが生成される。
3. まだエッジが生成されていないノードのペア (u, v) 間のエッジは p_{out} の確率で生成される。

各パラメータの数値は、Chaturvedi ら⁷⁾が用いたツイッターデータに我々のデータを近づけるために調整した。本研究はツイッターデータに関する研究ではないが、ツイッターデータはグループ構造を有する複雑なネットワークであるため、GIB レイアウトで可視化する対象として想定される。まず、Chaturvedi らが使用したグループ、ノード、エッジの数に対応するようにパラメータを調整した。次に、そのままではノードとエッジの数が多すぎてグラフの理解が困難なため、 v_{mean} と v_{stdev} に 0.4、 p_{in} , p_{bridge} , p_{out} に 0.3 を掛けてノードとエッジの数を減らした。表 3.1 に使用したパラメータを示す。生成されたデータは、各 GIB レイアウトを適用することによって可視化された。ただし、GIB レイアウトはボックスを配置するだけであるため、各ボックス内のノード座標を決める必要がある。ボックス内にノードを配置するために、フォースレイアウトを使用した。この方法は、グラフドローリングにおいてよく用いられる手法であり、ノード間の斥力、隣接ノード間の引力、それらが属するボックスの中心からの重力によってノードが配置される。ボックス内のレイアウト方法はタスクの結果に影響を及ぼすが、フォースレイアウトはエッジクロッシングを減らし、可読性を上げる²⁵⁾ことが知られているため、本研究において採用した。また、各グループの配色はランダムに設定した。色によって心理効果が異なり、その見え方がタスクパフォーマンスに影響する可能性があるが、本研究ではランダムに配色することでその効果を取り除いた。

3.4 実験設計

本実験では、4種類の GIB レイアウトに対して、4種類のタスクを設定し、各レイアウトの各タスクに対して 30 試行用意した。試行は合計で 4 レイアウト × 4 種類のタスク × 30 試行 = 480 試行である。試行の途中でタスクの種類を変更した場合、被験者が混乱する可能性があるので、一連の試行は同じ種類のタスクで構成されている。また、同じデータからなるグラフを表示すると、被験者はデータに慣れる可能性があるので、上記の方法で問題ごとに異なるデータを示した。3.3 節

の方法に基づき、タスクごとに 120 のデータが生成され、これらのデータは 4 種類の GIB レイアウトによって可視化された。試行の順番はランダムであったため、提示される GIB レイアウトの種類はランダムであった。慣れと疲労の影響を考慮して、被験者は 4 種類のタスクをランダムな順序で行った。20 試行を 1 セットとして、被験者は 1 セットごとに約 30 秒の短い休憩を取り、各タスク（6 セット）が終わると最大 5 分の比較的長い休憩を取った。

3.5 実験環境

実験は、人工的な照明で明るくした実験室で行われた。タスクは、ASUS 社製の解像度 $1,980 \times 1,080$ ピクセルの 24 インチモニターである VX24AH（以降、単にモニターと記す）に表示された。

3.6 実験手順

実験は準備と説明、休憩を含めて 1.5~2 時間かかった。最初に、被験者に GIB レイアウトに関する説明を行った。次に、モニターから 65 cm の位置に座ってもらい、チュートリアルを行いながらタスクに関する説明をした後、本番の実験で慣れの影響が出ないように十分に訓練してもらった。実験は 3.4 節で説明したように行った。被験者には、各試行を正確に遂行するよう指示し、被験者が正解を選ぶまでに十分な時間を与えるため、各試行に対して制限時間は設定しなかった。被験者が早く答えることに集中する場合、誤答率が高くなる可能性があり、これは本実験で意図しているものではない。参加者は、クリックによって解答を選び、エンターキーを押すこと次の試行に進んだ。

3.7 被験者

被験者は正常または矯正された正常な視力を有する健常成人 20 名で、うち 12 名が男性、8 名が女性であった。被験者の年齢は 18~24 歳で平均年齢は 20.8 歳であった。被験者は可視化の研究に従事していないが、図や表から情報を読み取ることに関するリテラシーを持っていた。すべての被験者は GIB レイアウトに対する事前知識を持っていなかった。全ての被験者から事前にインフォームドコンセ

ントを得た。また、被験者には 3000 円の報酬を支払った。

3.8 結果

実験の結果を表 3.2 に示す。タスク 1 では、全てのレイアウトでの正答率の平均は 98.3 % となり、高い正答率が得られた。タスク 1 はボックスの数を数えるタスクであり、時間をかければ正解できるため、高い正答率が得られたと考えられる。そのため、ボックスの数のみが正答率に影響を与えた要素と考えられる。タスク 2 では、すべてのレイアウトでの正答率の平均は 83.3 % で、比較的高い正答率が得られた。タスク 2 は最大のノード数を有するグループを見つけるタスクであり、ノードの数に比例するボックスの面積のみが正答率に影響を与えた要素と考えられる。タスク 3, 4 では、すべてのレイアウトでの正答率の平均がそれぞれ 71.6 %, 61.1 % であった。タスク 4 の正答率はすべてのタスクの中で最低の結果であった。タスク 1, 2 についてはタスクパフォーマンスに影響を与える要素が唯一であり、単純なタスクとして選ばなかった。また、タスク 4 はグループ間エッジの数に関するタスクであり、被験者はグループ間エッジに注目する。視線追跡データを分析する際、グループ間のエッジに対して領域を定義することが困難であることからタスク 4 を選択しなかった。タスク 3 では、グループ内エッジの数だけではなく、ボックスの数や面積、グループ内エッジの見え方がタスクパフォーマンスに影響を与える可能性と、視線追跡データによる分析の際にボックスを領域として定義可能なことから、タスク 3 を選択した。また、タスク 3 ではグループ内エッジが最も多いグループを選択するタスクと最も少ないグループを選択するタスクの 2 種類を設定したが、タスクをシンプルにするためにグループ内エッジが最も多いグループを選択するタスクのみ選択した。タスク 3 の結果を詳細にみると、他のレイアウトに比べて見た目が大きく異なる FD-GIB が最も良い結果で、続いて TR-GIB の正答率が高かった。ウィルコクソンの符号順位検定の結果、タスクパフォーマンスに有意差が確認された（正答率 : $p = 0.017$; 完了時間 : $p < 0.001$ ）。同じ可視化要素を含む可視化手法でも、その配置の仕方が異なれば、タスクパフォーマンスに影響を与える要素が変化する可能性があるため、この 2 つのレイアウトを実験の対象として選択した。

第4章 予備実験2：被験者の内部状態を捉える 生体指標の選択に関する検討

4.1 実験の目的

この予備実験の目的は、予備実験1で選択したGIBレイアウトとそれに対する評価タスクを行う被験者の内部状態を取得するのに有効な生体指標を特定することである。可視化技術は本質的にヒトの視覚認知に支えられており、可視化技術の設計及び評価のためにヒトを可視化システムの一部として組み込む必要がある。本研究では、GIBレイアウトを対象としているが、GIBレイアウトはこれまで計算実験によって評価されていたのみであり、これらのレイアウトを人間の認知プロセスの観点から評価した研究はなかった。そのため、どの生体指標が本タスクにおいて有効であるかを調べる必要があった。我々は、FD-GIBとTR-GIBに対するユーザ実験を行った。被験者は、最大のグループ内エッジを持つグループを探すという評価タスクを実行するよう求められた。そして、複数の生体信号と、タスクパフォーマンスを取得し、被験者の内部状態を分析するのに重要なモダリティを調査した。

4.2 データ・レイアウトの生成

本実験では、3.3節で説明した手法を用いて生成したランダムデータを使用した。各パラメータは表4.1のように設定した。ただし、グループの数については難易度を設定するために7と14に固定してデータを生成した。データを生成した後、FD-GIBとTR-GIBを用いてデータを可視化した。ただし、GIBレイアウトの各ポックス内のノード座標は、予備実験1と同様にフォースレイアウトで決定した。また、各グループのノードの色もランダムに決定した。

4.3 実験設計

本実験では、予備実験 1 の結果からユーザ実験の対象として FD-GIB と TR-GIB を選択し、それに対する評価タスクとして、Intra-GNT を選択した。2 種類の GIB レイアウトに対して 2 つの難易度を設定し、各レイアウトの 2 つの難易度に対してタスクをそれぞれ 60 試行用意した。試行は合計で 2 レイアウト × 2 難易度 × 60 試行 = 240 試行である。難易度はグループの数によって設定し、難易度が低い試行ではグループの数を 7、難易度が高い試行ではグループの数を 14 に設定した。各レイアウトに対して、難易度が低いタスクは 60 試行を 2 セットに分け、1 セットを 30 試行とし、難易度が高いタスクは 60 試行を 3 セットに分け、1 セット 20 試行とした。難易度ごとに 1 セットの試行数を変えているが、これは難易度の違いにより、1 試行にかかる時間が異なると予測されたためである。1 セットにおける疲労の影響を考慮し、セットごとの試行数を難易度に応じて変化させた。また、同じデータを異なる GIB レイアウトで可視化した場合に、慣れの影響が発生する可能性を考えて、4.2 節の方法に基づき、試行ごとに異なるデータを生成した。さらに、慣れと疲労の影響を考慮して、被験者には 10 セットをランダムな順序で実行してもらった。被験者は、各セットの後に約 30 秒の短い休憩を取り、3 セットを行った後に長めの休憩を取った。長めの休憩は最大 5 分であり、その後、視線追跡システムのキャリブレーションを行った。また、3 セットあるいは 4 セットのタスクを開始する前に基準信号を測定するため、2 分間の開眼安静状態を測定した。

4.4 実験環境

実験は人工的な照明で明るくした実験室で行われた。重要な生体指標を特定するために、被験者が上記のタスクを行っている間、脳波、脈波、瞬目、眼球運動を測定した。脳波は、10-20 法に基づいて、リファレンスを含めた 13 チャネル (FPZ, Fz, F3, F4, Cz, C3, C4, Pz, P3, P4, Oz, A1, A2) の電極を用いて測定した。電極は、外来ノイズや電極リードの揺れなどのアーチファクトに強いアクティブ電極を使用した。頭皮上の電極の位置は、キップタイプの電極ホルダーである g.tec 社製の g.GAMMAcap (以降、単に EEG キップと記す) を用いて固定し、電極は導電性ジェルを用いて頭皮と接触させた。脈波は、心拍に関連する

生体情報を評価するために用いられた。左手人差し指の第2関節にミユキ技研社製の光学反射型の脈波センサー AP-C 030 (A) (以降、単に脈波センサーと記す)を取り付け、脈波を測定した。脳波信号は、同じくミユキ技研社製の EEG アンプである Polymate Pro 6000 を用いて記録した。Polymate Pro 6000 は 24 bit で AD 変換を行う。被験者の瞬目は、Logicool 社製の商用ウェブカメラ C920r (以降、単にウェブカメラと記す) を用いて記録した。眼の動きは、角膜反射法を用いた視線追跡システムである Tobii 社製の Tobii Pro X3-120 (以降、単にアイトラッカーと記す) を用いて記録した。タスクは、予備実験 1 と同様のモニターに表示された。また、実験中に頭部が動いて視線追跡データにノイズが入ることを防ぐために、あご台を用いて被験者の頭部の動きを抑制した。実験環境の概要を図 4.1 に示す。

4.5 実験手順

実験は準備と説明、休憩を含めて 1.5~2 時間かかった。被験者に生体信号計測に関する説明と各 GIB レイアウトに関する説明を行った。特に脳波測定に関して、体動が脳波信号に及ぼす影響を最小限に抑えるために、実験中はできるだけ動かないように指示した。次に、脳波を測定するための 13 個の電極とそれらを固定する EEG キャップおよび脈波センサーを被験者に取り付け、モニターから 65 cm の位置に座らせた。その後、チュートリアルを行いながらタスクに関する説明を行い、本番の実験で慣れの影響が出ないように十分に訓練してもらった。これらの準備が終わると、彼らには本番の実験を行ってもらった。

実験は 4.3 節で説明したように行った。被験者には、正解を選ぶまでに十分な時間を与えるために、各試行を正確に解くように指示し、各試行に対して制限時間は設定しなかった。被験者が早く答えることに集中する場合、誤答率が高くなり、混沌とした視線軌跡が得られる可能性があり、これは本実験で意図しているものではない。被験者はクリックによって解答を選び、エンターを押すことで次の問題に進んだ。クリック動作が脳波信号に及ぼす影響を最小限にするために、正解を探索した後に少ない動きで解答を行うように指示した。実験の概要を図 4.2 に示す。

4.6 被験者

被験者は正常または矯正された正常な視力を有する健常成人 6 名でうち 5 名が男性、1 名が女性であった。被験者の年齢は 21~32 歳で平均年齢は 25 歳であった。被験者の内、GIB に詳しい被験者は 1 名、可視化に詳しい被験者は 1 名で残りの被験者は GIB に関する事前知識を持っていなかった。残りの被験者は可視化の研究に従事していないが、図や表から情報を読み取ることに関するリテラシーを持っていた。全ての被験者から事前にインフォームドコンセントを得た。

4.7 分析

4.7.1 信号処理

脳波と脈波はサンプリングレート 1 kHz で測定した。脳波信号は脳波処理に特化した Matlab の Toolbox である EEGLAB²⁶⁾ を用いて処理した。1-40 Hz でバンドパスフィルターを適用し、高周波ノイズ、基線のゆらぎ及び電源ノイズを除去した。また、左右の耳朶の信号の平均を基準電極として設定し、他のすべての電極のデータから差分した。また、瞬目や目の動きによるアーチファクト信号を除去するために独立成分分析 (ICA: Independent Component Analysis)^{27, 28, 29)} を用いた。また、脈波信号は、高周波ノイズを除去するため遮断周波数 35 Hz のローパスフィルターで平滑化した後、脈拍間隔を求めた。表情は 10 fps で記録しており、ウェブカメラで撮影した顔画像から顔や目の領域を検出するために pyFace-Tracker³⁰⁾ を用いた。顔の特徴的な 66 点の座標（図 4.3）が pyFaceTracker を用いることで得られる。このうち、目の領域に関する座標を用いることで瞬きの検出を行った。視線追跡データは 120 Hz で記録しており、視線データ分析ソフトの Tobii Pro Studio* を使用して、処理した。視線の動きが注視かサッケードかを判断するために I-VT フィルタ³¹⁾ を使用した。

* <https://www.tobiipro.com/product-listing/tobii-pro-studio/>

4.7.2 特徴量抽出

脳波

脳波は、その律動の周波数帯域に応じて、 δ 波 (1-4 Hz), θ 波 (4-8 Hz), α 波 (8-13 Hz), β 波 (13-30 Hz), γ 波 (30-47 Hz) と分けられている。それぞれに対応した生理的、心理的機能や大脳皮質における発生機序について研究が行われている³²⁾。本実験では、各試行の各周波数帯域のパワーと開眼安静状態でのパワーとの比を特徴量として抽出した。まず、各タスクの各試行について刺激提示の 100 ミリ秒後から解答を選択する 500 ミリ秒前までの脳波信号に対し、512 点のサンプルウィンドウの Welch 法³³⁾を用いてパワースペクトルを計算した。得られたパワースペクトルから各電極において δ 波, θ 波, α 波, β 波, γ 波の積分値を求め、それぞれの周波数帯域について開眼安静時のパワーとの比を計算した。

脈波

人間は、精神的、身体的負荷によって自律神経系の働きが変化する。自律神経系の働きの変化は交感神経系と副交感神経系の活動のバランスとして考えることができる。例えば、ストレスや、怒りによって交感神経が刺激されると心臓の動きが促進し、心拍や脈拍は早くなる。一方で深呼吸や、安静状態になることで副交感神経が刺激されて、心臓の動きは抑制され、心拍や脈拍は遅くなる³²⁾。このことから、心拍数や脈拍数がストレスの指標として用いられることが多い。また、心拍の R-R 間隔の変動（心拍変動）を自律神経の働きを見るために用いることもある。心拍変動の時系列変化に対してパワースペクトル分解を行い、そこで得られた周波数成分を 0.04~0.15 Hz の低周波成分（LF 成分）と 0.15~0.40 Hz の高周波成分（HF 成分）に分ける。LF 成分は交感、副交感神経の活動を反映し、HF 成分は副交感神経の活動を反映しているといわれている。このことから、0.04~0.15 Hz のパワーの総和を LF 成分、0.15~0.40 Hz のパワーの総和を HF 成分とし、LF 成分と HF 成分の比率を交感神経と副交感神経の活動バランスの指標として用いられることが多い。LF/HF が増加すると交感神経と副交感神経の活動バランスが交感神経側に傾いていることを示し、LF/HF の低下はその逆を示す³²⁾。本研究では、LF/HF を特徴量として取り出した。心拍変動や LF と HF の比 (LF/HF) は脈波処理に特化した Matlab の Toolbox である HRVAS³⁴⁾を用いて求めた。

瞬目

瞬きは認知処理を反映し、ランダムに発生しない³⁵⁾ことから、タスク遂行中の被験者の瞬きのパターンと瞳孔反応から集中度の予測が行われている³⁶⁾。さらに、Brookings らはタスクが難しくなると、瞬きの頻度が下がることを明らかにした³⁷⁾。Lee らの方法³⁶⁾を参考に左右の目の上瞼と下瞼の間の距離の平均を求め（図 4.4），各被験者について閾値を設定することで瞬きが発生したかどうかを検出し、瞬きの回数を求めた。

視線

タスク遂行時の被験者の戦略を調べるために視線追跡データを取得した。視線データに関する特徴量の抽出は、AOI (Area of Interest) を基準に行われた。AOI は各ボックスに対して定義され、グループ内エッジが多いボックスから順に AOI 1 から AOI 7 まで定義された。また、AOI 7 と定義されたボックスよりもグループ内エッジが少ないものは others として定義された。図 4.5 に AOI の例を示す。今回のタスクにおいては、グループ内エッジが多いボックスを答えさせるタスクのため、AOI 1 が正解のボックスに相当する。前処理を行った視線追跡データから各 AOI に対する注視時間と注視位置を計算した。

4.8 結果

全タスクを通してのタスクパフォーマンスは、正答率が 78.1 %であり、平均完了時間は 4.29 秒であった。タスクパフォーマンスについてウィルコクソンの符号順位検定を行ったところ、予備実験 1 では、可視化手法の違いによるタスクパフォーマンスの有意差が確認されたが、本実験では有意差は確認されなかった。しかし、難易度の違いによるタスクパフォーマンスの有意差は確認された（正答率： $p = 0.031$ ；完了時間： $p = 0.031$ ）。結果を図 4.6 に示す。タスクが難しい場合、完了時間が増加し、正答率が下がった。

4.8.1 タスクパフォーマンスに影響を与える可視化要素

我々はボックスの数を変更することでタスクの難易度を変更した。しかし、被験者の主観的フィードバックから、タスクの難易度はボックスの数以外の可視化

要素によって決まる可能性があることがわかった。以下の3つの要素がタスクパフォーマンスに影響を与える可能性があると仮定した。

要素1：ボックスの数（7か14） もしボックスの数が多くなると答えの候補が多くなって正解を選ぶのが難しくなる。そのため、ボックスの数が多いと正答率は下がり、完了時間も増加する。

要素2：正解のボックスが面積最大かどうか 被験者が迷ったときに面積最大のボックスを選ぶ傾向があるとすると、正解のボックスと面積最大のボックスが一致するときに正答率が上がると考えた。逆に一致しないと、正答率が下がると考えた。

要素3：正解のボックスと正解の候補とのグループ内エッジ数の差 正解のボックスと2番目に多いグループ内エッジを持つボックスとのグループ内エッジの差が大きいほど正解を選ぶのが簡単になるとと考えた。そのため、差が小さくなると正答率は下がり、解答時間も増加する。

これら3つの要素からタスクパフォーマンスに与える影響が大きいものを見つけるために各GIBレイアウトに対して重回帰分析を行った。これら3つの要素を説明変数とし、タスクパフォーマンスの各要素（正答率、完了時間）をそれぞれ目的変数として重回帰分析を行った。FD-GIBに対する結果を表4.2, 4.3に、TR-GIBに対する結果を表4.4, 4.5に示している。表4.2, 4.4から分かるように、どちらのレイアウトにおいても完了時間は要素1によって大きく影響される（FD-GIB: $t = 11.4$, TR-GIB: $t = 10.4$ ）ことが分かった。また、表4.3, 4.5から分かるように、どちらのレイアウトにおいても正答率は要素2によって大きく影響される（FD-GIB: $t = 8.41$, TR-GIB: $t = 7.06$ ）ことが分かった。また、各説明変数の係数から、ボックスの数が少なく、グループ内エッジ数最大のボックスが面積最大のボックスであり、正解のボックスと2番目に多いグループ内エッジを有するボックスとのグループ内エッジ数の差が大きい時に、タスクパフォーマンスが良くなる（正答率が高く、完了時間が減少する）ことが分かった。

4.8.2 生体信号の分析

視線

4.7.2 節で説明したように、視線追跡データの分析は、ボックスに対して定義した AOI に基づいて行った。図 4.7 は、正解のボックスが面積最大である場合と、そうでない場合の視線追跡データを示している。図 4.7 (a), (b), (c), (d) の 1 段目は FD-GIB と TR-GIB においてボックスが 7 個の場合の例を示している。枠が赤いボックスがグループ内エッジ数最大のボックスであり、このタスクでの正解を示す。赤に続いて青、緑の枠を持つボックスが 2 番目、3 番目にグループ内エッジが多いボックスを示す。また、面が赤いボックスが面積最大のボックスであり、青、緑が 2 番目、3 番目に面積が大きいボックスを示す。そのため、辺と面が赤いボックスはグループ内エッジ数が最大で面積も最大であることを示す。ただし、この色は説明のためであり、実際のタスクではボックスに色は付いていない。図 4.7 (a), (b), (c), (d) の 2 段目は被験者の視線遷移を示しており、各被験者には異なる色が与えられる。図 4.7 (a), (b), (c), (d) の 3 段目は、全被験者の注視時間から求めたヒートマップである。

注視位置から探索行動を調べるために各 AOI に対する注視遷移行列を求めた（図 4.7 (e), (f), (g), (h)）。各列は、各 AOI から他の AOI へ視線が遷移する相対的な確率を示す。この分析方法では、難易度の違いによるボックスの数を考慮していないため、“others” とラベル付けされた AOI は “outside” として扱った。この解析の主な目的は、正解のボックスが最大面積である場合とそうでない場合の 2 条件間での探索行動の主な違いを識別することであるため、“outside” を始点として “outside” を終点とするような多くの遷移をこの方法では無視している。

遷移行列を比較することにより、以下の洞察を得た。どちらのレイアウトにおいても正解のボックスが面積最大ではない場合、AOI 2 から AOI 3 への遷移、およびその逆の遷移が高くなる。また、正解のボックスが面積最大である場合に比べて、AOI 1 から AOI 2, AOI 3 への遷移確率はより高く、AOI 2, AOI 3 から AOI 1 への遷移確率は低くなる。

この結果を元に、AOI 1, 2, 3 に対する平均注視時間を求めた。図 4.8 は FD-GIB と TR-GIB において、正解のボックスが面積最大である場合とそうでない場合の注視時間の比較結果を示している。各 AOI に対する注視時間より、どちらのレイアウトにおいても、正解のボックスが面積最大のボックスである場合には、AOI 1

が一番長く注視されている。一方、正解のボックスが面積最大のボックスでない場合、FD-GIB では AOI 2 が注視される時間が一番長く、TR-GIB では AOI 1 と AOI 3 に対する注視時間がほぼ等しくなっている。

まとめると、正解のボックスが面積最大のボックスではない場合、被験者は正解である AOI 1 以外のボックスを頻繁に比較する傾向があることが分かった。このことから、グループ内エッジ数が最大のボックスを探すというタスクにおいて、グループ内エッジの数だけではなく、ボックスの大きさといった他の要素が被験者のタスク遂行に影響を与えていた。正答率にはボックスの大きさが影響を与えるということが、被験者の探索行動の分析からも分かった。

脳波

EEG は 4.7.2 節に示されている通り処理され、各試行における δ 波 (1-4 Hz), θ 波 (4-8 Hz), α 波 (8-13 Hz), β 波 (13-30 Hz), γ 波 (30-47 Hz) のパワーが求められた。各周波数帯域のパワーが条件ごとに異なるかどうかを調べるために統計分析を行った。これらの分析は、 p 値の閾値を 0.05 に設定した並び替え検定を使用して行った。並び替え検定の結果、ボックスの数が異なるときに前頭 (Fz) の γ 波において有意な違いが見られた。図 4.9 に並び替え検定の結果を示している。ボックスの数が増えると γ 波のパワーが増加し、ボックスの数が増えると γ 波が活性化した。 γ 波は高次の認知処理を反映していると言われており^{38, 39, 40)}、前頭の γ 波は、タスクの難しさを測定するための指標として使用できる可能性がある。

脈波

我々は、脈波を用いて閉眼安静時からの LF/HF の変化を求めた。図 4.10 は、完了時間と LF/HF の変化量との関係を示す。どちらの難易度においても、有意な正の相関 (difficult : $r = 0.43, p < 0.01$, easy: $r = 0.45, p < 0.05$) が確認された。ストレス指標として用いられる LF/HF が、完了時間が長くなるにつれて増加することから、タスクを完了するのに時間がかかった場合、被験者はより多くのストレスを感じたと考えられる。LF/HF は、タスク完了時間を反映する可能性がある。

瞬目

図 4.11 は、タスクパフォーマンスと開眼安静時からの 1 分間の瞬き回数の変化量の関係を示している。開眼安静時に比べると、瞬きの回数が減る被験者が多かったが、瞬きの回数の変化量とタスク完了時間、正答率との間に有意な相関は確認されなかった。また、全体の傾向としては安静時より瞬きの数が減少する被験者多かったものの、被験者によってばらつきがある。このため、瞬きの回数の本タスクにおける有効性は確認できなかった。

4.8.3 まとめ

本実験では、次章の実験で被験者の内部状態を取得するのに有効な生体指標を特定することを目的としていた。6名の被験者が実験に参加し、GIB レイアウトの評価タスクを行った。予備実験 1 の結果から選ばれた FD-GIB と TR-GIB に対して、グループ内エッジ数が最大のボックスを選択するというタスクを行ってもらった。タスク実行中は、タスクパフォーマンスに加え生体情報（脳波、脈波、瞬目、視線）を取得した。タスクパフォーマンスに影響を与える可視化要素を調べたところ、どちらのレイアウトにおいても正解のボックスが面積最大のボックスであるかどうか正答率に大きな影響を与えることが分かった。また、生体情報を分析した結果、視線追跡データより、ユーザの探索行動が明らかとなり、正答率に影響を与えた可視化要素が探索行動にも反映されることが分かった。脳波、脈波については、データ量やタスクにかかる時間といった本タスクにおける単純な認知負荷との関連は見られるが、可視化手法の違いによる認知の差や、探索行動の違いまでを捉えるのには不十分であると考えられる。また、瞬目については今回のタスクにおける有効性を確認することができなかった。そのため、次章の実験では視線追跡データのみを計測する。また、ボックスの数は完了時間には影響を与えるものの、正答率に対する影響は小さいため、5章の実験ではボックスの数は一定にして考慮に入れないものとする。

第5章 実験：視線追跡データを用いた被験者の探索行動分析に関する検討

5.1 実験の目的

予備実験 2 より、グループ内エッジ数が最大のボックスを探すタスクにおいて、グループ内エッジの差だけではなく、ボックスの大小関係が正答率に影響することが分かった。すなわち、複数の可視化図を提示するようなシステムにおいて、ある特徴を読み取る際のパフォーマンスにそれ以外の可視化要素が影響を与える。この効果を定量化し、タスクパフォーマンスをモデル化することで複雑なタスクにおけるヒトのパフォーマンス特性を理解し、適切な可視化利用への知見を得ることが目的である。

5.2 タスク

予備実験 2 の結果から、グループ内エッジ数が最大のボックスを探すタスクにおいて、グループ内エッジの数とボックスの大小関係が正答率に影響を与えると考えられる。これらの効果を定量化し、タスクパフォーマンスをモデル化するためには、次の 3 つの場合の正答率が必要と考えらえる。

1. ボックスの大きさが影響を与える場合
2. グループ内エッジの数 or 密度が影響を与える場合
3. ボックスの大きさ、グループ内エッジの数 or 密度が影響を与える場合

以上 3 種類のデータは本タスクにおいて正答率に影響を与え得る可視化要素の有無を考慮した場合の正答率であり、これらのデータからユーザの探索行動をモデル化できると考えられる。そこで我々は、以上 3 種類のデータを取得するために以下の 3 種類のタスクを作成した。

タスク 1 ノードとボックスのみ表示した状況で、どのグループが面積最大か？(ノード数が最大か？)

タスク 2 ノードとエッジのみ表示した状況で、どのグループのグループ内エッジ数が最大か？

タスク 3 ノード・エッジ・ボックスという本来の GIB の形式で表示した状況でどのグループのグループ内エッジ数が最大か？

各タスクの例を図 5.1 に示す。

5.3 データ・レイアウトの生成

本実験では、5.2 節で説明したタスクを実現するために、グループ内のノードの配置方法と配色方法以外は、予備実験とは異なる方法でデータを生成した。3 つのタスクにおいて、データの表示の仕方は異なるが、実験条件を統一するために、すべてのタスクに対して同じ手法を用いてデータを生成した。タスク 1 についてはボックスの大きさ、タスク 2, 3 についてはボックスの大きさとグループ内エッジの数を制御する必要がある。これらの条件を制御するために、GIB レイアウトを構成する各要素を以下のように設定した。

グループ数

今回はデータ量の違いによるタスクパフォーマンスの変化を見ることは目的としているため、グループ数は予備実験 2 で使用したのと同じ、7 に固定した。

ノード数

本タスクにおいて制御しなければならないボックスの大きさは、グループに所属するノードの数によって決まる。一連のタスクにおいて、ボックスの大きさはタスクパフォーマンスに影響を与えると考えられ、正解候補となる 2 つのボックスの大きさが比較しやすい場合と比較しにくい場合を作成する。そのうえで、TR-GIB と FD-GIB で条件を同一に設定する必要があるが、TR-GIB と FD-GIB では、前者の方が紙面利用効率が高いため、各レイアウトで正解候補となる 2 つのボックスのノード数をレイアウト間で揃えても、その辺の長さの差は同じにならない。また、TR-GIB と FD-GIB で正解候補となる 2 つのボックスの辺の長さの差を揃えると、紙面利用効率が高く、辺が長い TR-GIB より、紙面利用効率が低く、辺が短い FD-GIB の方が、その差を認識しやすい可能性が高い。そこで、正解候補

となる 2 つのボックスの辺の比を両レイアウトにおいて揃えた。正解候補となる 2 つのボックスの辺の比 (Er: Edge Ratio) は、 $1 : 0.98$, $1 : 0.91$ となるように設定した。ただし、TR-GIB ではレイアウトの特性上、正解候補となる 2 つの長方形のボックスが縦に並んで配置され、共通の長辺を持つため、残りの短辺の比を表す。また、FD-GIB では、正解候補となる 2 つのボックスはレイアウトの特性上どちらも正方形のボックスとなるため、全ての辺について前述した辺の比が適用される。全体のノード数については予備実験 2 で使用したデータの全体のノード数の平均である 185 とした。正解候補となる 2 つのボックスのノード数は、 N_1, N_2 、残りのボックスのノード数を $N_3 \sim N_7$ と定義した。それぞれのノード数は以下のように設定された。

$$(N_1, N_2) = \begin{cases} (55, 54) & (\text{TR-GIB, 辺の比} = 1 : 0.98) \\ (55, 50) & (\text{TR-GIB, 辺の比} = 1 : 0.91) \\ (55, 53) & (\text{FD-GIB, 辺の比} = 1 : 0.98) \\ (55, 46) & (\text{FD-GIB, 辺の比} = 1 : 0.91) \end{cases} \quad (5.1)$$

$$4 \leq N_3, \dots, N_7 \leq N_2 - 10 \quad (5.2)$$

図 5.2 にノード数 N_1, N_2 を有する正解候補のボックスの辺の比の例を示す。ただし、 $N_3 \sim N_7$ は正解の候補に入らないようにするために、2 番目に大きいボックスの辺との比が 0.91 より小さくなるように設定してランダムに決めた。

グループ内・グループ間エッジ数

グループ内エッジ数が最大のボックスを選択するタスクにおいて、比較対象のボックスのグループ内エッジ数が多いと、人はグループ内エッジの数を数えて選択するのではなく、多く見える方、つまり密度の高い方を選ぶと考えられる。ここで、ボックスの重心を中心とし、グループに属するすべてのノードを囲む円の面積を $C \text{ pixel}^2$ 、そしてグループ内エッジの数を L とすると、グループ内エッジの密度 D は次のように定義できる。

$$D = L/C \quad (5.3)$$

タスクを遂行する際に、グループ内エッジの密度の方がタスクパフォーマンスに直接影響すると考えられるため、本実験では、タスクの難易度をグループ内エッジの数の差で調整するのではなく、グループ内エッジの密度差で調整した。ここ

で、正解候補となる2つのボックスに関して、ノード数 N_1 を有するグループのグループ内エッジ数を L_1 、ノード数 N_2 を有するグループのグループ内エッジ数を L_2 とし、それぞれのグループに属するすべてのノードを囲う円の面積を C_1, C_2 とする。グループ内エッジの密度差 ΔD は、

$$\Delta D = L_1/C_1 - L_2/C_2 \quad (5.4)$$

と表され、予備実験のデータを参考に内部密度の差の範囲を

$$-9 \times 10^{-4} \leq \Delta D \leq 9 \times 10^{-4} \quad (5.5)$$

と設定した。ここで、 ΔD が正の場合と負の場合を考えているが、これは最大の面積を有するボックス（ノード数 N_1 ）のグループ内エッジ数が最大となる場合と、2番目に大きい面積を有するボックス（ノード数 N_2 ）のグループ内エッジ数が最大となる場合の2通りを考える必要があるためである。今回のレイアウトの生成方法では、各グループのすべてのノードを囲う円はノードの数とグループ内エッジ、グループ間エッジの数によって決まり、事前にその面積を決めることができないことから、ノード数だけを固定し、エッジの数はランダムに設定して作成したデータの中から、以下の条件を満たすものを今回の実験の対象データとして選んだ。

L1 が最大の内部エッジ数である場合

$$\Delta D \approx \begin{cases} 1 \times 10^{-4} \\ 3 \times 10^{-4} \\ 5 \times 10^{-4} \\ 7 \times 10^{-4} \\ 9 \times 10^{-4} \end{cases} \quad (5.6)$$

$$L_1 > L_2 > L_3, \dots, L_7 \geq 1 \quad (5.7)$$

$$L_1/C_1 > L_2/C_2 > L_3/C_3, \dots, L_7/C_7 \quad (5.8)$$

$$|L'_1/C_1 - L'_2/C_2| \approx 1 \times 10^{-4} \quad (5.9)$$

$$20 \geq L'_1, L'_2 > L'_3, \dots, L'_7 \geq 1 \quad (5.10)$$

L_2 が最大の内部エッジ数である場合

$$\Delta D \approx \begin{cases} -1 \times 10^{-4} \\ -3 \times 10^{-4} \\ -5 \times 10^{-4} \\ -7 \times 10^{-4} \\ -9 \times 10^{-4} \end{cases} \quad (5.11)$$

$$L_2 > L_1 > L_3, \dots, L_7 \geq 1 \quad (5.12)$$

$$L_2/C_2 > L_1/C_1 > L_3/C_3, \dots, L_7/C_7 \quad (5.13)$$

$$|L'_1/C_1 - L'_2/C_2| \approx 1 \times 10^{-4} \quad (5.14)$$

$$20 \geq L'_1, L'_2 > L'_3, \dots, L'_7 \geq 1 \quad (5.15)$$

ただし、ノード数 $N_1 \sim N_7$ に対応するボックスに対して、グループに属するすべてのノードを囲う円の面積を $C_1 \sim C_7$ とし、それぞれのグループ内エッジ数を $L_1 \sim L_7$ 、グループ間エッジ数を $L'_1 \sim L'_7$ とする。グループ間エッジ数を最大でも 20 としているが、これは GIB レイアウトにおいてグループ間エッジ数が多い場合に、可読性を上げるためにエッジバンドリングというエッジを 1 本にまとめる手法がよく用いられるためグループ間エッジが多すぎて全体の可読性を下げるという状況が少ないことを考慮して設定している。図 5.3 に FD-GIB における内部エッジ数 L_1, L_2 を有する正解候補のボックスの密度差の例を示す。

データ数

可視化手法 (TR-GIB, FD-GIB)、辺の比 (1 : 0.98, 1 : 0.91)、グループ内エッジの密度差 ($-9 \times 10^{-4} \sim 9 \times 10^{-4}$) の組合せから作成されるデータのタイプは全部で 20 種類あり、タスク 1 では各タイプに対して 1 データ、タスク 2、タスク 3 では各タイプに対して 5 データ作成した。すなわち、タスク全体では、 $20 \text{ タイプ} \times (1 \text{ データ} + 5 \text{ データ} + 5 \text{ データ}) = 440 \text{ データ}$ を作成した。

5.4 実験設計

本実験では、2 種類の GIB レイアウトに対して、タスク 1 では 2 種類の辺の比を設定したため、2 種類の辺の比 $\times 10$ 試行 = 20 試行、タスク 2, 3 では 2 種類の辺

の比と 10 種類の密度差を設定したため, 2 種類の辺の比 \times 10 種類の密度差 \times 5 試行 = 100 試行 ずつ用意した. ただし, タスク 2 はデータを可視化する際にボックスを表示しないので辺の比を目で確認することはないが, タスク 2, 3 での条件を揃えるため, このようにデータを生成した. 試行は合計で, 2 レイアウト \times (20 試行 + 100 試行 + 100 試行) = 440 試行 である. タスク 1 では, 40 試行を 2 セットに分け, 1 セット 20 試行とした. タスク 2, 3 では, 100 試行を 8 セットに分け, 1 セットを 25 試行とした. 各セットでの GIB レイアウトは固定している. また, 試行の順番による影響をなくすためにセットの順番, セット内の試行の順番はすべてランダムに設定した. 被験者は, 各セットのあとに約 30 秒の短い休憩を取り, 各タスク間とタスク 2, 3 の 4 セットと 5 セットの間に長めの休憩をとった. 長めの休憩は最大 5 分であり, その後視線追跡システムのキャリブレーションを行った.

5.5 実験環境

実験は, 人工的な照明で明るくした実験室で行われた. タスクは, 予備実験で用いたのと同様のモニターに表示され, 視線はモニタ下部に取り付けられたアイトラッカーによって測定された. また, 実験中に頭部が動いて視線追跡データにノイズが入ることを防ぐために, あご台を用いて被験者の頭部の動きを抑制した. 実験環境の概要を図 5.4 に示す.

5.6 実験手順

実験は準備と説明, 休憩を含めて 1.5~2 時間かかった. まず, 被験者に生体信号計測に関する説明と各 GIB レイアウトに関する説明を行った. 次に, モニターから 65 cm の位置に座ってもらい, それぞれの本番のタスクを行ってもらう前にチュートリアルを行いながらタスクに関する説明をした後, 本番の実験で慣れの影響が出ないように十分に訓練してもらった. 実験は 5.4 節で説明したように行った. 被験者には, 各問題を正確に解くように指示し, 被験者が正解を選ぶまでに十分な時間を与えるため, 各問題に対して制限時間は設定しなかった. 被験者が早く答えることに集中する場合, 誤答率が高くなり, 混沌とした視線軌跡が得られる可能性があり, これは本実験で意図しているものではない. 参加者は, クリッ

クによって解答を選び、エンターキーを押すこと次の問題に進んだ。実験の概要を図 5.5 に示す。

5.7 被験者

被験者は正常または矯正された正常な視力を有する健常成人 27 名で、うち 21 名が男性、6 名が女性であった。被験者の年齢は 21~33 歳で平均年齢は 24 歳であった。被験者の内、GIB に詳しい被験者は 3 名、可視化に詳しい被験者は 4 名で残りの被験者は GIB に関する事前知識を持っていなかった。残りの被験者は可視化の研究に従事していないが、図や表から情報を読み取ることに関するリテラシーを持っていた。すべての被験者から事前にインフォームドコンセントを得た。

第6章 実験結果と考察

6.1 結果

6.1.1 ボックスの大きさと正答率の関係

タスク1では、ノードとボックスのみ表示した状態で、どのグループが面積最大であるかというタスクを被験者に課した。図6.1にタスク1の結果を示す。どちらのGIBレイアウトでも辺の比が小さい場合、すなわち辺の違いが大きい場合に正答率が高くなることが確認された。ウィルコクソン符号順位検定の結果、FD-GIBとTR-GIBの間に有意差は確認されなかった。

6.1.2 グループ内エッジの密度と正答率の関係

タスク2では、ノードとエッジのみを表示した状態で、どのグループのグループ内エッジ数が最大であるかというタスクを被験者に課した。図6.2にタスク2の結果を示す。どちらのGIBレイアウトでも密度差が大きくなるほど正答率が高くなった。ここで、タスク2の結果に対して心理物理学でよく用いられる累積正規分布関数へのフィッティング（プロビット回帰分析）を行うことで本実験の結果をモデル化した。以下にプロビット回帰分析の原理を説明する。まず、累積正規分布関数は以下のようにあらわすことができる。

$$y \approx n(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} dt \quad (6.1)$$

今回の実験では、 x をグループ内エッジの密度差、 y を正答率としたときに、実データに最もよくフィッティングする未知パラメータ μ と σ を求める必要がある。累積正規分布関数は単調増加関数であり、逆関数が存在する。逆関数を n_{std}^{-1} とす

ると, n と n_{std}^{-1} の関係は以下のようにあらわすことができる.

$$n(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp \left\{ \frac{-(t-\mu)^2}{2\sigma^2} \right\} dt \quad (6.2)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\frac{x-\mu}{\sigma}} \exp \left\{ \frac{-u^2}{2} \right\} \sigma du \quad (6.3)$$

$$\left(\because u = \frac{t-\mu}{\sigma} \right) \\ = n_{std} \left(\frac{x-\mu}{\sigma} \right) \quad (6.4)$$

(6.5)

今, 比較刺激の値が $\mathbf{x} = [x_1, x_2, \dots, x_N]$, そのときの正答率が $\mathbf{y} = [y_1, y_2, \dots, y_N]$ だったとする. すると x, y の関係が正規累積分布関数でフィッティングできるということは,

$$y_1 \approx n(x_1) = n_{std} \left(\frac{x_1 - \mu}{\sigma} \right) \quad (6.6)$$

ということであるから,

$$n_{std}^{-1}(y_1) \approx \frac{x_1 - \mu}{\sigma} \quad (6.7)$$

となる. n_{std}^{-1} は z 値と呼ばれ, 数表から求められる値である. すべてのデータに対して同じことをすれば,

$$\begin{aligned} n_{std}^{-1}(y_1) &\approx \frac{x_1 - \mu}{\sigma} \\ n_{std}^{-1}(y_2) &\approx \frac{x_2 - \mu}{\sigma} \\ &\vdots \\ n_{std}^{-1}(y_N) &\approx \frac{x_N - \mu}{\sigma} \end{aligned}$$

となる. ここで, $n_{std}^{-1}(y)$ を w , $1/\sigma$ を a , $-\mu/\sigma$ を b とすれば, $w = ax + b$ と考えることができ, μ, σ は単純な線形回帰で求めることができる. この方法を用いてタスク 2 で得られたデータをモデル化すると, 図 6.2 の実線のように回帰曲線がプロットできる. しかし, 適合度検定の結果, どちらの GIB レイアウトにおいても, このモデルは実測データに適合しているとは言えない (FD-GIB: $p = 0.0376$; TR-GIB: $p = 0.0028$) .

6.1.3 ボックスの大きさ・グループ内エッジの密度と正答率の関係

タスク3では、ノード、エッジ、ボックスというGIBレイアウトが持つすべての要素を表示した状態で、どのグループのグループ内エッジ数が最大であるかというタスクを被験者に課した。図6.3, 6.4にタスク3の結果を示す。どちらのGIBレイアウトにおいても密度差の絶対値が大きくほど正答率が高くなる傾向が見られた。タスク3においてもタスク2と同様のモデル化を行ったところ、図6.3, 6.4の実線のように回帰曲線が得られた。適合度検定の結果、表6.1のようになり、TR-GIBの辺の比0.98と辺の比0.91の密度差 >0 を除く条件において、このモデルは実測データに適合していないとまでは言えない。

6.2 考察

6.2.1 ボックスの大きさが正答率へ与える影響

タスク1の結果より、辺の比が小さい、すなわち辺の違いが大きくなるほど正答率が高くなることが確認された。タスク3ではボックスを表示しているため、正答率に対してボックスの影響があると考えられる。図6.5, 6.6にタスク1とタスク3で得られた結果を示す。タスク3では、どちらのGIBレイアウトでも密度差 >0 の範囲においてボックスの辺の比が小さいほど正答率が高くなっている。一方、密度差 <0 の範囲においてはボックスの辺の比が大きいほど正答率は低くなっている。タスク3で得られた実測値に対して、辺の比の条件間で正答率に差があるのか確認したところ、表6.2のようになった。有意差のある密度差において、正答率の関係を見てみると、やはり、ボックスの辺の比が小さいほど、密度差 >0 の範囲において正答率が上がっており、密度差 <0 の範囲において正答率は下がっている。このことから、正答率には密度差だけでなくボックスの大きさも影響することが分かる。密度差 >0 の範囲では、正解のボックスが面積最大であり、一方、密度差 <0 の範囲では正解のボックスは2番目に大きいボックスであるため、被験者はボックスの大きい方を解答する傾向があると言える。

6.2.2 ボックスの有無が正答率に与える影響

タスク2とタスク3では、ボックスの表示の有無があった。そのため、タスク2とタスク3を比較すれば、ボックスの有無による効果を確認できる。我々は、ボッ

クスによる効果について次のように仮説を立てた.

- タスク 3において辺の比が 0.98 のとき, すなわちボックスの辺の違いがほとんどない場合, ボックスがあることによる効果は小さいと考えられるため, タスク 3 の辺の比が 0.98 のときの正答率とタスク 2 の正答率に差はない.
- タスク 3において辺の比が 0.91 のとき, すなわちボックスの辺の違いが大きい場合, ボックスがあることによる効果は大きいと考えられるため, タスク 3 の辺の比が 0.91 のときの正答率とタスク 2 の正答率には違いが現れる. 具体的には, 密度差 > 0 においてはタスク 3の方がタスク 2よりも正答率が高くなり, 密度差 < 0 においてタスク 3の方がタスク 2の正答率よりも低くなる.

上述の仮説において, タスク 2 とタスク 3 の正答率の関係は図 6.7 のようになる. 各 GIB レイアウトに対して, タスク 2 とタスク 3 の正答率を比較した結果を図 6.8, 6.9 に示す. また, タスク 2 とタスク 3 における正答率の有意差を ウィルコクソンの符号順位検定によって調べた結果を表 6.3 に示す. FD-GIB の辺の比が 0.98 のとき, 密度差 > 0 の範囲では, 密度差 $1 \times 10^{-4}, 5 \times 10^{-4}$ において有意差が確認され, タスク 2より高い正答率となった. 一方, 密度差 < 0 の範囲では, 密度差 $-9 \times 10^{-4}, -5 \times 10^{-4}, -3 \times 10^{-4}$ において有意差が確認され, タスク 2より低い正答率となった. これは, タスク 1において辺の比が 0.98 のときでも正答率が 0.71 であったことから, わずかな辺の違いでも認識できたと考えられ, ボックスの効果によって密度差 > 0 の範囲においてタスク 2より正答率が高くなり, 密度差 < 0 の範囲においてタスク 2より正答率が低くなったと考えられる. 一方, FD-GIB の辺の比が 0.91 のとき, 密度差 -3×10^{-4} においてのみ有意差が確認されなかった. それ以外の密度差では仮説通りの結果となった. また, TR-GIB の辺の比が 0.98 のとき, 密度差 $-7 \times 10^{-4}, -3 \times 10^{-4}, 1 \times 10^{-4}, 5 \times 10^{-4}, 7 \times 10^{-4}, 9 \times 10^{-4}$ において有意差が確認された. 密度差 $5 \times 10^{-4}, 7 \times 10^{-4}, 9 \times 10^{-4}$ においてタスク 2よりも正答率が高くなった原因是, FD-GIB と同じくボックスの大きさによる影響だと考えられる. タスク 1において, TR-GIB の辺の比が 0.98 のときの正答率は 0.67 であったため, こちらもわずかな辺の差を認識できたと考えることができる. 一方, 密度差 $-7 \times 10^{-4}, -3 \times 10^{-4}$ においては, タスク 2よりも正答率が高くなり, 密度差 1×10^{-4} においては, タスク 2よりも正答率が低くなつた. これはボックスによる効果とは逆であり, ボックス以外の可視化要素が影響を

与えていると考えられる。TR-GIB の辺の比が 0.91 のときは、密度差 -7×10^{-4} , -1×10^{-4} において有意差が確認されなかった。それ以外の密度差では仮説通りの結果となった。FD-GIB の辺の比 0.91 のときの密度差 -3×10^{-4} , TR-GIB の辺の比 0.91 のときの密度差 -7×10^{-4} , -1×10^{-4} において有意差が確認できなかつた原因についてもボックスの効果では説明できないため、ボックス以外の可視化要素の影響を受けたと考えられる。影響を与えた可視化要素として考えられるのが、グループに属するすべてのノードを囲む円の大きさである。円の大きさの効果については次項で説明する。

6.2.3 円の大きさが正答率に与える影響

ノードを囲む円の面積はノードの数に比例するボックスの面積とは独立に変化する。そのため、本タスクで扱うデータには、以下のような 6 つの状況が存在する。

1. 正解のボックスの面積が最大で、ノードを囲む円の面積も最大
2. 正解のボックスの面積が最大で、ノードを囲む円の面積は 2 番目に大きい
3. 正解のボックスの面積が最大で、ノードを囲む円の面積は正解候補の円の面積と同じ
4. 正解のボックスの面積は 2 番目に大きくて、ノードを囲む円の面積は最大
5. 正解のボックスの面積は 2 番目に大きくて、ノードを囲む円の面積も 2 番目に大きい
6. 正解のボックスの面積は 2 番目に大きくて、ノードを囲む円の面積は正解候補の円の面積と同じ

以上 6 つの状況とそれが正答率に及ぼすと考えられる影響をまとめたものを図 6.10 に示す。ボックスの大きさの影響は 6.2.1 項で述べたように、密度差 > 0 の範囲では正答率を上げる効果があるが、密度差 < 0 の範囲では正答率を下げる効果がある。これは被験者がボックスの大きい方を選択する傾向があるからであった。円の面積が正答率に与える影響についても同様のことが言えるはずである。今回使用したデータについて、正解のグループの円の半径を R1, 2 番目に多

いグループ内エッジを持つグループの半径を R_2 , 比較対象となるグループの円の半径の比 (Rr: Radius Ration) を R_2/R_1 と定義したところ, タスク 2 の結果に対して, 円の半径の比と正答率の関係は図 6.11, 6.12 のようになつた. この結果から, 半径の比 < 1 , すなわち正解のグループの円が大きいときの方が, 半径の比 > 1 , すなわち正解のグループの円が小さいときより正答率が高い傾向にあることが分かる. この傾向から, 正答率に対して円の大きさの効果があることが確認できる. 円の大きさの効果があることから, 6.2.1 項で述べたボックスの大きさによる効果は円の大きさの効果も含んでおり, タスク 3 での辺の違いによる正答率の差は, ボックスの大きさによるものと断定することはできない. そこで, 各 GIB レイアウトのタスク 2, 3 において円の半径の比によって試行を分類し, 各条件での正答率を求めた. 結果を図 6.13, 6.14 に示す. また, 各条件の試行数は制御できおらず条件によって試行数が異なるため, タスク 2, 3 それぞれにおける, 各条件の試行数を表 6.4, 6.5 に示す. タスク 2 において, どちらの GIB レイアウトでも半径の比 > 1 と半径の比 < 1 の場合を比べると, FD-GIB の密度差 1×10^{-4} を除いて, 半径の比 > 1 , すなわち, 正解のグループの円が大きいときの方が正答率が高くなつておる, 円の効果によって正答率が高くなる傾向にあることが分かる. FD-GIB の密度差 1×10^{-4} の試行数は表 6.4 から, 1 試行であることから, この試行で偶然正答率が高くなつたと考えられる. また, タスク 3 において同じ円の特徴を持つデータを比べると, 密度差 > 0 において, 辺の比が小さい方, すなわち辺の違いが認識しやすい場合に正答率が高くなる傾向があり, 密度差 < 0 において, 辺の比が小さい方が正答率が低くなる傾向があるため, ボックスの大きさによる効果がみられる. タスク 2 とタスク 3 の結果を比べてボックスの有無による違いを見ると, 表 6.5 の赤くハイライトされた部分でボックスの有無の効果の傾向が確認できるが, 確認できなかつたデータも多くあつた. 今回の実験では, 円の大きさを制御できていないため, 各データのサンプリング数が異なり, 統計的にボックスの大きさの効果, 円の大きさの効果, ボックスの有無の効果について議論することはできないが, データの傾向としてはそれぞれの効果が確認できる部分があつた. 今後, 円の大きさの条件も制御してタスクを設計する必要がある.

6.2.4 GIB レイアウトの違いによる正答率の比較

タスク 3 において, FD-GIB と TR-GIB の正答率に有意差があるのか調べた. 図 6.15 にタスク 3 における各 GIB レイアウトの正答率を比較した結果を示す. ま

た, 表 6.6 に各密度差において, GIB レイアウトの違いによる正答率の有意差を ウィルコクソンの符号順位検定によって調べた結果を示す. 密度差 < 0 の範囲では, 辺の比が 0.98 で密度差が -1×10^{-4} の場合と辺の比が 0.91 で密度差が -3×10^{-4} の場合を除いて, どちらの辺の比においても, TR-GIB の方が FD-GIB よりも正答率が有意に高くなった. 一方, 密度差 > 0 の範囲では, 辺の比が 0.98 で密度差が 1×10^{-4} のときに FD-GIB の方が TR-GIB よりも正答率が高くなつたが, それ以降は TR-GIB の方が正答率が高くなつた. また, 辺の比が 0.91 のときは, 密度差が 1×10^{-4} の場合のみで有意差が確認され, FD-GIB の方が正答率が高くなつた. 全体としては TR-GIB の方が正答率が高い傾向にあるが, これは, TR-GIB の方が紙面利用効率が高く, ボックスの面積が FD-GIB に比べると大きくなるため, グループ内エッジが見やすいからだと考えられる. 一部の辺の比, 密度差の組合せにおいて, FD-GIB の方が TR-GIB に比べて正答率が高くなつたが, これは 6.2.3 項で述べた円の大きさの効果による影響だと考えられる. また, 予備実験 1においては FD-GIB の正答率が高く, 予備実験 2においては可視化手法間での正答率の差が確認できなかつた. 予備実験 1, 2 では, 本実験で制御している密度差, ボックスの面積が制御できていないことに加えて, ボックスの数, 円の大きさによる効果が加わってくるため, 実験ごとに可視化手法間の優劣が変化したものと考えられる.

6.2.5 視線追跡データの分析

本実験では, タスクを遂行する被験者から視線追跡データを取得した. 予備実験 2 と同様にボックスに対して AOI を設定する必要がある. 本実験は予備実験 2 とは異なり, 正解の候補を 2 つに絞れるようにタスクを設計しているため, 2 つの正解候補に対する視線追跡データのみを分析の対象とした. AOI は図 6.16 のよう に決めた. それぞれのボックスに対して “in” と “out” という AOI を定義する. “in” は, グループに属するノードをすべて囲む円の領域であり, “out” はボックスの辺に対して 13.5 pixel 余裕を持たせた領域の内, “in” を除く範囲となつて いる. “out” の範囲をボックスよりも少し大きくしているが, これはタスク 1 において明らかにボックスの辺を比べているときに, 注視位置が辺の周辺に分布してお り, ボックスの領域を “out” としまうと辺の長さを比べているときの視線データが取得できなくなるためである. 本実験では, 視線追跡データから各タスクを行つ被験者がボックスの大きさに注目するのか, 内部エッジの数もしくは密度に

注目するのかを分析する。被験者がボックスの大きさを判断基準とする場合には，“out”の領域に対する注視時間が増加し、グループ内エッジの数あるいは密度に注目する場合は、“in”の領域に対する注視時間が増加すると考えられる。27名の被験者の内、6名のデータは実際の注視位置との誤差が大きかったため、この分析の対象外とした。各タスクの各試行において、開始100ミリ秒後から解答を確定して次の問題に移るまでの時間を対象として、被験者が“in”と“out”を見た時間に対する“in”を見た時間の割合を求めた。そして全被験者で平均値を求めたところ、図6.17と図6.18のようになった。どちらのGIBレイアウトでもタスク2,3に比べるとタスク1での“in”を見た時間の割合が下がっており、被験者がボックスの大きさを比較するときに、辺に注目することが分かる。タスク1において、辺の比の違いによる“in”を見た時間の割合の有意差は確認されなかった。タスク2,3において“in”を見た時間の割合は高くなっている、グループ内エッジの数もしくは密度に注目していることが分かる。どちらのGIBレイアウトにおいても、タスク2とタスク3の結果の有意差は確認できなかった。また、タスク3において辺の比の違いによる有意差はTR-GIBの密度差 -3×10^{-4} , 3×10^{-4} のみで確認された($-3 \times 10^{-4} : p = 0.0437$; $3 \times 10^{-4} : p = 0.0437$)。タスク3ではボックスが表示されるため、タスク2に比べると“in”を見た割合が下がると予想していたが、そのようにはならなかった。タスク2,3の結果から、被験者はボックスの有無にかかわらずグループ内エッジの密度あるいは数に注目していることが分かる。また、タスク3において一部の条件を除いて辺の比による有意差が確認されなかったことから、グループ内エッジに注目していてもボックスの大きさの影響を受けることが分かる。被験者はボックスに集中するという戦略をとっているのではなく、グループ内エッジの数、あるいは密度に集中する戦略をとっているが、周辺視野によってボックスの大きさの情報が情報として得られるため、ボックスの大きさの効果が正答率に現れたと考えることができる。

第7章 結論

本研究では、生体情報を基に、複雑な可視化の評価タスクを行う被験者の探索行動について分析した。可視化技術の評価タスクを遂行する被験者の行動分析に関する研究はこれまで行われてきたが、可視化図から得られる複数の情報がタスクパフォーマンスに影響を与える複雑な評価タスクに対する、被験者の探索行動について定量的に分析した研究はない。複数の可視化要素から構成される2種類のGIBレイアウトに対して、グループ内エッジが最大のボックスを探すというタスクを27名の被験者に行ってもらった。予備実験から有効性を検証した視線追跡データと正答率を基に被験者の探索行動を分析した。

予備実験から、本来読み取るべき可視化の特徴以外の可視化要素が正答率に影響を与えていることが示唆された。そのことを検証するために、本実験において3種類のタスクを行った。グループ内エッジ数が最大のボックスを探すというタスクにおいて、被験者はボックスが表示された場合でも、グループ内エッジの数または密度に焦点を当ててタスクを行っていることが分かった。それにも関わらず、ボックスの大きさによる正答率の変化が傾向として確認された。そのため、被験者は本タスクにおいてボックスの大きさには焦点を当てていなかったが、周辺視野によってボックスの大きさに関する情報を得て、その情報が正答率に影響したことが明らかとなった。また、ボックスの有無による正答率を比較したところ、グループに属するすべてのノードを囲う円の大きさが正答率に影響することが示唆された。

GIBレイアウトのような複数の可視化要素から構成される可視化図から、目的の情報を抽出する際に、タスク達成のための一番重要な要素にユーザが注目していたとしても、探索時に他の可視化要素の情報も視覚情報として取得される。そして、可視化図の設計者が意図していなかったこれらの可視化要素が正答率に影響するということが本研究から明らかとなった。そのため、可視化図の設計者は、ヒトの視覚認知の特性をよく考えて可視化図を設計しなければならない。本研究で扱ったGIBレイアウトに対して、意図していない可視化要素からの影響を避けるための改善案としては、ユーザインタラクションを加えて、ユーザが知りたい情報に応じてデータの見せ方を変えることが考えられる。そして、ユーザがグルー

内エッジの数を見ることを支援するために、各グループのノードの色をグループ内エッジの数に応じたヒートマップによって付けることが一例として挙げられる。これらの改善は、ツイッターデータにおいてはユーザの交流が多いグループの発見、生命科学分野での表現型特徴の可視化においては、同一細胞期内で相関する表現型特徴のペアが多い細胞期を発見するのに役立つだろう。

今後は、円の大きさを制御したタスクを行い、ボックスの有無による効果、ボックスの大きさによる効果を統計的に分析したいと考えている。本研究で対象としたGIBレイアウトは、複雑な可視化の一例であり、GIBレイアウト以外の可視化図に対してタスクを設計することができれば、様々なケースでヒトの多くの行動やパフォーマンスを定量化でき、可視化技術の設計に有益な知見が得られるであろう。

謝 辞

本研究を行うにあたって、京都大学学術情報メディアセンターの小山田耕二教授から熱心な研究指導を賜りました。研究の問い合わせの立て方から論文の書き方にいたるまで、丁寧にご指導いただいたことを心より感謝いたします。実験環境の整備や生体情報の扱い方等、多岐にわたる指導をいただき、研究指導のために多くの時間を割いていただきました京都大学学術情報メディアセンターの夏川浩明特定助教に深く感謝致します。また、本研究を進めるにあたり有益な御指導、御助言を頂いた京都大学学術情報メディアセンターの江原康生特定准教授、そして年末報告会にて有益な御助言をいただきました日本大学文理学部情報科学科の尾上洋介助教に深く感謝致します。本研究で行った実験のデータ作成を手伝っていただきました青山望くんに深く感謝致します。そして、本実験の被験者として手伝っていただいた方々に感謝致します。また、日々の研究を進めていく上で多くの協力をしていただいた小山田研究室の皆様に感謝の意を表します。

参 考 文 献

- 1) Sung-Hee Kim, Zhihua Dong, Hanjun Xian, Benjavan Upatising, and Ji Soo Yi, Does an eye tracker tell the truth about visualizations?: findings while investigating visualizations for decision making, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, (2012), pp. 2421–2430.
- 2) Michael Burch, Natalia Konevtsova, Julian Heinrich, Markus Hoeferlin, and Daniel Weiskopf, Evaluation of traditional, orthogonal, and radial tree diagrams by an eye tracking study, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17, No. 12, (2011), pp. 2440–2448.
- 3) Rudolf Netzel, Michel Burch, and Daniel Weiskopf, Comparative eye tracking study on node-link visualizations of trajectories, *IEEE transactions on visualization and computer graphics*, Vol. 20, No. 12, (2014), pp. 2221–2230.
- 4) Tanja Blascheck, Markus John, Kuno Kurzhals, Steffen Koch, and Thomas Ertl, Va 2: a visual analytics approach for evaluating visual analytics applications, *IEEE transactions on visualization and computer graphics*, Vol. 22, No. 1, (2016), pp. 61–70.
- 5) Weidong Huang, Peter Eades, and Seok-Hee Hong, Beyond time and error: a cognitive approach to the evaluation of graph drawings, *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, (ACM, 2008), p. 3.
- 6) Eduarda Mendes Rodrigues, Natasa Milic-Frayling, Marc Smith, Ben Shneiderman, and Derek Hansen, Group-in-a-box layout for multi-faceted analysis of communities, *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, (IEEE, 2011), pp. 354–361.
- 7) Snigdha Chaturvedi, Cody Dunne, Zahra Ashktorab, R Zachariah, and Ben Shneiderman, Group-in-a-box meta-layouts for topological clusters and

- attribute-based groups: Space-efficient visualizations of network communities and their ties, *Computer Graphics Forum*, Vol. 33, (Wiley Online Library, 2014), pp. 52–68.
- 8) Yosuke Onoue and Koji Koyamada, Optimal tree reordering for group-in-a-box graph layouts, *SIGGRAPH Asia 2017 Symposium on Visualization*, (ACM, 2017), p. 13.
 - 9) Hong Zhou, Panpan Xu, Xiaoru Yuan, and Huamin Qu, Edge bundling in information visualization, *Tsinghua Science and Technology*, Vol. 18, No. 2, (2013), pp. 145–156.
 - 10) Rudolf Netzel, Marcel Hlawatsch, Michel Burch, Sanjeev Balakrishnan, Hansjrg Schmauder, and Daniel Weiskopf, An evaluation of visual search support in maps, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 23, No. 1, (2017), pp. 421–430.
 - 11) Gennady Andrienko, Natalia Andrienko, Michael Burch, and Daniel Weiskopf, Visual analytics methodology for eye movement studies, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, (2012), pp. 2889–2898.
 - 12) D. A. Szafir, Modeling color difference for visualization design, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 24, No. 1, (2018), pp. 392–401.
 - 13) Ulrich Engelke, Daniel P Darcy, Grant H Mulliken, Sebastian Bosse, Maria G Martini, Sebastian Arndt, Jan-Niklas Antons, Kit Yan Chan, Naeem Ramzan, and Kjell Brunnström, Psychophysiology-based qoe assessment: a survey, *IEEE Journal of Selected Topics in Signal Processing*, Vol. 11, No. 1, (2017), pp. 6–21.
 - 14) Seong-Eun Moon and Jong-Seok Lee, Perceptual experience analysis for tone-mapped hdr videos based on eeg and peripheral physiological signals, *IEEE Transactions on Autonomous Mental Development*, Vol. 7, No. 3, (2015), pp. 236–247.
 - 15) Mojtaba Khomami Abadi, Jacopo Staiano, Alessandro Cappelletti, Massimo Zancanaro, and Nicu Sebe, Multimodal engagement classification for affective

cinema, *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, (IEEE, 2013), pp. 411–416.

- 16) Mark Bruls, Kees Huizing, and Jarke J Van Wijk, Squarified treemaps, *Data visualization 2000*, (Springer, 2000), pp. 33–42.
- 17) Richard A. Becker, Stephen G. Eick, and Allan R. Wilks, Visualizing network data, *IEEE Transactions on visualization and computer graphics*, Vol. 1, No. 1, (1995), pp. 16–28.
- 18) Helen Purchase, Which aesthetic has the greatest effect on human understanding?, *International Symposium on Graph Drawing*, (Springer, 1997), pp. 248–261.
- 19) Helen C Purchase, Performance of layout algorithms: Comprehension, not computation, *Journal of Visual Languages & Computing*, Vol. 9, No. 6, (1998), pp. 647–657.
- 20) Helen C Purchase, David Carrington, and Jo-Anne Allder, Empirical evaluation of aesthetics-based graph layout, *Empirical Software Engineering*, Vol. 7, No. 3, (2002), pp. 233–255.
- 21) Emden R Gansner and Yifan Hu, Efficient node overlap removal using a proximity stress model, *International Symposium on Graph Drawing*, (Springer, 2008), pp. 206–217.
- 22) 橋村勝, 飯塚博実, 李軍, 特集3 人間工学のための計測手法 第4部:生体電気現象その他の計測と解析(2), 人間工学, Vol. 51, No. 6, (2015), pp. 406–410.
- 23) Corinna Vehlow, Fabian Beck, and Daniel Weiskopf, Visualizing group structures in graphs: A survey, *Computer Graphics Forum*, Vol. 36, (Wiley Online Library, 2017), pp. 201–225.
- 24) Bahador Saket, Paolo Simonetto, and Stephen Kobourov, Group-level graph visualization taxonomy, *arXiv preprint arXiv:1403.7421*, (2014).
- 25) Stephen G. Kobourov, 12 force-directed drawing algorithms, (2004).

- 26) Arnaud Delorme and Scott Makeig, Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis, *Journal of neuroscience methods*, Vol. 134, No. 1, (2004), pp. 9–21.
- 27) Anthony J Bell and Terrence J Sejnowski, Blind separation and blind deconvolution: an information-theoretic approach, *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, Vol. 5, (IEEE, 1995), pp. 3415–3418.
- 28) Te-Won Lee, Mark Girolami, and Terrence J Sejnowski, Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources, *Neural computation*, Vol. 11, No. 2, (1999), pp. 417–441.
- 29) S. Amari, A. Cichocki, and H. H. Yang, A new learning algorithm for blind signal separation, *Proceedings of the 8th International Conference on Neural Information Processing Systems*, NIPS'95, (MIT Press, Cambridge, MA, USA, 1995), pp. 757–763.
- 30) Jason M Saragih, Simon Lucey, and Jeffrey F Cohn, Deformable model fitting by regularized landmark mean-shift, *International Journal of Computer Vision*, Vol. 91, No. 2, (2011), pp. 200–215.
- 31) Anneli Olsen, The tobii i-vt fixation filter, *Tobii Technology*, (2012).
- 32) 市原茂, 阿久津洋巳, 石口彰, 視覚実験研究ガイドブック, (朝倉書店, 2017).
- 33) Peter Welch, The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, *IEEE Transactions on audio and electroacoustics*, Vol. 15, No. 2, (1967), pp. 70–73.
- 34) John T Ramshur, *Design, evaluation, and application of heart rate variability analysis software (HRVAS)*, PhD thesis, (University of Memphis, 2010).
- 35) John A Stern, Donna Boyer, and David Schroeder, Blink rate: a possible measure of fatigue, *Human factors*, Vol. 36, No. 2, (1994), pp. 285–297.

- 36) Giyoung Lee, Amitash Ojha, and Minho Lee, Concentration monitoring for intelligent tutoring system based on pupil and eye-blink, *Proceedings of the 3rd International Conference on Human-Agent Interaction*, (ACM, 2015), pp. 291–294.
- 37) Jeffrey B Brookings, Glenn F Wilson, and Carolyne R Swain, Psychophysiological responses to changes in workload during simulated air traffic control, *Biological psychology*, Vol. 42, No. 3, (1996), pp. 361–377.
- 38) Lucia Melloni, Carlos Molina, Marcela Pena, David Torres, Wolf Singer, and Eugenio Rodriguez, Synchronization of neural activity across cortical areas correlates with conscious perception, *Journal of neuroscience*, Vol. 27, No. 11, (2007), pp. 2858–2865.
- 39) Friedemann Pulvermüller, Hubert Preissl, Werner Lutzenberger, and Niels Birbaumer, Brain rhythms of language: nouns versus verbs, *European Journal of Neuroscience*, Vol. 8, No. 5, (1996), pp. 937–941.
- 40) HT Tiitinen, J Sinkkonen, K Reinikainen, Kimmo Alho, J Lavikainen, and R Näätänen, Selective attention enhances the auditory 40-hz transient response in humans, *Nature*, Vol. 364, No. 6432, (1993), p. 59.
- 41) 富田豊, 特集 3 人間工学のための計測手法, 人間工学, Vol. 52, No. 1, (2016), pp. 1–5.
- 42) John Allen, Photoplethysmography and its application in clinical physiological measurement, *Physiological measurement*, Vol. 28, No. 3, (2007), p. R1.
- 43) David Cristinacce and Timothy F Cootes, Feature detection and tracking with constrained local models., *Bmvc*, Vol. 1, (Citeseer, 2006), p. 3.

付録 A 生体計測の原理

A.1 脈波

生体組織を透過する光は皮膚、骨、脂肪、血液などの物質によって吸収されるが、このうち、皮膚、脂肪、骨での光吸収は一定値を示し、吸収量が時間変化することがない。血流のほとんどの変化は動脈で起こり、動脈中の血液量は心拍によって変動する。動脈の血液内には、ヘモグロビンが存在し、入射光を吸収する特性があるため、反射光、あるいは透過光の時間変化を計測することで、脈波を間接的に測定することができる^{41, 42)}。本研究では、脳波用の EEG アンプに対応した、脈波センサー（図 A.1）を利用し、サンプリング周波数 1 kHz で測定した。この脈波センサーは発行部と受光部が一体となった脳波電極サイズのもので、波長 525 nm の緑色波長の光を当て、その反射光を受光することで、脈波信号を得る。

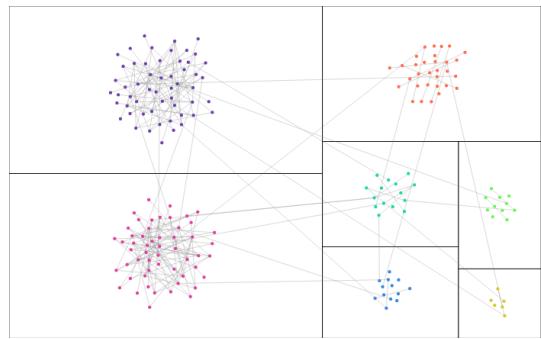
A.2 脳波 (EEG)

脳波は、大脳の神経細胞群の電気活動を体表に装着した電極を用いて記録する。頭皮上で観察しているものは、主に大脳皮質の電気活動であるが、深部の活動も一部反映される。電極で得られた信号は、生体アンプにかけて増幅し、ノイズを除去し、フーリエ解析によって δ 波、 θ 波、 α 波、 β 波といった基礎律動に分解することができる。本実験で使用した Polymate Pro MP6000 は、サンプリング周波数 1 kHz で、24 bit の AD 変換、外来ノイズや電極リードの揺れなどのアーチファクトに強いアクティブ電極を使用している。

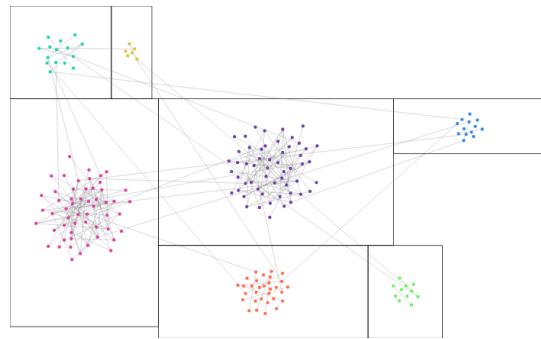
A.3 瞬目

本研究では、ウェブカメラから得られた画像から顔の特徴点を検出し、目の周辺の情報を用いることで瞬目を検出した。顔の特徴点をとらえるために、pyFace-Tracker^{30, 43)} を用いた。この手法は、Constrained Local Model(CLM) を用いた顔特徴追跡手法の一つである。CLM はモデルベースの特徴点検出手法であり、統計情

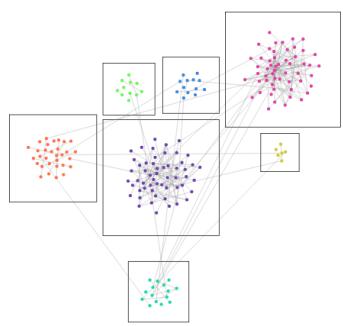
報を用いて検出対象をモデル化することで変形に伴う対象も検出可能となる。CLMは人物ごとにモデルを作成することなく、不特定の人物に対して特徴点検出が可能である。



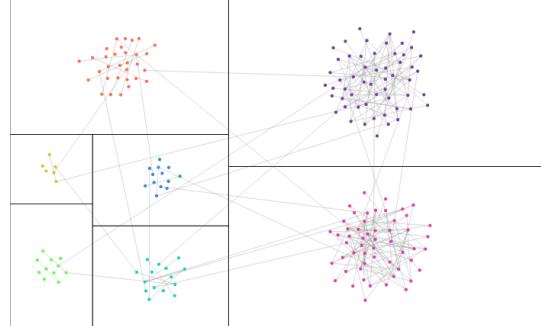
(a) ST-GIB



(b) CD-GIB



(c) FD-GIB



(d) TR-GIB

図 2.1: Group-In-a-Box(GIB) レイアウトの例

表 3.1: 予備実験 1 のデータ生成に用いたパラメータ

m_{mean}	m_{stdev}	m_{min}	m_{max}	v_{mean}	v_{stdev}	v_{min}	p_{in}	p_{group}	p_{bridge}	p_{out}
11.4	5.4	6	17	21	14.12	4	0.0858	0.06	0.015	0.0006

表 3.2: 実験結果：正答率（平均完了時間）

	Task 1	Task 2	Task 3	Task 4
ST-GIB	98.1% (3.987 s)	89.6% (2.535 s)	67.4% (3.606 s)	62.8% (5.387 s)
CD-GIB	98.2% (4.527 s)	76.9% (2.761 s)	67.2% (3.843 s)	59.3% (5.588 s)
FD-GIB	98.7% (4.919 s)	82.9% (2.427 s)	78.8% (3.382 s)	59.3% (5.518 s)
TR-GIB	98.1% (4.493 s)	83.6% (2.713 s)	72.8% (3.845 s)	62.8% (5.128 s)
AVERAGE	98.3% (4.48 s)	83.3% (2.61 s)	71.6% (3.67 s)	61.6% (5.41 s)

表 4.1: 予備実験 2 のデータ生成に用いたパラメータ

v_{mean}	v_{stdev}	v_{min}	p_{in}	p_{group}	p_{bridge}	p_{out}
21.0	14.12	4	0.0858	0.06	0.015	0.0006



図 4.1: 予備実験 2 の実験環境. 電極を取り付けた EEG キャップと脈波センサーが被験者に取り付けられた. また, 視線の動きはモニター下部に取り付けられたアイトラッカーを用いて記録した. さらに, モニター上部に取り付けられたウェブカメラを用いて被験者の顔を記録した.

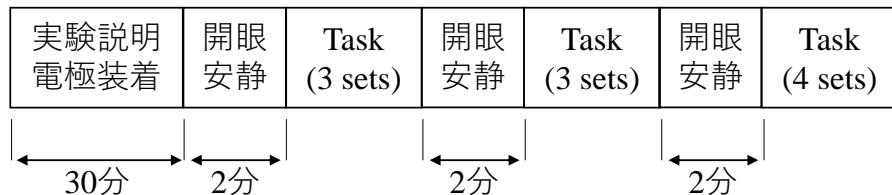


図 4.2: 予備実験 2 の概要

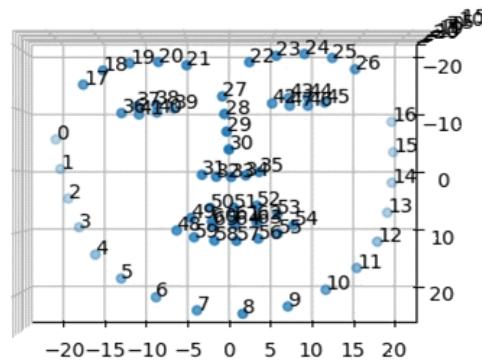


図 4.3: 顔に関する特微量の座標

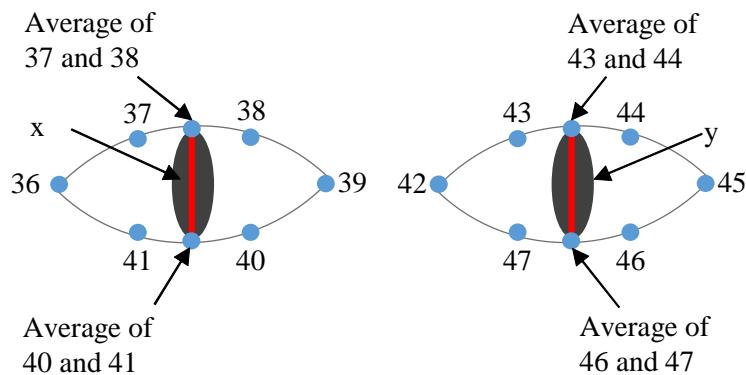


図 4.4: 瞳き回数の抽出. 図中 x は左目, y は右目の上瞼と下瞼の距離を示している. x, y の平均値を計算し, 瞳きの検出を行った.

表 4.2: FD-GIB における完了時間に対する重回帰分析の結果

completion time Variables	Unstandardized Coefficients		Standardized Coefficients	t-value	p-value
	B	Std. Error	Beta		
constant	3.04	0.2888	-	10.5	0
factor1	0.205	0.018	0.717	11.4	0
factor2	-0.575	0.236	-0.575	-2.43	0.0165
factor3	-0.0141	0.00233	-0.369	-6.06	0

$$R^2 = 0.611, R^2_{adj} = 0.601, SEE = 0.689$$

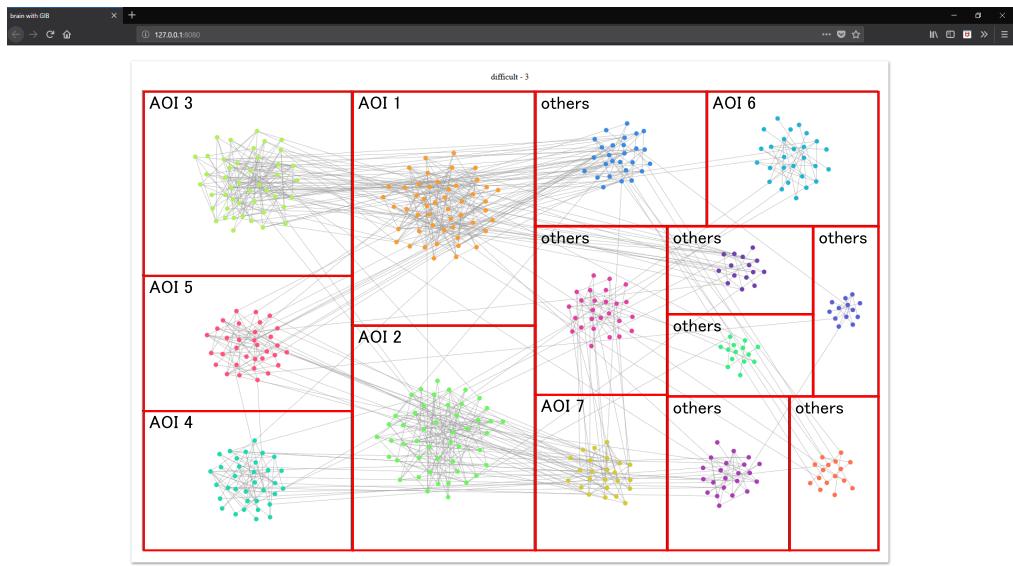


図 4.5: AOI の定義. グループ内エッジの数が多い順に AOI 1 から AOI 7 が各ボックスに対して定義され, それ以外のボックスは others として定義された.

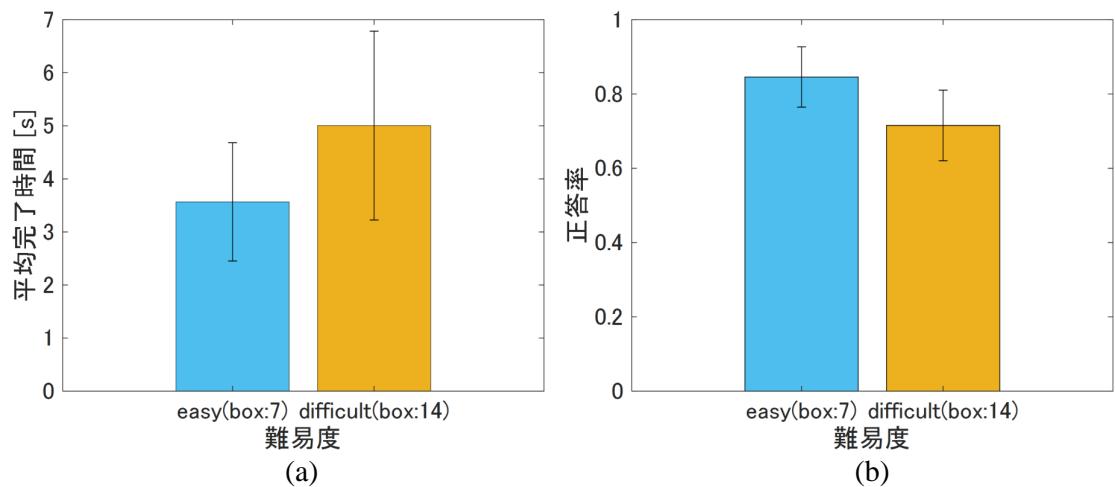


図 4.6: (a) 難易度と平均完了時間の関係 (b) 難易度と正答率の関係. ウィルコクソンの符号順位検定の結果, 難易度の違いによるタスクパフォーマンスの有意差が確認された ($p < 0.01$).

表 4.3: FD-GIB における正答率に対する重回帰分析の結果

Variables	correct answer rate	Unstandardized Coefficients		Beta	t-value	p-value
		B	Std. Error			
	constant	0.455	0.0759	-	5.99	0
	factor1	-0.0225	0.00473	-0.079	-4.76	0
	factor2	0.523	0.0622	0.523	8.41	0
	factor3	0.00293	0.000613	0.077	4.78	0

$R^2 = 0.558$, $R^2_{adj} = 0.547$, $SEE = 0.181$

表 4.4: TR-GIB における完了時間に対する重回帰分析の結果

Variables	completion time	Unstandardized Coefficients		Beta	t-value	p-value
		B	Std. Error			
	constant	3.05	0.277	-	11	0
	factor1	0.204	0.0196	0.715	10.4	0
	factor2	-0.281	0.206	-0.281	-1.36	0.0176
	factor3	-0.0237	0.00307	-0.621	-7.73	0

$R^2 = 0.625$, $R^2_{adj} = 0.615$, $SEE = 0.749$

表 4.5: TR-GIB における正答率に対する重回帰分析の結果

Variables	correct answer rate	Unstandardized Coefficients		Beta	t-value	p-value
		B	Std. Error			
	constant	0.512	0.066	-	7.75	0
	factor1	-0.0131	0.00466	-0.046	-2.8	0.005
	factor2	0.347	0.0492	0.347	7.06	0
	factor3	0.00407	0.000732	0.107	5.57	0

$R^2 = 0.543$, $R^2_{adj} = 0.531$, $SEE = 0.179$

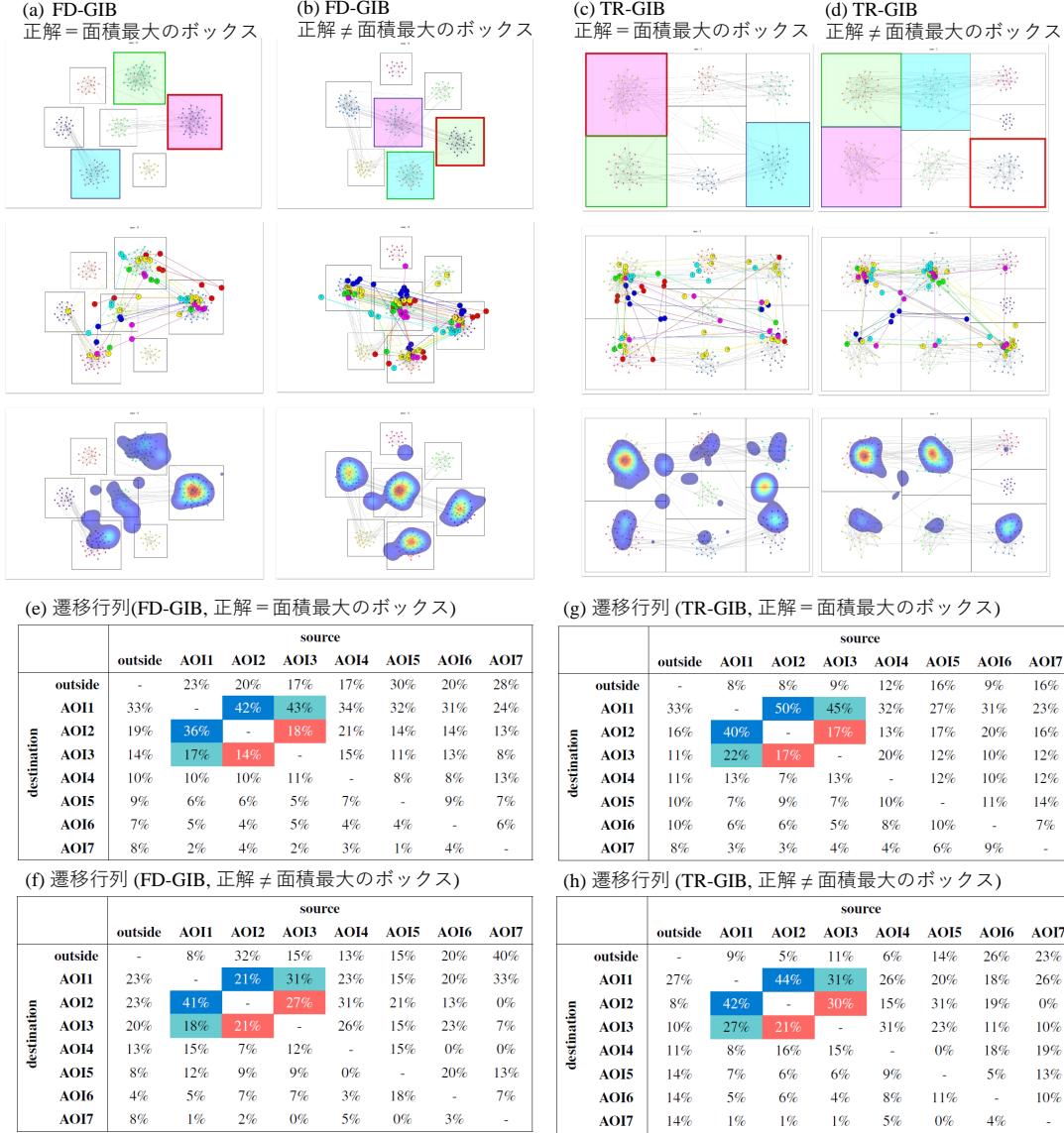


図 4.7: FD-GIB と TR-GIB において正解のボックスが面積最大である場合とそうでない場合の視線追跡データの比較. (a), (b), (c), (d) の 1 行目: タスクの例. (a), (b), (c), (d) の 2 行目: 各被験者の注視位置. 各被験者に異なる色が割り当てられている. (a), (b), (c), (d) の 3 行目: 注視時間のヒートマップ. 長い時間見られたところが赤く、短い時間見られたところが青く表示される. (e), (f), (g), (h): 各 AOI から各 AOI への遷移行列. 同じ色で塗られているところは、同じ AOI のペア同士の視線の遷移確率を示す.

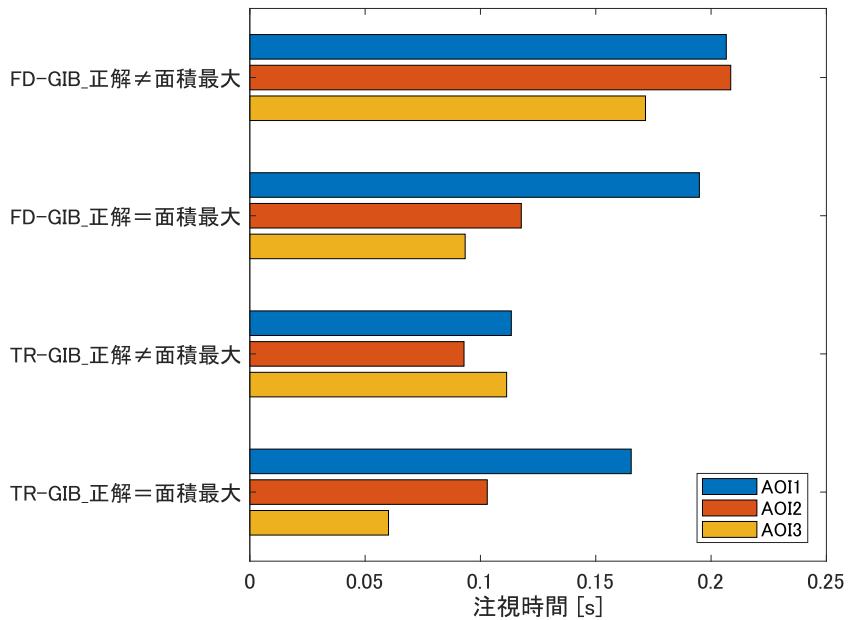


図 4.8: FD-GIB と TR-GIB において正解のボックスが面積最大である場合とそうでない場合の AOI 1~AOI 3 に対する注視時間の比較

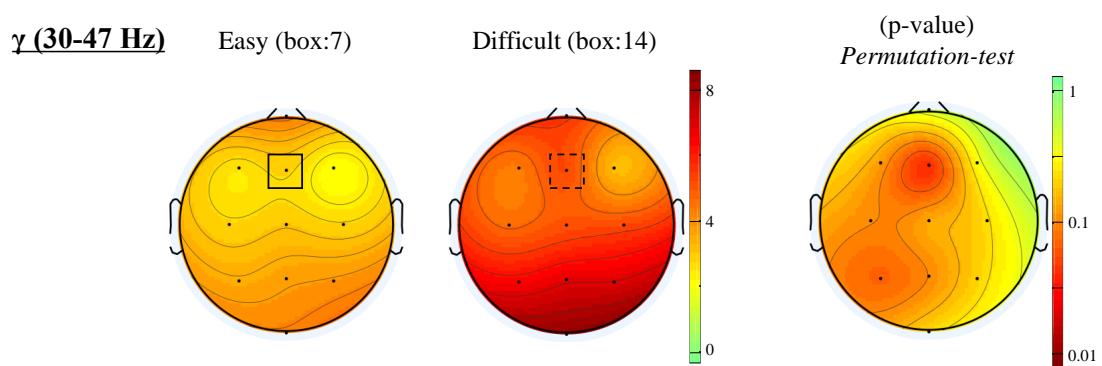


図 4.9: タスクの難易度と γ 波 (30-47 Hz) のパワーの関係。左側の 2 つのトポグラフィーは開眼安静状態の γ 波に対する各条件における γ 波のパワーの比率を表す。右側の p 値は並び替え検定における結果を示している。

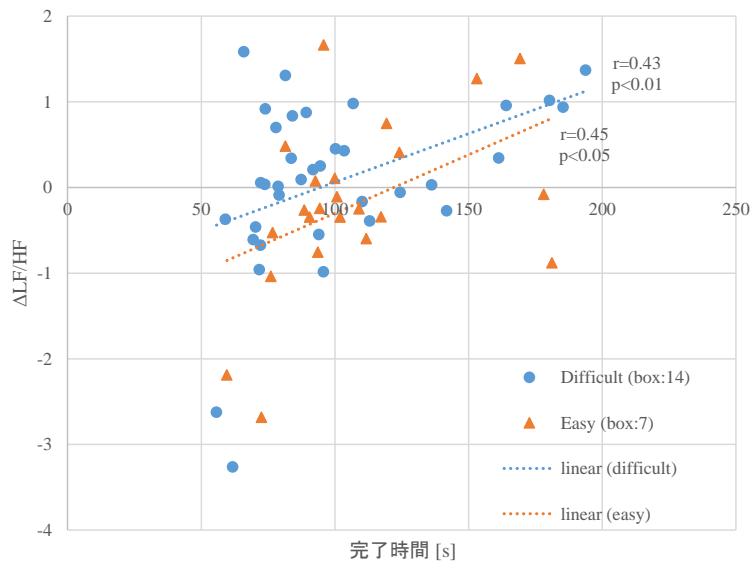


図 4.10: タスク完了時間と LF/HF の変化量の関係

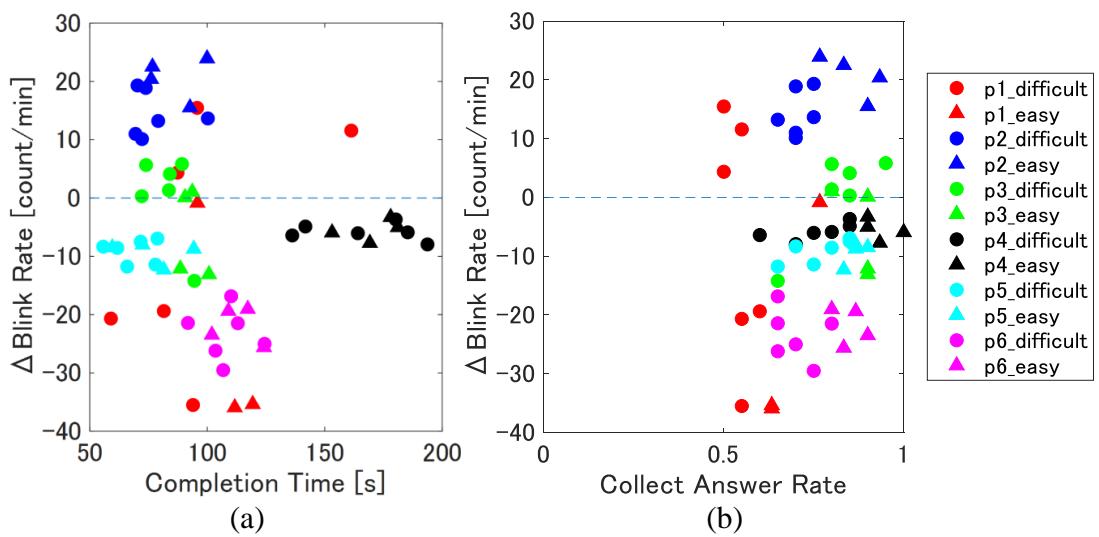
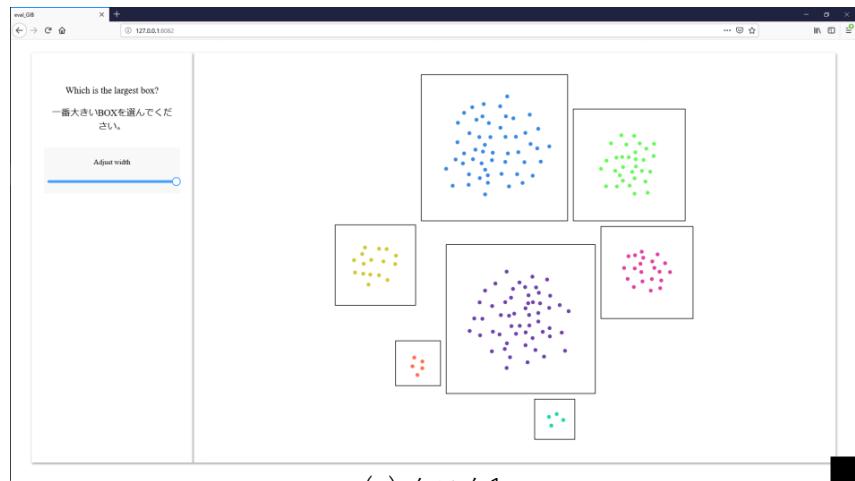
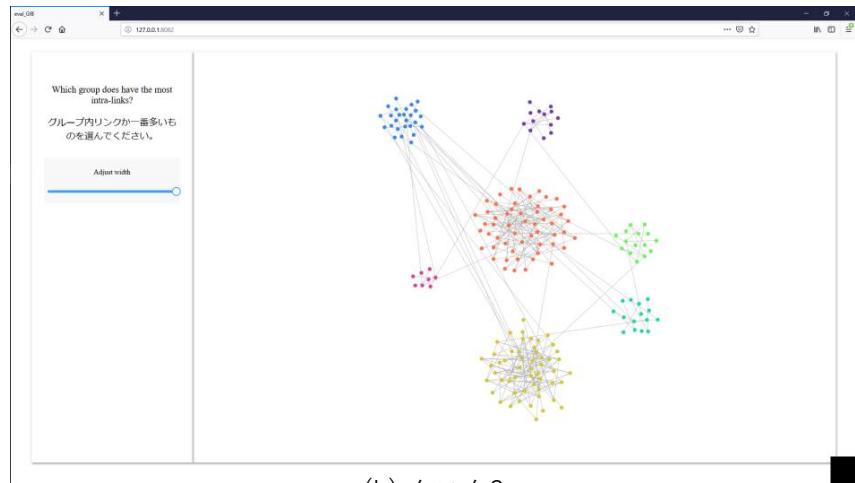


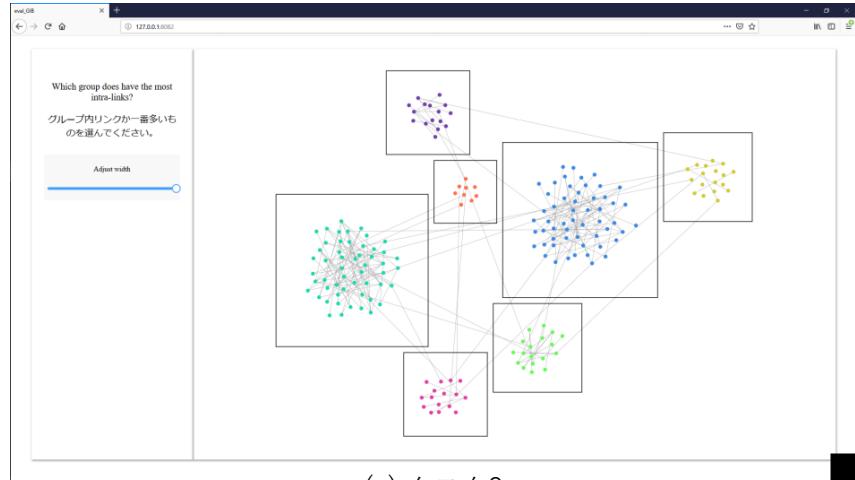
図 4.11: タスクパフォーマンス ((a) タスク完了時間, (b) 正答率) と 1 分間の瞬き回数の変化量の関係



(a) タスク1



(b) タスク2



(c) タスク3

図 5.1: 各タスクの例. (a) タスク 1 : どのグループが面積最大か? (b) タスク 2: どのグループのグループ内エッジ数が最大か? (c) タスク 3 : どのグループのグループ内エッジ数が最大か?

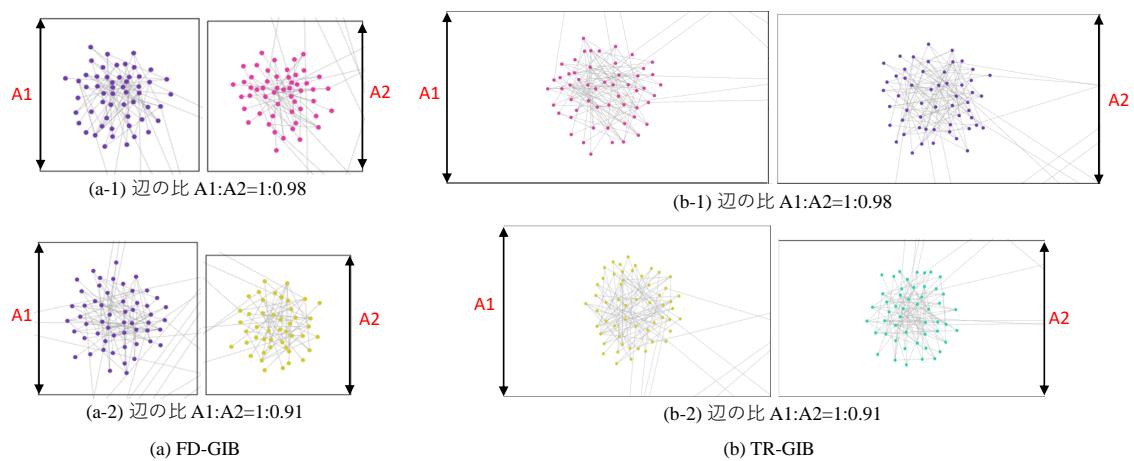


図 5.2: 各 GIB レイアウトにおける正解候補のボックスの辺の比の例

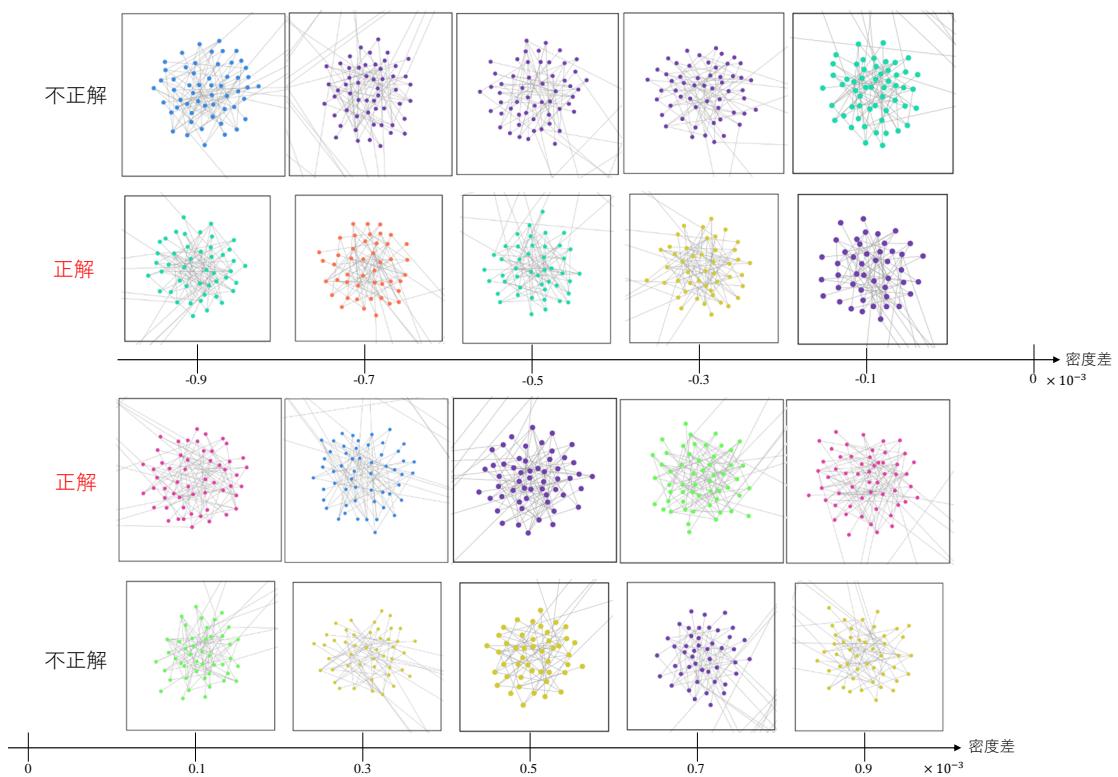


図 5.3: FD-GIB の辺の比 0.91 における正解候補のボックスの密度差の例

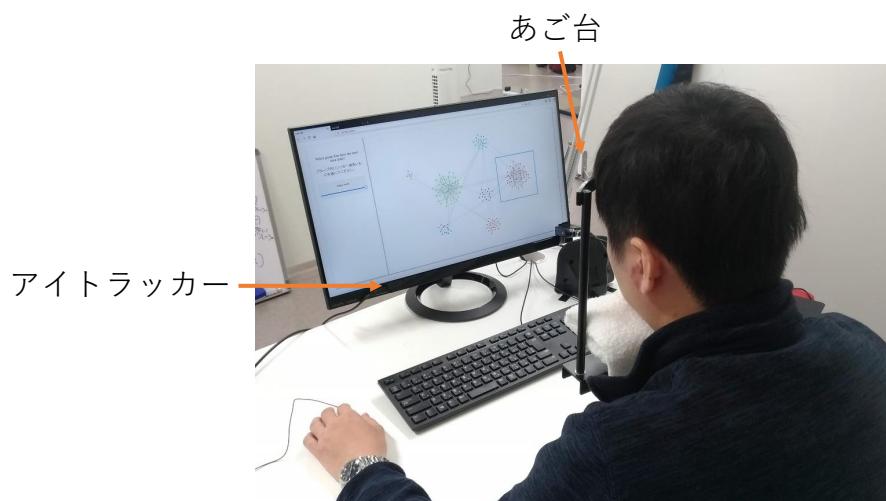


図 5.4: 実験環境. 視線の動きはモニタ下部に取り付けられたアイトラッカーを用いて記録した.

実験説明	休憩	Task 1 (2 sets)	休憩	Task2 (4 sets)	休憩	Task2 (4 sets)	休憩	Task3 (4 sets)	休憩	Task3 (4 sets)
	10分	5分		5分		5分		5分		5分

図 5.5: 実験の概要

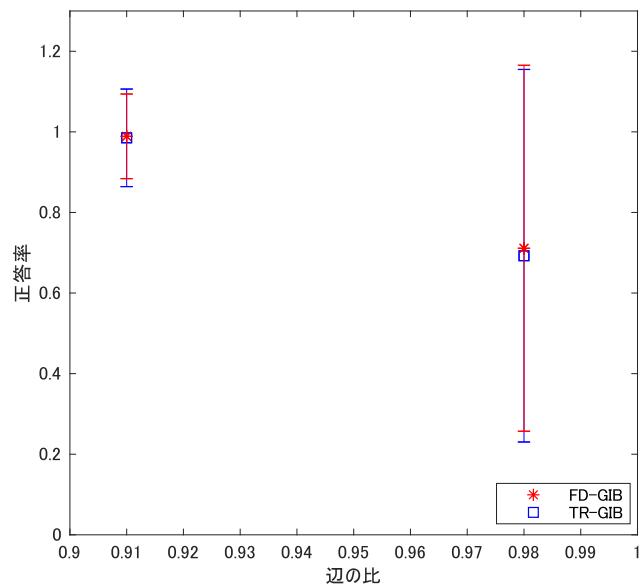


図 6.1: タスク 1 での各 GIB レイアウトにおける辺の比と正答率の関係

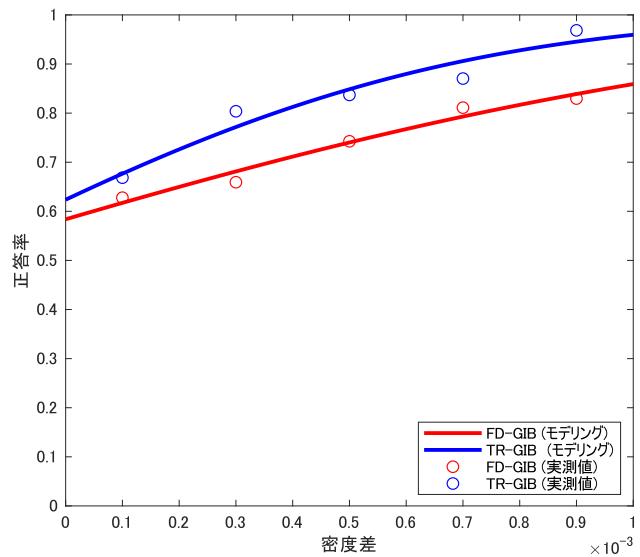


図 6.2: タスク 2 の FD-GIB, TR-GIB における密度差と正答率の関係

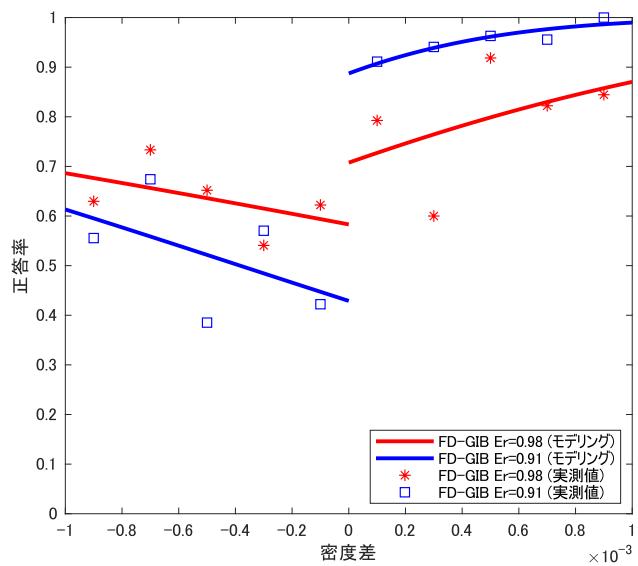


図 6.3: タスク 3 の FD-GIB における密度差と正答率の関係をモデル化した結果

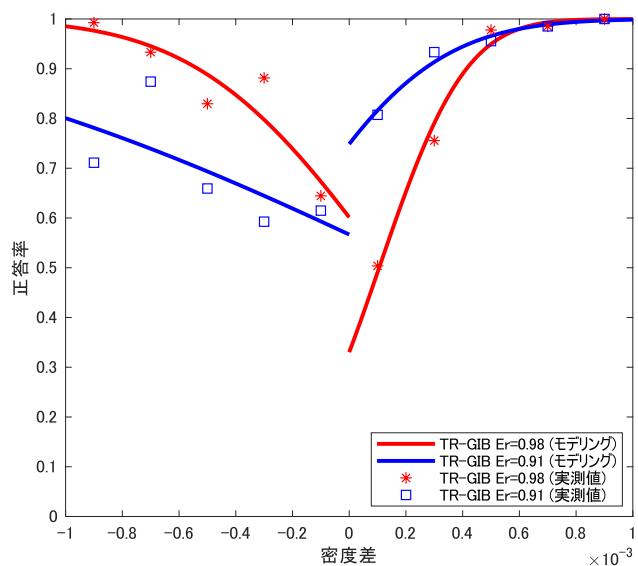


図 6.4: タスク 3 の TR-GIB における密度差と正答率の関係をモデル化した結果

表 6.1: タスク 3 のモデル化に対する適合度検定の結果. p 値 > 0.05 の条件においてこのモデル式が実測値に適合していないとまでは言えない

可視化手法	FD-GIB				TR-GIB			
	$\Delta D < 0$		$\Delta D > 0$		$\Delta D < 0$		$\Delta D > 0$	
辺の比	0.98		0.91		0.98		0.91	
密度差	$\Delta D < 0$	$\Delta D > 0$	$\Delta D < 0$	$\Delta D > 0$	$\Delta D < 0$	$\Delta D > 0$	$\Delta D < 0$	$\Delta D > 0$
p 値	0.479	0.182	0.222	0.119	2.57×10^{-4}	5.99×10^{-10}	9.18×10^{-2}	1.65×10^{-3}

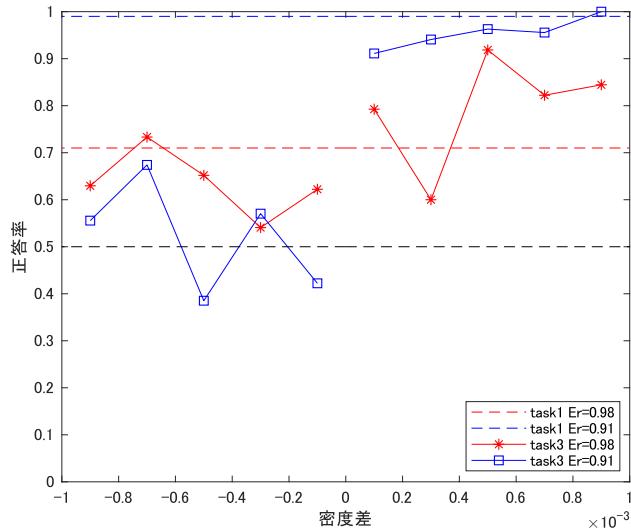


図 6.5: FD-GIB におけるタスク 1 の結果とタスク 3 の密度差と正答率の関係

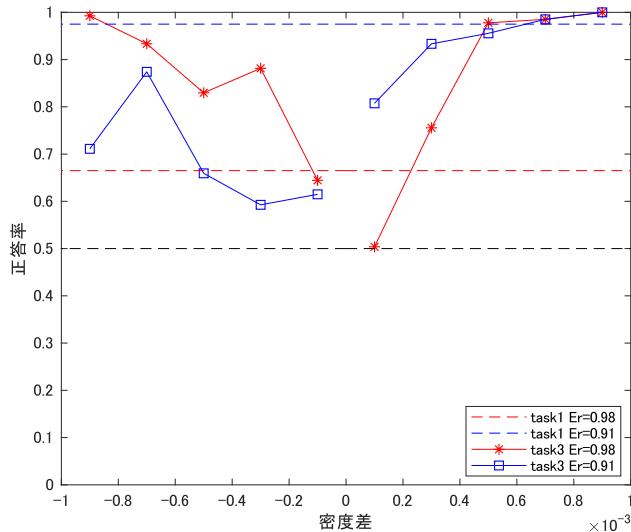


図 6.6: TR-GIB におけるタスク 1 の結果とタスク 3 の密度差と正答率の関係

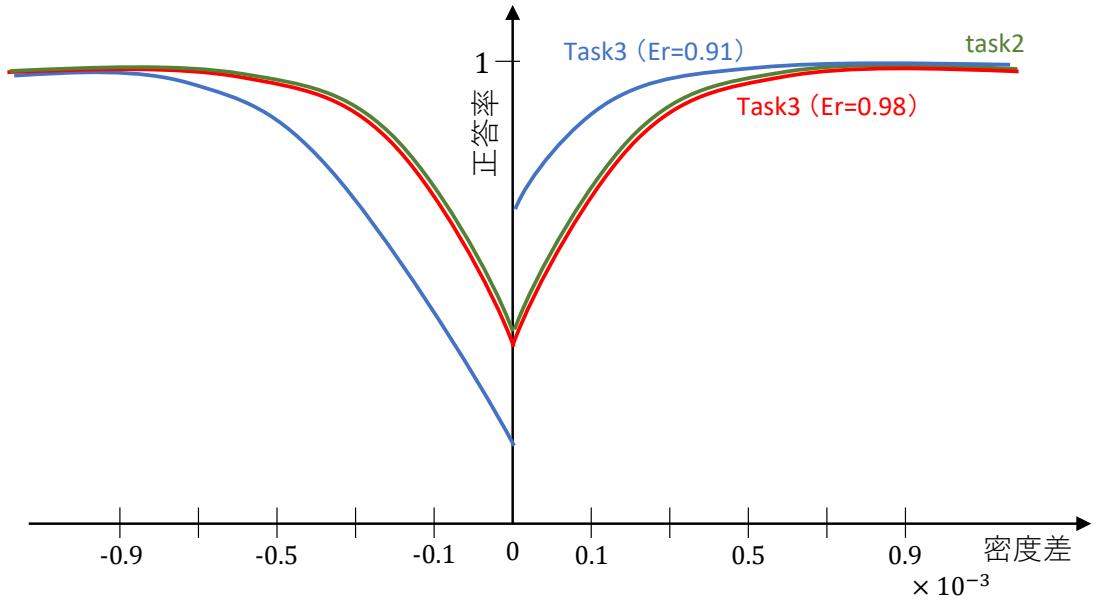


図 6.7: 仮説におけるタスク 2 とタスク 3 の正答率の関係

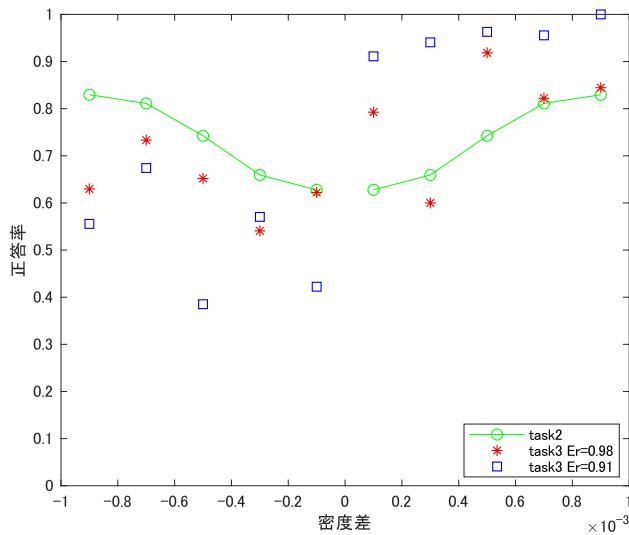


図 6.8: タスク 2, 3 の FD-GIB における密度差と正答率の関係

表 6.2: タスク 3 において各 GIB レイアウトの密度差ごとに辺の比による正答率の有意差があるかを wilcoxon の符号順位検定で調べた結果. 赤くハイライトされた部分が有意差あり ($p < 0.05$) と判断されたデータである.

密度差	-9×10^{-4}	-7×10^{-4}	-5×10^{-4}	-3×10^{-4}	-1×10^{-4}	1×10^{-4}	3×10^{-4}	5×10^{-4}	7×10^{-4}	9×10^{-4}
FD-GIB	0.0709	0.265	2.05×10^{-4}	0.497	6.90×10^{-4}	9.05×10^{-4}	1.57×10^{-5}	0.146	1.10×10^{-3}	6.10×10^{-5}
TR-GIB	5.57×10^{-5}	0.194	9.46×10^{-3}	4.88×10^{-4}	0.547	5.50×10^{-5}	1.24×10^{-4}	0.508	1.00	1.00

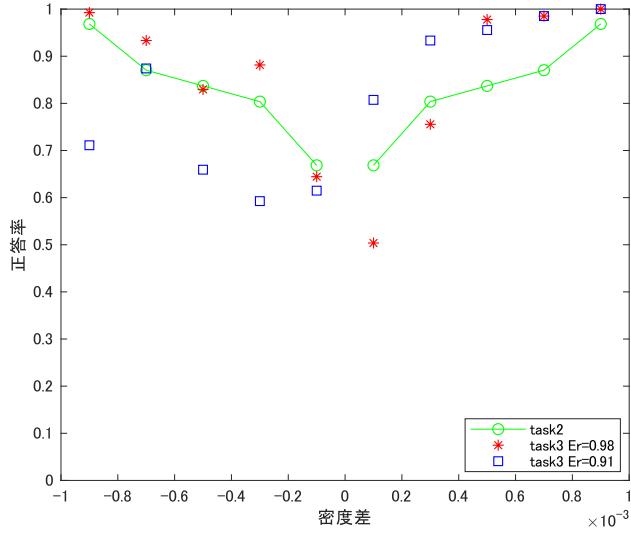


図 6.9: タスク 2, 3 の TR-GIB における密度差と正答率の関係

表 6.3: タスク 2 とタスク 3 の各条件での正答率に対してウィルコクソンの符号順位検定を行った結果。赤くハイライトされた部分が有意差あり ($p < 0.05$) と判断されたデータである。

可視化手法		FDGIB		TRGIB	
辺の比		0.98	0.91	0.98	0.91
密度差	-9×10^{-4}	8.94×10^{-4}	9.21×10^{-6}	0.0625	5.92×10^{-4}
	-7×10^{-4}	0.0881	7.70×10^{-3}	0.0189	0.245
	-5×10^{-4}	0.0182	1.20×10^{-5}	0.961	1.17×10^{-3}
	-3×10^{-4}	3.89×10^{-3}	0.0756	0.0121	3.67×10^{-4}
	-1×10^{-4}	0.942	2.71×10^{-4}	0.521	0.167
	1×10^{-4}	6.50×10^{-4}	1.12×10^{-5}	2.83×10^{-4}	8.17×10^{-3}
	3×10^{-4}	0.126	1.17×10^{-5}	0.285	1.95×10^{-4}
	5×10^{-4}	2.54×10^{-5}	1.73×10^{-5}	1.01×10^{-5}	1.28×10^{-4}
	7×10^{-4}	0.524	8.15×10^{-5}	1.75×10^{-5}	1.67×10^{-5}
	9×10^{-4}	0.645	6.36×10^{-6}	0.0313	0.0313

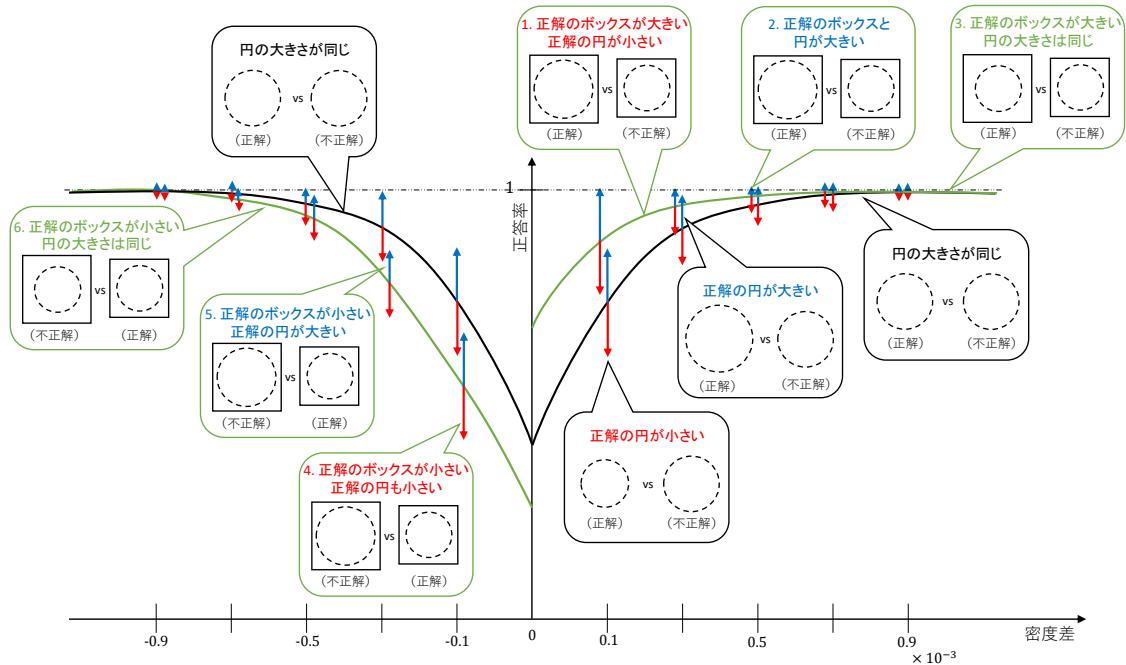


図 6.10: ボックスの大きさ, 円の大きさが正答率に及ぼす影響

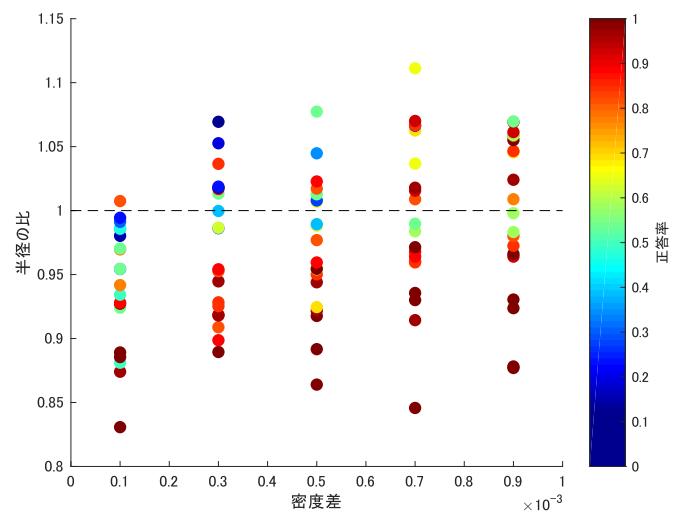


図 6.11: FD-GIB における半径の比と正答率の関係

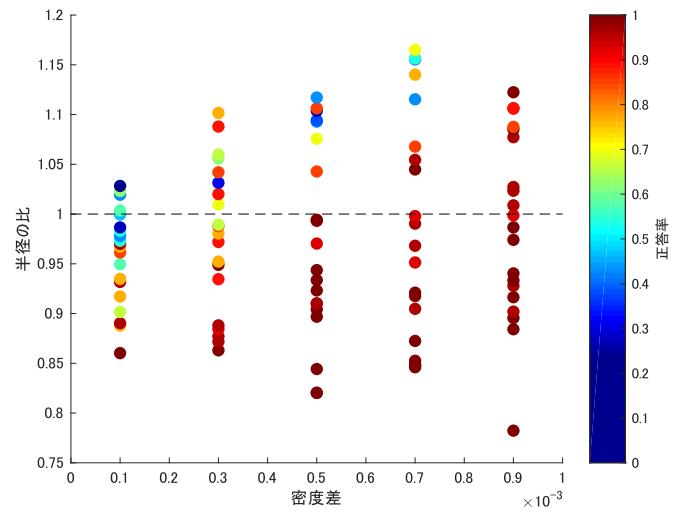


図 6.12: TR-GIB における半径の比と正答率の関係

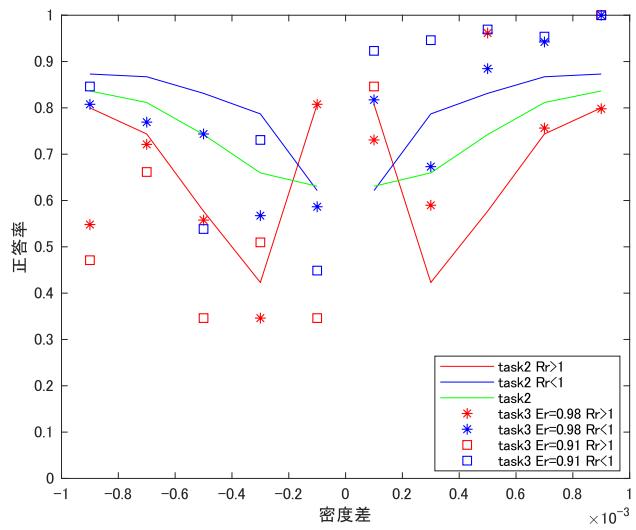


図 6.13: FD-GIB における半径の比を考慮したタスク 2, 3 の比較

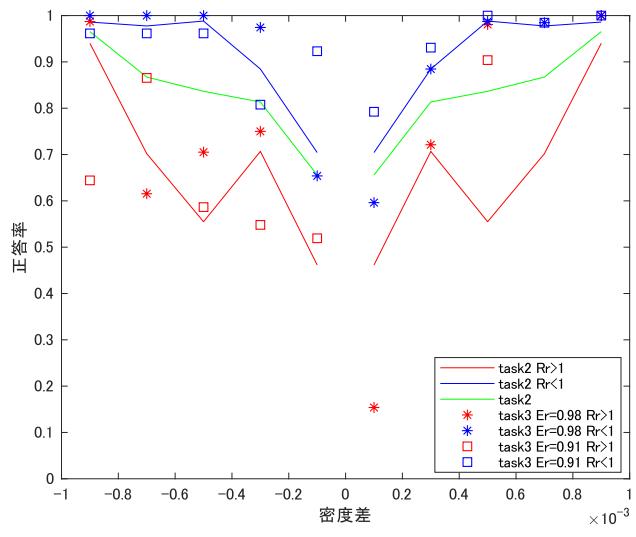


図 6.14: TR-GIB における半径の比を考慮したタスク 2, 3 の比較

表 6.4: タスク 2 における各条件の試行数

可視化手法		FD-GIB		TR-GIB	
半径の比		$Rr > 1$	$Rr < 1$	$Rr > 1$	$Rr < 1$
密度差	1×10^{-4}	1	19	4	16
	3×10^{-4}	7	13	8	12
	5×10^{-4}	7	13	7	13
	7×10^{-4}	9	11	8	12
	9×10^{-4}	10	10	9	11

表 6.5: タスク 3 における各条件の試行数. 赤くハイライトされた部分はボックスの有無による効果の傾向が確認できるデータである. ただし, 辺の比 0.98 において (タスク 2 とタスク 3 の正答率の差 ≤ 0.03) を満たす条件, 辺の比 0.91 において 密度差 < 0 では (タスク 2 の正答率 - タスク 3 の正答率の差 > 0.03), 密度差 > 0 において (タスク 2 の正答率 - タスク 3 の正答率 < -0.03) を満たす条件をハイライトした.

可視化手法		FD-GIB				TR-GIB			
辺の比		0.98		0.91		0.98		0.91	
半径の比		$Rr > 1$	$Rr < 1$						
密度差	-9×10^{-4}	4	1	4	1	3	2	4	1
	-7×10^{-4}	4	1	5	0	1	4	4	1
	-5×10^{-4}	2	3	4	1	3	2	4	1
	-3×10^{-4}	1	4	4	1	2	3	4	1
	-1×10^{-4}	1	4	2	3	0	5	4	1
	1×10^{-4}	1	4	1	4	1	4	0	5
	3×10^{-4}	3	2	0	5	4	1	0	5
	5×10^{-4}	3	2	0	5	2	3	2	3
	7×10^{-4}	3	2	0	5	0	5	0	5
	9×10^{-4}	4	1	2	3	1	4	0	5

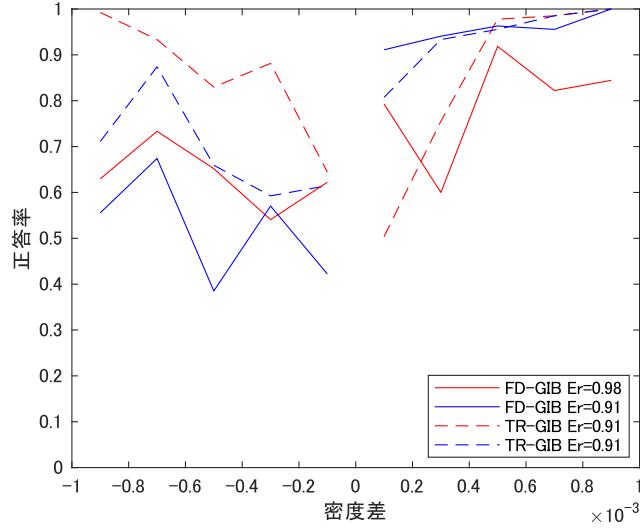
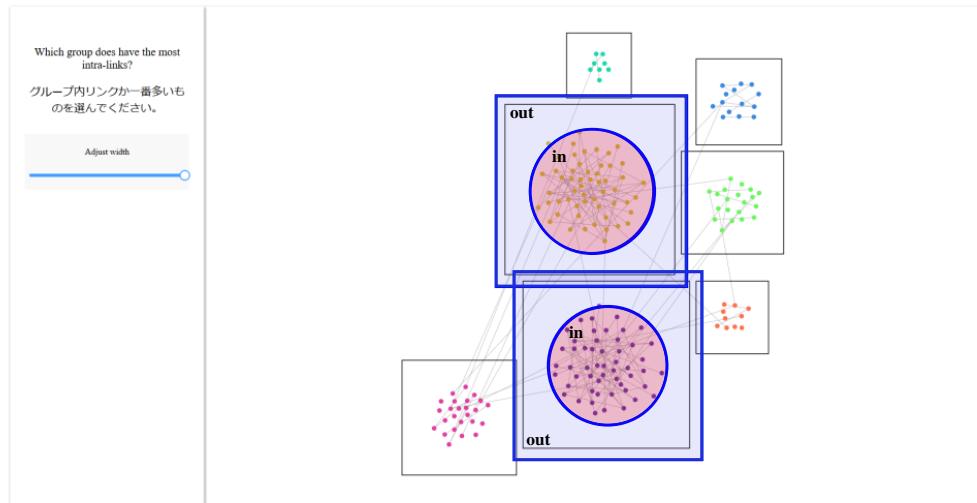


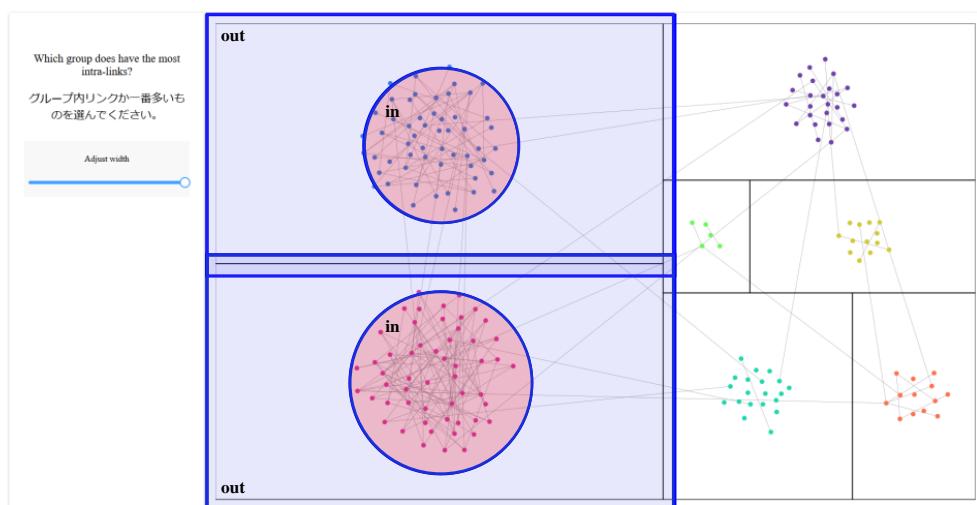
図 6.15: タスク 3 における FD-GIB と TR-GIB の比較

表 6.6: タスク 3 において, GIB レイアウトの違いによる正答率の有意差があるかどうかを ウィルコクソンの符号順位検定で調べた結果. 赤くハイライトされた部分が有意差あり ($p < 0.05$) と判断されたデータである.

密度差	-9×10^{-4}	-7×10^{-4}	-5×10^{-4}	-3×10^{-4}	-1×10^{-4}	1×10^{-4}	3×10^{-4}	5×10^{-4}	7×10^{-4}	9×10^{-4}	
辺の比	0.98	9.56×10^{-6}	8.22×10^{-4}	2.99×10^{-4}	1.37×10^{-5}	0.639	1.12×10^{-4}	9.62×10^{-3}	0.0352	8.21×10^{-4}	6.10×10^{-5}
	0.91	0.0128	2.51×10^{-3}	1.75×10^{-4}	0.740	3.80×10^{-3}	0.0357	1.00	1.00	0.289	1.00



(a)FD-GIB



(b)TR-GIB

図 6.16: AOI の定義. (a) FD-GIB, (b)TR-GIB. グループに属するすべてのノードを囲む円の領域を “in”, ボックスよりわずかに大きい領域の内 “in” を除く領域を “out” を定義した.

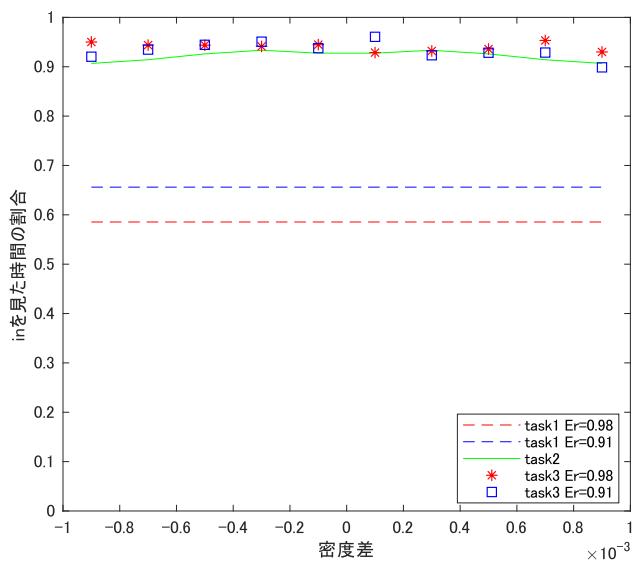


図 6.17: FD-GIB の各タスクにおいて, AOI の “in” と “out” を見た時間の合計に対する “in” を見た時間の割合

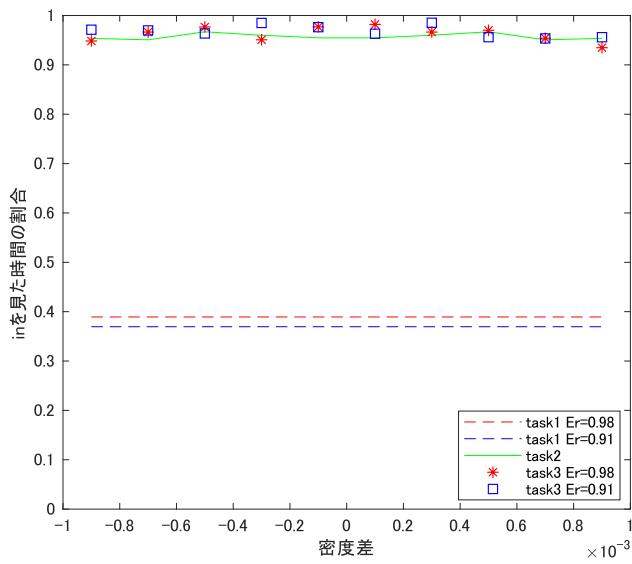


図 6.18: TR-GIB の各タスクにおいて, AOI の “in” と “out” を見た時間の合計に対する “in” を見た時間の割合



図 A.1: 実験に使用した脈波センサー