

筑波大学 情報学群 情報メディア創成学類

卒業研究論文

タグクラウドにおいて頻度の周期性を
視覚的に表現する手法に関する研究

長谷川 貴紀

指導教員 三末 和男，志築 文太郎，田中 二郎

2014年2月

概要

タグクラウドとは、ブログや写真・動画投稿サイト等の Web 上で用いられるタグの視覚的表現であり、一般的にアルファベット順や五十音順で並べたり、頻度の表現にフォントサイズや色を用いて表現する。このタグクラウドで Web 上のタグに限らない時刻情報付きデータ、特に頻度の周期性の特徴を表現できれば、人の行動や事象の生起パターンを把握することに役立つ。本研究ではこの頻度の周期性の特徴について、特に直感的に把握するということが目的とした。タグクラウドの拡張手法として広く行われている、他の視覚的表現を組み合わせるという方向性から表現を設計し、被検者実験により速さと正確さから直感性を評価するというアプローチをとった。また、頻度の周期性の特徴を表現することに付随して表現されることが期待される事象間の共起性についても読み取られ方の評価を行った。これらの結果が今後の可視化表現設計の指針の 1 つとなることが期待される。

目次

第1章	序論	1
1.1	タグクラウド	1
1.2	時間共起性	1
1.3	本研究で取り組む課題	1
1.4	研究の流れ	2
1.5	本論文の構成	2
第2章	関連研究	3
2.1	タグクラウドの拡張に関する研究	3
2.1.1	様々な視覚的表現と組み合わせる手法	3
2.1.2	折れ線グラフと組み合わせる手法	3
2.1.3	色を用いて時間共起性を強調する手法	3
2.2	頻度の周期性表現に関する研究	4
2.2.1	カレンダー表現を用いる手法	4
2.2.2	螺旋状の時間軸を用いる手法	4
2.3	視覚変数に関する研究	4
第3章	表現の設計	5
3.1	設計手順	5
3.1.1	タグクラウドと組み合わせる視覚的表現の設計	5
3.1.2	タグクラウドとの組合せ	6
3.2	作成した手法	6
3.2.1	扇形の半径の長さに頻度を割り当てる表現	6
3.2.2	扇形の弧の長さに頻度を割り当てる表現	6
3.2.3	扇形の明度に頻度を割り当てる表現	6
3.2.4	カレンダー表現	7
3.2.5	折れ線表現	7
第4章	直感性と共起性に関する被検者実験	9
4.1	実験の目的	9
4.2	実験の設計	9
4.2.1	観点1の実験の設計	9

4.2.2	観点 2 の実験の設計	10
4.2.3	データの作成	17
4.3	実験結果	17
4.3.1	観点 1	18
4.3.2	観点 2	25
4.4	考察	25
4.4.1	観点 1	25
4.4.2	観点 2	28
第 5 章	追実験	30
5.1	実験の目的	30
5.2	表現の再設計	30
5.2.1	頻度を表現する視覚変数として中心からの長さ (半径) を用いる表現 .	30
5.2.2	折れ線表現	30
5.3	実験の設計	31
5.4	実験結果	31
5.4.1	考察	35
第 6 章	議論	37
第 7 章	結論	38
	謝辞	39
	参考文献	40

図目次

3.1	扇形の半径の長さに頻度を割り当てる表現の例	6
3.2	扇形の弧の長さに頻度を割り当てる表現の例	7
3.3	扇形の明度に頻度を割り当てる表現の例	7
3.4	カレンダー表現の例	7
3.5	折れ線表現の例	8
4.1	扇形の半径の長さに頻度を割り当てる表現の実験画面例	10
4.2	扇形の弧の長さに頻度を割り当てる表現の実験画面例	11
4.3	扇形の明度に頻度を割り当てる表現の実験画面例	11
4.4	カレンダー表現の実験画面例	12
4.5	折れ線表現の実験画面例	12
4.6	分布 1 の例	13
4.7	分布 2 の例	13
4.8	分布 3 の例	13
4.9	分布 4 の例	14
4.10	扇形の半径の長さに頻度を割り当てる表現の共起性評価実験	15
4.11	扇形の弧の長さに頻度を割り当てる表現の共起性評価実験	15
4.12	扇形の明度に頻度を割り当てる表現の共起性評価実験	15
4.13	カレンダー表現の共起性評価実験	16
4.14	折れ線表現の共起性評価実験	16
4.15	頻度を半径の長さに割り当てる表現の回答時間のヒストグラム	18
4.16	頻度を半径の長さに割り当てる表現の NFE の真値のヒストグラム	19
4.17	頻度を半径の長さに割り当てる表現の FE の真値のヒストグラム	19
4.18	頻度を弧の長さに割り当てる表現の回答時間のヒストグラム	20
4.19	頻度を弧の長さに割り当てる表現の NFE の真値のヒストグラム	20
4.20	頻度を弧の長さに割り当てる表現の FE の真値のヒストグラム	21
4.21	頻度を明度に割り当てる表現の回答時間のヒストグラム	21
4.22	頻度を明度に割り当てる表現の NFE の真値のヒストグラム	22
4.23	頻度を明度に割り当てる表現の FE の真値のヒストグラム	22
4.24	カレンダー表現の回答時間のヒストグラム	23
4.25	カレンダー表現の FE の真値のヒストグラム	23
4.26	折れ線表現の回答時間のヒストグラム	24

4.27	折れ線表現の NFE の真値のヒストグラム	24
4.28	折れ線表現の FE の真値のヒストグラム	25
5.1	扇形の半径の長さに頻度を割り当てる表現 (改良版) の例	30
5.2	折れ線表現 (改良版) の例	31
5.3	扇形の半径の長さに頻度を割り当てる表現 (改良版) の実験画面例	31
5.4	折れ線表現 (改良版) の実験画面例	32
5.5	頻度を半径の長さに割り当てる表現 (改良版) の回答時間のヒストグラム . . .	32
5.6	頻度を半径の長さに割り当てる表現 (改良版) の NFE の真値のヒストグラム .	33
5.7	頻度を半径の長さに割り当てる表現 (改良版) の FE の真値のヒストグラム . .	33
5.8	折れ線表現 (改良版) の回答時間のヒストグラム	34
5.9	折れ線表現 (改良版) の NFE の真値のヒストグラム	34
5.10	折れ線表現 (改良版) の FE の真値のヒストグラム	36

表 目 次

4.1	観点 1 の実験結果	18
4.2	観点 2 の実験結果 (負の相関考慮)	26
4.3	観点 2 の実験結果 (負の相関非考慮)	27
4.4	分布の組合せの分類	28
5.1	追実験結果	35

第1章 序論

本章では本研究で扱う分野の基本的な説明及び、背景となる課題を述べ、本論文の理解を促すものとして研究の流れと論文の構成を述べる。

1.1 タグクラウド

タグクラウドとは、ブログや写真・動画投稿サイトをはじめとする Web 上で用いられるタグを複数個並べて提示する視覚的表現である。タグの並べ方はアルファベット順や五十音順に従うことが多く、頻度の総量をフォントサイズや色に割り当てることが多い。

情報可視化において、可視化手法が表現するデータを読み取るためにはラベル等のテキスト情報が必要な場合が少なくない。対して、テキスト自体を情報可視化の視覚的表現として利用できるタグクラウドは、視覚的表現に加えラベルを提示しなければならない可視化手法に比べ空間効率の良い表現である可能性がある。また、多数の重複を含むような大規模データに対しては、タグクラウドでは同件事象の頻度を蓄積して1つのタグで表現するため、個々の事象を全て画面上に表示する可視化手法よりも多くの事象を一度に俯瞰できる可能性があると考ええる。

1.2 時間共起性

本研究では時間共起性というものに注目する。本論文中使用される「共起性」は、この時間共起性を指すものとする。

本論文での時間共起性の定義は、本研究で想定する時刻情報付きデータにおいて、ある事象が生起すると、それに伴って他の別の事象も生起するという状況、その有無や強弱といった性質とする。

1.3 本研究で取り組む課題

タグクラウドは時間経過に伴う頻度変化により変化する視覚的表現であり、その時刻情報には事象の持つ生起パターン等が含まれると考えられる。この時刻情報をタグクラウドで表現することができれば、タグクラウドの空間効率や同一事象の頻度の蓄積という特徴により、多数の時刻情報付きデータについて時刻情報の特徴と共に俯瞰できると考えられる。しかし、

タグクラウドで時刻情報を表現することを目的とした先行研究においては、その頻度の周期性の特徴の表現に特化しておらず、特徴の把握が直感的でないという課題があった。

そこで本研究では、タグクラウドと組合せて頻度の周期性の特徴を直感的に把握できる表現を設計するための指針作りを目的とした。そのためのアプローチとして、複数の視覚的表現のバリエーションについて被験者実験により比較を行うこととした。

1.4 研究の流れ

本研究は以下のような流れで行った。

まず、タグクラウドを拡張するためのアプローチを決定するため、先行研究の調査を行った。その結果、タグクラウドと他の視覚的表現を組み合わせる方法が広く行われていることが確認されたため、本研究でもこれを採用した。次に、タグクラウドと組み合わせる視覚的表現の設計のため、周期性の可視化表現について調査し、用いる形状と頻度を割り当てる視覚変数を決定した。形状と視覚変数の組み合わせのバリエーションから、被験者実験により評価を行うための表現を設計した。設計した表現を用いて、頻度の周期性の特徴を読み取るためのPC上のプログラムによる実験と、共起性を読み取るための質問紙上での実験を被験者に実施した。実験結果より、表現の形状と視覚変数について、頻度の周期性の特徴の読み取られ方と共起性の読み取られ方を考察した。

1.5 本論文の構成

本論文は、以下第2章で関連研究を紹介し、第3章で実験に用いた表現の設計について述べる。第4章では1度目の被験者実験について目的・設計・結果・考察を、第5章では2度目の被験者実験について目的・表現の再設計・実験の設計・結果・考察を述べる。第6章で全体を通しての議論を述べ、第7章を結論とする。

第2章 関連研究

2.1 タグクラウドの拡張に関する研究

以下に挙げるタグクラウドの拡張に関する研究より，タグクラウドで時刻情報を表現するための拡張方法として，タグクラウドに他の視覚的表現を組み合わせる方法が広く行われていることがわかった．

2.1.1 様々な視覚的表現と組み合わせる手法

Nguyen らは，時刻情報を表現するために様々な視覚的表現を用いてタグクラウドと組合せ，また地図上に配置することで空間情報も表現した [1]．連続時間経過に伴うタグの頻度変化について TOI(Time Of Interest) の観点も混じえて評価した．しかし，頻度の周期性の特徴の観点からは評価を行っていなかった．

2.1.2 折れ線グラフと組み合わせる手法

Lee らは，Sparkline [2] という折れ線グラフをタグクラウドと組み合わせる手法を提案した [3]．手法を設計する際の試行錯誤の段階で Sparkline の配置などを検討した．また，完成した手法について棒グラフなどの伝統的な手法とともに Parallel Tag Clouds [4] との比較評価を行った．彼らも頻度の周期性の特徴の観点からは評価を行っていなかった．

彼らの手法 (SparkClouds) はタグクラウドに僅かな描画領域の視覚的表現を付加するだけで，時刻情報を効果的に表現できたと述べられている．

2.1.3 色を用いて時間共起性を強調する手法

Lohmann らはマイクロブログサービスに投稿されるワードをタグクラウド化し，各タグの背景部分に色を用いたヒストグラムを描くことによって時間共起性を強調した [5]．他の視覚変数を用いて共起性を強調する試みは行っていなかった．

2.2 頻度の周期性表現に関する研究

2.2.1 カレンダー表現を用いる手法

Nguyen らの研究において，タグクラウドと組合せられた視覚的表現の中に，カレンダー表現があった [1]．我々の生活の中で広く用いられ，慣習的に周期性を想起させる形状の 1 つであると考えられた．

2.2.2 螺旋状の時間軸を用いる手法

連続時間での頻度変化の様子と，年や月などの周期性の両方を同時に提示する形状として，螺旋状の時間軸を用いた手法を Carlis ら [6] や Weber ら [7]，Tominski ら [8] が開発・改良した．

連続時間の表現のために時間軸の始点と終点を繋げていない一方で，周期性の表現のために同じ周期性の特徴が半径の直線上に重ねている．したがって，環状の時間軸を持つという形状が周期性を想起させると考えられた．

2.3 視覚変数に関する研究

Mackinlay は，効果的な視覚的表現設計のために，視覚変数がデータをどのように表現するのか，どのように読み取られるのかを研究した [9]．本研究で作成した表現において視覚変数の割り当ての参考とした．

第3章 表現の設計

本研究ではタグクラウドと他の視覚的表現を組み合わせる表現を計5つ実装し、被検者実験により比較を行った。本章では表現の設計手順及び各表現を解説する。

3.1 設計手順

本研究で用いた表現は、タグクラウドと組み合わせる視覚的表現の設計と、設計した視覚的表現とタグクラウドとの組合せという2段階で設計を行った。

3.1.1 タグクラウドと組み合わせる視覚的表現の設計

タグクラウドと組み合わせる視覚的表現の設計は、周期性を表現する形状と頻度を割り当てる視覚変数という2つの観点から設計した。

まず、周期性を表現する形状について、時間軸を環状に繋げるものとカレンダー状のものの2種類に大別した。次に、頻度を割り当てる視覚変数について検討を行い、既存の周期性の表現から、頻度を中心からの長さに割り当てるもの、頻度を外周の長さに割り当てるもの、頻度を色に割り当てるものの3種類に分類した。

この3種類の頻度を割り当てる視覚変数がすべて適用でき、且つ頻度を割り当てる視覚変数以外の条件を公平にするための時間軸が環状な形状として扇形を採用した。ここで、扇形を採用することにより、中心からの長さ(半径)を変化させて頻度を表現する場合と外周(弧)の長さを変化させて頻度を表現する場合で、面積比を対応させるという設計案もあった。しかし、Mackinlay [9]によると、量的データを表現する視覚変数として面積よりも長さの方が精度が良いということから、あくまで長さの比を対応させることとした。また、カレンダー表現ではその形状から自ずと頻度を色で表すことになる。

色には使用できる視覚変数が複数考えられる(例:色相, 彩度, 明度など)が、これも Mackinlay [9]によると、量的データを表現する色の視覚変数としては明度 (Density) が最も精度が良い。したがって、扇形の色及びカレンダー表現の色は、共に視覚変数として明度を採用することとした。

以上の検討から4つの周期性の特徴を表現する表現を設計した。また、周期性の特徴を表現することに特化していないと考えられる頻度変化の視覚的表現との比較のため、Lee らの開発した SparkClouds [3] を参考とした折れ線を付加する表現も実装した。

3.1.2 タグクラウドとの組合せ

前節の通り作成した周期性の特徴を表現する表現をタグクラウドと組み合わせるにあたり、配置場所と頻度の総量の表現という2点を検討した。

まず、配置場所について各タグの近傍あるいは背景部分への配置が考えられたが、今回採用したカレンダー表現を背景部分へ配置すると、タグにより遮蔽されて読み取ることができない箇所が生じると考えられた。そのため、本研究で作成する表現はすべて、周期性の特徴の視覚的表現を各タグの近傍にタグへの重畳なく配置することで統一した。

また、各タグのもつ頻度の総量をどのように表現するかということについて、周期性の特徴の視覚的表現に割り当ててタグのフォントサイズを統一する場合と、タグのフォントサイズに割り当てて周期性の特徴の視覚的表現を描画する領域を統一する場合を考えた。本研究の被検者実験では、周期性の特徴の視覚的表現に注目することの重要度が高いタスクを課すことを予定していた。したがって、頻度の総量はタグのフォントサイズに割り当て、周期性の特徴の視覚的表現を描画する領域を統一することとした。

3.2 作成した手法

前節の通り設計・作成し、被検者実験に用いた手法を実際の図を混じえて紹介する。

3.2.1 扇形の半径の長さに頻度を割り当てる表現

周期性を表現する形状として扇形を採用し、頻度を表現する視覚変数として半径の長さを採用した表現。



図 3.1: 扇形の半径の長さに頻度を割り当てる表現の例

3.2.2 扇形の弧の長さに頻度を割り当てる表現

周期性を表現する形状として扇形を採用し、頻度を表現する視覚変数として弧の長さを採用した表現。

3.2.3 扇形の明度に頻度を割り当てる表現

周期性を表現する形状として扇形を採用し、頻度を表現する視覚変数として明度を採用した表現。



図 3.2: 扇形の弧の長さに頻度を割り当てる表現の例



図 3.3: 扇形の明度に頻度を割り当てる表現の例

3.2.4 カレンダー表現

周期性を表現する形状としてカレンダー形状を採用した表現。頻度は明度で表現する。



図 3.4: カレンダー表現の例

3.2.5 折れ線表現

水平方向の時間軸上に頻度変化の様子を折れ線で表現する表現。周期性の特徴の視覚的表現との比較のために作成。

Tag5



図 3.5: 折れ線表現の例

第4章 直感性と共起性に関する被検者実験

第3章で設計した表現を用いて被検者実験を行った。本章では実験の目的・設計・結果・考察を順に述べていく。

4.1 実験の目的

本実験では各表現に用いられる形状と視覚変数の観点から

1. 頻度の周期性の特徴が読み取られる際の直感性 (観点 1)
2. 読み取られる共起性の強さ (観点 2)

を評価することを目的とする。

4.2 実験の設計

4.2.1 観点 1 の実験の設計

観点 1 の実験では直感性を速さと正確さから評価する。具体的なタスクは、1 度の試行ごとに 9 つのタグを提示し、各タグにおける頻度の総量を 100% として問題文で指定する曜日が 25% 以上であれば回答対象であると見なし選択するものである。これを連続 20 試行、全表現について各被検者が行うものとした。

被検者に対する各表現の実施順は、順番による影響をなくすため以下の通り行った。

1. Radius → Arc → Line Chart → Brightness → Calendar (3 名)
2. Brightness → Calendar → Radius → Arc → Line Chart (3 名)
3. Line Chart → Brightness → Radius → Calendar → Arc (3 名)

上記の実施順を決定した基準として、読み取る視覚変数が長さのもの (Radius, Arc) と明度のもの (Brightness, Calendar) に分類し、一方を連続して先に実施するものと後に実施するもの、交互に実施するものとした。Line Chart に関しては長さに含まれる可能性も考慮した順序とした。

評価のための指標として、速さは探索 (回答) 時間、正確さは再現性と精度をそれぞれ用いることとした。再現性と精度は情報検索における性能評価の一般的な指標である。これらの

指標を得るための実験プログラムの出力は以下の通りである。

- 各タグの回答時間
- 各タグの各曜日の頻度の割合 (元データより算出)

再現性及び精度の算出式は以下の通りである。

$$(\text{再現性}) = (\text{正答数}) / (\text{回答対象数}) \quad (4.1)$$

$$(\text{精度}) = (\text{正答数}) / (\text{正答数} + \text{FE 数}) \quad (4.2)$$

尚、FE とは回答対象ではなかったが回答が為されたエラー (False Error) を指す。

実際の実験画面を図 4.1～図 4.5 に示す。

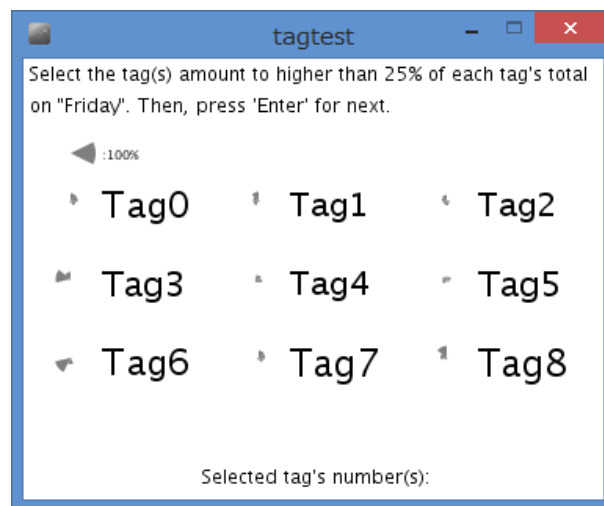


図 4.1: 扇形の半径の長さに頻度を割り当てる表現の実験画面例

4.2.2 観点 2 の実験の設計

観点 2 の実験では、周期性の特徴が表現されることに付随して表現されることが期待される共起性について、タグを生成する元データの組合せごとに共起性の読み取られ方の強弱を評価する。この元データは、1つのタグを生成するデータの分布を、概形が以下の図 4.6～図 4.9 となるような 4 種類に定めて作成した。

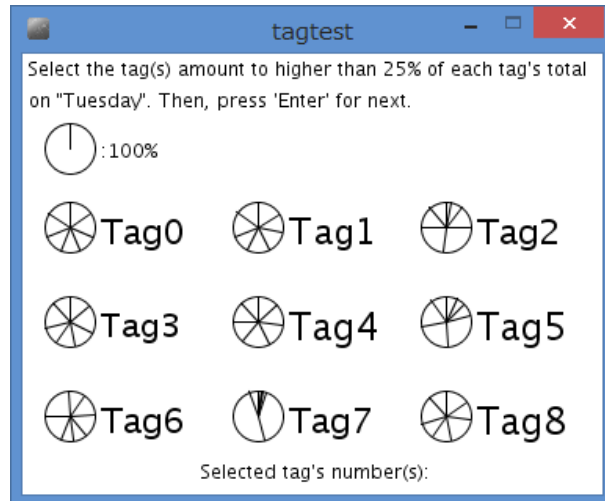


図 4.2: 扇形の弧の長さに頻度を割り当てる表現の実験画面例

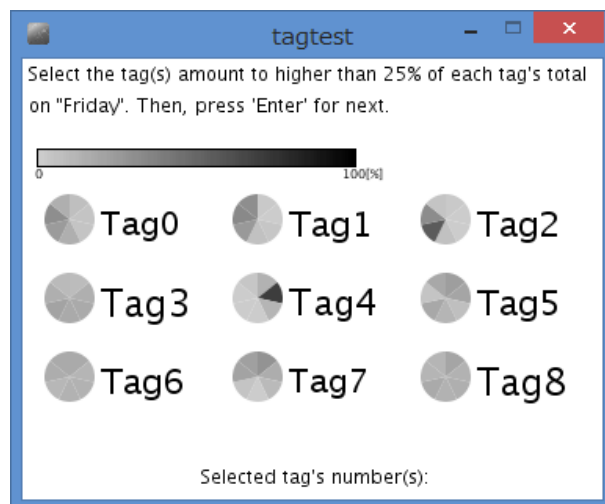


図 4.3: 扇形の明度に頻度を割り当てる表現の実験画面例

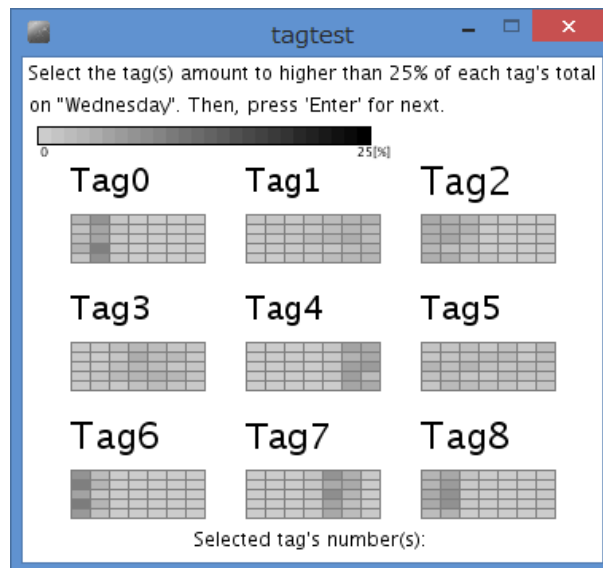


図 4.4: カレンダー表現の実験画面例

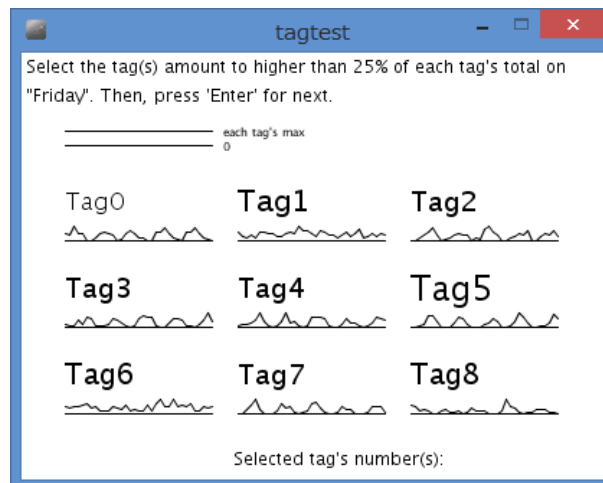


図 4.5: 折れ線表現の実験画面例

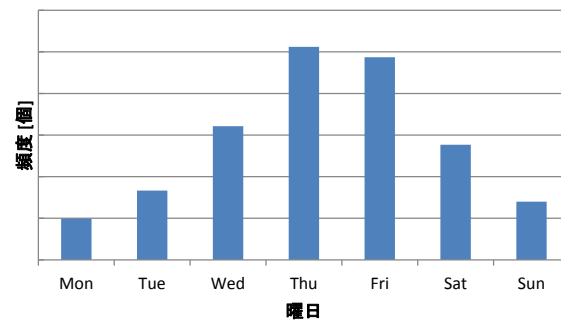


図 4.6: 分布 1 の例

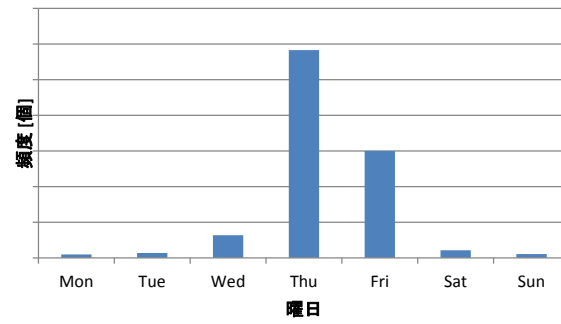


図 4.7: 分布 2 の例

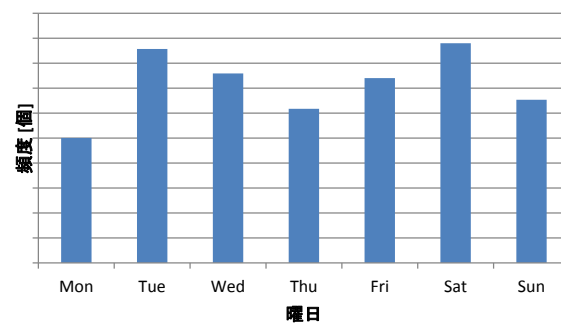


図 4.8: 分布 3 の例

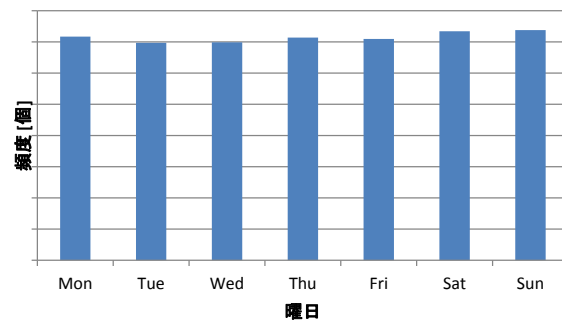


図 4.9: 分布 4 の例

また、この分布に基づく元データの組合せを以下の 13 通りとした。

1. 分布 1 同士で最頻値の位置も同じ
2. 分布 1 同士で最頻値の位置が 1 つずれている
3. 分布 1 と分布 2 で最頻値の位置が同じ
4. 分布 1 と分布 3 で最頻値の位置が 1 つ同じ
5. 分布 1 と分布 4
6. 分布 2 同士で最頻値の位置も同じ
7. 分布 2 同士で最頻値の位置が 1 つずれている
8. 分布 2 と分布 3 で最頻値の位置が 1 つ同じ
9. 分布 2 と分布 4
10. 分布 3 同士で最頻値の位置も両方とも同じ
11. 分布 3 同士で最頻値の位置が片方のみ 1 つずれている
12. 分布 3 同士で最頻値の位置が両方 1 つずつずれている
13. 分布 3 と分布 4

具体的なタスクは、上記の組合せについて同一の元データを各表現で可視化したものを紙の印刷物で提示し、タグの組合せ 13 組それぞれについて共起性を

1. とても弱い
2. 弱い
3. どちらでもない
4. 強い
5. 非常に強い

の 5 段階 Lickert 尺度で回答するものである。

実際に実験で提示した各表現を図 4.10～図 4.14 に示す。

実験で提示した図中のタグの組合せと、評価する元データの組合せの対応は以下の通りで



図 4.10: 扇形の半径の長さに頻度を割り当てる表現の共起性評価実験



図 4.11: 扇形の弧の長さに頻度を割り当てる表現の共起性評価実験



図 4.12: 扇形の明度に頻度を割り当てる表現の共起性評価実験

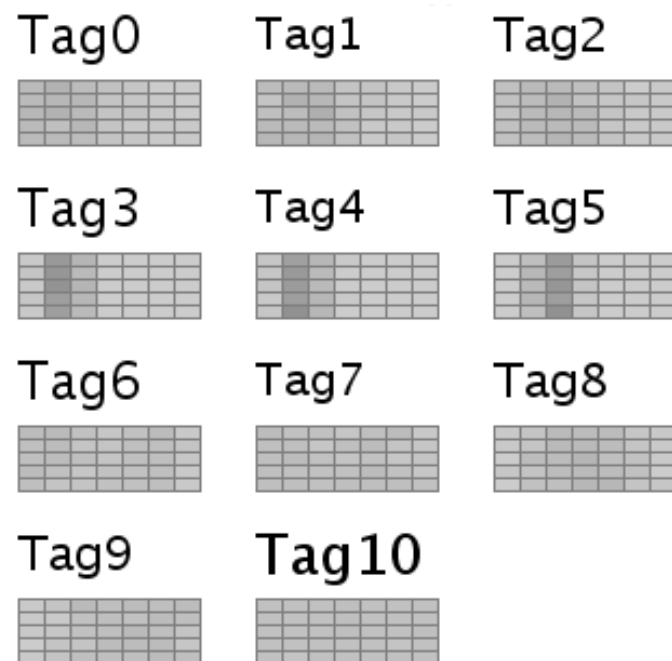


図 4.13: カレンダー表現の共起性評価実験

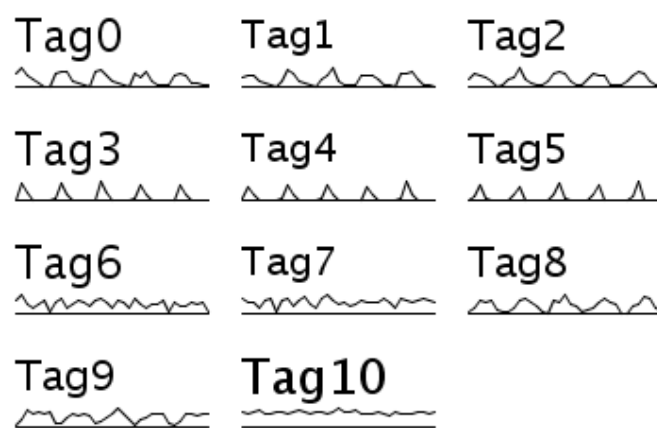


図 4.14: 折れ線表現の共起性評価実験

ある。

1. 元データの組合せ 1 — Tag0 と Tag1
2. 元データの組合せ 2 — Tag0 と Tag2
3. 元データの組合せ 3 — Tag0 と Tag3
4. 元データの組合せ 4 — Tag0 と Tag6
5. 元データの組合せ 5 — Tag0 と Tag10
6. 元データの組合せ 6 — Tag3 と Tag4
7. 元データの組合せ 7 — Tag3 と Tag5
8. 元データの組合せ 8 — Tag3 と Tag6
9. 元データの組合せ 9 — Tag3 と Tag10
10. 元データの組合せ 10 — Tag6 と Tag7
11. 元データの組合せ 11 — Tag6 と Tag8
12. 元データの組合せ 12 — Tag6 と Tag9
13. 元データの組合せ 13 — Tag6 と Tag10

4.2.3 データの作成

実験で提示する表現を生成するためのデータは、観点 1 の実験及び観点 2 の実験のいずれにおいても前述の 4 種類の分布に基づき作成した。

観点 1 の実験用データは、表現する周期の特徴を用いる表現の 1 つにカレンダー形状があることから曜日とした。1 つのタグを生成するデータは 5 週間分とし、頻度の総和を 1 週間分ごとに、正規分布の最頻値の位置を各タグごとにランダムで決定し生成した。

実験プログラムが読み込むファイルとしては、35 行× 756 列の CSV ファイルで、1 列が 1 つのタグの生成に対応している。756 列の内訳は

$(4 \text{ 分布}) \times (\text{最頻値の位置の組合せ } 21 \text{ 通り}) \times (1 \text{ 試行で提示するタグの数 } 9) = 756$

である。この 756 列の中のランダムな先頭位置から 180 列を読み込み、20 試行分のタグを生成する。

観点 2 の実験用データは、表現する周期の特徴を限定せず作成した。周期を構成する単位時間は 7 としたため、形状としては表現する周期の特徴を曜日とした場合と同様である。観点 1 の実験用データとの相違は、各表現で同じ元データを可視化するために、分布の種類と最頻値の位置を指定したタグ 11 個分 1 セットのみのデータを作成した。

4.3 実験結果

前節までの設計による被検者実験を 21～22 歳(当時)の男性 9 名に実施した。実験実施環境は蛍光灯下の明るい室内、観点 1 の実験を実施したディスプレイは以下のものである。

- 型番 : MultiSync LCD 2490WUXi
- 製造会社 : NEC

4.3.1 観点 1

頻度の周期性の特徴を読み取る実験の結果として、各表現ごとに回答対象であったタグの数、被験者の正答数、平均回答時間、回答対象であったが回答が為されなかったエラー (NFE: Not Found Error) の数、FE 数、再現性、精度、F 値 (再現性と精度の平均) を表 4.1 に示す。平均回答時間は正答の回答のみを採択して算出した。

表 4.1: 観点 1 の実験結果

	Radius	Arc	Brightness	Calendar	Line Chart
回答対象数 [個]	219	237	206	56	66
正答数 [個]	88	184	178	56	37
平均回答時間 [ms]	10023	11864	9071	7058	15950
回答時間の分散	49703658.68	59332487.31	21134846.14	25702240.25	198498299.1
NFE 数 [個]	131	53	28	0	29
FE 数 [個]	41	71	279	179	283
再現率	0.402	0.776	0.864	1.000	0.561
精度	0.682	0.722	0.389	0.238	0.116
F 値	0.542	0.749	0.627	0.619	0.338

また、各表現について回答時間、NFE の真値、FE の真値のヒストグラムを図 4.15～図 4.28 に示す。

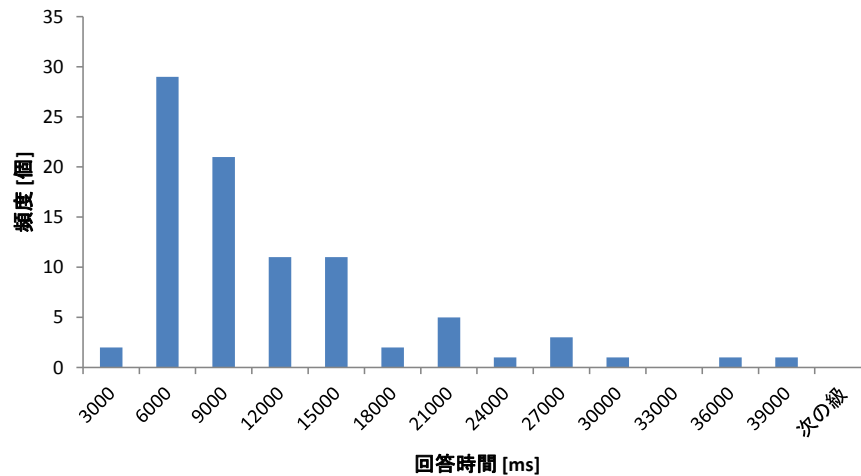


図 4.15: 頻度を半径の長さに割り当てる表現の回答時間のヒストグラム

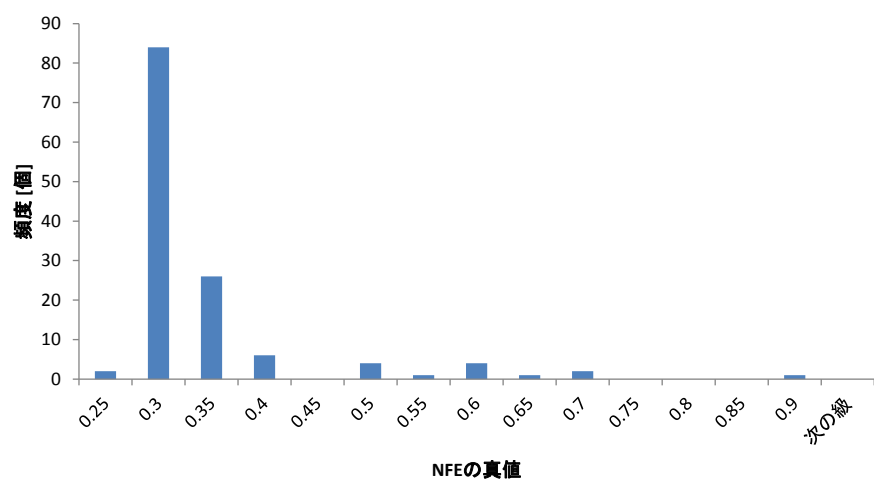


図 4.16: 頻度を半径の長さに割り当てる表現の NFE の真値のヒストグラム

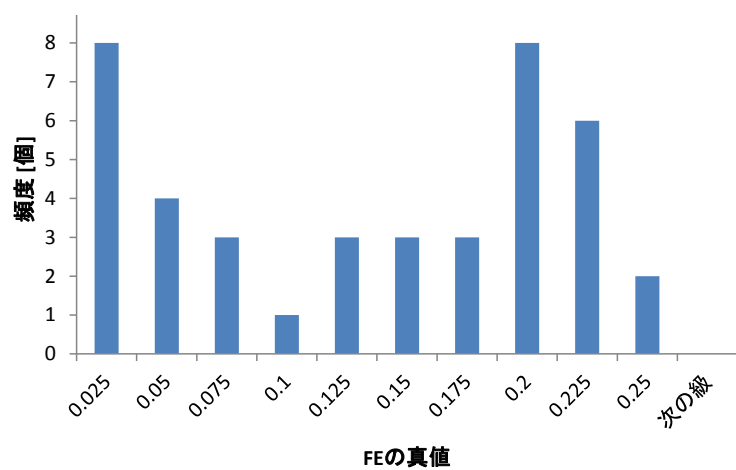


図 4.17: 頻度を半径の長さに割り当てる表現の FE の真値のヒストグラム

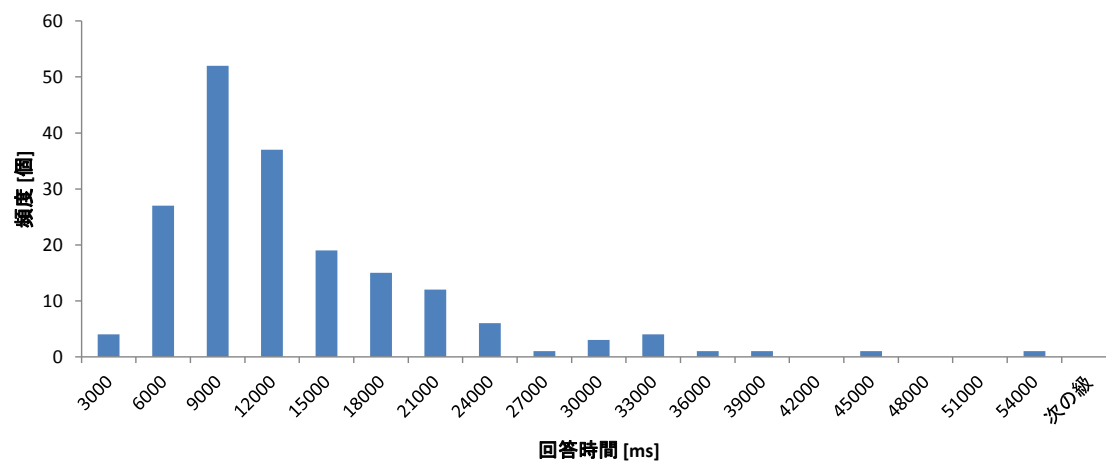


図 4.18: 頻度を弧の長さに割り当てる表現の回答時間のヒストグラム

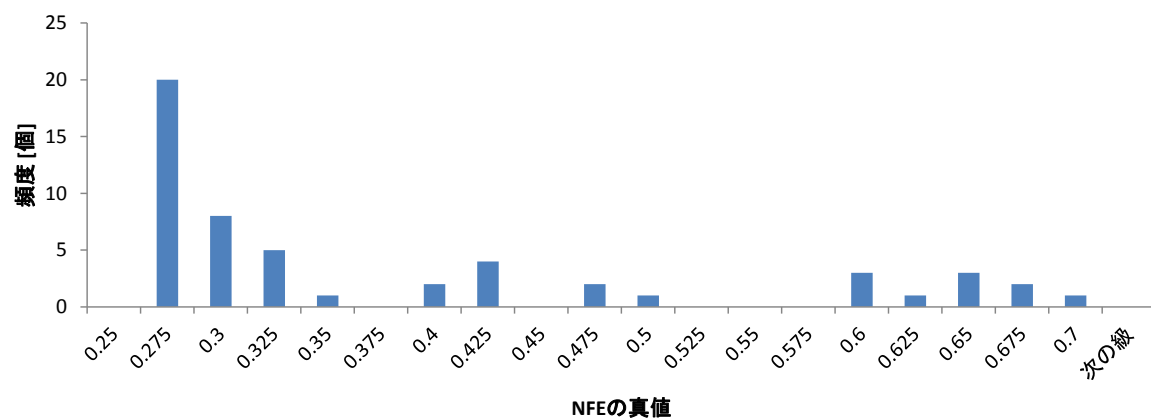


図 4.19: 頻度を弧の長さに割り当てる表現の NFE の真値のヒストグラム

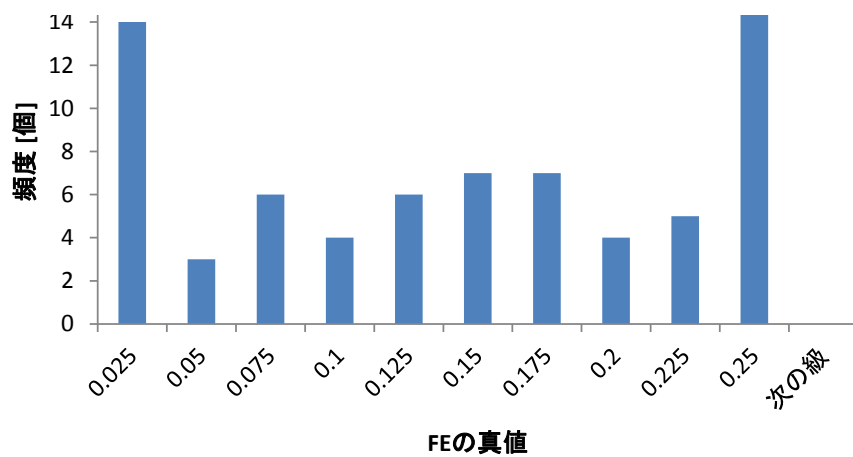


図 4.20: 頻度を弧の長さに割り当てる表現の FE の真値のヒストグラム

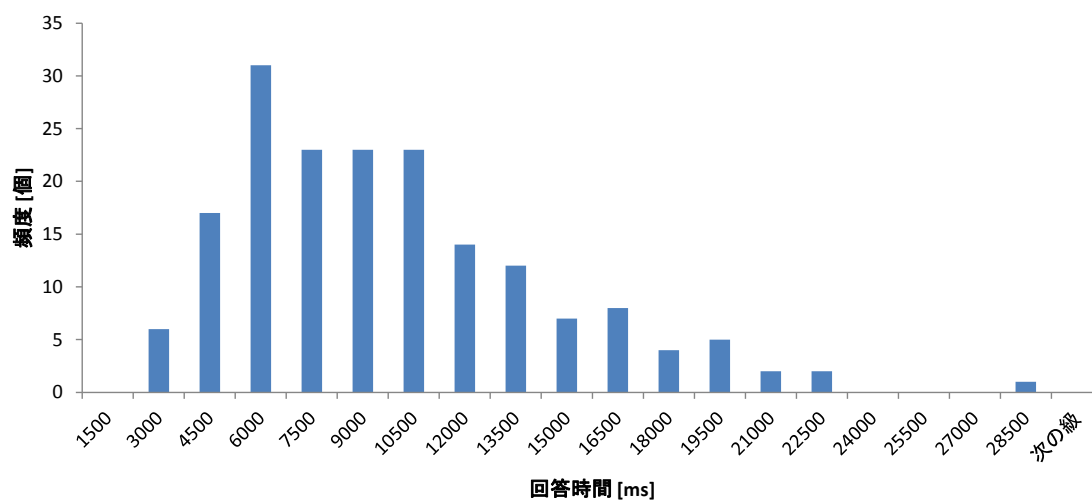


図 4.21: 頻度を明度に割り当てる表現の回答時間のヒストグラム

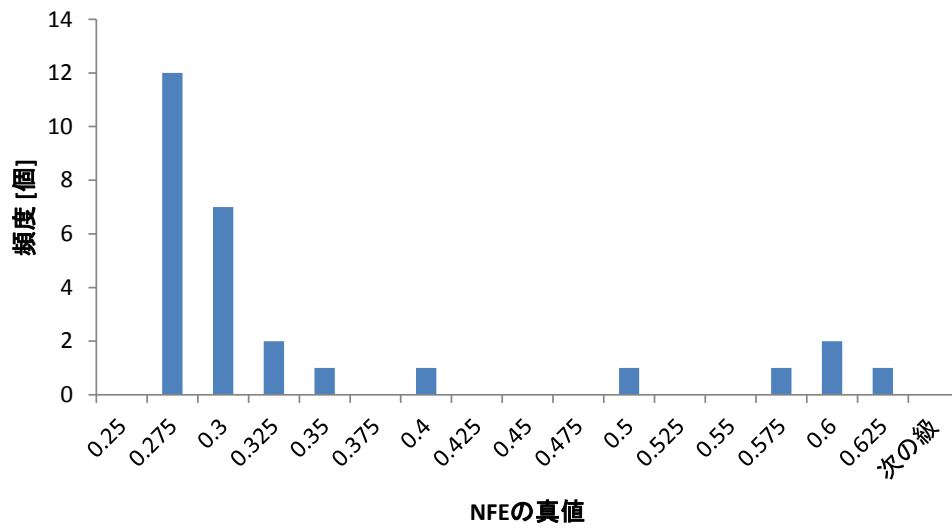


図 4.22: 頻度を明度に割り当てる表現の NFE の真値のヒストグラム

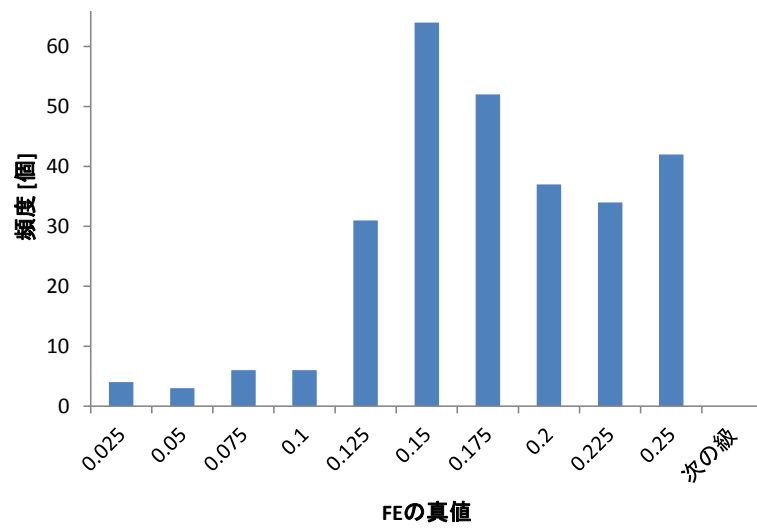


図 4.23: 頻度を明度に割り当てる表現の FE の真値のヒストグラム

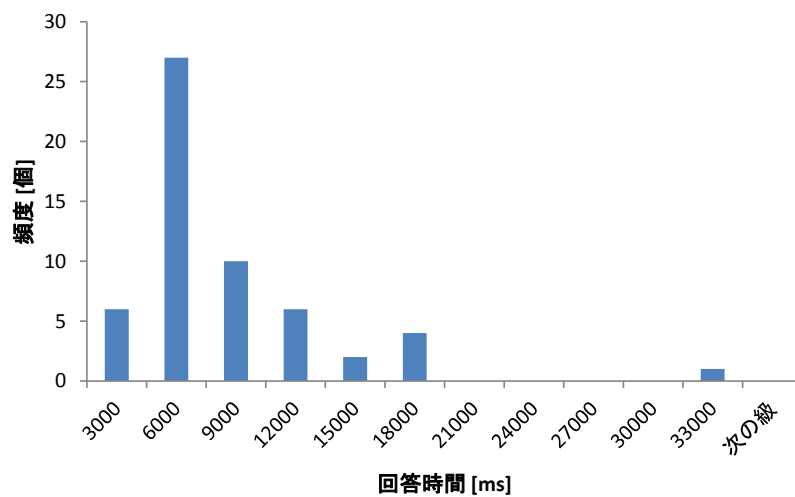


図 4.24: カレンダー表現の回答時間のヒストグラム

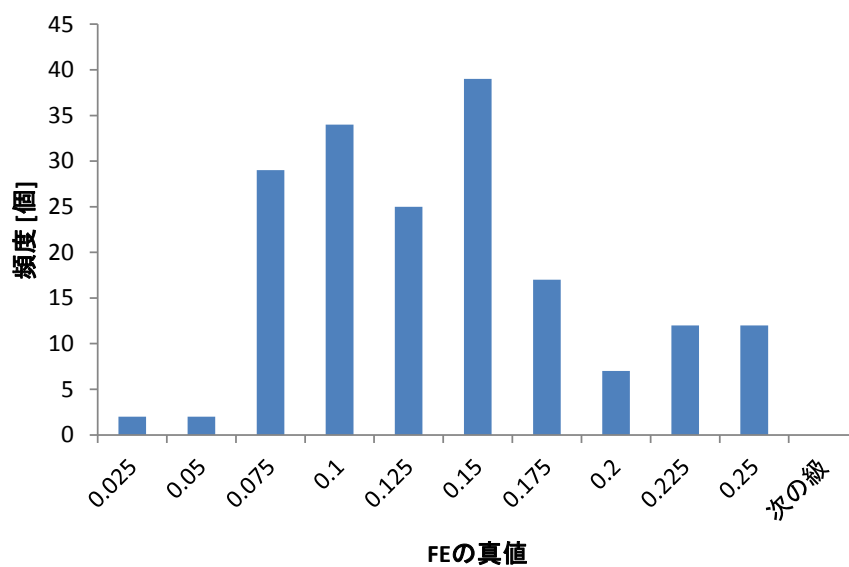


図 4.25: カレンダー表現の FE の真値のヒストグラム

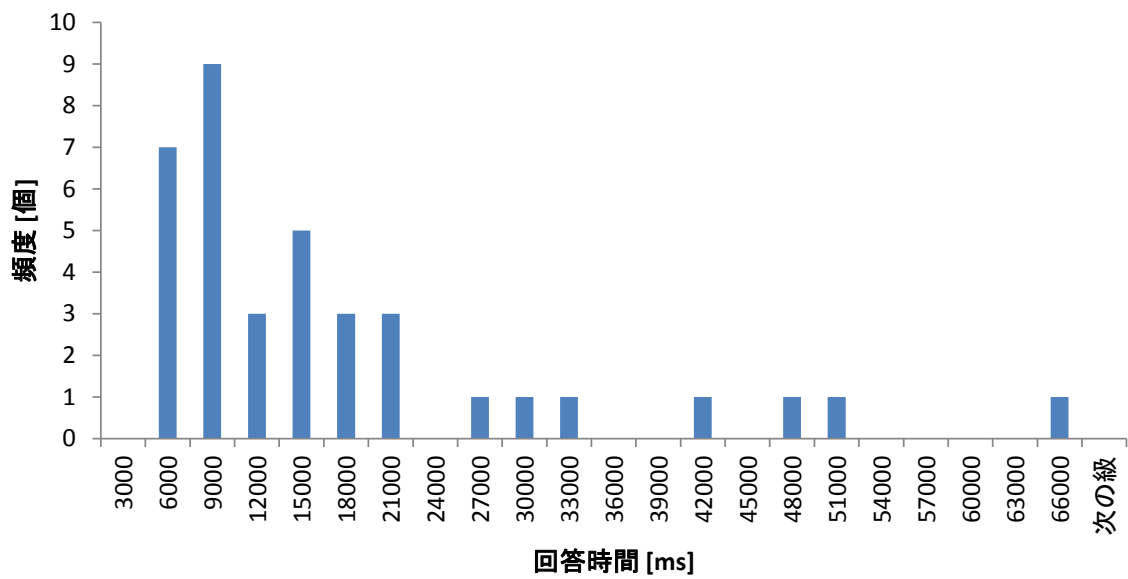


図 4.26: 折れ線表現の回答時間のヒストグラム

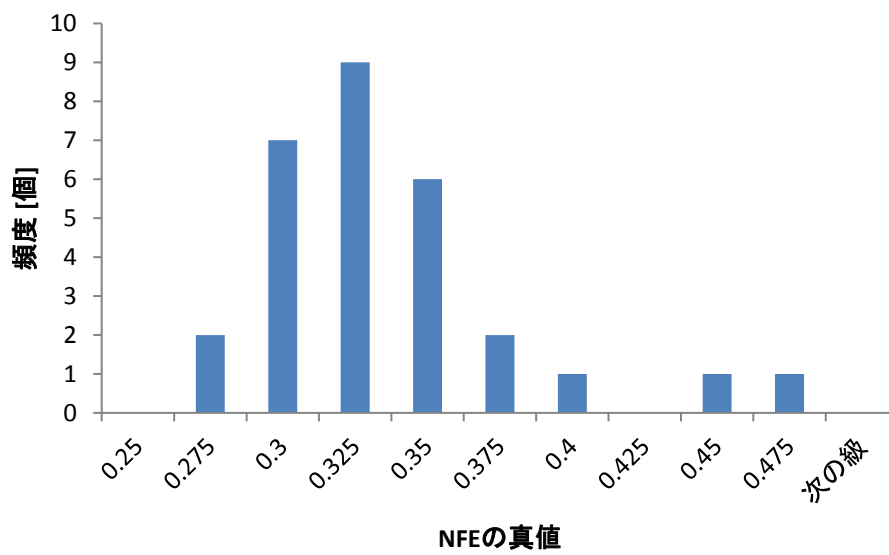


図 4.27: 折れ線表現の NFE の真値のヒストグラム

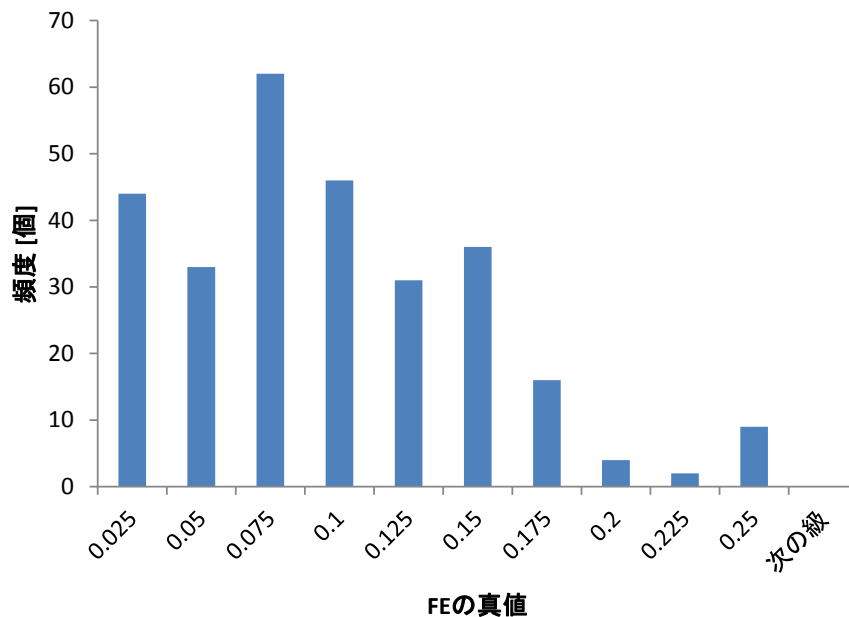


図 4.28: 折れ線表現の FE の真値のヒストグラム

4.3.2 観点 2

まず、結果を示す前に次の点を考慮する。観点 2 の実験実施の際、タグのペアにおける共起性の強弱を「相関」という言葉で教示した。この「相関」という言葉について、被験者間で負の相関を考慮したかどうかという差異が生じてしまった。これについて認識を統一しての再実験を行う時間的猶予がなかったため、今回は負の相関を考慮したグループと考慮しなかったグループに分けて結果を提示し、次節以降の考察や議論を行うこととする。

前述の点を踏まえて、観点 2 の実験結果を表 4.2、表 4.3 に示す。結果の数値は各組合せに対する被験者の回答を、とても弱い=1, 弱い=2, どちらでもない=3, 強い=4, 非常に強い=5 として中央値をとったものである。

4.4 考察

各実験について、実験結果のデータ及び被験者の意見から考察を行った。

4.4.1 観点 1

まず直感性のうち速さの指標とした回答時間から考察する。表 4.1 より、平均回答時間の上位 2 つが Calendar と Brightness であった。この 2 つとその他の表現との特徴的な差異は頻度の表現に色 (明度) を割り当てたかそうでないかであるといえる。したがって、頻度の周期性

表 4.2: 観点 2 の実験結果 (負の相関考慮)

組合せ	Radius	Arc	Brightness	Calendar	Line Chart
1	5	5	5	5	4
2	4	4	4	4	4
3	1	1	4	3	3
4	2	3	3	4	3
5	3	2	2	4	2
6	5	5	5	5	5
7	1	2	4	3	4
8	1	2	3	3	2
9	1	1	2	3	2
10	4	5	4	5	4
11	4	4	3	4	3
12	3	3.5	2	5	3
13	3	4	4	5	3

の特徴を表現し、それを直感的に読み取って欲しいという要求に対しては、頻度の表現には色(明度)を用いることが適していると考えする。また、周期性の特徴を表現することに特化していない視覚的表現として比較した Line Chart は、表 4.1 より平均回答時間が最下位であり、次いで遅い Arc と比較しても約 4 秒の差がある。これが有意差である確認するため、Line Chart と各表現について t 検定をおこなう。

まず、Radius に対して、帰無仮説は「Radius の平均回答時間は Line Chart の平均回答時間以上である」となる。Microsoft Excel を用いて、棄却域 5% の片側 t 検定を行った結果、 $t = -2.43$, $P(T \leq t) = 0.0095$, (t 境界値) = 1.68 であった。P は棄却域 5% より小さく、 t 境界値も t の絶対値より小さいため、帰無仮説は棄却される。よって、Radius の平均回答時間と Line Chart の平均回答時間は統計的に有意である。

Arc に対して、帰無仮説は「Arc の平均回答時間は Line Chart の平均回答時間以上である」となる。同様に棄却域 5% の片側 t 検定を行った結果、 $t = -1.71$, $P(T \leq t) = 0.047$, (t 境界値) = 1.68 であった。P は棄却域 5% より小さく、 t 境界値も t の絶対値より小さいため、帰無仮説は棄却される。よって、Arc の平均回答時間と Line Chart の平均回答時間は統計的に有意である。

Brightness に対して、帰無仮説は「Brightness の平均回答時間は Line Chart の平均回答時間以上である」となる。同様に棄却域 5% の片側 t 検定を行った結果、 $t = -2.94$, $P(T \leq t) = 0.0028$, (t 境界値) = 1.69 であった。P は棄却域 5% より小さく、 t 境界値も t の絶対値より小さいため、帰無仮説は棄却される。よって、Brightness の平均回答時間と Line Chart の平均回答時間は統計的に有意である。

Calendar に対して、帰無仮説は「Calendar の平均回答時間は Line Chart の平均回答時間以上である」となる。同様に棄却域 5% の片側 t 検定を行った結果、 $t = -3.69$, $P(T \leq t) = 0.00032$,

表 4.3: 観点 2 の実験結果 (負の相関非考慮)

組合せ	Radius	Arc	Brightness	Calendar	Line Chart
1	4	4	4.5	5	4
2	3.5	3.5	4	4.5	4
3	2	2	4	2.5	3.5
4	3	3	2.5	4.5	2.5
5	2.5	2.5	2.5	4.5	1.5
6	5	5	5	4.5	5
7	2	2.5	3.5	2.5	5
8	3.5	1	2.5	2	2
9	2	1	2.5	2	1
10	4	5	4	4.5	3.5
11	3	4	3	4	2
12	2	3	2.5	4.5	2
13	2	4.5	3	4.5	2

(t 境界値) = 1.68 であった。P は棄却域 5% より小さく、 t 境界値も t の絶対値より小さいため、帰無仮説は棄却される。よって、Calendar の平均回答時間と Line Chart の平均回答時間は統計的に有意である。

Line Chart と他の 4 つの表現全てとの間で有意差が存在した。したがって、本実験で周期性の特徴の表現に特化させるねらいで設計・実装した表現はすべて、Line Chart よりも読み取られる速さの点で直感的であるといえる。

次に、正確さについて考察する。正確さの指標としては再現性と精度を設けた。これらそれぞれが意味するところとしては、再現性は読み取って欲しい特徴が漏れなく発見させることができるか、精度は如何に読み取って欲しい特徴のみを発見させることができるかということである。

表 4.1 より、再現率が最も高いものは Calendar、最も低いものは Radius である。上位 3 つの表現では 70% 以上、特に Calendar では 100% となり、読み取って欲しい特徴は高い確率で発見させることができると考えられる。また、第 4 位の Line Chart でも 50% 以上と、Radius 以外では半分以上程度は発見させることができると考えられる。

次に、精度について表 4.1 より、上位 2 つの Arc と Radius が約 70%、その他の 3 表現が 40% 未満と大きな差が読み取れる。Arc と Radius、その他の表現との差異は頻度の表現に長さを割り当てたかそうでないかであると考えられる。Line Chart では時間軸と折れ線との間隔に頻度が表現され、これを長さとして捉えることもできるかもしれない。しかし、実線分の長さを読み取るか間隔 (空間の幅) を読み取るか、5 週間分を足し合わせた数値を表現しているかそうでないかという点で区別できると考える。この結果より、頻度の周期性の特徴を表現し、読み取らせたい特徴のみを発見させたい場合には、頻度の表現に長さ (実線分) を用いることが適していると考えられる。

再現率と精度のどちらか一方のみを重視するという設計を行わない場合には、再現率と精度の平均という指標が利用できる。表 4.1 より、再現率と精度の平均が最も高いものは Arc であり、以下 Brightness, Calendar, Radius, Line Chart と続く。再現率では 100% で最高位であった Calendar が第 3 位になることから、色 (明度) による表現は特徴を発見し易くはするが、必要以上に強調しすぎるといった側面があると考えられる。また、再現率では Calendar を下回っていた Brightness が精度との平均を取ることで Calendar よりも上位となることから、色 (明度) の和を考えることの難易度が高いということも考えられる。Line Chart は精度との平均では全表現中唯一 50% を下回った。この結果より、再現率と精度の両方を平均的に重視した場合に頻度の表現に用いることが適しているのは、外周 (弧) の長さであると考えられる。

NFE と FE について、25%(0.25) を基準として探索課題を行っているため、頻度が正しく表現されていればヒストグラムの頻度は 25% 近傍の頻度が高く、25% から離れるほど頻度が低くなることが推測された。しかし、Radius の FE (図 4.17), Arc の FE (図 4.20), Calendar の FE (図 4.25), Line Chart の FE (図 4.28) において FE の真値が 0.00 近傍でも高い頻度が見受けられた。このうち、Arc については区間 (0.025, 0.25) において増加の傾向が見て取れるにも関わらず区間 [0.00, 0.025] の頻度が高い。よって、問題の曜日の頻度がごく小さく扇形が潰れてしまい、且つ隣接する曜日の頻度が回答対象程度以上の頻度であった場合に回答が為されたと考えられる。Calendar については先で考察した通り、色 (明度) の和を考えることの困難さに起因するエラーだと考えられる。

残る Radius と Line Chart に対しては特に被験者から以下のような意見が寄せられていた。

- Radius は中心 (扇形の向き) がわかりにくく難しかった (2 名)
- Line Chart は曜日の基準もなく非常に難しかった (2 名)

上記の指摘通り、Radius は中心 (扇形の向き) がわからなければ問題の曜日に隣接する曜日以外を誤認する可能性があり、Line Chart は曜日の判別を補助するような基準を描画することもしていなかったため、この 2 表現については表現の設計が不適切であったと考えられる。したがって、5 章で述べる追実験について検討を行った。

4.4.2 観点 2

共起性の評価について、まず分布の組合せを以下のように分類することができる。この分

表 4.4: 分布の組合せの分類

分類	分布の組合せ (4.2.2 節)
分布と最頻値の位置が両方同じ	1, 6, 10
分布が同じで最頻値の位置が 1 つずれている	2, 7, 11
分布が同じで 2 つの最頻値の位置が 1 つずつずれている	12
分布が異なるが最頻値の位置が同じ	3, 4, 8
無関係な分布同士	5, 9, 13

類を元に、表 4.2 及び表 4.3 に示した結果から各表現が表す共起性について考察する。

まず、分布と最頻値の位置が両方同じ組合せについて、負の相関の考慮/非考慮に拠らず各表現で 4 点以上と、共起性が強いと判断される傾向にある。

分布が同じで最頻値の位置が 1 つずれている組合せについて、負の相関の考慮/非考慮に拠る大きな影響はないと思われる。特徴的な部分として、Line Chart は分布の組合せ 2 と 7 において共起性が強く読み取られ、分布の組合せ 11 においては共起性がないように読み取られている。この点について、分布の組合せ 2 及び 7 においては Line Chart が周期性の特徴の表現に特化していないことから、最頻値の位置が 1 つずれていても同じ位置に生起したと読み取られたか、ずれが僅かなものとして読み取られたのではないかと推測する。また、分布の組合せ 11 との差として、元データの分布が最頻値を 1 つしか持たないか 2 つ持つかということが挙げられる。最頻値を 2 つ持つ分布同士の組合せである 11 に対しては、Line Chart における山の間隔が異なるという点で共起性が弱く読み取られたのではないかと推測する。

分布が同じで 2 つの最頻値の位置が 1 つずつずれている組合せは、最頻値の位置が異なる正規分布を 2 つ重ね合わせた分布同士のみのものである。これについて、負の相関の考慮/非考慮に拠らず、概ね共起性が弱い或いはないと読み取られている中で、Calendar のみ強く読み取られている。

分布が異なるが最頻値の位置が同じ分布の組合せについて、負の相関の考慮/非考慮によって、分布の組合せ 8 における Radius の結果が大きく異なる。負の相関非考慮の場合の 3.5 という中央値は Lickert 尺度での「どちらでもない」～「強い」にあたり、負の相関考慮の場合の 1 という中央値にあたる「とても弱い」とは大きな差がある。これについて、実際の表現は図 4.10 中 Tag3 と Tag6 であるが、この差の要因を推測することはできなかった。

無関係な分布同士とは、一様分布とその他の各表現との組合せである。この組合せについて、負の相関の考慮/非考慮に拠る大きな影響はないと思われる。特徴的な部分としては、負の相関の考慮/非考慮に拠らず、分布の組合せ 5 における Calendar、分布の組合せ 13 における Arc 及び Calendar が強く読み取られている。分布の組合せ 13 については、分布 3 と一様分布の組合せであり、分布 1～3 のうち最頻値と最も小さい頻度との差が小さい分布であることが要因であると考えられる。

全体的に見た表現ごとの傾向として、Calendar が共起性を最も強く読み取らせる傾向があり、次いで Brightness が共起性を強く読み取らせる傾向にある。この 2 つの表現はともに頻度の表現に色 (明度) を用いており、色 (明度) という視覚変数が共起性を強調する可能性が考えられる。しかし、分布の組合せと照らし合わせて考えると、Calendar では必要以上に共起性を想起させる可能性も考えられる。

Radius と Arc の 2 つ、即ち頻度の表現に長さ (実線分) を用いるものは共起性の強弱を両極端に読み取られる傾向がある。強く読み取られる方は Calendar や Brightness と同等以下であるため、相対的に弱く読み取らせる傾向がある。

Line Chart に関しては分布の組合せの細かな差異が、共起性の強弱の読み取られ方にも表れていると考えられる。即ち、頻度を周期によって重ねずにそのまま表現するというアプローチが最も中立的であり、過剰に強弱を強調しないということを示しているのかもしれない。

第5章 追実験

4.4.2 節での考察により，頻度の表現に中心からの長さ (半径) を用いる表現と折れ線の表現について，表現を再設計し追実験を行った．

5.1 実験の目的

頻度の周期性の特徴を表現する視覚的表現として

- 頻度を表現する視覚変数として中心からの長さ (半径) を用いる表現
- 折れ線表現

のより良い表現を設計し，頻度の周期性の特徴について直感性と正確さを正しく評価出来るデータを得ることを目的とする．

5.2 表現の再設計

5.2.1 頻度を表現する視覚変数として中心からの長さ (半径) を用いる表現

被験者からの意見より，扇形の中心がわからず方向がわからないことが問題であると解釈した．そこで，すべての扇形について 100% の大きさの枠線を描き，100% の扇形の半径対して何% の半径まで内部が塗りつぶされているかをわかりやすくする改良を施した．改良した表現を図 5.1 に示す．



図 5.1: 扇形の半径の長さに頻度を割り当てる表現 (改良版) の例

5.2.2 折れ線表現

被験者からの意見より，週や曜日の基準が描かれていないことが問題であると解釈した．そこで，折れ線表現の時間軸の下側に，週の先頭を示す目印を加える改良を施した．改良した表現を図 5.2 に示す．

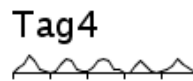


図 5.2: 折れ線表現 (改良版) の例

5.3 実験の設計

基本的には 4 章の観点 1 同様に実施した。但し、使用した表現は再設計した 2 表現のみであり、実施順は以下の通り行った。

1. Radius → Line Chart (5 名)
2. Line Chart → Radius (4 名)

実際の実験画面を図 5.3 及び図 5.4 に示す。

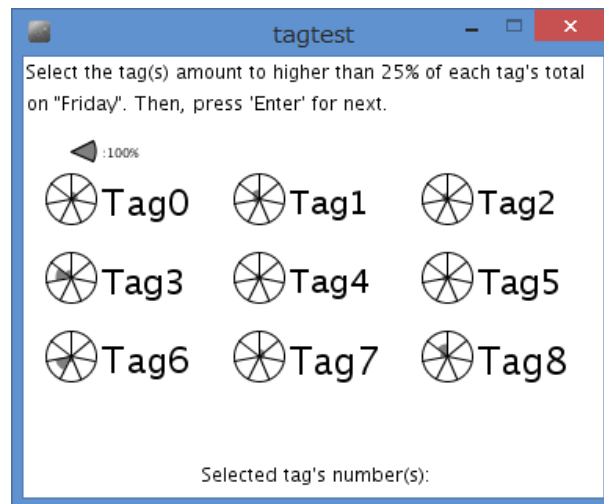


図 5.3: 扇形の半径の長さに頻度を割り当てる表現 (改良版) の実験画面例

5.4 実験結果

他の 3 表現と同時に実験を実施しなかったため厳密に比較することはできないが、4 章の観点 1 の実験と同様に、回答対象数、正答数、平均回答時間、NFE 数、FE 数、再現率、精度、再現率と精度の平均を表 5.1 に示す。

また、各表現について回答時間、NFE の真値、FE の真値のヒストグラムを図 5.5～図 5.10 に示す。

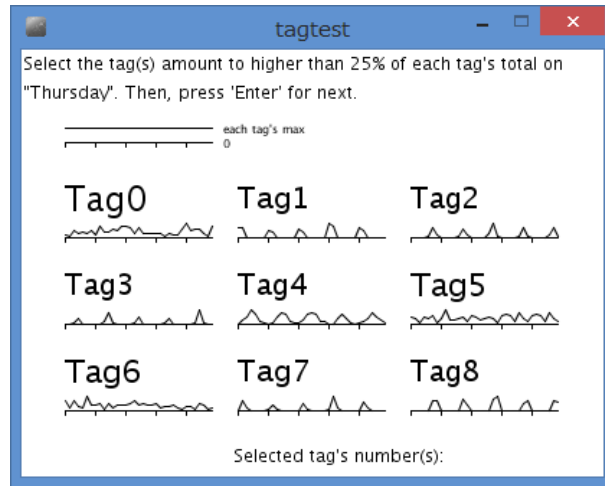


図 5.4: 折れ線表現 (改良版) の実験画面例

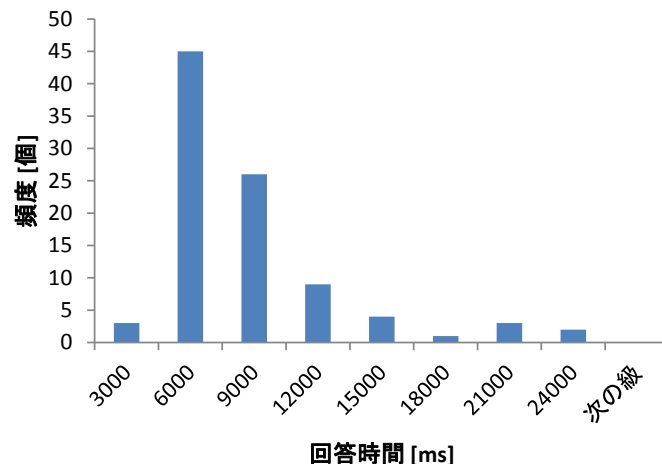


図 5.5: 頻度を半径の長さに割り当てる表現 (改良版) の回答時間のヒストグラム

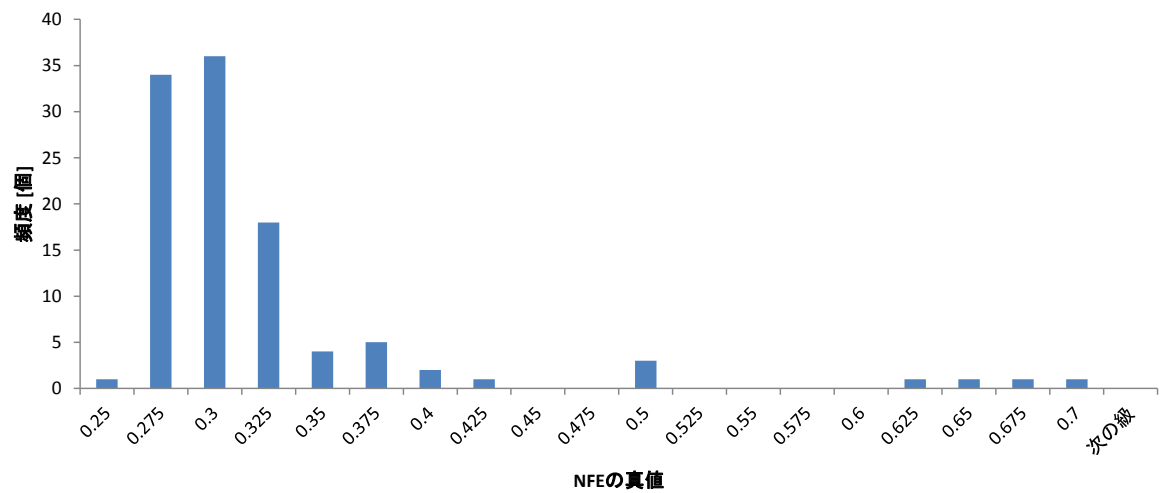


図 5.6: 頻度を半径の長さに割り当てる表現 (改良版) の NFE の真値のヒストグラム

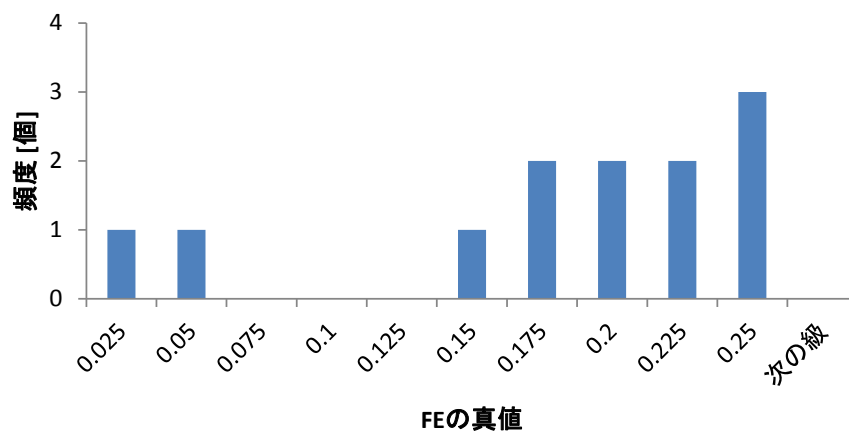


図 5.7: 頻度を半径の長さに割り当てる表現 (改良版) の FE の真値のヒストグラム

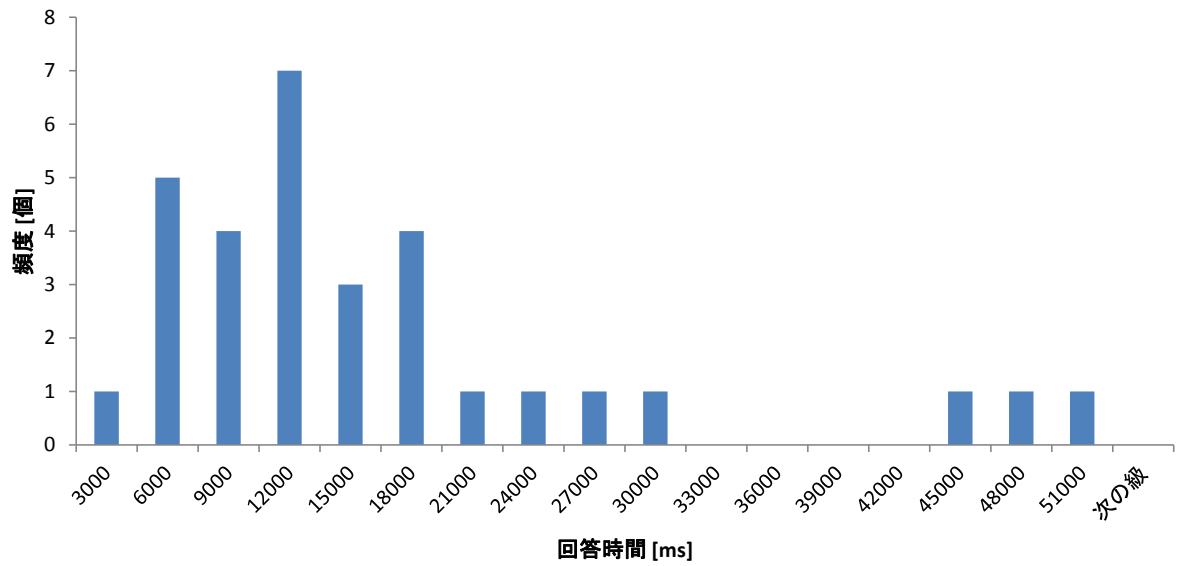


図 5.8: 折れ線表現 (改良版) の回答時間のヒストグラム

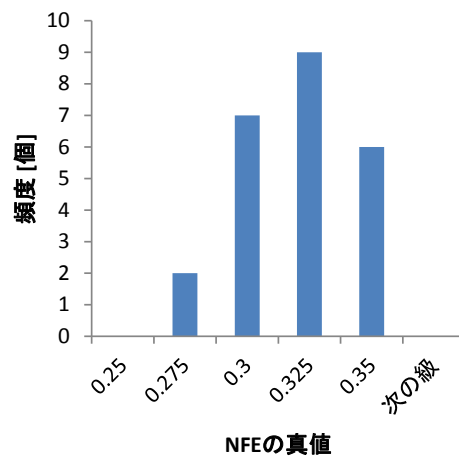


図 5.9: 折れ線表現 (改良版) の NFE の真値のヒストグラム

表 5.1: 追実験結果

	Radius'	Line Chart'
回答対象数 [個]	201	55
正答数 [個]	93	31
平均回答時間 [ms]	7199	14956
回答時間の分散	17984321.3	149653245.7
NFE 数 [個]	108	24
FE 数 [個]	12	248
再現率	0.463	0.564
精度	0.886	0.111
再現率と精度の平均	0.674	0.337

5.4.1 考察

実験結果 (表 5.1) より、参考までに 4 章の観点 1 の実験結果 (表 4.1) と比較を行う。

まず、平均回答時間について、Radius' (頻度の表現に半径の長さを用いる表現の改良版) の方が改良前の Radius よりも約 3 秒早くなった。Line Chart' (折れ線表現の改良版) においても、約 1 秒の向上が見られた。

次に、再現率について、Radius' では約 6% の向上が見られた。Line Chart' についてはほとんど変化が見られなかった。

精度については、Radius' では約 20% の向上が見られた。Line Chart' についてはほとんど変化が見られなかった。

再現率と精度の平均は、Radius' では約 13% の向上が見られた。Line Chart' についてはほとんど変化が見られなかった。

また、NFE 及び FE の真値のヒストグラムを確認すると、特に Radius' の FE の真値のヒストグラム (図 5.7) において、FE の真値が 0.00 近傍の頻度が大幅に減少したという変化が見られた。Line Chart' の FE の真値のヒストグラム (図 5.10) においても頻度分布の変化は見られたが、依然として 0.25 近傍の頻度は低く、改良により期待されたものではなかったといえる。

以上の変化を踏まえると、頻度の表現に半径の長さを用いる表現の改良には成功した可能性が高い。他の 3 表現と同時に実験を実施し、追実験と同等の結果が得られていたならば、回答時間に関しては最も早かったカレンダー表現とほぼ並び、精度に関しては全表現の中で最も高いと評価することができた可能性がある。

一方で折れ線表現の改良版については改良に失敗した可能性が高い。もしくはこれ以上の改良には、例えば周期単位で折れ線を重ねるなど根本的な形状の変更などが必要である可能性があると考えられる。

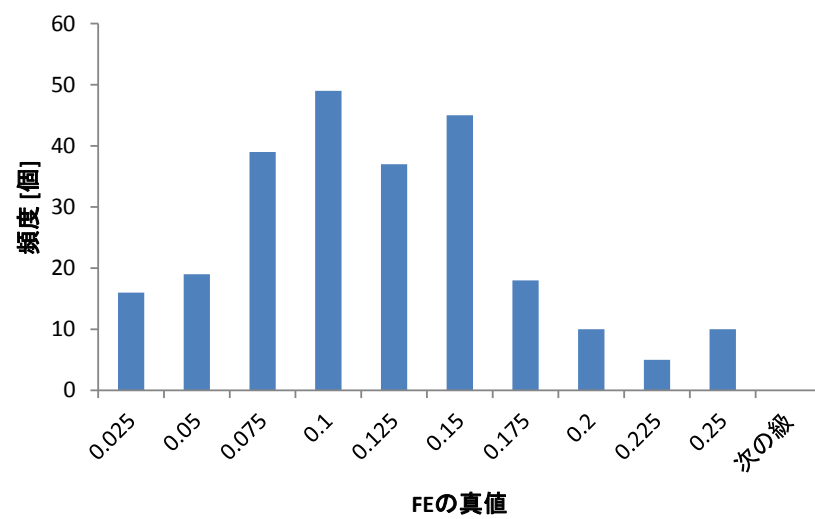


図 5.10: 折れ線表現 (改良版) の FE の真値のヒストグラム

第6章 議論

周期性の特徴を表現することに特化していない時刻情報付きデータの視覚的表現とタグクラウドの組合せ表現において、頻度の周期性の特徴を直感的に把握できないことが課題であった。本研究において、頻度の周期性の特徴を表現する表現をタグクラウドと組合せたことにより、時刻情報付きデータの頻度の周期性の特徴をより直感的 (速く、正確に読み取られる) に表現することができた。また、頻度の周期性の特徴を表現したことに付随して、それぞれの表現がタグ間の共起性をどのように表現するかということを評価することができた。タグクラウドにおいてタグ間の共起性を見たいという要求は、複数の事象を俯瞰的に見せるというタグクラウドの性質から効果的な要求であり、共起性を提示するタグクラウドの設計において1つの指針となり得る知見を得られたことは有益であったと考える。

一方で、表現や実験の設計には改善の余地があると考えられ、周期性の特徴を把握することにより適している表現のバリエーションを検討する、実験タスクを見直す、といった必要性があると考えられる。特に、共起性の実験においては元データの分布や組合せについての知見が曖昧な設計となってしまったため、より緻密な実験計画に基づく共起性の表現の特徴分析の必要性があると考ええる。

また、得られた知見から具体的に有効なユースケースを提示することも必要であると考ええる。

今後の大きな課題としては、大規模なデータを対象とした際に、他の可視化手法と比較してタグクラウドとしての長所を生かせるのかという検証や、連続時間での頻度変化の様子と周期性の特徴を両立して表現できるタグクラウドの開発が考えられる。

第7章 結論

本研究では、一度に複数の事象を俯瞰させるというタグクラウドの性質から、時刻情報付きデータに対して頻度の周期性の特徴を直感的に表現することを目指した。タグクラウドを拡張する方法としては他の視覚的表現との組合せを採用し、形状や視覚変数、タグクラウドとの組合せの面から、周期性の特徴を表現する手法を4つ設計・実装した。周期性の特徴を表現する手法及び頻度の周期性の特徴の表現に特化していない頻度変化の可視化手法が、頻度の周期性の特徴を如何に直感的に表現するのかを被験者実験により評価した。

その結果、頻度の周期性の特徴を読み取られる時間の面から直感的に表現するには、カレンダー表現、あるいは向きが判別できるように設計された図形の中心からの長さ(半径)に頻度を割り当てる表現が適していると考えられた。正確さの面では、カレンダー表現を筆頭に頻度を扇形の色(明度)で表現する表現において再現率が高く、中心からの長さ(半径)や外周(弧)の長さのように実線分の長さで表現する表現において精度が高く表現された。また、再現率と精度の平均では外周(弧)の長さで表現する表現が最も優れている。周期性の特徴を表現するために設計・実装した表現は、ほぼ全ての指標において周期性の特徴の表現に特化していない表現(折れ線)に優っていた。但し、カレンダー表現では再現率に対して精度が極端に低いなど、色(明度)を足し合わせることに関する認知の幅の広さが窺えた。

本研究ではまた、頻度の周期性の特徴を表現することに付随して表現されることが期待された、時間共起性についても評価した。13組の分布の組合せについて評価を行ったが、全体の傾向としてカレンダー表現では他の表現に比べて共起性が強く読み取られていた。また、次いで扇形の色(明度)による頻度表現でも共起性がやや強めに読み取られており、共起性は色(明度)という視覚変数により強調されやすいと考えられる。中心からの長さ(半径)や外周(弧)の長さで頻度を表現した場合には、他の表現に比べて若干、共起性の弱いものがより弱く読み取られている可能性がある。

謝辞

本研究を進めるにあたり，指導教員である三末和男准教授には非常に多くの丁寧なご指導・ご助言を賜りました．この場をお借りして心より感謝を申し上げます．また，インタラクティブプログラミング研究室教員である志築文太郎准教授，高橋伸准教授，田中二郎教授におかれましても，ゼミや研究活動の中で多くの学びを与えてくださいました．大変感謝しております．

インタラクティブプログラミング研究室の皆さまとは，まだ1年足らずの付き合いである方がほとんどではありますが，互いに刺激し合い，尊敬できる仲間と環境に恵まれたと実感しております．特に，NAIS チームの皆さまには日頃の研究室生活から懇意にいただき，研究活動の大きな支えでありました．1年間の感謝をここに申し上げます．

最後に，家族や友人をはじめ，公私に亘り大学生活の中でお世話になった全ての方々に，心より感謝を申し上げます．ありがとうございました．

参考文献

- [1] Dinh Quyen Nguyen, Christian Tominski, Heidrun Schumann, and Tuan Anh Ta. Visualizing tags with spatiotemporal references. In *Proceedings of the 15th International Conference Information Visualisation (IV 2011)*, pp. 32–39, 2011.
- [2] Edward R Tufte. *Beautiful evidence*, Vol. 23. Graphics Press Cheshire, CT, 2006.
- [3] Bongshin Lee, Nathalie Henry Riche, Amy K Karlson, and Sheelagh Carpendale. Sparkclouds: Visualizing trends in tag clouds. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 16, No. 6, pp. 1182–1189, 2010.
- [4] Christopher Collins, Fernanda B Viegas, and Martin Wattenberg. Parallel tag clouds to explore and analyze faceted text corpora. In *Visual Analytics Science and Technology. IEEE Symposium on VAST 2009.*, pp. 91–98, 2009.
- [5] Steffen Lohmann, Michael Burch, Hansjörg Schmauder, and Daniel Weiskopf. Visual analysis of microblog content using time-varying co-occurrence highlighting in tag clouds. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp. 753–756, 2012.
- [6] John V Carlis and Joseph A Konstan. Interactive visualization of serial periodic data. In *Proceedings of the 11th annual ACM symposium on User interface software and technology*, pp. 29–38, 1998.
- [7] Marc Weber, Marc Alexa, and Wolfgang Müller. Visualizing time-series on spirals. In *Infovis*, Vol. 1, pp. 7–14, 2001.
- [8] Christian Tominski and Heidrun Schumann. Enhanced interactive spiral display. In *Proceedings of the Annual SIGRAD Conference, Special Theme: Interactivity*, pp. 53–56, 2008.
- [9] Jock Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics (TOG)*, Vol. 5, No. 2, pp. 110–141, 1986.