

fNIRS で取得した脳活性に基づく文書理解度推定法

吉村 和代[†] Kai Kunze^{††} 黄瀬 浩一[†]

[†] 大阪府立大学大学院工学研究科

〒 599-8531 大阪府堺市中区学園町 1-1

^{††} 慶應義塾大学大学院メディアデザイン研究科

〒 223-8526 横浜市港北区日吉 4-1-1

E-mail: †yoshimura@m.cs.osakafu-u.ac.jp, ††kai@kmd.keio.ac.jp, †††kise@cs.osakafu-u.ac.jp

あらまし 本稿では、fNIRS で取得した脳内の酸素化ヘモグロビンの情報を用いて、文書に対する理解度を推定する手法を提案する。本研究の最終目的は、アイトラッカで取得した視点情報を用いて、文書に対する理解度を推定する手法の提案である。以前、我々は視点情報から直接理解度を推定する手法を提案したが、十分な精度を得るには至っておらず、眼球運動と理解度の関連も明らかになっていない。そこで我々は、次の手段として脳の活性の情報を用いることで、眼球運動と理解度の関連を明らかにすることを旨とする。脳の活性は文書に対する理解度と密接に関連していると仮定する。この仮定が成り立つならば、眼球と理解度の関係を明らかにするという問題は脳の活性と眼球運動の関係を明らかにする問題へと変換される。そこで、脳の活性と理解度の関連性を明らかにし、さらに脳の活性と眼球運動の関連性を明らかにすることで、眼球運動と理解度の関連性を調査する。本稿ではその第一歩として、脳の活性と理解度の関連を明らかにする。本手法ではまず、fNIRS で得られた酸素化ヘモグロビンの情報に対し、特徴量を抽出し、機械学習を用いることで理解度を推定する。被験者 4 人、5 文書を用いた実験の結果、3 段階で理解度を表現したところ、識別率 80 %、平均誤差 0.25 を得た。

キーワード アイトラッカ、視点情報、文書理解度、fNIRS

1. はじめに

人は日常的に読むことで多くの情報を得ている。そのため、ユーザが読んだ情報を記録、分析することができれば非常に有益な情報の取得が可能になる。これを実現するものとして、リーディングライフログが提案されている [1]。リーディングライフログとは、人がどのような文書を読んでいるかなど、人と文書との関わりを記録するものである。

文書との関わりを表す手がかりとして、文書を読む際の人の目の動きが考えられる。換言すれば、文書を読む際の眼球運動を測定し、分析することで文書との関わりを得ることが考えられる。具体的に得られるものとしては、読んだ単語数といった低次の情報から文書に対する理解度といった高次の情報まで様々なものが考えられる。本研究では、このうち最も高レベルの文書理解度に焦点をあてる。

理解度を容易に推定できれば、自分や他者の理解度を簡単に確認することができ、自分の学習の見直しや理解度の統計データをとるなどに利用することができる。更に、従来のペーパーテストによる理解度測定では対象の内容全てを問題に反映させることは不可能であり、学習活動全体を評価の対象にすることはできない。そのため、テストの問題に反映されていない学習内容に関しての理解度はわからないままである。しかし、眼球運動を利用した理解度推定では、文書全体を読む際の眼球運動

を対象とすることで、文書全体の内容を反映することができる。

以前、我々は眼球運動から理解度を推定する手法を提案した [2]。この手法では、視点を入力とし、理解度を出力とする。以下にこの手法の概要を説明する。まず、アイトラッカと呼ばれる眼球運動を測定する装置を用いて、ユーザの文書を読む際の眼球運動を取得する。そして、得られた眼球運動を数値化し、理解度ごとに異なる眼球運動のパターンを得る。それらを学習することで理解度を推定する。しかし、この手法では十分な精度を得るに至っていない。その理由としては、眼球運動と理解度の関係は、学習に用いたデータから把握できるほど単純でなかったことが考えられる。

それでは何があれば推定可能であろうか。我々は手がかりとして脳の活性に着目することを考える。これは、脳の活性は、文書の理解度とともに眼球運動にも密接に関連していると考えられるからである。

そこで、我々は文書を読んでいる際の被験者の脳の活性を利用することで、この推定を実現することを目指す。つまり、眼球運動と理解度との間に脳の活性を挟むことにより、間接的に眼球運動と理解度の対応関係を把握するというのである。具体的には、1) 脳の活性と理解度の関係を明らかにする、2) 脳の活性と眼球運動を明らかにするという 2 つのステップとなる。本稿ではその第一歩として、脳の活性と理解度との関連を明らかにすることを目指す。脳の活性状態の計測には、

近赤外光イメージング装置 (fNIRS: Functional Near-Infrared Spectroscopy) [3] を用いる。fNIRS は近赤外分光法を利用することで、酸素化ヘモグロビン (oxyHb) の濃度変化を算出し、脳活性を計測するものである。この装置を用いて読書時の脳の活性状態を計測し、得られた脳活性から特徴量を抽出し、理解度を推定する。なお、本研究は大阪府立大学大学院工学研究科の倫理委員会で承認済みであることを付記しておく。

2. 関連手法

ここでは、文書と眼球運動に関連する研究と人間の言語理解のメカニズムとそれに関する研究について述べる。

2.1 文書と眼球運動に関連する研究

眼球運動の測定にはアイトラッカと呼ばれる装置が利用される。アイトラッカはカメラを用いてユーザの眼球運動を測定するものや、電極を装着することで電位差を利用して眼球動きを測定するものがある。アイトラッカはマーケティングの分野等によく使用されており、ユーザの興味の推定等に利用される。

文書と眼球運動に関連する研究として、ユーザの読むという行動を認識する研究がある。Bulling らは眼球運動から、ユーザが室内で椅子に座って行う行動を認識する手法を提案した [4]。この手法は、頻出する眼球運動のパターンを符号化することで、ユーザの行動を認識するものである。また、Kunze らは眼球運動を用いて、ユーザの読んだ単語数を推定する手法を提案した [5]。この手法は、一行当たりの平均単語数といった文書の情報と眼球運動を組み合わせることにより、単語数を推定するものである。人が一日に読む単語の数とその人の学力や一般常識には相関関係があることが知られている [6]。そこで、この手法を用いることでユーザの読んだ単語数を自動で記録することができれば、学力をはかる有効な手がかりとなる。また、どんな単語をどれくらい読んだかを知ることができれば、ユーザの興味を推定することもできる。更に、自分が読んだ単語数を把握する、もしくは単語数を他者と比べることで、自分の読書習慣を客観的に解析することが可能になる。しかしながら、これらの研究はいずれも理解度などの高レベルの情報を引き出すものではない。

2.2 人間の言語理解のメカニズムとそれに関する研究

次に、人間の言語理解のメカニズムとそれに関する研究について説明する。言語機能に関連する代表的な領域としては、ウェルニック野、角回、ブローカ野などがわかっている [7]。しかし、これらの領域以外に脳の様々な領域が言語処理に関連していると考えられており、詳しい機能分布などは明らかになっていない。ところが近年の技術革新のより、新しい脳機能イメージング法が開発され、人間の内面的なメカニズムを客観的に観察することが可能となっている。そのため、人間の言語理解と脳の活性に関する研究が盛んに行われるようになってきている。

脳の活性と第二言語習得に関する包括的な研究として大石の研究がある [8]。大石はウェルニック野・角回・聴覚野・縁上回におけるリーディングとリスニング課題遂行時の活性パターンを調査した。その結果から、言語の習熟度と課題の難易度によって特定の脳活性のパターンがあることを明らかにした。特

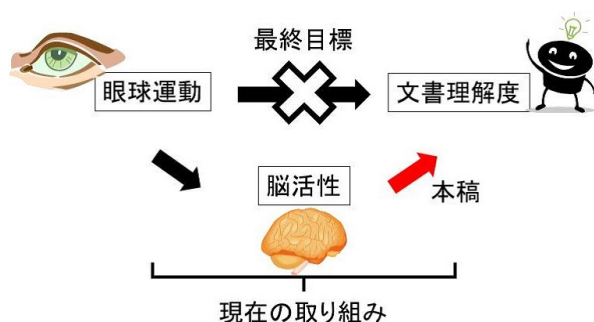


図 1 提案手法の概要

に習熟度と脳活性の関係に着目すると、初級レベルの被験者では脳の活性が見られず無活性型という状態になり、習熟度が上がるほど脳の活性が増していく (過剰活性型)。更に上級レベルになると活性増加の割合が減少し、必要な部位のみが活性する選択的活性型になり、母語話者の脳活動のパターンに近づく。しかし、この研究では理解度と脳活性については十分には言及されていない。

3. 提案手法

図 1 に提案手法の概要を示す。本研究の最終目標は、眼球運動から文書に対する理解度を推定する手法の提案である。その手法の提案のために、脳の活性と理解度の関連性を明らかにし、さらに脳の活性と眼球運動の関連性を明らかにすることで、眼球運動と理解度の関係を把握することを目指す。本稿はその第一歩として、脳の活性と理解度の関連を明らかにする。まず、ユーザの脳の活性状態を取得するために、fNIRS を用いてユーザの脳内の酸素化ヘモグロビンの濃度変化を記録する。次に得られた脳の活性状態を分析し、特徴を抽出する。このようにして得られた特徴量を用いて、機械学習を使用することにより、文書に対する理解度を推定する。

3.1 fNIRS を用いた脳の活性状態の取得

脳の活性状態の取得には、fNIRS を用いる。本稿で使用している fNIRS は脳イメージング装置の 1 つである。fNIRS は騒音もなく、体の位置や向きに制約が少なく自然な状態での計測が可能となっており本研究に適している。この装置は近赤外分光法を利用することで、脳血流量の変化を測定し、図 2 のような酸素化ヘモグロビンの濃度変化を算出する。計測結果は絶対値ではなく、単位時間前と比較した酸素化ヘモグロビン濃度の変化量となっている。そのため、fNIRS の計測の際に短い休憩タスクと実験タスクを交互に試行する必要がある。

以下に計測方法について説明する。図 3 に測定デバイスであるファイバホルダーを装着した様子を示す。被験者の頭部に測定デバイスであるファイバホルダーを取り付けることにより計測する。ファイバホルダーは数種類提供されており、測定部位を選択することができる。

次に得られる計測結果について説明する。図 4 に計測結果とその分析の概要を示す。測定ポイント (チャンネル) ごとに酸素化ヘモグロビンの濃度変化が計測される。更に、酸素化ヘ

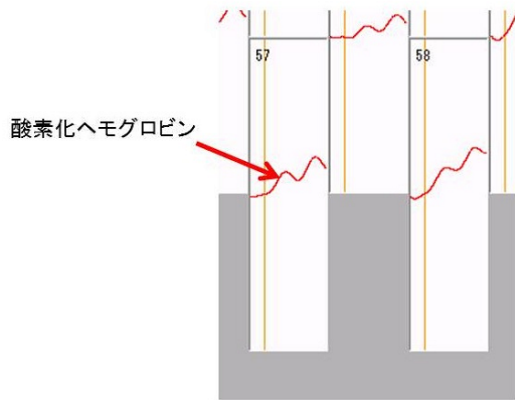


図 2 酸化ヘモグロビンの濃度変化の様子



図 3 装着の様子

ヘモグロビンの濃度変化の分析機能としてマッピング解析の機能が島津製作所によって提供されている。これは、得られた計測結果を基に作成され、脳の活性状態を視覚的に捉えることができるものである。マッピング図はチャンネル位置に対応しており、これを利用することで、脳活性の時間変化を容易に観察することができる。また、時間軸に沿って積算し、タスクごとのマッピング図を作成することもでき、各タスク間の比較も可能となっている。

3.2 特徴抽出

この節では、理解度推定のための特徴抽出について説明する。fNIRS で得られる酸化ヘモグロビンの濃度変化は個人差が大きく、一様に他者と比較することができない。そこで、「どのくらい活性が得られたのか」、ではなく「どこに活性が見られたのか」に注目する。文書を理解するときに、脳のどの部位が活性化するかということはわかってはいない。本手法では、「理解度の違いは活性部位の違いに表れる」という仮定のもと、特徴を抽出する。

脳が活性化するという事は、脳内の血流量の変化が大きいということである。そこで、酸化ヘモグロビンの濃度変化の分散を調べることで、活性化している部位を特定する。更に、酸化ヘモグロビンの濃度変化の平均が正のものだけに限定す

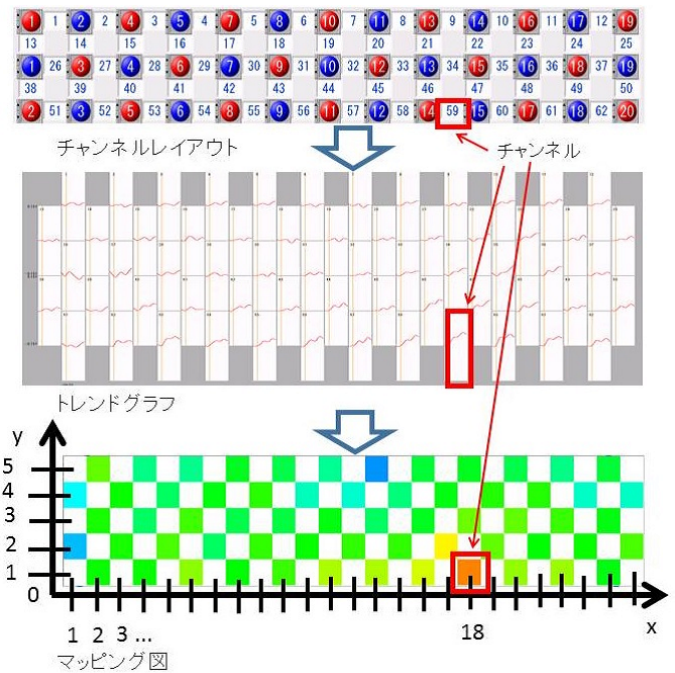


図 4 fNIRS から得られるデータとその分析の概要

ることで、血流量の上昇が見られたものだけを特定することができる。よって分散値、平均値を求めることで、脳の活性が激しい脳の部位を特定しその座標を特徴量とする。以下に具体的な手順を説明する。まず、分散値が大きく、平均値が正であるチャンネルを求めるために、文書を読んでいる間の酸化ヘモグロビンの濃度変化に対し、チャンネルごとに分散値と平均を求める。そして、チャンネルごとの平均値が正であり、かつ 1 番目と 2 番目に分散値の大きいチャンネルを特徴として利用する。次に、得られたチャンネルの位置を 2 次元の座標で表す。座標は図 4 のように設定する。この座標軸を用いると、図 4 のマッピング図中のオレンジのチャンネル 59 は座標 (18,1) となる。このような座標軸を用いて、1 番目に分散の大きなチャンネルの x 座標 (特徴量 1) と y 座標 (特徴量 2)、2 番目に分散の大きなチャンネルの x 座標 (特徴量 3) を求め特徴量とする。以上の手順で、3 次元の特徴ベクトルを得る。

3.3 理解度推定

得られた特徴ベクトルに対し、機械学習を用いることで理解度を推定する。機械学習には SVM(Support Vector Machine) と SVR(Support Vector Regression) を用いる。カーネルは線形カーネルを使用する。

次に、理解度の定義について説明する。本研究で着目する理解度とは文書に対する理解度であり、「文書の内容をどのくらい理解したか」ということである。しかし、理解度には絶対的な指標はなく、定義することが難しい。そこで、本研究では文書を読んだ後に文書に対するテストをし、その正答数により理解度を定義する。

4. 実験

本章では、理解度推定の性能評価のための実験について述べる。

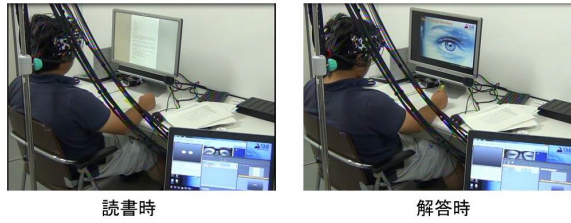


図5 実験風景

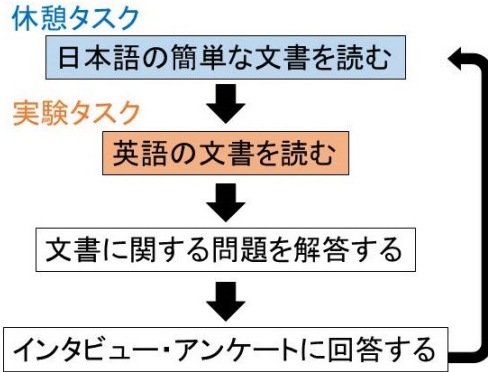


図6 実験の手順

4名の被験者を用いて、データセットを作成し、Leave one outによって評価する。

4.1 実験条件

図5に実験風景を示す。被験者は、全員大学院生である。実験では、fNIRSを用いて脳内の酸素化ヘモグロビンの変化量と、アイトラッカを用いて眼球運動を測定した。前者は島津製作所のLABNIRS [9]を使用し、後者にはSMI社の非装着型アイトラッカ iViewXTM RED [10]を使用した。ファイバホルダーに関しては、言語に関する機能は脳の前方にあるため、今回は頭部の前方を覆うものを使用した。

実験では、被験者にタスクを与え、その際の酸素化ヘモグロビンの変化量と眼球運動を記録した。図6に実験の手順を示す。タスクは、英語の文書を読み、文書に関する問題を解くというものであり、データは文書を読む間のみ記録する。また被験者に、椅子に座った状態で頭部をできるだけ動かさないようにという指示をした。

次に、タスクの内容について以下に示す。まず、休憩タスクとして、被験者に目の前のモニターに表示された簡単な日本語の文書を読むように指示した。そして、文書を読み終えたとき被験者は実験者に合図をした。次に、実験タスクとして、被験者はモニターに表示された英語の文書を読み、読み終わると同様に合図をする。それ以降、読んでいた文書は被験者に提示されない。同時に問題文が実験者によって提示され、被験者は提示された問題を解答した。被験者が問題に対し解答を終えたとき、被験者は実験者に合図をし、問題は回収される。そして被験者はタスクや文書に関するインタビューに回答し、アンケートを記入した。

表1 実験結果-SVM

被験者		文書 1	文書 2	文書 3	文書 4	文書 5	平均誤差
A	推定値	2	2	3	3	2	0.0
	正解	2	2	3	3	2	
	誤差	0.0	0.0	0.0	0.0	0.0	
B	推定値	1	1	2	1	2	0.8
	正解	3	2	2	1	1	
	誤差	2.0	1.0	0.0	0.0	1.0	
C	推定値	1	2	3	2	1	0.0
	正解	1	2	3	2	1	
	誤差	0.0	0.0	0.0	0.0	0.0	
D	推定値	2	2	1	1	3	0.4
	正解	2	2	1	1	1	
	誤差	0.0	0.0	0.0	0.0	2.0	
	平均誤差	0.5	0.25	0	0	0.75	0.3

被験者は実験に慣れるために、実験前にサンプル問題を1問解いた。実験に用いた文書数は、5文書であり、TOEICのPART7の練習問題を使用した。文書の順番における影響を考慮し、被験者ごとに文書の読む順番は変更した。問題数は1文書につき4問であり、それぞれ4つの選択肢が与えられている。今回の実験では、理解度は問題の正答数を用いて3段階で表す。問題は各文書につき4問設定されているため最低で0問正解、最高で4問正解となる。サンプル数を揃えるために、正答数が0~2問のものを理解度1、3問のものを理解度2、4問のものを理解度3と定義する。

被験者の主観評価を計測するために2種類の方法を用いた。1つ目は、主観的に作業負荷を評価するNASA-TLXを用いたアンケートである。NASA-TLXについては後で詳しく述べる。2つ目に、文書問題に関するインタビューである。インタビューでは文書の理解度、文書の難易度、問題の難易度、集中度について、実験者がそれらを5段階で評価した。さらに、全ての文書を読み終わってから、文書中のわからない単語、文そして文書中で一番難しかった段落に印をつけてもらった。

NASA-TLXはタスクに対する負荷仕事量を示すための手法である[11]。6つの評価尺度を持ち、それぞれの各評価尺度に対して主観的評価をすることにより、タスク全体の負荷仕事量を数値化する。評価尺度はMENTAL DEMAND(精神的な要求)、PHYSICAL DEMAND(身体的な要求)、TEMPORAL DEMAND(時間的な要求)、EFFORT(努力)、PERFORMANCE(効率)、FRUSTRATION(不満)である。被験者は1文書を読み終えるごとにNASA-TLXに回答する。

4.2 実験結果

fNIRSの計測結果は酸素化ヘモグロビンの絶対値ではなく濃度変化である。そのため、測定をする際には、被験者が落ち着いている状態との差を見ることで、脳の活性状態を観察する。測定の際には被験者が落ち着いている状態(レスト)と実験タスクをする状態(タスク)を繰り返す必要がある。この際、レストの際に被験者が落ち着いた状態でなければ、正しい計測結果が得られない。そこで、得られたデータにノイズが含まれてい

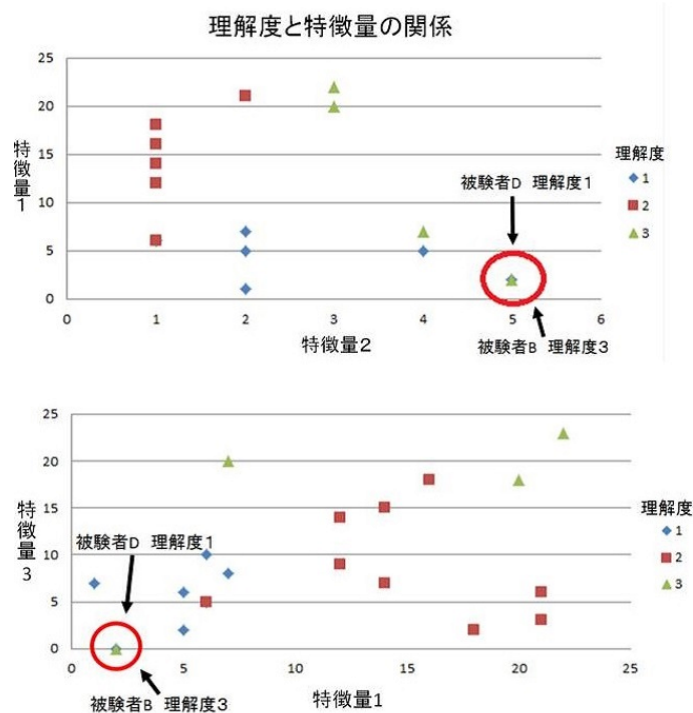


図 7 特徴量と理解度の関係

表 2 実験結果-SVR

被験者		文書 1	文書 2	文書 3	文書 4	文書 5	平均誤差
A	推定値	1.53	2.37	2.76	1.99	1.58	0.2
	正解	2	2	3	2	2	
	誤差	0.0	0.0	0.0	1.0	0.0	
B	推定値	0.90	1.02	1.78	1.35	1.15	0.6
	正解	1	1	2	1	1	
	誤差	2.0	1.0	0.0	0.0	0.0	
C	推定値	1.43	1.55	3.27	2.0	0.97	0.0
	正解	1	2	3	2	1	
	誤差	0.0	0.0	0.0	0.0	0.0	
D	推定値	2.27	2.40	1.25	1.23	1.54	0.2
	正解	2	2	1	1	1	
	誤差	0.0	0.0	0.0	0.0	1.0	
	平均誤差	0.5	0.25	0	0.25	0.25	0.25

ないかを確認するために、タスクがレストに比べて有意であることを統計的に検定した。fNIRS では、チャンネルごとに酸素化ヘモグロビンの濃度変化の時系列データが得られる。この時系列データを休憩タスク時と実験タスク時に分け、レスト点群とタスク点群とする。チャンネルごとにレスト点群とタスク点群から無作為に数点抽出し、t 検定をした。t 検定の結果より、タスク点群はレスト点群に比べて有意差が得られた (有意水準 5%)。この結果より、データには大幅なノイズが含まれていないという仮定のもと考察を続ける。

文書理解度推定の結果を表 1 と表 2 に示す。表 1 は SVM によって推定された理解度と正解、推定結果と正解との誤差を示す。識別率は 80 %、平均誤差は 0.3 という結果であった。被験者別の結果では最も誤差が小さいもので誤差 0.0、最も誤差の大きい被験者で誤差が 0.8 であった。また、表 2 に SVR によって推定された理解度と正解、推定結果と正解との誤差を示す。SVR では推定値は小数で表されているため、SVR で得られた結果を四捨五入し、整数としたものを推定結果として誤差を算出した。その結果識別率は 80 %、平均誤差は 0.25 となった。

最も誤差が大きくなった被験者 B について考察する。被験者 B の文書 2、文書 4、文書 5 の特徴量は全て同一のものが得られた。しかし、文書 2 は理解度 2、文書 4 と文書 5 は理解度 1 となっており、異なる理解度となっている。このことが被験者 B の結果が悪くなった原因であると考えられる。

次に、異なる理解度で同一の特徴量が得られたケースについて考察する。図 7 に示すように、異なる理解度で同一の特徴量が被験者 B の文書 1 と被験者 D の文書 5 で見られた。この原因について考察する。これらのマッピング図を図 8 に示す。被験者 B の文書 1 と被験者 D の文書 5 のマッピング図は非常に似かよっており、このように全体が黒と青と緑で占められるようなマッピング図は全被験者全文書中でこの 2 つだけに見られた。このようなマッピング図は被験者がリラックスした状態を示している。表 3 に被験者 B と被験者 D の NASA-TLX とインタビューの一部を示す。両者とも精神的負担、不満ともに低くなっており、文書に対しては簡単であったと回答している。このことが図 8 のようなマッピング図が得られた原因であると考えられる。今回の手法は、この 2 つのケースのようにほぼ全てのチャンネルで酸素化ヘモグロビンの濃度変化の平均が負と

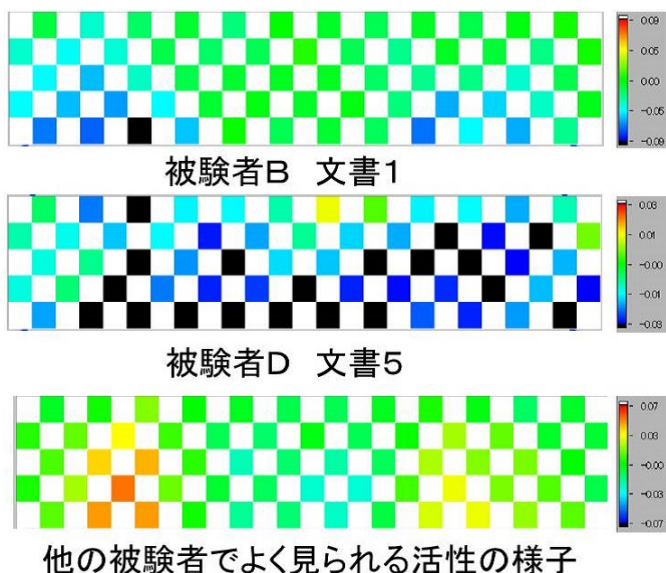


図8 被験者Bと被験者Dの脳の活性の様子

なるような場合には適さない。その理由について以下に考察する。本手法は「読書時には休憩時と比較して脳の活性が見られる」ことを前提に設計されており、その上で脳のどの部位が最も活性化しているかを特徴量にしている。このケースのように「休憩時と比較して読書時に脳の活性が見られない」場合には、本手法は対応していない。このことが異なる理解度で同一の特徴量が得られた原因であると考えられる。

次に、本手法の精度を検証する。今回の実験の理解度の平均は2であった。そこで、この平均値を用いて理解度推定を行った場合と比較して、本手法の精度について考察する。全ての理解度推定の結果が2の場合、識別率45%、平均誤差は0.55となる。この結果と比較すると、本手法の方が識別率も高く、平均誤差も小さい。よって、今回のケースにおいてfNIRSから得られた情報は理解度推定に有効であることが示された。

5. まとめ

本稿では、fNIRSで取得した脳内の酸素化ヘモグロビンの濃度変化の情報を用いて、理解度を推定する手法を提案した。性能評価によって、fNIRSから得られた情報は理解度推定に有効であることが示された。この結果により、脳の活性と理解度の関連が明らかにされた。

今後の課題としては、脳の活性と眼球運動の関連性を明らかにすることで、眼球運動と理解度の関連性を調査する。今回の実験では、fNIRSによる酸素化ヘモグロビンの濃度変化の計測と同時にアイトラッカによる眼球運動も計測している。そこで、その実験データを用いて、脳の活性と眼球運動の関連を明らかにすることを目指す。

謝辞 本研究の一部は、JST CREST および日本学術振興会科学研究費補助金基盤研究(A)(25240028)の補助による。

文 献

- [1] 川市仁史, K. Kunze, 黄瀬浩一, “文書画像検索を用いたリーディングライフログの提案,” PRMU2012-223 495, 電子情報通信学会技術研究報告, March 2013.

表3 被験者Bと被験者DのNASA-TLXとインタビューの一部

被験者 B	文書 1	文書 2	文書 3	文書 4	文書 5	平均
精神的負担	45	60	55	70	35	53
不満	40	70	55	60	40	53
難易度	1	2	3	4	2	2.4
被験者 D	文書 1	文書 2	文書 3	文書 4	文書 5	平均
精神的負担	20	50	45	60	15	38
不満	15	50	50	25	20	32
難易度	2	4	2	5	2	3

精神的負担：0-小さい ... 100-大きい

不満：0-小さい ... 100-大きい

難易度：1-簡単 ... 5-難しい

- [2] 吉村和代, 川市仁史, K. Kunze, 黄瀬浩一, “アイトラッカで取得した視点情報と文書理解度の関係,” PRMU2012-224 495, 電子情報通信学会技術研究報告, March 2013.
- [3] <http://www.med.shimadzu.co.jp/products/om/> 2011.
- [4] A. Bulling, J.A. Ward, H. Gellersen, and G. Troster, “Eye movement analysis for activity recognition using electrooculography,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol.33, no.4, pp.741–753, 2011.
- [5] K. Kunze, H. Kawaichi, K. Yoshimura, and K. Kise, “The wordometer—estimating the number of words read using document image retrieval and mobile eye tracking,” Document Analysis and Recognition (ICDAR), 2013 12th International Conference on IEEE, pp.25–29 2013.
- [6] P.T. Terenzini, L. Springer, E.T. Pascarella, and A. Nora, “Influences affecting the development of students’ critical thinking skills,” Research in higher education, vol.36, no.1, pp.23–39, 1995.
- [7] 坂井建男, 久光正, “ぜんぶわかる脳の事典,” 2011.
- [8] 大石晴美, 脳科学からの第二言語習得論: 英語学習と教授法開発, 昭和堂, 2006.
- [9] <http://www.an.shimadzu.co.jp/bio/nirs/nirs.htm> 2011.
- [10] <http://www.smivision.com/> 2013.
- [11] S.G. Hart and L.E. Staveland, “Development of nasa-tlx (task load index): Results of empirical and theoretical research,” Advances in psychology, vol.52, pp.139–183, 1988.