

修士論文

コンピュータ使用時における
ユーザの前傾姿勢の検知手法の予備的検討

37-206490 高島 諒

指導教員：矢谷 浩司 准教授

東京大学大学院
工学系研究科 電気系工学専攻

令和4年1月27日提出

Acknowledgements

本研究を進めるにあたり、厚いご指導をいただきました指導教員の矢谷浩司准教授に心より感謝いたします。自分はテーマ決めが難航し、最終的には研究室に入った時とは全く方向性の異なるテーマでの研究となりました。そのテーマ決め過程でも自分の本当に興味があることを見つけることが大事だと議論しながら設定まで辛抱強く待ってくださったおかげで意欲的に研究を続けることができました。また、毎週 1on1 ミーティングで時間を作ってください、進み方に迷った時や間違った方向に舵を切りそうになったときに適切に指導していただき、リズムを保ちながら自分の研究を適切に進めることができました。それだけでなく、自身が精神的に不安定になった時も献身的にそばにいてくださり温かい言葉をかけていただき心の支えとなりました。修士課程における 2 年間、ご指導いただき誠にありがとうございました。

また、所属する矢谷研究室のメンバーにも大変お世話になりました。同期である平井雄太君、佐野翔子さん、林裕嵩君には研究の方向性に関する議論はもちろん、プライベートに関しても様々な場面で協力してもらいました。全員が社会的でお互いのテーマを相談し合ったり切磋琢磨することのできる良き仲間だったと感じています。特に平井雄太君には、同じ工学系研究科として授業の選択や輪講等のスケジュールを互いに確認したりするだけでなく、研究の相談に乗ってもらい、忙しい中でコード等の助言やバグに関する相談等をしていただき大変感謝しております。

また、先輩方にも大変お世話になりました。コロナ禍で研究室に直接行けない中でも先輩方が初期に研究室のルールや年間の大まかな流れ等を積極的にかつ丁寧に教えていただいたことで順調に修士生活を始めることができました。また、大きな発表や論文提出の前には資料に目を通していただき、適切なアドバイスをいただきました。特に Arissa Janejera Sato さんは忙しい中時に自分で時間を作ってまで abstract, スライドの見直し、研究内容等のアドバイスをいただきました。

自分が研究に詰まったと感じ、Arissa さんに時間を作っていただいて Discussion したことでそこから前に進むことができたことは忘れません。本当にありがとうございました。また、Zhongyi Zhou さんには研究する分野に共通するところもあり、使うツールの剪定や研究の方向性について助言をいただきました。

後輩のみなさんにも発表後のフィードバックや論文の添削など様々な場面でお世話になりました。被験者募集にも快く参加していただいたこと感謝しています。

秘書の元岡みさ子さんには学会申し込み、必要機材購入等各種手続きの際に常に迅速な対応をしていただき、大変お世話になりました。

最後に大学生活を金銭的にも精神的にも支えてくださった家族に感謝して謝辞の結びとしたいと思います。本当にありがとうございました。

Abstract

With the increase of remote activities to promote safe social distancing, there has been an increase in the number of users of video display terminals(VDTs) such as computer screens. Prolonged exposure to VDTs can lead to VDT Syndrome, consisting of eye dryness, neck pain, shoulder pain, headaches, etc. To prevent this, we identified and evaluated an approach to recognize posture while using a laptop PC. In this paper, we explained the background and details of VDT syndrome, presented our literature survey about VDTs. We then detailed our approach for designing an experimental setup for posture detection, and shared findings from the preliminary study to more precisely detect posture using a depth camera. Then we used face angles to detect proper posture and constructed models to presume face angles from 3d coordinates of 68 face feature points. We compared and analyzed these models. Finally, we concluded with the system overview using this model to detect improper posture and discussed future directions of our research.

Abstract

近年スマートフォンやインターネットの普及，コロナウイルスによるテレワークの増加に起因して，IT 機器に触れる人口や 1 人当たりの IT 機器使用時間は益々増加している．IT 機器の中でも特にコンピューターのディスプレイ等の画像表示端末を Visual Display Terminals (VDT) と呼ぶが，VDT 使用の際には，それらの画面を注視し続ける，同じ姿勢を取りつづける，猫背など前屈みの姿勢を維持するなどの使用中の行動が人体の健康に及ぼす悪影響が懸念されている．以上のような症状は総称されて VDT 症候群と言われている．この VDT 症候群を検知し，予防を支援する研究は広く行われてきたが，画角が限定されまた平面的な情報で検知が難しいという課題があった．そこで本論文では深度カメラを用いてユーザの顔の特徴点を抽出し顔の角度を推定するモデルを構築することで上半身の前傾姿勢を検知する方法を提案し，その結果及び分析について述べ，得られた知見をもとに議論する．

Table of contents

List of figures	viii
List of tables	ix
1 はじめに	1
1.1 背景	1
1.2 貢献	3
2 関連研究	4
2.1 VDT 症候群	4
2.1.1 VDT 症候群の具体的な症状	4
2.1.2 望ましい姿勢	4
2.1.3 VDT 作業時における休憩支援アプリケーションの開発	5
2.2 カメラを用いた VDT 症候群検知またはその予防に関する先行研究	6
2.2.1 Web カメラ内蔵型ノート PC 等の携帯端末における VDT 症候群予防のための瞬きを促すアプリケーションの提案	6
2.2.2 Web カメラを用いた PC ユーザー作業時間の自動測定	6
2.2.3 内蔵 Web カメラを用いた VDT 作業者の姿勢モニタリング	7
2.2.4 VDT 作業における姿勢意識向上アプリケーションの開発	7
2.2.5 VDT 機器を使用した学習環境における疲労検知のための自己接触動作認識	7
2.2.6 ロボットを活用した VDT 症候群予防における効果	8
2.3 まとめ	8
3 ウェブカメラを用いた予備実験	9
3.1 実験手法	9

3.2	実験のセットアップ	10
3.3	結果	11
3.4	考察	13
4	深度カメラを用いた顔の特徴点抽出	16
4.1	実験手法	16
4.2	結果	17
4.3	考察	17
5	深度カメラを用いた顔の角度検出による姿勢推定手法の検討	20
5.1	はじめに	20
5.2	実験手法及び取得したデータ	20
5.3	本実験のセットアップ	21
5.4	顔の角度の推定方法	21
5.5	特徴量抽出	22
5.6	変数増減法による重回帰分析	22
5.7	説明変数の与え方	23
5.7.1	グループ A、B を用いた分析	23
5.7.2	グループ C、D を用いた分析	26
5.7.3	全ての特徴量を用いた分析	29
6	考察	32
6.1	各分析結果に対する考察	32
6.2	出現頻度の分析	33
6.3	用いるモデルの決定	35
7	まとめと今後の展望	36
7.1	今後のこのモデルを活かした姿勢推定	36
7.1.1	適切な姿勢における顔の角度の範囲を設定	37
7.2	このモデルを用いた姿勢推定システムの利点	37
7.3	モデルとモデルを用いた姿勢推定システムの問題点及び改善案	38
7.4	フィードバック方法の提案	38
8	おわりに	40
	Publications	41

List of figures

2.1	望ましい姿勢の例.	5
3.1	実験における実験参加者, 画面及びウェブカメラの配置	10
3.2	顔と画面の距離が 50cm の場合の各角度におけるウェブカメラから取得した画像.	11
3.3	出力結果画像の例.	12
4.1	顔の特徴点の番号の対応.	17
4.2	画面との距離が 20cm における顔の特徴点の分布.	18
4.3	画面との距離が 30cm における顔の特徴点の分布.	18
4.4	画面との距離が 40cm における顔の特徴点の分布.	18
4.5	画面との距離が 50cm における顔の特徴点の分布.	18
5.1	顔の特徴点の番号と AB のグループ分け.	24
5.2	顔の特徴点の番号と CD のグループ分け.	27

List of tables

3.1	取得できる 18 の関節名.	11
3.2	0° における各関節識別の信頼度.	12
3.3	30° における各関節識別の信頼度.	13
3.4	60° における各関節識別の信頼度.	13
3.5	90° における各関節識別の信頼度.	14
3.6	各角度における信頼度の平均値.	15
5.1	特徴量抽出により得られた 54 の点の組み合わせ. 各組み合わせの 2 点を結 んだ線分に対して計算される仰角を説明変数とする.	24
5.2	重回帰分析結果.	26
5.3	特徴量抽出により得られた 54 の点の組み合わせ. 各組み合わせの 2 点を結 んだ線分に対して計算される仰角を説明変数とする.	27
5.4	重回帰分析結果.	28
5.5	特徴量抽出により得られた 62 の点の組み合わせ. 各組み合わせの 2 点を結 んだ線分に対して計算される仰角を説明変数とする.	29
5.6	重回帰分析結果.	30
6.1	各モデルの AIC 値及び決定係数.	32
6.2	全てのまとめ.	33

Chapter 1

はじめに

1.1 背景

近年スマートフォンやインターネットの普及，コロナウイルスによるテレワークの増加に起因して，IT 機器に触れる人口や 1 人当たりの IT 機器使用時間は益々増加している [24]. IT 機器の中でも特にコンピューターのディスプレイ等の画像表示端末を Visual Display Terminals (VDT) と呼ぶが，そのような VDT を日常的に使用する VDT 労働者を対象とした厚生労働省の調査¹ [13–15] でもその増加傾向は顕著に表されている [19].

VDT 使用の際には，それらの画面を注視し続ける，同じ姿勢を取りつづける，猫背など前屈みの姿勢を維持するなどの使用中の行動が人体の健康に及ぼす悪影響が懸念されている [22]. 同調査においては VDT の使用時間に応じた症状の有無も調査され，6 時間以上の作業を行う場合に，首や肩のこり・痛みを発症する割合は 80% を超えることが示される等特に長時間の作業においては高い割合で症状を発症してしまうことが調査された. 以上のような症状は総称されて VDT 症候群（別名，テクノストレス眼症）と言われている [25]. VDT 症候群は現代病に含まれ社会的にもその発見及び解決が重要な大きな問題となっているため，その検知や対策が急務である [9, 11, 12]. VDT 使用時には実際に眼の痛み，ドライアイ，疲労感，手首や腕の疲れ，首や肩の痛みなどに繋がることがわかっている [7, 21, 28]. このように VDT 症候群は大きく分けて精神，眼機能，筋骨格系の三種類に悪影響を及ぼす.

このような VDT 症候群の症状の中でも眼機能の症状であるドライアイを検知し予防の支援を行う研究は広く行われてきた [8, 10, 27]. しかし，筋骨格系の症状の検知及び予防の研究は，数が少なく課題を残す. その理由としては適切でない姿勢には様々な条件があ

¹平成 20 年まで実施

り、また定量的指標も少なく個人差や環境の影響が大きいことがある。さらに、筋骨格系の症状の検知は眼のみではなく時に全身の情報が必要であり、通常のコンピュータ使用時にコンピュータがユーザの姿勢全体に対して収集できるデータは限定的であることとそこから的確に姿勢の歪みを推定できるとは限らないという問題点がある。

そこで本研究は、VDT 症候群につながる適切でない姿勢の状態を検知しその姿勢を適切なものにするという観点で VDT 症候群の予防を試みて、この実現を目指す。

その実現のため、3章では、まず一般的なコンピュータを使用する状況において収集できるデータの限界や問題点について調べた。この特別なセンサー等は使えない状況において、姿勢の状態をコンピュータが得るためにはウェブカメラを用いることが考えられる。多くの使用者がいるウェブカメラから得られた画像を用いて姿勢推定を行うため、その画像に対して姿勢推定のための骨格推定を試みる予備実験を行った。ここで画像を用いたのは、画像で行うことのできる処理を動画のフレーム単位で行うことで動画にも応用できるからである。実験では、顔と画面の距離が適切な姿勢かどうかの判定の指標の一つであることから、コンピュータの画面と顔の距離を 10 cm 単位で遠ざけて行った時の各骨格推定の信頼度を見た。さらに、この実験においてはウェブカメラの位置を変更することで取得できるデータにどのような変化が生じるかも調べた。その結果、多くのユーザがウェブカメラを置く位置である正面に置いた時顔が 30 cm 以内の場合に骨格が全く推定できないことや首及び肩の信頼度が低くなる等の結果を得られた。2.1.2の節でも示している通り、顔と画面の距離が近い（40 cm 未満の）時はその体勢が VDT 症候群につながる適切でない姿勢となっているため、この状態を検知できないのは大きな問題である。

そこで3章の結果を踏まえた上で4章では、顔と画面との距離が近い場合にも顔を認識できず、さらには肩や首の認識も困難であるという問題を解決するために、深度の情報を得ることのできる深度カメラを用いて顔の特徴点の座標を取得することができるか調査する予備実験を行なった。深度センサーの分野は、2019 年発売の Google 社の Pixel 4²のフロントカメラに uDepth と言われるデプスセンサーが搭載される、2020 年発売の Apple 社の第四世代の iPad Pro³に光を用いて奥行きの情報を得る Light Detection and Ranging (LiDAR) センサが搭載される、そして自動運転技術への応用等近年注目を浴びている。深度カメラを用いた結果ウェブカメラで困難であった顔と画面の距離が 20 cm, 30 cm, 40 cm の時の特徴点の 3 次元座標取得に成功した。

5章では、4章で取得できた 3 次元座標を適切でない姿勢の判別方法として用いるために顔の角度に注目し、顔の 68 点の特徴点 3 次元座標を組み合わせた仰角を説明変数、実際の顔の角度を目的変数として特徴量抽出を行いモデルを作成した。

²<https://store.google.com/jp/category/phones?hl=ja>

³<https://www.apple.com/jp/ipad-pro/>

6章でその分析を行い結果をまとめ、7章にて、本研究を総括した上で今後取り組むべき課題、フィードバック方法や実際のシステム構築に向けた展望についてまとめる。

1.2 貢献

本研究の最終的な目的は、VDT 症候群につながる適切でない姿勢をユーザがしている場合にその状態を的確に検知し、適切な姿勢に是正するためのフィードバックを行うシステムを構築することである。本研究の貢献は以下の通りである。

- ウェブカメラを用いた姿勢推定のための関節推定の信頼度と画角的限界の評価
- 深度カメラを用いた顔の特徴点の3次元座標の取得及びその座標から計算される仰角を用いた実際の顔の角度を推定するモデルの構築

Chapter 2

関連研究

2.1 VDT 症候群

この章では VDT 症候群に関連した参考文献を参照しながら VDT 症候群について整理し、その具体的な症状や予防となる理想的な姿勢について述べる。

2.1.1 VDT 症候群の具体的な症状

VDT は長期使用することによって目の痛み及び疲労、首や肩のこり及び痛み、手や腕の痛み及び疲労、腰のこり及び疲労、めまいなどを引き起こす。これらの症状は、

1. 視機能に関するもの：眼疲労，ドライアイなど
2. 筋骨格系に関するもの：猫背，腰痛，肩こり
3. 精神・心理的なもの：疲労感，うつなど

の 3 つに大別される。

中でも [19] の中でも表されているように目の疲れなどの1. 視機能に関するもの及び2の首や肩こりなどの筋骨格系に関するものの症状が主要的でありそれらを対象とした研究も多く行われている。

2.1.2 望ましい姿勢

VDT 作業における労働衛生管理のためのガイドライン [16] においては VDT 症候群を予防する観点で望ましい姿勢も以下のように明示されている。

1. 画面の上端が目線と同じか、やや下の高さであること。



図 2.1: 望ましい姿勢の例.

2. 画面から 40 cm 以上離すこと.
3. 適当な背もたれを有していること, (傾きを調整できるのが望ましい.)
4. 椅子が安定して容易に移動できること.
5. 足裏全体が床に接していること.

の 5 点が主な注意点である.

望ましい姿勢の例を図2.1に示す.

2.1.3 VDT 作業時における休憩支援アプリケーションの開発

厚生労働省が発表した VDT 作業における労働衛生管理のためのガイドライン [16] においては, 作業中「1 時間に 1~2 回ほどの小休止を設けること」及び「1 時間ごとに 10 分から 15 分の休憩を取る」ことが定められている. しかしこれは, VDT 作業内容が多様化していることや個々人の疲れ具合などから個別性が高い. そのため, 作業内容と疲労度に応じた休憩を取ることが理想と考えられる. そこで平間らは, 精神疲労の主観評価を使用し明確な数値ではなく境界がぼんやりとはっきりしない場合に用いられるファジィ推論を用いて適切な休憩間隔を導出した [1, 6].

このように VDT 症候群の予防には姿勢の観点だけでなく、休憩時間や連続作業時間を用いたアプローチも存在するが、本研究では姿勢を用いたアプローチに焦点を当てる。

2.2 カメラを用いた VDT 症候群検知またはその予防に関する先行研究

本節では2.1節を踏まえた上で、VDT 症候群の検知及び予防にカメラ等から取得した画像及び映像を用いている先行研究の紹介を行う。

2.2.1 Web カメラ内蔵型ノート PC 等の携帯端末における VDT 症候群予防のための瞬きを促すアプリケーションの提案

鶴岡らは内蔵カメラの映像から瞬きを検出することで VDT 症候群を予防するアプリケーションを提案した [29]。具体的な手法としては VDT 作業者の顔の連続映像データを取得し、OpenCV¹による画像処理を行うことでデータ間のフレーム間差分を計算し、瞬きを検出した。システムでは設定時間内の瞬き回数が閾値を下回った時警告音やポップアップウィンドウにより作業者に注意を促すものである。ドライアイの検査方法として、10 秒間目を開け続けられるかをチェックし、開け続けられない場合はドライアイの疑いがあるとする方法を提案した。

このようなドライアイの有無によって VDT 症候群を検知する研究は多数存在する。

2.2.2 Web カメラを用いた PC ユーザー作業時間の自動測定

宮崎らは厚生労働省が発表した VDT 作業における労働衛生管理のためのガイドライン [16] において VDT 作業時間は 1 時間で区切ることが推奨されていることに注目し、連続作業時間が 1 時間を超えるところで警告を出すシステムを想定し、ウェブカメラからの画像から作業者がいるかどうかを認識する方法を提案した [18]。宮崎らは作業者の存在は画像に対する色の変化と考え、10 人の実験参加者のウェブカメラ画像の各ピクセルごとの RGB 値に対して白と赤の 2 値化の処理を行い、赤領域の最大独立面積を変数としてサポートベクターマシンによる判別を行った。結果作業者がいる場合の適合率は 95% 以上であったものの、いない場合の適合率は 45.1% から 100% とばらつきが見られた。作業者がうつむく、後ろを向く、席をたつ、横にはみ出る、後ろに椅子をひき距離が離れる等の想定しない状況で間違った判別が見られた。この先行研究はウェブカメラ画像内での作業者の有無を画像の差分から求めているため、背景や場所等の環境による影響を大きく受け

¹<https://opencv.org/intel/>

る問題点がある。また、対象者が動く以上作業者のさまざまな状況を想定する必要性が示唆された。

2.2.3 内蔵 Web カメラを用いた VDT 作業者の姿勢モニタリング

守屋らはテンプレートマッチング法を用いて VDT 作業者の姿勢を改善するシステムを実装した [17]。このシステムではテンプレートマッチング法が用いられており、正しい姿勢の画像を基準として取得し、リアルタイムの画像において理想の姿勢と現在の姿勢に対して相関演算を行うことにより相関値を画像から計算して設定した閾値を下回った場合警告を出す。正しい姿勢は 2.1.2 も考慮され顔と画面との距離が 50 cm の時とされた。また、ここでの閾値の決定の際に実験参加者の主観評価により姿勢の崩れを感じる顔と画面との距離は前傾で 35 cm 以内、後傾で 65 cm～70 cm 以上であった。ここから顔と画面の距離で主観的にも姿勢が崩れると感ずることが示唆された。

2.2.4 VDT 作業における姿勢意識向上アプリケーションの開発

杉本らは Microsoft 社の製品である Kinect²を用いて姿勢の意識を向上させるアプリケーションの開発を行った [20]。Kinect は身体の頭や腰などの 20 点の骨格情報を取得できるスケルトン認識機能を備えていることに加えて深度センサを備えている利点がある。アプリケーションでは頭と肩中央の位置の座標を姿勢判断材料として使用し判断を行う。悪い姿勢を感知すると Kinect で撮影した RGB 画像を示すとともに姿勢不良通知を行った。アプリケーション使用前後で 5 段階主観的評価実験も行い、「作業中の姿勢は改善されたか」という項目では 4.1 のスコアを得ている。この実験では頭と肩中央のジョイント座標を姿勢判断材料として使用したと書かれているが、その具体的な数値等は書かれておらず、姿勢が悪いか良いかを判断するのみであった。

2.2.5 VDT 機器を使用した学習環境における疲労検知のための自己接触動作認識

延らは疲労の蓄積を検知すると自発的かつ有効な休息を促すようなシステムを開発する前段階として、疲労度を測定するための指標の決定とその自動的な認識手法の提案を試みた [23]。まず延らは疲労が蓄積している場合に現れる動作を調べるために髪や顔を触る、腕を組む等の自己接触動作、腕を伸ばす等の伸び動作、座り直す等の脱力動作の 3 種類の動作を定義した上でそれらの出現回数と疲労度との相関を計算した。ここでの疲労度は 25 個の疲労に関する質問に対し 5 段階で答えるアンケートによって計算されたものである。その結果、自己接触動作が最も高い相関を示したため延らはその自己接触動作の自

²<https://developer.microsoft.com/ja-jp/windows/kinect/>

動認識を試みた。この際、初めに Kinect センサによる骨格点推定の利用を試みたが上半身のみのデータであることから認識が不安定となったため画像処理に切り替えた。色相ヒストグラムの類似度とオプティカルフローを特徴量として用いて手領域の設定を行い、MeanShift [5] と CAMShift [2] の二つを組み合わせることで手の追跡を行うことで、全体の追跡検知は 90% 以上の高い精度を示した。

2.2.6 ロボットを活用した VDT 症候群予防における効果

長谷川らはコミュニケーションロボットを用いて一定時間ごとに姿勢改善を促すシステムの開発を行なった [26]。その結果姿勢改善動作の増加が観測されると同時にその介入を作業の妨げと感じた実験参加者はおらずコミュニケーションロボットを用いたフィードバックには煩わしさを抑えるという点で一定の効果があることが示唆された。

2.3 まとめ

以上のように、先行研究ではカメラを用いて姿勢推定を試みるものはあるものの本研究の目指す VDT 症候群につながる適切でない姿勢をユーザがしている場合にその状態を的確に検知し、適切な姿勢に是正するためのフィードバックを行うシステムとしてはどれも不十分であった。

具体的な問題点としては二つある。一つ目は環境にロバストでないものが多いということである。理想的な姿勢とのずれのみで判定するというアプローチは、姿勢以外の環境が変わることで大きな影響を受けてしまう。また、ユーザが変わることでも基準となる理想の姿勢の画像は変わってしまう。二つ目は、姿勢がどのように適切でないかまでの評価は行っていない点である。理想的な姿勢から崩れていることを判定していても具体的にどのように崩れているかについて注目した研究はなかった。

以上二つの問題を解決するため、次章では実際にウェブカメラを用いて姿勢推定の予備実験を行うことで、その制約と限界を見る。

Chapter 3

ウェブカメラを用いた予備実験

VDT 症候群を対策するため、まずは最も一般的なコンピュータを使用する際の姿勢を想定し、その使用時にウェブカメラから得られる情報を基に姿勢検知を行うことを考える。上記のような場合、ウェブカメラはその画角に全身が映らない場所に配置されている場合がほとんどであるため、その限られた画角から理想でない姿勢を示す情報を全て取れるか、取れないとしたらどこまで取れるのかということは重要である。

そこで、そのウェブカメラから得られる情報にはどのようなものがあるのか、その制約や限界はどこにあるのか、そしてウェブカメラの配置を変え画角に入る情報を変えることでその制約や限界は変わるのかについて調べるために予備実験を行った。

3.1 実験手法

椅子に座って VDT 端末を使用する際の姿勢に関してウェブカメラで取得されるデータにおいて、VDT 使用時の姿勢に関連する指標として顔の画面からの距離があるため、この距離を変更することで姿勢推定にどのような影響があるかを調べた。具体的には顔から画面までの距離を 10 cm 刻みで 10 cm, 20 cm, 30 cm, 40 cm, 50 cm, 60 cm, 70 cm の 7 つの長さで変化させ姿勢の推定への影響を見ることとした。

さらに、ウェブカメラの位置を変え画角を変更することで画像から取得できるデータはどのように変わるのかも調べた。これも具体的には、ウェブカメラの位置による影響も見られるためウェブカメラの場所を椅子の中心から画面までの角度を 0° としたときに 0° , 30° , 60° , 90° の 30° 刻みの 4 つの角度に配置してその影響を見た。

実験の概要を図3.1に示す。

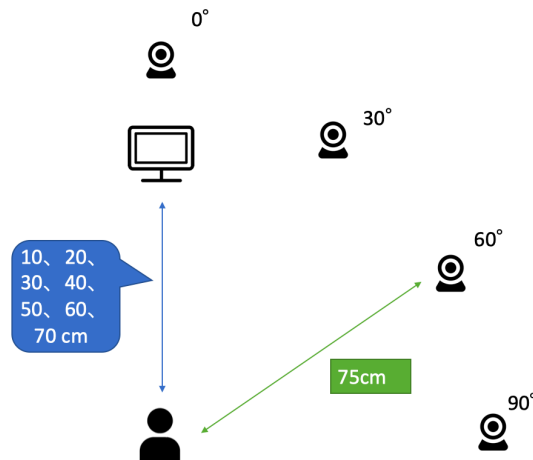


図 3.1: 実験における実験参加者, 画面及びウェブカメラの配置

3.2 実験のセットアップ

実験には ELECOM 社の UCAM-C980FBBK¹を用いた。また、実験参加者は学生 1 名である。

他の条件は揃えるために床を基準として、

- 椅子の高さ : 41 cm
- 画面の上端の高さ : 107 cm
- カメラのレンズの中心の高さ : 107 cm
- 目線の高さ : 117 cm
- 画面から椅子の中心 : 70 cm

と設定した。これらは2.1.2を参考にし、画面から 50 cm 付近で理想的な姿勢となるように設定した数値である。

また、実際にウェブカメラから取得した画像を図3.2に示す。

姿勢推定には人物の骨格を深層学習（ディープラーニング）を用いて推定することができる OpenPose [4, 3] を元にし、Google 社によって開発されたオープンソース機械学習プラットフォーム TensorFlow²上で動作する tf-pose-estimation³を用いた。

¹<https://www.elecom.co.jp/products/UCAM-C980FBBK.html>

²<https://www.tensorflow.org/>

³<https://github.com/iloonet/tf-pose-estimation>



図 3.2: 顔と画面の距離が 50 cm の場合の各角度におけるウェブカメラから取得した画像.

tf-pose-estimation では 18 の関節の (x 座標, y 座標, 信頼度) を求めることができる. 18 の関節はそれぞれ「鼻, 首, 右肩, 右肘, 右手首, 左肩, 左肘, 左手首, 右尻, 右膝, 右足首, 左尻, 左膝, 左足首, 右目, 左目, 右耳, 左耳」である. ただし, 本実験では全身を映すことができるわけではないためこれらの中から画角内にある可能性のある関節である「鼻, 首, 右肩, 左肩, 右目, 左目, 右耳, 左耳」の 8 つの関節を用いる.

その様子を表 3.1 にまとめる.

表 3.1: 取得できる 18 の関節名.

関節名	鼻	首	右目	左目	右耳	左耳
本実験での使用	○	○	○	○	○	○
関節名	右肩	左肩	右肘	左肘	右手首	左手首
本実験での使用	○	○	×	×	×	×
関節名	右尻	左尻	右膝	左膝	右足首	左足首
本実験での使用	×	×	×	×	×	×

3.3 結果

tf-pose-estimation の標準出力結果の一つとして得られる画像の一例を図 3.3 に示す.

左上の画像が結果, 右上が検出対象である各関節の信頼度のヒートマップ, 下側にある二つの画像が検出した関節同士を繋ぐために求められたベクトルマップを示している. こ

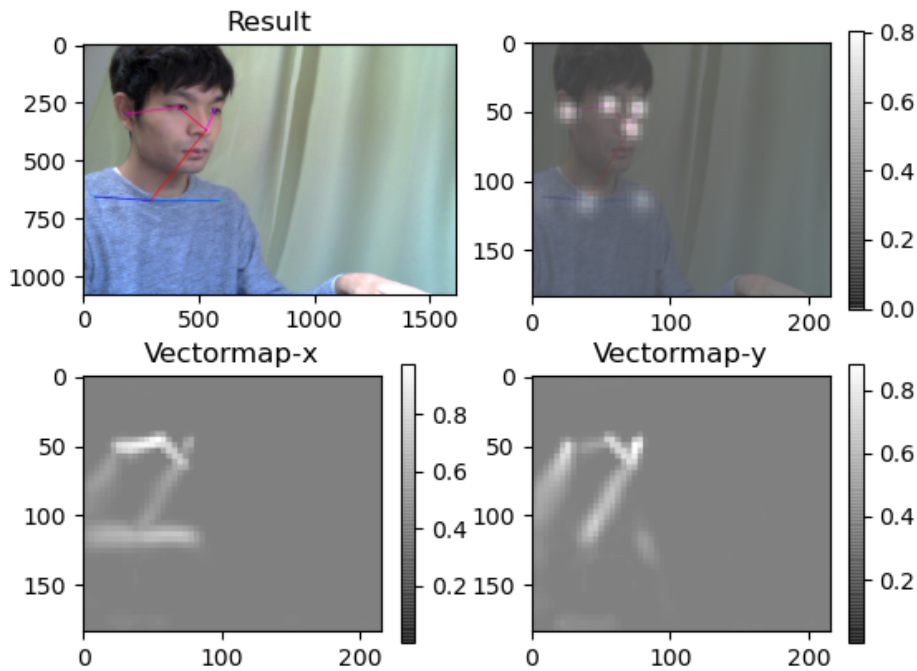


図 3.3: 出力結果画像の例.

の例では「鼻, 首, 右肩, 左肩, 右目, 左目, 右耳」の認識に成功している. また, 実験による認識の結果を信頼度 (0 から 1) の数値を用いて表3.2, 表3.3, 表3.4, 表3.5にそれぞれ示す.

表 3.2: 0° における各関節識別の信頼度.

顔と画面の距離	鼻	首	右目	左目	右耳	左耳	右肩	左肩
10cm	0	0	0	0	0	0	0	0
20cm	0	0	0	0	0	0	0	0
30cm	0	0	0	0	0	0	0	0
40cm	0.69	0	0.79	0.78	0.45	0.33	0	0
50cm	0.63	0	0.82	0.8	0.66	0.48	0	0
60cm	0.7	0.28	0.74	0.8	0.64	0.7	0	0.49
70cm	0.81	0.17	0.82	0.84	0.66	0.57	0	0.29

それぞれの表3.2, 表3.3, 表3.4, 表3.5における「0」表記はその関節の認識に失敗したことを示している.

表3.2に示すように, ウェブカメラの位置が0° (正面) である時, 画面と顔の距離が30 cm 以内においては全ての関節を認識することができていなかった. またこの時, 画面との距離に関わらず右肩を認識することができていなかった.

表 3.3: 30° における各関節識別の信頼度.

顔と画面の距離	鼻	首	右目	左目	右耳	左耳	右肩	左肩
10cm	0.18	0.11	0.8	0.46	0.41	0	0.14	0
20cm	0.2	0	0.78	0.61	0.34	0	0	0
30cm	0.47	0.09	0.85	0.73	0.28	0	0.18	0
40cm	0.57	0	0.85	0.62	0.62	0	0	0
50cm	0.69	0.34	0.81	0.74	0.71	0	0.11	0.3
60cm	0.72	0.28	0.76	0.78	0.65	0	0	0.22
70cm	0.77	0.39	0.72	0.74	0.59	0	0.09	0.6

表 3.4: 60° における各関節識別の信頼度.

顔と画面の距離	鼻	首	右目	左目	右耳	左耳	右肩	左肩
10cm	0.34	0.14	0.25	0.27	0.66	0	0.28	0
20cm	0.33	0.1	0.68	0	0.67	0	0.14	0
30cm	0.74	0.14	0.74	0.24	0.72	0	0	0
40cm	0.68	0	0.77	0.71	0.62	0	0	0
50cm	0.74	0	0.88	0.71	0.7	0	0	0
60cm	0.72	0	0.84	0.6	0.74	0	0	0
70cm	0.75	0.25	0.89	0.59	0.7	0	0.14	0

表3.3に示すように、ウェブカメラの位置が30°である時、0°で認識できなかった画面と顔の距離が30cm以内の時でも関節の認識に成功した。ただし、左耳の認識はされず、首や右肩で認識できない距離も存在した。

表3.4に示すように、ウェブカメラの位置が60°の際には、左耳及び左肩の認識はされず、右肩の認識も30cm~60cmにおいてはできていなかった。

最後に表3.5に示すようにウェブカメラの位置が90°である時、左耳や左肩、左目は数カ所を除き認識されず、首は全ての距離で認識ができていた。

3.4 考察

ウェブカメラの位置は標準的には0°（正面）である。この場合において画面と顔の距離が30cm以内のとき全ての関節を認識することができていなかったことは大きな問題であると考えられる。これは、顔がウェブカメラの画角に対して大きすぎることで輪郭が取れ

表 3.5: 90° における各関節識別の信頼度.

顔と画面の距離	鼻	首	右目	左目	右耳	左耳	右肩	左肩
10cm	0.63	0.1	0.8	0.35	0.74	0	0	0
20cm	0.7	0.11	0.73	0	0.64	0	0	0
30cm	0.72	0.16	0.82	0	0.71	0	0.13	0
40cm	0.67	0.17	0.83	0	0.63	0	0.11	0.08
50cm	0.53	0.15	0.83	0	0.67	0	0.12	0
60cm	0.54	0.12	0.81	0.13	0.68	0	0.1	0
70cm	0.72	0.15	0.83	0.27	0.74	0	0.08	0

ず関節が認識できなかったからであると考えられる。顔と画面の距離が近い場合は画面内に顔が収まっていない場合や見切れている場合も多く、関節として認識できていなかった。特に画面からの距離が 40 cm より小さくなると悪い姿勢の一つの指標となるため認識できないことは改善が必要である。

また、右肩が認識できなかったことに関しては実験参加者の右肩がなで肩であり左右で形が全く異なるためだと考えられる。また、角度を変更し 30°, 60°, 90° にすることで 10 cm から 30 cm の距離で認識が行えるようになっていた。これによりウェブカメラの位置を変えることで認識率を高められる可能性が示唆された。ただし、その信頼度は例えば 30° における 20 cm の鼻の関節の信頼度は 0.2 となるなど、低い傾向は変わらなかった。

90° の角度においては首及び右肩の検出が他三つの角度と比較して最も多く行うことができた。また全ての角度において鼻及び右目の検出は安定していた。画面に対して垂直方向の姿勢の歪みが適切でない姿勢に直結するため、90°, すなわち真横から撮影した画像が最も姿勢判定に用いるには有効である。さらに今回の結果では各関節の認識においてもこの角度が最も安定した結果を出していた。ただし、常に真横にウェブカメラを固定して置くという状況は一般的なく実現が難しい場合があるという問題点がある。しかし、ここで「奥行き」という情報の取得が大いに重要であることがわかった。

それぞれの角度における各関節ごとの信頼度の平均値を整理して表3.6に示す。ただし、ここでは検出できなかった場合の信頼度を 0.0 として計算した。

そのため、3つの距離地点で検出結果が得られなかった角度 0° においては信頼度が他の角度と比較し低くなっている傾向が見られる。一方でウェブカメラの位置の角度が 60°, 90° の際には左目を除いて信頼度の平均に大きな差は見られなかった。30° においては鼻や右耳の検出の信頼度で劣ったものの、首の検出は全ての角度で最も高い数値を出した。全体の傾向として鼻や耳、目の検出の信頼度が高く、首や肩の信頼度が低いことが分かる。

表 3.6: 各角度における信頼度の平均値.

ウェブカメラの位置	鼻	首	右目	左目	右耳	左耳	右肩	左肩
0°	0.4	0.06	0.45	0.46	0.34	0.3	0	0.11
30°	0.51	0.17	0.8	0.67	0.51	0	0.074	0.16
60°	0.61	0.09	0.72	0.45	0.69	0	0.08	0
90°	0.64	0.14	0.81	0.11	0.69	0	0.077	0.01

これはやはりウェブカメラから取得できる画像は上半身の限られた部分であり，ユーザの動きによっては見切れたり，完全に隠れる等の状況が生み出されてしまうことが原因であると考えられる．

以上の分析結果により，ウェブカメラと骨格推定を用いた姿勢検知には大きく二つの問題があることが分かった．一つ目は，ウェブカメラを置く位置に関わらず首や肩の認識は困難であるということである．二つ目は画面との距離が近すぎる場合に顔の関節が全く認識できないということである．特に二つ目の距離が近い場合は VDT 症候群を解決する上では必ず認識に成功しなければいけない．そこでこれら二つの問題点を解決するためのアプローチを次章で示す．

Chapter 4

深度カメラを用いた顔の特徴点抽出

3章で行なった実験では首や肩が認識されないことがあるという問題点やウェブカメラは通常正面に配置されているため立体的な関節の座標の入手が困難であるという問題点があった。そこでまず、想定しているコンピュータ使用時におけるカメラからの画像から姿勢推定する場合において首や肩が映らない場合や見切れる場合がありこれらを推定の材料とすることは安定しない。そこで検出が安定している顔の情報から姿勢推定できないかと考えた。ただし、顔の関節も 40 cm 以内でできなかつたことと、顔と画面との距離がとても重要、すなわちカメラから見た顔の奥行きが大きな意味をもつため、正面からでも深さの情報を得ることができる深度カメラを用いればこの問題は解決できるのではないかと考え、まず認識ができるかどうかを確認するために深度カメラを用いた顔の情報認識を試みた。具体的には顔の 68 点の特徴点の 3 次元座標の取得を試みた。

4.1 実験手法

実験には intel 社の Depth Camera RealSense D435i¹を用いた。RealSense カメラには 2 台の赤外線カメラが搭載されており、またドットパターンを照射することも可能となっている。これらの赤外線カメラの視差と投影したパターンの変化により物体の形状や距離を認識することができる。また、3章の実験より肩の認識が困難であることが示唆されたため、深度というより多くの情報量を持った顔の情報であればそこから姿勢を推定できるのではないかと考え、関節の推定ではなく顔の特徴点の推定を行うこととした。顔の特徴点の推定には機械学習ライブラリ Dlib²を用い、顔の特徴点計 68 点の 3 次元座標を得た。

ここで、特徴点の番号と実際の顔の対応を図4.1に示す。

¹<https://www.intelrealsense.com/depth-camera-d435/>

²<http://dlib.net/>

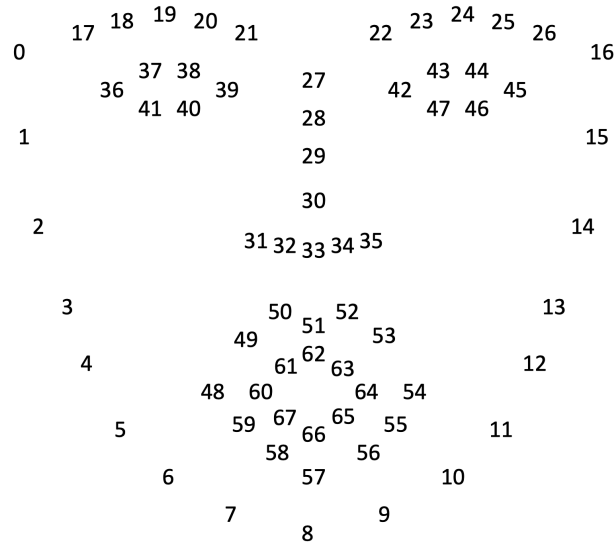


図 4.1: 顔の特徴点の番号の対応.

また、顔と画面の距離が 10 cm の時は、想定する場合の距離として極端に近いだけでなく、目のみなどごく一部の部位しか映らないことから、画面からの距離は 20 cm～50 cm の間で顔の特徴点検出が行えるかどうかを試行した。

4.2 結果

結果は、以下図4.2から図4.5に示す通りに全ての場合で顔の認識には成功した。

尚、図4.2から図4.5内の単位はそれぞれ X 軸と Z 軸が mm, Y 軸が m となっている。

4.3 考察

画面との距離が 50 cm のとき顔の形がはっきりと現れていることがわかる。また、3章で行なった実験と比較すると顔の関節を全く認識できなかった 20 cm と 30 cm の時にも顔の認識に成功していることがわかる。画面との距離は2.1.2でも示したように 40 cm 以上が望ましいとされているため、40cm 未満の状態の顔の検知を行えることは大きな利点であると考えられる。それだけでなく、深度の情報も加えて得ることができているため、顔と画面の距離も情報として得ることができている。ここで、VDT 症候群の原因としては顔

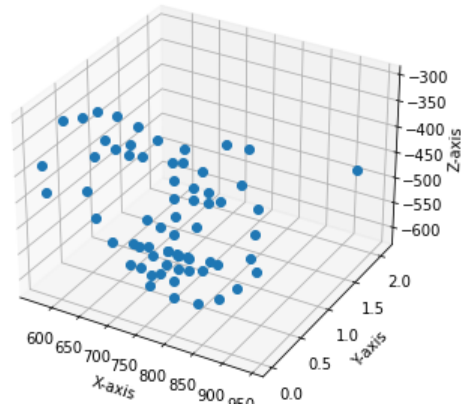
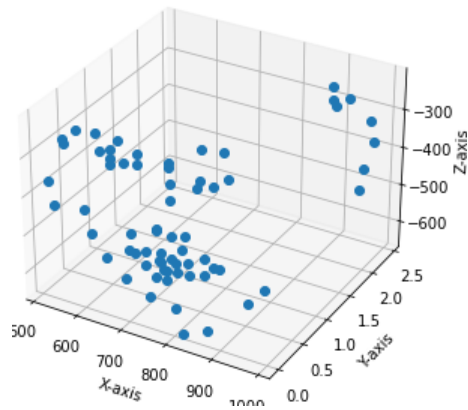


図 4.2: 画面との距離が 20cm における顔の特徴点の分布. 図 4.3: 画面との距離が 30cm における顔の特徴点の分布.

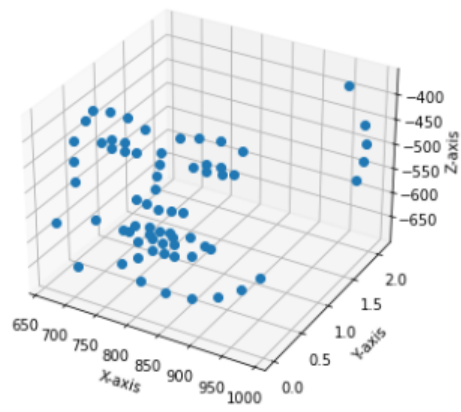
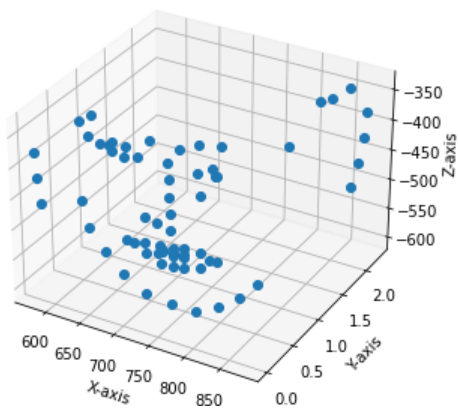


図 4.4: 画面との距離が 40cm における顔の特徴点の分布. 図 4.5: 画面との距離が 50cm における顔の特徴点の分布.

と画面の距離だけでなく、首の曲がり方や背筋の曲がり方に連動する顔の角度も一つの大きな指標であると考えられるため、本章で得られた顔の特徴量の三次元座標を用いて顔の角度を推定できれば、その角度を姿勢推定の一つの情報として用いることができるため、次章で顔の角度を推定する予備実験を行なった。

Chapter 5

深度カメラを用いた顔の角度検出による姿勢推定手法の検討

5.1 はじめに

3章や4章から得た知見も考慮し，ウェブカメラから得られる情報は2次元かつ全身が映らない限定的な画角であるため，実験には正面からでも深さの情報を得ることができる深度カメラを用いることとした。

さらに，姿勢に関連する関節の座標推定は，全身が映らず肩が一部しか映らない等の条件により良い精度が見込めない．そこで深度の情報を含む顔の特徴量であればそこから姿勢を推定できるのではないかと考え，3次元座標の推定を行なった．本章では，顔の特徴点の3次元座標を用いて実際の顔の角度を推定するため，まず学生6名に対し顔の角度を推定する実験を行った。

5.2 実験手法及び取得したデータ

顔の3次元座標から顔の角度を推定することで悪い姿勢を特定するため，まず学生6名に対し顔の角度を推定する予備実験を行った。

具体的には，顔の角度を 0° から 10° 刻みで変更していき，その時の顔を深度カメラで撮影しつつ顔の68点の特徴点検出を行った．その上でその3次元座標を取得し，その特徴点の3次元座標から顔の角度が推定できるかを実験した。

実験では一人当たり顔の角度を 0° から 60° の範囲で 10° 刻みの計7つの角度について3セットずつデータを取得し，分析には $6 \times 7 \times 3$ の126個のうち，角度が極端（ 0° や 60° ）な場合に顔として認識できなかった22個を除いた104個を使用した。

また、実際の顔の角度は地面に対し垂直に立てた場合にその角度となる模範となる平面板を用意し、調整を行った。

各角度ごとに模範となる平面板を用いて調整し、各角度における3回の試行の中でも適時修正を行った。なお、ここで収集したデータは30fpsの30フレーム、つまり1秒間の動画データ(.bagフォーマット)である。なお、本章の実験の目的のためには画像で十分なため1フレーム目の画像を取得データとして用いている。その1フレーム目の顔の各特徴点の座標を計算したものを取得した。ただし、上述の通り顔の角度が極端(0° や 60°)な場合に画像を顔として認識できなかった場合があったため、その場合は動画データのみを保持した。また、本実験の詳細なセットアップについては5.3章に示す。

5.3 本実験のセットアップ

深度カメラにはIntel社のRealSense™D435iを使用し、顔の特徴点の推定には顔の特徴点計68点を得ることができる機械学習ライブラリDlibを用いた。

さらに顔の角度について、鼻の頂点と顎の先を結んだ線分に関して、地面に対して垂直方向を 0° とし、そこから顔が下を向く方向に角度が正に増加するものとした。

実験は全て同じ場所で行い、他の条件は揃えるために床を基準として、

- 地面からの椅子の高さ：41 cm
- 深度カメラの地面からの高さ：93 cm
- カメラの中心から椅子の距離：60 cm

と設定した。これらは2.1.2節の厚生労働省によるガイドラインに基づく理想の姿勢を参考にした。

5.4 顔の角度の推定方法

ある3次元座標上の2点に対して、その仰角をここでは、その2点の角度とする。実際の顔の角度を推定する方法としては、角度の推定に鼻の頂点と顎の先の情報を用いたためこれらに対応する特徴点2点を用いてその角度を用いることが一番単純な方法である。しかし、その2点の角度のみを用いてしまうとそれら2点の測定部分の誤りや誤差がそのまま角度推定結果に反映されてしまう恐れが存在し、他の特徴点の情報を一切用いないこととなる。そこで、顔の68点のどの特徴点を結んだ線分から計算される角度が実際の顔の角度に寄与するのかを特徴量抽出等を用いて分析することとした。

5.5 特徴量抽出

角度の推定に鼻の頂点と顎の先の情報のみを用いてしまうと、それら2点の測定部分の誤りや誤差がそのまま角度推定結果に反映されてしまうため、顔の68点のどの特徴点を結んだ線分から計算される角度が実際の顔の角度に寄与するのかを分析することとした。なお、ここでの顔の角度は2点を結んだ線分の仰角と定義する。ただし、68点の特徴点から2点選ぶ方法は ${}_{68}C_2$ の2278通りも存在する。

これらを全て特徴点として用いるのは数が多すぎるかつ例えば右眉毛の右端（図4.1における17番）とその隣の点（図4.1における18番）を結んだ線分の角度など、実際の顔の角度の推定に対し明らかに影響力の小さいと思われる組み合わせも多数含まれている。そこで、その組み合わせの中から目的変数推定に影響力の大きい組み合わせを選ぶ特徴量抽出を行うこととした。

5.6 変数増減法による重回帰分析

本実験では説明変数が多いため1つの目的変数を複数の説明変数で予測する重回帰分析を用いることとする。ただし、本実験における説明変数は2278と多すぎるため、これを減らすことが望ましい。変数を選択する方法には知識や科学的理論によって最適な説明変数を指定する変数指定法や全ての説明変数の組み合わせを計算する総当たり法、一定の規則に従って説明変数を逐次選択していく逐次選択法の3つが主にあるが、変数指定法では事前にある程度最適な組み合わせを決定する必要があり本実験の趣旨と外れる。また、総当たり法は説明変数が多く、また効率も悪いため逐次選択法を用いることとした。そこで有用な特徴量の選択を行う事のできる逐次選択法の一つである変数増減法（ステップワイズ法）を用いて特徴量抽出を行うこととした。

この変数増減法内で用いる赤池情報量規準 (AIC) についてまず説明する。AIC は最大尤度 L と説明変数の数 k を用いて次の5.1式で定義される値である。

$$AIC = -2 \ln L + 2(k + 1) \quad (5.1)$$

尤度は得られたデータに対しモデルのパラメータの確らしさを示す値であり、最大尤度はその値が最大となるもの。この最大尤度が大きいほど良いモデルであることを示す。ただし与える説明変数の数が多いほどオーバーフィッティングしてしまうためペナルティ項として $2(k + 1)$ を与えている。すなわち AIC は値が小さいほどモデルの性能の良さを示す指標である。

さて、変数増減法の詳細を以下に示す。選択された特徴量がない状態から始め、そこからまず全ての特徴量に対してそれぞれ一つの特徴量のみでモデルを作った場合の AIC を計算し、AIC はその値が小さい方が良いモデルであることを示す指標なため、その AIC 値が最も小さくなるモデルを作る一つの特徴量を最初の有用な特徴量として選択する（仮にこの特徴量を A とする.）。そして、次にそこで選ばれた特徴量 A に選ばれなかった他の特徴量（仮に B, C, D, E, ... とする.）1つを代わり代わりに加えていき、全ての新たに加えた特徴量と前に選ばれた特徴量 A との二つの特徴量の組み合わせで構成されたモデルの AIC の値が A のみで構成されたモデルの AIC の値を上回った場合は一つの特徴量のみを抽出し終了する。逆に二つの特徴量を用いたモデル (例えば A と B) で計算される AIC 値が一つ目のループの特徴量一つを用いたモデル中で最も小さかった (A のみを用いた場合の) AIC の値を下回った場合はその値を保持し、そのループで最も小さくなる組み合わせを作る新たな変数を選択し加える。（例えば A と F が最も小さい場合、F を新たに選択する特徴量として採択する.）そして次に新たに変数を加えて同様のループを開始する。これを、AIC がそれ以上小さくなくなるとループが終了するまで繰り返した。

5.7 説明変数の与え方

まずはじめに、5.5節で示した通り、そもそもの説明変数の数が多いことから、特徴量抽出を行う前に実際の顔の角度を目的変数として、説明変数として与えるものをフィルタリングしようと考え、以下のように顔の特徴点をグループ分けすることで特徴点同士の組み合わせの数を減らして説明変数の数を減少させることとした。

5.7.1 グループ A、B を用いた分析

得られた 68 点の特徴点を 17 番から 47 番の 31 点 (グループ A) と 0 番から 16 番及び 48 番から 67 番の 37 点 (グループ B) に 2 分割し、 $31 \times 37 = 1147$ の組み合わせの 2 点を結んだ線分に対して計算される仰角を説明変数、目的変数を実際の顔の角度として変数増減法による最小二乗回帰の特徴量抽出を行った。このグループ分けは、全ての点ではなく部位ごとにまとめ、違う部位同士の組み合わせを説明変数として与えることでフィルタリングできるのではないかと考え、顔の部位に基づいて行い、グループ A を顔の上半分を構成する両眉毛、両目、鼻を構成する特徴点とし、グループ B は顔の輪郭及び口とした。特徴点の番号と実際の顔の対応を図5.1に示す。ここでは各顔の部位を色分けした上で、A グループ及び B グループに含めた特徴点の範囲を囲って示している。

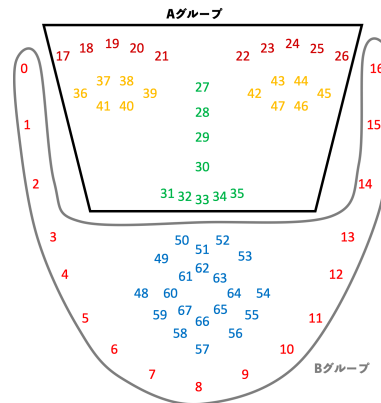


図 5.1: 顔の特徴点の番号と AB のグループ分け。

結果

上述の特徴点をグループ A とグループ B に分けた上で組み合わせ、説明変数として用いて重回帰分析した結果を以下に示す。まず、変数増減法によって選ばれた 54 個の説明変数の詳細を表 5.1 に示す。

説明変数名において、定数項の const 以外の、4 つの数字の並びは前から二つ目までの数字と、三つ目以降の数字を特徴点番号としてそれらを組み合わせた仰角を表している。例えば、表 5.1 の二行目の 2614 は、図 4.1 における顔の 26 番目の特徴点と 14 番目の特徴点を結んだ線分の仰角を表している。

次の列名の coef は回帰係数を表している。これは、回帰分析モデルにおける各説明変数の係数の値である。

次の列名の std err は標準誤差を表している。これは、推定量と母集団のばらつきを示し、小さいほど推定値の精度の高さを表す。

次の列名の t は t 値を表し、値が大きいほど有意性を示す値である。

次の列名の $P > |t|$ は、p 値を表し、t 値とは逆に値が小さいほど有意性が高いことを示し、一般的に 0.05 以下であれば良いとされている。

最後の 95% 信頼区間は母集団の平均が 95% で含まれている区間を示している。

表 5.1: 特徴量抽出により得られた 54 の点の組み合わせ。各組み合わせの 2 点を結んだ線分に対して計算される仰角を説明変数とする。

説明変数名	coef	std err	t	$P > t $	95% 信頼区間	
const	60.6706	1.938	31.308	0.000	56.776	64.565
2614	0.0314	0.010	2.998	0.004	0.010	0.052
2900	-0.2348	0.026	-9.140	0.000	-0.286	-0.183

表は次ページに続く

前ページからの続き

説明変数名	coef	std err	t	P> t	95% 信頼区間	
2612	0.0082	0.012	0.687	0.495	-0.016	0.032
4250	-0.1380	0.025	-5.558	0.000	-0.188	-0.088
2516	-0.0537	0.013	-4.096	0.000	-0.080	-0.027
3500	-0.4867	0.034	-14.139	0.000	-0.556	-0.417
1716	0.0618	0.011	5.828	0.000	0.040	0.083
4004	-0.3052	0.026	-11.755	0.000	-0.357	-0.253
2902	0.0697	0.014	5.065	0.000	0.042	0.097
2506	-0.1917	0.020	-9.828	0.000	-0.231	-0.153
1715	-0.0696	0.014	-5.138	0.000	-0.097	-0.042
4461	0.0532	0.014	3.896	0.000	0.026	0.081
3558	0.0508	0.025	2.021	0.049	0.000	0.101
1958	-0.0434	0.032	-1.357	0.181	-0.108	0.021
1759	0.0924	0.012	7.492	0.000	0.068	0.117
2104	0.2002	0.025	8.050	0.000	0.150	0.250
3567	0.3155	0.035	9.106	0.000	0.246	0.385
2565	0.1347	0.017	8.063	0.000	0.101	0.168
2756	-0.0127	0.010	-1.296	0.201	-0.032	0.007
3351	0.1861	0.018	10.410	0.000	0.150	0.222
3267	-0.2113	0.030	-7.077	0.000	-0.271	-0.151
3365	-0.0974	0.021	-4.702	0.000	-0.139	-0.056
3665	0.1002	0.021	4.666	0.000	0.057	0.143
4615	-0.0029	0.013	-0.232	0.818	-0.028	0.023
3466	0.1732	0.023	7.698	0.000	0.128	0.218
1951	-0.0615	0.033	-1.847	0.071	-0.128	0.005
4108	0.0706	0.018	3.866	0.000	0.034	0.107
4005	-0.0115	0.022	-0.524	0.603	-0.056	0.033
3656	-0.0887	0.016	-5.666	0.000	-0.120	-0.057
3401	-0.0794	0.017	-4.608	0.000	-0.114	-0.045
2613	0.0367	0.010	3.857	0.000	0.018	0.056
2066	-0.1349	0.025	-5.426	0.000	-0.185	-0.085
4648	0.0307	0.012	2.596	0.012	0.007	0.054
4059	0.1822	0.030	6.153	0.000	0.123	0.242
3960	-0.0453	0.018	-2.498	0.016	-0.082	-0.009
3751	-0.0748	0.028	-2.707	0.009	-0.130	-0.019
2648	0.0735	0.013	5.568	0.000	0.047	0.100
4448	-0.0458	0.010	-4.402	0.000	-0.067	-0.025
2455	-0.0642	0.014	-4.728	0.000	-0.091	-0.037
2306	0.0529	0.033	1.586	0.119	-0.014	0.120
3462	-0.0187	0.014	-1.352	0.183	-0.047	0.009
3403	-0.0299	0.013	-2.248	0.029	-0.057	-0.003
3614	0.0256	0.011	2.306	0.025	0.003	0.048
2257	0.0048	0.017	0.277	0.783	-0.030	0.040
3207	0.0682	0.024	2.820	0.007	0.020	0.117
2950	0.0320	0.025	1.265	0.212	-0.019	0.083

表は次ページに続く

前ページからの続き

説明変数名	coef	std err	t	P> t	95% 信頼区間	
4516	-0.0424	0.013	-3.280	0.002	-0.068	-0.016
4558	0.0317	0.011	2.931	0.005	0.010	0.054
3159	-0.0521	0.021	-2.524	0.015	-0.094	-0.011
2010	0.0354	0.019	1.851	0.070	-0.003	0.074
4307	-0.0388	0.025	-1.578	0.121	-0.088	0.011
3006	0.0289	0.021	1.388	0.172	-0.013	0.071
3753	0.0196	0.016	1.229	0.225	-0.012	0.052
1854	0.0154	0.013	1.176	0.245	-0.011	0.042

これで終わり

表 5.2: 重回帰分析結果.

R-squared	Adj. R-squared	F-statistic	Prob(F-statistic)
0.983	0.965	53.01	3.94e-30
Log-Likelihood	AIC	BIC	Covariance Type
-234.26	578.5	724.0	nonrobust

特徴量抽出の結果, 1147 の角度情報のうち, 表5.1にあげる 54 の角度情報を用いるモデルが最も低い AIC (578.5) となり, 決定係数が 0.983 となった. モデルの分析結果を表5.2に示す. なお表5.1について, 変数増減法では変数を追加していき AIC が最も低くなる変数を選んでいくがその選ばれた順番となっている.

5.7.2 グループ C、D を用いた分析

次に5.7.1節では部位ごとにグループ分けしたが, 顔の輪郭が広範囲であったことを加味し, 顔の上半分と下半分でグループ分けすることを試した. すなわち, 得られた 68 点の特徴点を 0 番から 3 番及び 13 番から 47 番の 39 点 (グループ C) と 4 番から 12 番及び 48 番から 67 番の 29 点 (グループ D) に 2 分割し, $39 \times 29 = 1131$ の組み合わせの 2 点を結んだ線分に対して計算される仰角を説明変数, 目的変数を実際の顔の角度として変数増減法による最小二乗重回帰の特徴量抽出を行った. 特徴点の番号と実際の顔の対応を図5.2に示す. ここでは, C グループ及び D グループに含めた特徴点の範囲を囲って示している.

結果

上述の特徴点をグループ C とグループ D に分けた上で組み合わせ, 説明変数として用いて重回帰分析した結果を以下に示す. まず, 変数増減法によって選ばれた 54 個の説明変数の詳細を表5.3に示す.

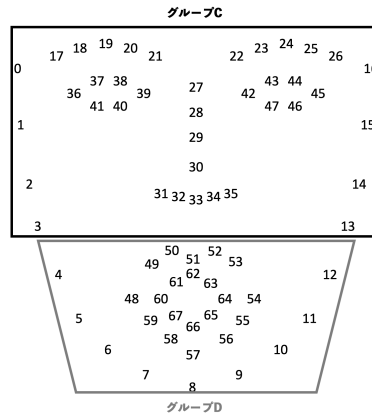


図 5.2: 顔の特徴点の番号と CD のグループ分け.

表 5.3: 特徴量抽出により得られた 54 の点の組み合わせ. 各組み合わせの 2 点を結んだ線分に対して計算される仰角を説明変数とする.

説明変数名	coef	std err	t	P> t	95% 信頼区間	
const	44.5524	2.182	20.417	0.000	40.167	48.937
2612	0.0056	0.019	0.300	0.765	-0.032	0.043
0061	0.2489	0.037	6.655	0.000	0.174	0.324
4261	0.0355	0.033	1.087	0.282	-0.030	0.101
1657	-0.1058	0.022	-4.818	0.000	-0.150	-0.062
4050	-0.0536	0.022	-2.462	0.017	-0.097	-0.010
1448	-0.1135	0.012	-9.376	0.000	-0.138	-0.089
1558	0.0738	0.013	5.793	0.000	0.048	0.099
2961	-0.1046	0.024	-4.396	0.000	-0.152	-0.057
2563	-0.1037	0.012	-8.657	0.000	-0.128	-0.080
4748	0.0646	0.013	5.090	0.000	0.039	0.090
4458	0.0973	0.034	2.888	0.006	0.030	0.165
3260	-0.1327	0.019	-7.001	0.000	-0.171	-0.095
0062	0.1435	0.042	3.451	0.001	0.060	0.227
0210	-0.0956	0.017	-5.547	0.000	-0.130	-0.061
3152	-0.0665	0.034	-1.969	0.055	-0.134	0.001
1759	0.1338	0.020	6.853	0.000	0.095	0.173
1949	-0.0892	0.032	-2.762	0.008	-0.154	-0.024
3465	-0.1532	0.017	-9.063	0.000	-0.187	-0.119
1665	0.1698	0.022	7.603	0.000	0.125	0.215
3048	0.0429	0.018	2.412	0.020	0.007	0.079
4352	-0.3239	0.039	-8.258	0.000	-0.403	-0.245
1367	0.0275	0.012	2.343	0.023	0.004	0.051
4662	-0.0267	0.015	-1.778	0.082	-0.057	0.003
0363	0.0822	0.016	5.281	0.000	0.051	0.113
3962	-0.0861	0.030	-2.870	0.006	-0.146	-0.026
2357	0.1014	0.023	4.504	0.000	0.056	0.147

表は次ページに続く

前ページからの続き

説明変数名	coef	std err	t	P> t	95% 信頼区間	
4504	-0.0176	0.011	-1.564	0.124	-0.040	0.005
3452	0.2317	0.029	8.057	0.000	0.174	0.290
4405	-0.0384	0.031	-1.219	0.229	-0.102	0.025
3611	-0.1197	0.023	-5.285	0.000	-0.165	-0.074
3063	0.0349	0.020	1.737	0.089	-0.005	0.075
4309	0.0816	0.030	2.764	0.008	0.022	0.141
4306	0.1720	0.036	4.843	0.000	0.101	0.243
2058	-0.0230	0.031	-0.740	0.463	-0.086	0.039
3407	-0.1067	0.025	-4.189	0.000	-0.158	-0.055
1755	-0.0511	0.019	-2.701	0.009	-0.089	-0.013
2664	0.0312	0.015	2.123	0.039	0.002	0.061
2855	0.0582	0.017	3.346	0.002	0.023	0.093
3151	0.1495	0.032	4.638	0.000	0.085	0.214
0104	-0.0642	0.017	-3.710	0.001	-0.099	-0.029
3554	0.0362	0.013	2.691	0.010	0.009	0.063
3156	-0.0523	0.015	-3.566	0.001	-0.082	-0.023
2851	-0.1035	0.029	-3.520	0.001	-0.163	-0.044
4448	-0.0116	0.015	-0.769	0.446	-0.042	0.019
3859	0.0829	0.030	2.740	0.009	0.022	0.144
4412	-0.0335	0.016	-2.050	0.046	-0.066	-0.001
2811	0.0961	0.027	3.588	0.001	0.042	0.150
1953	-0.0608	0.026	-2.356	0.023	-0.113	-0.009
3667	0.0573	0.019	2.973	0.005	0.019	0.096
3759	-0.0601	0.027	-2.249	0.029	-0.114	-0.006
3266	0.0205	0.015	1.326	0.191	-0.011	0.052
2449	0.0286	0.019	1.536	0.131	-0.009	0.066
2754	-0.0179	0.012	-1.450	0.153	-0.043	0.007
3551	-0.0199	0.016	-1.212	0.231	-0.053	0.013

これで終わり

表 5.4: 重回帰分析結果.

R-squared	Adj. R-squared	F-statistic	Prob(F-statistic)
0.976	0.949	36.49	2.55e-26
Log-Likelihood	AIC	BIC	Covariance Type
-253.29	616.6	762.0	nonrobust

特徴量抽出の結果, 1131 の角度情報のうち, 表5.3にあげる 54 の角度情報を用いるモデルが最も低い AIC (616.6) となり, 決定係数が 0.976 となった. モデルの分析結果を表5.4に示す. なお表5.3について, 変数増減法では変数を追加していき AIC が最も低くなる変数を選んでいくがその選ばれた順番となっている.

5.7.3 全ての特微量を用いた分析

上記の AB グループ及び CD グループのグループ分けによる説明変数の削減がモデルの性能にどのような影響を与えたかを調べるためにグループ分けを行わず、68 点から 2 点を選ぶ全ての組み合わせである ${}_{68}C_2$ の 2278 の組み合わせの 2 点を結んだ線分に対して計算される仰角を同様に説明変数、目的変数を実際の顔の角度として変数増減法による最小二乗回帰の特微量抽出を行った。

結果

全ての特徴点の組み合わせを説明変数として用いて重回帰分析した結果を以下に示す。まず、変数増減法によって選ばれた 54 個の説明変数の詳細を表 5.5 に示す。

表 5.5: 特微量抽出により得られた 62 の点の組み合わせ。各組み合わせの 2 点を結んだ線分に対して計算される仰角を説明変数とする。

説明変数名	coef	std err	t	P> t	95% 信頼区間	
const	51.3861	1.078	47.686	0.000	49.210	53.562
1426	-0.0834	0.009	-9.341	0.000	-0.101	-0.065
0050	0.4634	0.017	27.080	0.000	0.429	0.498
4250	-0.4919	0.023	-21.397	0.000	-0.538	-0.445
0526	-0.1135	0.012	-9.816	0.000	-0.137	-0.090
1942	-0.3901	0.019	-20.328	0.000	-0.429	-0.351
1929	0.2678	0.029	9.287	0.000	0.210	0.326
1761	0.0942	0.011	8.500	0.000	0.072	0.117
2125	0.1116	0.008	13.501	0.000	0.095	0.128
0219	-0.1379	0.013	-10.340	0.000	-0.165	-0.111
2940	0.0872	0.019	4.497	0.000	0.048	0.126
1635	-0.0061	0.008	-0.772	0.444	-0.022	0.010
2744	-0.0234	0.006	-3.684	0.001	-0.036	-0.011
5160	-0.1119	0.011	-9.790	0.000	-0.135	-0.089
1167	0.1005	0.017	6.051	0.000	0.067	0.134
3940	-0.0773	0.011	-7.268	0.000	-0.099	-0.056
1358	0.0563	0.009	6.255	0.000	0.038	0.075
2038	0.0082	0.011	0.733	0.468	-0.014	0.031
0414	0.1454	0.013	11.530	0.000	0.120	0.171
1141	-0.1190	0.017	-6.879	0.000	-0.154	-0.084
3043	0.1236	0.012	10.592	0.000	0.100	0.147
0234	0.0896	0.015	5.786	0.000	0.058	0.121
0359	0.0406	0.009	4.521	0.000	0.022	0.059
1248	-0.0012	0.010	-0.124	0.902	-0.021	0.019
1465	0.1074	0.010	10.293	0.000	0.086	0.128
3246	-0.0628	0.008	-7.440	0.000	-0.080	-0.046
0966	0.0680	0.012	5.814	0.000	0.044	0.092

表は次ページに続く

前ページからの続き

説明変数名	coef	std err	t	P> t	95% 信頼区間	
1213	0.0831	0.011	7.895	0.000	0.062	0.104
1044	0.0506	0.010	5.062	0.000	0.030	0.071
3452	0.0756	0.010	7.814	0.000	0.056	0.095
1727	0.0450	0.006	7.223	0.000	0.032	0.058
0865	0.0169	0.011	1.568	0.125	-0.005	0.039
1315	0.0353	0.006	5.878	0.000	0.023	0.047
1725	-0.0870	0.009	-9.775	0.000	-0.105	-0.069
2254	-0.0878	0.012	-7.075	0.000	-0.113	-0.063
3057	-0.1237	0.015	-7.989	0.000	-0.155	-0.092
5258	0.0409	0.008	4.816	0.000	0.024	0.058
0531	0.0361	0.012	3.025	0.004	0.012	0.060
3751	-0.0734	0.015	-5.024	0.000	-0.103	-0.044
0024	-0.0389	0.010	-3.909	0.000	-0.059	-0.019
4347	0.0524	0.008	6.182	0.000	0.035	0.070
0166	0.0343	0.010	3.277	0.002	0.013	0.055
1617	-0.0088	0.007	-1.336	0.189	-0.022	0.005
3659	0.0057	0.006	0.964	0.341	-0.006	0.018
3233	-0.0366	0.008	-4.680	0.000	-0.052	-0.021
5564	0.0307	0.008	3.741	0.001	0.014	0.047
0945	-0.0244	0.007	-3.324	0.002	-0.039	-0.010
1517	0.0169	0.007	2.561	0.014	0.004	0.030
5465	-0.0631	0.011	-5.514	0.000	-0.086	-0.040
1255	0.0321	0.008	3.855	0.000	0.015	0.049
0922	-0.0471	0.015	-3.210	0.003	-0.077	-0.017
3035	0.0579	0.012	4.808	0.000	0.034	0.082
0608	0.0630	0.015	4.243	0.000	0.033	0.093
3553	0.0360	0.011	3.142	0.003	0.013	0.059
4448	0.0145	0.009	1.691	0.098	-0.003	0.032
0653	-0.0211	0.013	-1.589	0.120	-0.048	0.006
5162	-0.0147	0.009	-1.644	0.108	-0.033	0.003
0767	0.0334	0.015	2.175	0.035	0.002	0.064
4546	-0.0082	0.006	-1.488	0.144	-0.019	0.003
0761	-0.0186	0.012	-1.599	0.118	-0.042	0.005
3141	-0.0162	0.013	-1.228	0.227	-0.043	0.010
2547	0.0067	0.006	1.073	0.290	-0.006	0.019
4049	0.0132	0.015	0.901	0.373	-0.016	0.043

これで終わり

表 5.6: 重回帰分析結果.

R-squared	Adj. R-squared	F-statistic	Prob(F-statistic)
0.995	0.987	128.5	6.97e-34
Log-Likelihood	AIC	BIC	Covariance Type
-172.38	470.8	637.4	nonrobust

特徴量抽出の結果、2278 の角度情報のうち、表5.5にあげる 62 の角度情報を用いるモデルが最も低い AIC (470.8) となり、決定係数が 0.995 となった。モデルの分析結果を表5.6に示す。なお表5.5について、変数増減法では変数を追加していき AIC が最も低くなる変数を選んでいくがその選ばれた順番となっている。

Chapter 6

考察

6.1 各分析結果に対する考察

5章の結果により，全ての 2278 通りの特徴点の組み合わせに変数増減法を適用しモデルを構築した方が，事前に自身でグループ分けした上で与える組み合わせを減らしたモデルよりも最終的な AIC 値，決定係数共に良いモデルを示す値となった。

各モデルの AIC 値及び決定係数を表6.1に示す。

表 6.1: 各モデルの AIC 値及び決定係数.

モデル名	AIC 値	決定係数
AB	578.5	0.983
CD	616.6	0.976
全て	470.8	0.995

これは，二つの理由があると考えられる。

一つ目は，自身で場合分けし作成した組み合わせでない組み合わせの中に目的変数に寄与するものが存在したということである。例えば全ての組み合わせで用いたモデルにおいて一番最初に選ばれた 14 番と 26 番の組み合わせや，六番目に選ばれた 19 番と 29 番のような想定していた顔全体に対する上半分と下半分の組み合わせではなく想定していない局所的な上半分と下半分の組み合わせが目的変数に大きな影響力を持っていたことが挙げられる。

二つ目は，最終的に選ばれた説明変数の数が全てを用いたモデルでは 62 個であったのに対し，AB グループ及び CD グループに分けた場合のモデルの説明変数数は 54 個と，選ばれた説明変数が多かったことである。このため，変数が多い方がより与えたデータに適合しやすいことでモデルの精度が上がったことが考えられる。

また、自身が AB の組み合わせを改良する目的で新たにグループ分けした CD のモデルの性能は予想に反し、AB を下回ることとなった。

この理由も自身の想定していない組み合わせが大きな影響力を持っていたことが挙げられる。それは、例えば全ての組み合わせの場合の特徴量抽出で一番最初に選ばれた「14番と26番」という組み合わせが、AB グループの組み合わせでも一番最初に選ばれた組み合わせとなっていてモデルに大きな影響を与える説明変数であることが示唆されるが、CD グループでは14番と26番が同じ C グループに属しこの組み合わせの説明変数は存在しないことから裏付けされる。

6.2 出現頻度の分析

次に構築した三つのモデルにおいて選択されたそれぞれの特徴点の傾向を見るために、モデルの説明変数として用いられた角度における組み合わせに用いられていた特徴点の出現回数を調べた。AB グループで分けた場合のモデル、CD グループで分けた場合のモデル、全ての組み合わせで構築したモデルのそれぞれについて特徴点ごとに出現回数をまとめたものを表6.2に示す。

表 6.2: 全てのまとめ。

特徴点番号	AB	CD	全て	3つの平均
0	2	2	2	2.0
1	1	1	1	1.0
2	1	1	2	1.33
3	1	1	1	1.0
4	2	2	1	1.67
5	1	1	2	1.33
6	3	1	2	2.0
7	2	1	2	1.67
8	1	0	2	1.0
9	0	1	3	1.33
10	1	1	1	1.0
11	0	2	2	1.33
12	1	2	3	2.0
13	1	1	3	1.67
14	2	1	3	2.0
15	2	1	2	1.67
16	3	2	2	2.33
17	3	2	5	3.33
18	1	0	0	0.33
19	2	2	3	2.33
20	2	1	1	1.33

表は次ページに続く

前ページからの続き

特徴点番号	AB	CD	全て	3つの平均
21	1	0	1	0.67
22	1	0	2	1.0
23	1	1	0	0.67
24	1	1	1	1.0
25	3	1	3	2.33
26	4	2	2	2.67
27	1	1	2	1.33
28	0	3	0	1.0
29	3	1	2	2.0
30	1	2	3	2.0
31	1	3	2	2.0
32	2	2	2	2.0
33	2	0	1	1.0
34	4	3	2	3.0
35	3	2	3	2.67
36	3	2	1	2.0
37	2	1	1	1.33
38	0	1	1	0.67
39	1	1	1	1.0
40	3	1	3	2.33
41	1	0	2	1.0
42	1	1	2	1.33
43	1	3	2	2.0
44	2	4	3	3.0
45	2	1	2	1.67
46	2	1	2	1.67
47	0	1	2	1.0
48	3	4	2	3.0
49	0	2	1	1.0
50	2	1	2	1.67
51	3	3	3	3.0
52	0	3	2	1.67
53	1	1	2	1.33
54	1	2	2	1.67
55	1	2	2	1.67
56	2	1	0	1.0
57	1	2	1	1.33
58	3	3	2	2.67
59	3	3	2	2.67
60	1	1	1	1.0
61	1	3	2	2.0
62	1	3	1	1.67
63	0	3	0	1.0
64	0	1	1	0.67

表は次ページに続く

前ページからの続き

特徴点番号	AB	CD	全て	3つの平均
65	3	2	3	2.67
66	2	1	2	1.67
67	2	2	2	2.0

これで終わり

表6.2から用いられた平均回数が2.0を上回る点を抽出すると、16(左上輪郭に位置する特徴点)、17、19(右眉毛)、25、26(左眉毛)、34、35(鼻)、40(右目)、44、48、51、58、59、65(口)の14点となった。これにより、眉毛と口を構成する特徴点が多く使われ、目や輪郭の大部分を構成する特徴点の出現頻度は少ないことが分かった。

また、表6.2における出現頻度を合計の出現回数で割ったものを各モデルの出現頻度ベクトルとしてそれらの類似度をユークリッド距離を用いて計算を行った。その結果、それぞれABと全てのユークリッド距離が0.07908576859556622、CDと全てのユークリッド距離が0.08653334878500264、ABとCDの距離が0.09711193038612514であった。

このことから全てを用いた場合に近いのはCDグループではなく、ABグループであり、ABとCDグループの距離は他と比べ離れていることがわかる。この結果からもCDグループのモデルが一番性能を示せなかったことが裏付けできる。

6.3 用いるモデルの決定

3つのモデルの分析の結果、全ての特徴点の組み合わせから計算された仰角を説明変数として特徴量抽出した結果である62の角度情報を用いたモデルが、最も低いAICの値(470.8)と、最も高い決定係数(0.995)を示し、最も高い性能を示したため今後はこのモデルを実際の顔の角度を推定するモデルとして用いていくこととする。

Chapter 7

まとめと今後の展望

5章の結果により，深度カメラを用いて全ての特徴点の組み合わせから，変数増減法による特徴量抽出を行った結果，62の組み合わせを変数とした，実際の顔の角度を推定するモデルを構築することに成功した．

7.1 今後のこのモデルを活かした姿勢推定

今後はこのモデルを利用し，ユーザの悪い姿勢を継続的にモニタリングし，必要に応じてユーザに姿勢を正すフィードバックを提示するシステムを構築する．

まず，このモデルをユーザの適切でない姿勢の判定に用いる方法について以下のように考えている．

一つ目に適切でない姿勢判定の数少ない定量的な指標としては，厚生労働省のガイドライン [16] に示すように「顔と画面の距離を 40 cm 以上離すこと」の条件があるが，これは顔の特徴点の 3次元座標を用いれば奥行きの情報をもつ利点を最大限に活かして容易にその条件を満たしているかどうかの判定ができる．顔が近すぎる場合を判別し適切でない姿勢と推定する．

二つ目に，顔の角度から姿勢全体の曲がり具合を推定する方法であるが，顔の角度は首にも連動し，顔の角度が大きい場合には首が大きく曲がり負荷がかかっているストレートネック等につながる悪い姿勢であることが推察できるため，顔の角度は適切でない姿勢の指標である．ただし，適切な顔の角度については同じく厚生労働省のガイドライン [16] には「視線はやや下向き」とだけあり，具体的に書かれてはいない．システム上でこの「やや下向き」を顔の角度の範囲を例えば 10° から 20° のように設定する必要があると考えられる．

7.1.1 適切な姿勢における顔の角度の範囲を設定

適切な姿勢であると判定する顔の角度の範囲を設定する方法は二つあると考えている。一つ目は、大規模な実験を実施し、図2.1の姿勢での角度を収集することで、理想の角度及びその幅を決定する方法である。データの数が必要であり、統計的な方法ではあるがデータの数に伴って精度を上げていくことが可能である。二つ目は、一人のユーザがシステムを使うことを想定し、システムの初期設定でそのユーザの図2.1の姿勢における顔の角度を検出し保存しておき、その角度に幅を持たせる方法である。この方法ではユーザが姿勢推定結果がおかしい場合にそれを還元することで、ユーザ単位の精度が上がっていく利点がある。

以上のように顔と画面との距離及び顔の角度の二つの値を組み合わせ推定することでこのモデルを適切な姿勢判定に用いることが可能であると考え提案する。

7.2 このモデルを用いた姿勢推定システムの利点

利点としては三つある。

一つ目は、これまで明確に具体的な指標がなかった「適切な姿勢」を定量的に判定できることが挙げられる。定量的に判定できることでユーザに対し、その時点の姿勢に合わせた適切なフィードバックが可能だけでなく、「あと 15° 顔をあげてください」「もう 10 cm 画面に近づいてください」等の具体的な助言を行うことができる。その長さや角度を精密に人間が調整できることはないが、調整の結果正解である時に知らせることを繰り返すことで、システムを使うほど感覚的に具体的な数値を実際の調整と関連付けさせることができるようになるため、最終的には大きな意味をもつ。

二つ目は、あくまで顔の認識のみを用いているため、背景や場所に影響されることが少ないという点である。肩や首の状況を用いると、画角や服の色等の影響も考慮しなければならないが、顔の認識はロバスト性が高く、頭が切れていても問題ない。

三つ目は、コンピュータ使用時の環境自体に助言できる可能性があることである。適切な姿勢を維持するためには、ユーザの意識だけでなくコンピュータ使用時の椅子やコンピュータの位置、机の高さ等も重要な要素となっている。本システムでは姿勢判定の際、顔の角度だけでなく三次元座標も取得しているため、座標の高さが標準より平均的に高い位置にあり、継続的に顔の角度が大きいことを検知した場合は、座標等から計算した上で端末画面の位置を具体的に何 cm 高くすべき等の助言を行うことが可能である。

7.3 モデルとモデルを用いた姿勢推定システムの問題点及び改善案

まず、より大規模なユーザ実験を行うことで各データへの依存度が小さくなることでモデルの精度が高くなるため、行いたい。さらに、実験参加者の実際の顔の角度の調整を平板を用いて行ったが実際の顔の角度を正確に固定するための眼科機器のような専用の装置を作成し、精密に測定することでも精度が上がると考えられる。さらに本実験では、メガネをつけていた場合は正常に顔認識ができていたことは確認できたものの、今世界中で標準となりつつあるマスク着用時の顔認識は行っていない。今後本格的に運用していくにあたってはマスク着用時でも顔認識が行えるか、そして行える場合その精度を検証する実験も必要であると考えられる。

また、本システムについて、ある程度深度カメラが一般的に標準装備される、または身近に入手できるようになることを想定しているため、この分野が衰退していく場合は深度カメラというハードウェアも含めたシステムとして運用していくこととなる。さらに、画面との距離と顔の角度だけでは測れない適切でない姿勢をしている場合（例えば首から先は適切だが、背筋のみが曲がっている等の極端な場合）に関しては、システムの運用を行いつつ継続的にモニタリングを行っていき、対策を行っていく必要があると考えている。また、2画面で作業等の想定していない場合にも応用させていきたい。そのような特殊な場合を対策するために首等の他の部位の情報を組み込むことで対応すればモデル自体の精度が上がる可能性もある。そして、このシステムをスマートフォンに応用すればまさに顔の角度が大きな問題となっているスマートフォンのストレートネック問題への解決の大きな糸口となるだろう。

7.4 フィードバック方法の提案

モデルの精度を高め、適切でない姿勢を検知する一方で、検知したのちにどうそれをユーザにフィードバックするかも重要な課題である。

「適切な姿勢」の具体的な数値等はユーザによって変わるものではあるものの、図2.1のようなユーザがとるべき姿勢が示されているため、その姿勢での顔の角度及び位置を基準としてそこからのずれを姿勢の歪みとして扱うことが可能である。この方法は同じユーザが継続的にシステムを使うことを想定した場合は、一度の初期設定ですむこととなり、負担も軽減される。さらに、得る情報が「座標」「角度」といった定量的なものであるため、そこに基づく数値を用いた適切なフィードバックが可能であると考えられる。

さて、そのフィードバック方法についても様々な方法が考えられる。文字をポップアップウィンドウで表示することや通知音等の音声でフィードバックする方法が一般的である

が、これらは時に煩わしくまた断続的である。それを解決するフィードバック方法としては例えば以下が考えられる。その一つが顔の角度や画面との距離に応じて画面の明るさを連続的に変更することである。具体的には顔が下を向くほど、または画面から顔の角度が離れるほど明るさを暗くし、逆に顔が上を向く、または顔が近くなるほど明るくするようなシステムを構築する。これによってユーザは通知等で作業を邪魔されることなく視覚的に自分の姿勢が崩れていることが認識できる。さらに横方向の崩れに対して画面のウィンドウを傾けること等も考えられる。

また、別の視点として、システムで具体的な姿勢の崩れ方を測定できることを活かし、具体的に「どこがどう崩れているか」を伝え、その改善案を提示する、キャラクターを用いてユーザの姿勢を模倣して自覚させるなどのアプローチもできると考えられる。

Chapter 8

おわりに

本論文では、まず VDT 症候群を解決するための姿勢を用いたアプローチを検証し、深度カメラを用いた顔角度推定モデルを比較検討した。6名の参加者から取得したデータにより顔の特徴点を深度カメラを用いて検出し、それらの中から2つの3次元座標の組み合わせで計算される仰角を説明変数として、目的変数を実際の顔の角度とすることによって変数増減法を用いて特徴量抽出を行うことで、最終的には62の変数を持つモデルを構築した。今後はこの予備実験の結果をもとにユーザの姿勢を継続的にモニタリングし、必要に応じてユーザに姿勢を正すフィードバックを提示するシステムを構築したいと考えている。

Publications

国内研究会

- 高島諒, 矢谷浩司, コンピュータ使用時のユーザの前傾姿勢の検知に向けた予備的検討, 電子情報通信学会総合大会 (2022年3月)

References

- [1] Based, P. A., Syndrome, V. D. T., HIRAMA, D., KAMOZAWA, K., and MINADUKI, A. (2013). Vdt 症候群における予防アプリケーションの開発. pages 431–432.
- [2] BRADSKI, G. R. (1998). Computer vision face tracking as a component of a perceptual user interface. Workshop on Applications of Computer Vision, 1998, pages 214–219.
- [3] Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., and Sheikh, Y. (2019). Openpose: Realtime multi-person 2d pose estimation using part affinity fields.
- [4] Cao, Z., Simon, T., Wei, S.-E., and Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [5] Comaniciu, D. and Meer, P. (2002). Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619.
- [6] HIRAMA, D. and MINADUKI, A. (2014). Vdt 作業時における休憩支援アプリケーションの開発. pages 125–126.
- [7] Iwakiri, K., Mori, I., Sotoyama, M., Horiguchi, K., Ochiai, T., Jonai, H., and Saito, S. (2004). Survey on visual and musculoskeletal symptoms in VDT workers. *Sangyō eiseigaku zasshi = Journal of occupational health*, 46(6):201–212.
- [8] Murata, A., Uetake, A., Otsuka, M., and Takasawa, Y. (2001). Proposal of an index to evaluate visual fatigue induced during visual display terminal tasks. *International Journal of Human-Computer Interaction*, 13(3):305–321.
- [9] Parihar, J., Jain, V. K., Chaturvedi, P., Kaushik, J., Jain, G., and Parihar, A. K. (2016). Computer and visual display terminals (vdt) vision syndrome (cvdts). *Medical Journal Armed Forces India*, 72(3):270–276.
- [10] Tsuruoka, K., Toda, T., Ozaki, S., and Tei, H. (2011). Visual display terminal (vdt) 利用時におけるドライアイ防止システムの提案 a proposal of dry eye prevention system for vdt user 鶴岡浩平 1 戸田健 2 尾崎信耶.
- [11] Tsuzuku, T. (2016). Vertigo induced by visual display terminal work. *Equilibrium Research*, 75(6):483–488.
- [12] 伸予, . (1992). テクノストレス症候群に関する研究 (第 1 報) : テクノ症的傾向の検査尺度用の質問項目群の作成. *心身医学*, 32(5):383–390.
- [13] 労働大臣官房政策調査部 (1998). 技術革新と労働に関する実態調査報告平成 10 年.
- [14] 厚生労働省 (2005). 技術革新と労働に関する実態調査結果の概況. 旬刊労働実務, (1450):2–16.

- [15] 厚生労働省 (2009). 平成 20 年技術革新と労働に関する実態調査報告の概況 (平成 21 年 9 月 29 日). <http://www.mhlw.go.jp/toukei/itiran/roudou/saigai/anzen/08/dl/20kaku-10.pdf>.
- [16] 厚生労働省 (2020). 情報機器作業における労働衛生管理のためのガイドラインについて (令和元年 7 月 12 日). <https://www.mhlw.go.jp/content/000539604.pdf>.
- [17] 守屋優, 戸田健, and 劉欣欣 (2014). 内蔵 web カメラを用いた vdt 作業者の姿勢モニタリング (知覚情報研究会・人間生活を支援する観測・支援・評価技術).
- [18] 宮崎亮輔, 阿部孝司, 南昌秀, et al. (2012). Web カメラを用いた pc ユーザ作業時間の自動測定. 平成 24 年度情報処理学会関西支部支部大会講演論文集, 2012.
- [19] 岩崎明夫 (2017). 労働衛生対策の基本 (13)vdt 作業とその対策. 産業保健 21 : 産業医・産業看護職・衛生管理者の情報ニーズに応える, 23(1):12-15.
- [20] 杉本拓也, 伊與田光宏, and 初谷拓郎 (2015). Vdt 作業における姿勢意識向上アプリケーションの開発. In 第 77 回全国大会講演論文集, volume 2015, pages 439-440.
- [21] 橋内勇 and 大塚吉則 (2008). 大学生における猫背、腰痛・肩凝りの発現率とその対策についての調査. 北海道大学大学院教育学研究紀要, (104):205-211.
- [22] 田中良平 and 野呂影勇 (2000). ノート型 pc とデスクトップ型 pc 使用時の作業姿勢の比較・検討. 人間工学, 36(Supplement):430-431.
- [23] 知奈美, ., 正夫, ., and 由貴恵, . (2018). Vdt 機器を使用した学習環境における疲労検知のための自己接触動作認識. 教育システム情報学会誌, 35(2):145-150.
- [24] 総務省 (2020). 令和元年通信利用動向調査.
- [25] 荒木奈緒, 小田和美, et al. (2020). Vdt 症候群. サービスロジー web マガジン.
- [26] 長谷川隼平 and 真嶋由貴恵 (2020). ロボットを活用した vdt 症候群予防における効果——作業中の意識喚起を通して——. 教育システム情報学会誌, 37(1):50-55.
- [27] 雅胤, . (2010). ドライアイ研究の最前線. 日本薬理学雑誌, 135(4):138-141.
- [28] 鳥山結加, 鈴木亨, 松澤明黎, and 松原貴子 (2017). Vdt 作業従事者の肩凝り関連因子の検討. 理学療法学 Supplement, 2016:0551.
- [29] 鶴岡浩平, 戸田健, and 鄭一 (2012). Web カメラ内蔵型ノート pc 等の携帯端末における vdt 症候群予防のための瞬きを促すアプリケーションの提案. 電気学会研究会資料. MBE, 医用・生体工学研究会, 2012(107):35-40.