

画像処理を用いた牡蠣の選別

大能彪駿 (情報工学科)

指導教員 大槻典行

1 はじめに

牡蠣の選別は手作業で行われており、大きな労力と技術を必要としている。そのため、自動選別装置が望まれていた。本研究では、安価で導入のしやすい選別装置の開発を目指す。

2 装置の概要

牡蠣は台座の上に置き、台の上に設置した RaspberryPi に接続したカメラモジュールを用いて撮影する。RaspberryPi を用いたので、画像処理には多数の処理用関数の実装された OpenCV を用いた。

3 選別方法の選択

牡蠣の選別には次のような手法を考案した。1つ目は、外接矩形を用いて長辺と短辺のサイズにより選別を行うものである。外接矩形とは、図形に接する四角形である。

牡蠣に接する最小の外接矩形を取ることにより、大まかなサイズが得られ比較できると考えられる。

2つ目は、牡蠣の輪郭に対して上下左右斜めの8方向に牡蠣の重心から距離を求める方法である。この方法では、画像に対して牡蠣の重心から輪郭まで8方向8個の距離を使って判定した。

判定は、重心から距離の最小値と最大値の差によって判定した。最大値と最小値の差が小さいということは、距離の変化が小さいということを表していると考えられ、牡蠣の形の良さを表していると見ることができる。

4 判別処理の流れ

どちらの方法も、まず始めにカメラで画像を撮影保存し処理を開始する。次に保存した画像をグレースケール化、二値化し収縮や膨張といったノイズ処理を行いそれぞれの判別に移る。外接矩形を使った方法では、長辺と短辺の和を取ったものを基準として判定した。

8方向の重心から輪郭までの距離を使った方法では、ノイズ処理後ラプラシアンフィルタによるエッジ検出を行う。エッジ検出をすると図1のようになる。



図1: ラプラシアンフィルタ適用後の画像

そして、Bresenham のアルゴリズムを使って重心からの線と輪郭との交点を求め距離を取っている。この操作を上下左右斜めの8回繰り返し、得られたの線の長さを基に選別を行う。

5 実験

実験は牡蠣を撮影した画像とパソコンを使って行った。画像は実験装置を使い撮影した。牡蠣は選別済みで、規格品の画像は20枚、規格外品は23枚の計43枚を使った。

その結果外接矩形を用いた方法の結果は、正答率はサイズの和が280を閾値とした時が最大で、規格品のうち正答数は10枚で、規格外品のうち正答数は6枚だった。正答率は、37パーセントだった。

上下左右斜めの8方向に重心から輪郭の距離を使った方法で結果は、差が105の時が最大で、規格品のうち正答数は10枚で、規格外品のうち正答数は21枚だった。正答率は72パーセントであった。

今回の選別済みのデータではサイズが大きいが規格外品とされていたりサイズが小さいが規格品とされているものがあった。このように、サイズだけではなく他のパラメータを必要としていることがわかる。なので外接矩形を使った方法では、サイズだけを見ているので正答率が35パーセントと低くなってしまったと考えられる。このような考え方から他のパラメータとして、形状を扱える方法を考えた。

その方法が、牡蠣の重心から輪郭まで8方向の距離を使った方法である。外接矩形を使った方法だと2方向だったのをこの方法では8方向の距離を扱っているので形状といったパラメータも見えてくると考えた。最大値と最小値の差が小さいということは、距離の変化が小さいということを表していると考えたからである。

6 おわりに

本研究では、牡蠣の画像から重心から輪郭までの距離といった特徴データを抽出することが出来た。さらに、正答率を上げるために扱う方向をさらに増やせばいいだろう。

ほかにも単純に距離ではなく変化量を扱うといった方法や、8方向のデータ全体を用いたデータ群で比較するといったことが考えられる。

また、本研究で得られた、データベースを教師データとして与えて学習させることも可能だろう。

AIを用いた牡蠣の選別機の製作

菅野 匠 (情報工学分野)

指導教員 大槻典行

1 はじめに

厚岸のブランド牡蠣である「かきえもん」はブランドを維持するため熟練した人が出荷基準を定め、目視で選別している。これでは特定の人のみしか判別できない。これらに対して選別を容易にする装置の開発が望まれている。そのため本研究室では牡蠣の画像から選別を行う装置を開発している。本研究は、装置にかかるコストを低くし、手軽に誰でも扱える装置を開発することを目的としている。

2 選別手法

選別手法に関しては、画像処理だけでは選別が困難であると推測される。理由としては画像処理だけでは単純に牡蠣のサイズや大きさ及び形状の情報しか得られないためである。またその判断基準が標準となる大きさや形状だけで決められている訳ではない。そこでAIを用いることでその解決を試みる。AIによって画像処理では捉えることのできない特徴を人間のように発見し、それを元に分類を行えると考えたからである。方法として、指標となる牡蠣のデータを元に学習モデルを作成し、これを元に牡蠣の分類を行う。

3 牡蠣選別装置の構成

牡蠣選別装置はバックライトを備えたテーブルとその上から牡蠣を撮影するカメラモジュール付きのマイコンを設置している。このマイコンに学習モデルを読み込ませることで牡蠣を撮影した際に基準を満たしているかを判別するようにしている。この時に使用する牡蠣は表面を撮影し、それを入力データとする。本研究では判別を行う上で学習にはCNN(畳み込みニューラルネットワーク)を利用した。これはニューラルネットワークの中でも特に画像処理に適しているため採用した。^[1]



図 1: 選別の流れ

4 実験と考察

実験は基準を満たす牡蠣と満たない牡蠣から作成した学習モデルを使って牡蠣の正解率を測った。学習モデルを作成するにあたり基準を満たす、満たさない牡蠣からそれぞれ10個程度ずつ選別する。これをOpenCVの機能で

データの水増しを行いデータの数をそれぞれ300個ずつに増やす。これをTensorFlowの機能でニューラルネットワークを利用し学習を行う。学習回数は学習画像のデータセットを今回はバッチサイズ32に従って約9のサブセットに分けた。このサブセットをそれぞれ10000回ずつ学習させた。その後作成した学習モデルを利用し、実際に牡蠣を撮影することでその個体の基準を調べた。表1は5つの学習モデルとそれに対するモデルの正解率と実際に牡蠣を撮影した際の結果からどの程度正しく識別できているかを示した。

表1: 各モデルと正解率とテストデータでの結果

学習モデル	学習データ	テストデータ
model1	68.12%	57.89%
model2	57.31%	47.29%
model3	54.39%	51.23%
model4	58.12%	49.11%
model5	82.13%	41.54%

実験結果から今回の手法では学習データでの判別もテストデータでの結果から精度は良いものとは言えない。精度が低い理由として学習させるデータの種類不足やクラスの分け方と与える画像データに問題があったと考えらる。当初は牡蠣は6つのクラスがあり、1クラスあたりのデータが足りないと考えそれを基準を満たす、満たさないのクラスに分けて学習を行った。これを各クラスに分け、更にそのクラスの中で良悪の判断を行えば良かったのではないかと考えた。もう一つの問題点に関しては与える画像データ自体が白黒に近いものであり、大きさと形でしか学習を行っていないのではないかと考える。これに関しては画像からより多くの特徴が取れるように画像を処理してデータとして与えることが解決に繋がると考える。

5 最後に

今回の実験により牡蠣の撮影から基準を満たしているかの判定を行うことはできた。しかし精度はあまりよいものとは言えない。また考察でもあるように牡蠣自体のクラスや個数を増やし、より多くの学習を行うことで精度の向上に繋がると考えた。しかし、実際は牡蠣のクラスは有限ではなく、より多くの種類が存在すると考えるとこの手法では不向きではないかと思う。今後の課題として教師あり学習での精度向上とパターン数がわからない場合の対処法が最大の難点になると考える。

参考文献

[1] TensorFlow Core <https://www.tensorflow.org/guide/keras?hl=ja>

カードを用いたプログラミング学習 読み取り精度の向上

木澤涼花 (情報工学科)

指導教員 大槻典行

1 はじめに

2016年度よりQRコードが印刷されたカードを並べるだけでプログラミング学習を行える環境が開発されてきた。この教材で子どもたちがカードを並べ、指導者がタブレットなどでそれを撮影し、送信するだけで、2輪走行ロボットが命令を実行し駆動できる。つまり、PCを操作せずにパズル感覚でプログラミングに触れることが可能である。しかし昨年度は写真撮影における位置合わせやカードの並べ方を正確にする必要があるという問題があった。そこで多少のズレを許容し、実用しやすいようにする。

2 研究概要

昨年度までの本環境の開発では、画面左上に画像を合わせ、一定の回数のみ画像分割をする手法が考案された。劇的な処理時間の高速化に繋がったが、カードの配置と撮影方法が制限が多く実用的ではなかった。昨年度は、最も左上にあるQRコードを起点にそれ以外のQRコードを読み取るとい手法が考案された。結果、読み取り時間の高速化が達成された。しかし、画面左上にカードを揃える必要が無くなったが、カードの並べ方を正確にする必要があった。本研究は、撮影した画像中のQRコードに多少のズレや歪みが含まれていたとしても、ユーザーにとって実用的な速度で読み取ることができる手法を提案する。

3 QRコードを含む画像への分割

これまでの研究では、QRコードを画像の中から検出し、これをデコードする手法をとってきた。そのため処理に時間がかかっていた。QRコードは元々破損や、回転に強い2次元コードである。なので、1枚の画像に1つのQRコードを含むように画像を分割できれば、安定して正確に読み込むことができる。そこで1枚の画像に1つのQRコードを含む画像に領域を分割できれば正確かつ処理速度が早くなることが期待できる。本手法はQRコードの切り出しシンボルを利用しQRコードが存在する領域に分割する。本研究の切り出しシンボルの特定にはテンプレートマッチングを用いる。用意した切り出しシンボルのテ

ンプレート画像と、撮影した複数のQRコードを含む画像で最も類似する箇所を探索する処理である。今回、テンプレートマッチングには相関係数を利用した方法を用いた。最も良いマッチング結果は大域的最大値で決定した。大域的最大値とは全体で最大となっている値である。大域的最大値で評価された類似度の高い周辺の領域を塗りつぶし、カードの枚数分繰り返した。そして塗りつぶした範囲を画像分割することによって、昨年度までの不自由な配置ルールに縛られない読み取りができる。

4 実験

考案した手法の有効性を確かめるため実験を行った。複数枚のカードを間隔を開けて配置するパターンやカードを回転させて配置するパターンなどで実際の使用場面を想定して撮影したPNGファイルまたはJPEGファイルを用いて、WebブラウザはGoogle Chrome(バージョン87.0.4280.141)を使用した。数種類の並べ方を実験した所、間隔を取る程度では順調に画像を分割した。しかし大きく回転した場合や、カメラとカードの距離によって分割できないことがある。改善策としては、撮影した画像のサイズに合わせて用意したテンプレートを変更する等が考えられる。



図1: 領域の分割に成功した例

おわりに

カードを用いたプログラミング学習を行うにあたって、画像分割は精度の向上させることができた。しかし処理速度が遅いこと、カメラとカードの距離によっては精度が落ちること、カードの枚数を指定する必要があることなどが今後の課題である。

カードを用いたプログラミング学習 ～学習環境の構築～

木村 紹 (情報工学科)

指導教員 指導教員 大槻 典行

1 はじめに

QRコードの付いたカードを並べてタブレット端末などで写真を撮りプログラミングを行える環境の開発が2016年度より行われてきた。この環境はPCの操作に慣れていなくてもプログラミングの学習を行えるという特徴がある。

2 研究概要

前年度に開発されたシステムでは、カードのサイズを揃え、カードを丁寧に並べて写真を撮ることによって解析するQRコード同士の距離が一定になるため、QRコード同士の距離によってQRコードの位置を把握していた。しかしこの手法ではカードを丁寧に並べなければいけないという制約があり、実際この環境を使用するのが小さな子供や几帳面ではない人だった場合に、不便な制約と言える。そのため、カードを並べた際に、カードが斜めになっていたり丁寧に並んでいないという状況でも誤差を許してQRコードの解析を行える環境にしたい。そのため、写真の中からQRコードの存在する範囲を判別し、その範囲を用いて矩形を抽出してQRコードの解析を行っていくシステムを作成する。本研究では特に、写真から矩形を抽出しQRコードを解析して、プログラミングに直すための中間言語までの解析を行うシステムの開発を行う。

3 QRコードの読み取り手法

前年度のQRコードを読み取る手法は、QRコードが均一に並んでいることを前提に、一枚の画像からQRコード同士の距離を調べ、他のQRコードも同じ間隔で順次検出していき、プログラムとして出力していた。本研究で用いた手法は、一枚の写真の中から、矩形ごとに抽出されたQRコードの位置によってプログラムのコードを出力するシステムを作成した。矩形の座標をもとにQRコードの解析結果を、上に位置する結果から順に出力する。その際、矩形の右側にまた別のQRコードの矩形が存在していれば、先にそちらの結果を出力する。矩形領域ごとに解析を行うことによって、QRコードの大きさや位置、角度にずれが生じてもQRコードを解析し出力することを可能にした。また、出力結果を並べる際に考慮

するQRコードの高さと幅のずれは、矩形領域の大きさの半分までの範囲とした。

4 実装した手法

タブレット端末などで写真を撮る作業でプログラミング学習を行っていくため、このシステムをWEBブラウザ上で動作させたい。そのため、JavaScriptを用いてブラウザ上で動作する環境を実装した。JavaScriptでは写真上のQRコードを一度に解析できるようなライブラリは確認できなかった。そのため、今回のように写真からQRコードを判別し矩形を抽出するという手法を採った。

5 システムの実行結果

前年度までに用いられていたカードを並べた写真を用意し、作成したシステムで解析する。



図1: 実際に用いた写真



canvasOutput
読み取り結果の表示 if_which 280 60
if_right 20 290
go 20 560
f 290 560
2 0 850
b 10 40

図2: 出力結果

図1は実際に解析に用いた写真、図2は作成したシステムの出力画面である。図2のように解析した結果が、QRコードが存在する範囲の矩形の左上端の座標と一緒に出力されている。しかし、右側に位置するカードが、左に位置するカードの中間にいると、エラーではなく順番を変えて解析してしまう事があった。そのため、想定外の動きを起こさないために、読み取り範囲を厳格にする必要がある。

6 おわりに

本研究では、写真の中に存在するQRコードの位置が判明すれば、全て解析することが可能なシステムの開発に成功した。このシステムに、QRコードの存在する範囲を判別できるシステムと前年度までのシステムを組み込むことで実用的なシステムが実現可能だと言える。

自動作曲システム開発のための 楽曲の印象調査と分析

淋代晴菜 (情報工学分野)

指導教員 鈴木未央

1 はじめに

自動作曲を利用する目的の1つとして、「作曲技能を持つユーザが他者の要望に合った楽曲を手間をかけずに得る」が挙げられる[1]。本研究では、上田らの手法[2]を応用し、ユーザの入力した印象語に応じて作曲家の作風を継承した楽曲を生成する作曲システムを検討する。上田らの手法ではコード進行・リズム・音の長さ・音符同士の高低差が同じなら印象は変わらないことを前提に作曲しているが、作曲された楽曲の印象が本当に変わらないのかは実際には検証されていない。そこで本研究では、作曲システム検討のために、コード進行・リズム・音の長さを固定し、音符同士の高低差の関係を極力変えずに作曲した場合、楽曲の印象に変化が見られるかどうかを調査する。

2 調査手法

3章で述べる手順で生成した楽曲と、その元になった楽曲の印象をSD(Semantic Differential)法で評価することで調査する。そして、評価の平均値を元に、生成された楽曲と元になった楽曲の印象を比較、考察する。尺度は5段階、評価に用いる形容詞対は先行研究[2]から流用し、データの集計にはgoogle フォームで作成したWebアンケートを用いる。

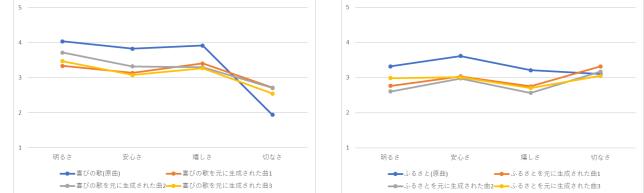
3 調査用楽曲の作曲手順

元となる楽曲はmidiファイルとして用意する。文献[2]を参考に、1音目は元となる楽曲におけるコードの構成音からランダムに選択し、2音目以降は元となる楽曲における音符同士の高低差から決定する。ここで、コード進行・リズム・音の長さ・音符同士の高低差は元の楽曲のものをそのまま流用する。なお、元となる楽曲のコード進行は判明しているものとする。

4 調査結果と考察

本校の1年生から専攻科2年生までの男女81名を対象に、喜びの歌とふるさとについて、原曲とそれをもとに生成したメロディを提示し、明るさ、安心さ、嬉しさ、切なさの各項目に、「とても当てはまる」を5、

「全く当てはまらない」を1とした尺度で評価させたところ、各項目の評価の平均値は図1のようになった。



(a) 喜びの歌をもとにした場合 (b) ふるさとをもとにした場合

図1: 各項目の評価の平均値

ある項目において、原曲での平均値が中点の値である3を上回る(下回る)場合、原曲をもとに生成したメロディでの平均値が同様に3を上回る(下回る)なら同じ印象を持ち、反対に3を下回る(上回る)なら逆の印象を持つと考えると、図(a)から、喜びの歌をもとにして生成されたメロディはいずれの項目についても原曲とおおむね同じ印象を持つと評価されたが、図(b)から、ふるさとにおいては、切なさ以外の項目について、生成されたメロディは原曲と逆の印象を持つと評価された、と考えることができる。しかし、項目ごとの評価を細かく見ると、ふるさとの原曲は安心さ以外の項目で回答者の評価が割れており、特に、明るさについては4と評価した回答者が31人と最も多かったことに対し、逆の印象と言える2と評価した回答者が21人とその次に多かった。このことから、図(b)のような結果が得られたのは、評価のばらつきが大きい楽曲であったためだと推察できる。評価のばらつきが大きいことは、この楽曲に対する印象の個人差が誤差として無視できないほど大きいことを示唆している。本研究の目的は原曲と生成された楽曲の印象に変化が見られるかどうかを調査することであるため、このような楽曲を原曲として用いるのは適当でなかったと考えられる。よって、今回用いた手法が、本研究で検討したい作曲システムに適しているかどうかは、予備実験により印象の評価に差が出にくい楽曲を事前に調べ、それを用いて再度実験を行った上で検証する必要があると思われる。

参考文献

- [1] 松原 正樹, 深山 覚, 奥村 健太, 寺村 佳子, 大村 英史, 橋田 光代, 北原 鉄朗:創作過程の分類に基づく自動音楽生成研究のサービス. コンピュータ ソフトウェア. 2013, 30, p.104
- [2] 上田 明頌, 西川 敬之, 福井 健一, 森山 甲一, 栗原 晃, 沼尾 正行:メロディモチーフを含む楽曲構造を考慮した自動作曲システム. 人工知能学会全国大会論文集 25, 1-4, 2011

映像の印象に適した音楽を推薦するシステムの開発

堀内 琉郁 (情報工学分野)

指導教員 鈴木 未央

1 はじめに

子供のなりたい職業として動画投稿者が挙げるほど動画投稿サイトが人気を博し、動画を投稿する者が増加している。動画投稿において映像編集と音声編集は動画の品質を高める重要なものでありながら時間と労力を要する。

本研究では、その内の音声編集に要する時間を短縮させることを目指し、映像の印象に適した音楽素材の選定を支援するシステムの開発を目的とする。

2 システムの概要

本研究で提案するシステムは、利用者が動画の中から音楽を流したい時点の画像を受け取り、その画像の代表色を決定する。決定した代表色から印象を推定し、あらかじめ印象ごとに分けられた音楽素材から印象が最も近いものをシステムが推薦する。

画像から受ける印象の要因は様々である。そのため、物体の色と表情情報をを利用して画像の印象を推定し、音楽を推薦する手法 [1] が提案されている。本研究で提案するシステムは人物や風景などの撮影された対象を限定しないために、その画像の代表的な色のみから印象を推定する。

3 代表色の決定方法

代表色の決定方法は人間の代表色知覚モデルに基づいた抽出法 [2] がある。この手法は人間が画像から瞬時に代表色を知覚する仕組みを領域分割と階層型クラスタリングを用いて再現したものである。しかし、複雑化を避けるために色の面積が考慮されていない。本研究では色の面積も代表色の決定の要素として扱う。

4 印象の推定方法

4.1 画像

色から印象を推定する方法として配色イメージチャート [3] を参考にする。文献 [3] には、印象語とその印象を与える配色が掲載されている。本研究では3色以下の組み合わせの配色を用いる。

3章で決定した代表色を最も類似している配色で用いられている色に変換し、3色以下の配色と比較する。一致した配色に対応する印象語を画像の印象とする。

4.2 音楽

音楽素材はフリーの音楽素材サイトで配布されているものを想定している。サイトの中には利用者の希望する音楽を検索しやすくするために「明るい」や「悲しい」といった印象を表す形容詞がタグ付けされていることがある。このタグとして利用されている形容詞を音楽素材の印象として扱う。

また、音楽素材の題名を word2vec を用いて印象語に対応させる。

5 音楽素材の選定方法

4章で用いた印象語と音楽素材のタグの類似度を word2vec を用いて求める。そして、類似度が高い順に一つの印象語に対して3つの音楽素材のタグを対応させる。これにより、画像や音楽素材の題名に当てはめた印象語から音楽素材のタグを決定する。

音楽素材ごとに付けられたタグと画像の印象から決定したタグを比較し、得点を付ける。タグが一致したときに加点し、面積が大きい色が使われた配色から決定したタグであるほど、加点を大きくする。これにより、色の面積を印象の推定に関連付ける。

得点が高い5曲を画像の印象に適した音楽素材としてシステムの利用者に推薦する。

6 今後の課題

動画の音声や音楽素材そのものから印象を推定することで精度を向上させる。

参考文献

- [1] 追木智明, 檻惇志, 宮崎純, 物体の色や表情情報を用いた画像の印象にあった音楽推薦の手法, 研究報告情報基礎とアクセス技術(IFAT), 2018-IFAT-132, 25, pp. 1-6, 2018
- [2] 高橋直己, 坂本隆, 加藤俊一, 知覚モデルとデータ分析に基づく画像からの代表色抽出及び評価実験, 日本色彩学会誌, 第42巻, 第4号, pp. 170-179, 2018
- [3] 南雲治嘉, 配色イメージチャート, グラフィック社, 2000

Raspberry Pi とウェブカメラを用いた ドライアイ監視システムの開発

葛西 彪斗 (情報工学分野)

指導教員 土江田 織枝

1 はじめに

近年、パソコンやスマートフォンなどの電子機器の発展により、ディスプレイを見る時間が増えている。そのため、ドライアイや目の疲れを訴える人の割合が増加している。そこで何らかの方法で瞬きをユーザに促せばドライアイの防止になると考えた。本研究では、瞬きと次の瞬きの時間情報を用いたドライアイ監視システムの開発を目的とする。

2 システムの概要

本システムは、Raspberry Pi 4 とウェブカメラでシステムを構築した。ウェブカメラはカメラ画像にユーザの顔全体が収まる位置に調整して設置する。

2.1 瞬きの判定

カメラ画像から瞬きの検出に必要な顔の部分を、OpenCV のカスケード分類器を使って抜き出す。その顔情報に対して、汎用クロスプラットフォームライブラリである Dlib を用いて顔のランドマークを検出することで、目の部分の検出を行う。

瞬きの判定には、EAR(Eye Aspect Ratio) という値を用いている [1]。EAR は、式 (1) で求めることができ、図 1 に示した p1 から p6 に対応している。目を閉じると、図 1 のように目の横と縦の比率が変わるので EAR の値が小さくなる。本システムでは EAR の値に閾値を指定することで瞬きを判定することとした。

$$\text{EAR} = \frac{\|p_2 - p_3\| + \|p_4 - p_5\|}{2\|p_1 - p_6\|} \quad (1)$$

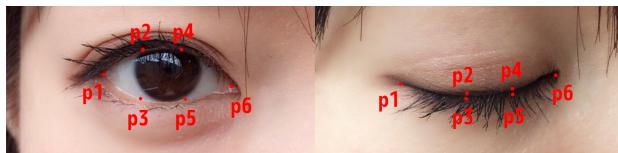


図 1: 瞬き判定に必要な各要素の位置

2.2 ユーザへの通知の基準

瞬きは、一般的には 1 分間に平均 20 回程度で、コンピュータなどを使用時は、1 分間で平均 5 回程度まで回数が減少すると言われている [2]。これらを元に、本システムでは 30 秒間で 10 回を下回った場合にユーザに通知を行う。

2.3 顔検出方法の改善

OpenCV のカスケード分類器を用いて顔の検出を行っていたが、検出領域が小刻みに震えることがあった。そこで、プログラム起動時にカスケード分類器で顔の矩形を取得した後に、その矩形を OpenCV の物体追跡アルゴリズムの一種である MedianFlow を用いて追跡する方法に変更した。これにより、検出領域が小刻みに震える現象を防ぐことができた。

3 評価実験

文章を閲覧した場合に 1 分間で実際の瞬き回数と本システムが検知した瞬きの回数を比較し、システムの性能の評価を行った。実験結果を表 1 に示す。表 1 から人によって瞬きの検出精度に大きく差があることがわかる。これは、目の開き具合に個人差があることが原因と考える。閾値の調整機能は用意していたが、プログラムの実行ごとに設定する必要があり、不便であることが今回の実験を通じて分かった。

表 1: 1 分間の実際の瞬き回数と検知した回数

	実際の回数	検知した回数
1 人目	13	16
2 人目	2	28
3 人目	12	12
4 人目	22	1
5 人目	13	14

4 まとめ

マスクの着用がない場合の瞬き検出、ユーザへの通知を行うことができた。しかし、検出精度に個人差がでたため、予め閾値を決めるのではなく、プログラム実行時に平均値を算出し、閾値を自動的に決めるなどの手段で瞬きを検出できるようにしていきたい。

参考文献

- [1] Eye blink detection with OpenCV, Python, and dlib, <https://www.pyimagesearch.com/2017/04/24/eye-blink-detection-opencv-python-dlib/> (参照 2020-10-16)
- [2] 目の健康を考える —参天製薬, <https://www.santen.co.jp/ja/healthcare/eye/eyecare/lecture/lecture05.jsp> (参照 2020-10-16)

必要な情報のみを保存する板書支援システムの開発

佐藤天音 (情報工学分野)

指導教員 土江田織枝

1 はじめに

現在でも黒板やホワイトボードによる板書で授業を行っている学校が多い。しかし、これらには板書内容を保存する機能がないので、消してしまうと再提示することはできない。本研究室では、通常のホワイトボードで板書した際に、板書内容の指定した部分だけを保存できるシステムの開発を行っている[1]。従来のシステムは、保存範囲を細かく指定できないため、必要なない部分についても保存してしまうなどの問題があった。そこで本研究では、保存範囲の指定方法の検討を行うとともに、保存した画像から必要な情報削減するシステムの開発を目的とする。

2 システムの概要

本システムでは、ホワイトボードを使用した板書の保存を対象とする。システム開発は Mac Air で行い、内蔵のカメラ (1280×720 ピクセル) で板書画像を取得した。また、カメラで撮影した画像に対しての保存処理や、保存範囲の指定などの画像処理は OpenCV により行い Python で実装した。保存範囲の指定には、直径 3 cm の赤色のマグネットを使った。図 1 は、横 1.2m、縦 1 m の小型のホワイトボードで本システムを使用している様子である。カメラ画像に板書全体がおさまる位置にパソコンをおくが、図 1 ではホワイトボードから約 1 m 離れた位置に設置している。

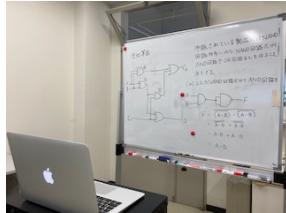


図 1: 本システムを使用中の様子

3 保存情報の指定

保存範囲の指定は、その範囲の対角となる 2 点の位置にマグネットを置くことで決定する(図 2)。その範囲内に不要な情報がある時には、さらにマグネットをもう一つ使

うことでの情報を削除範囲として指定できる。削除範囲は、保存範囲を示す 4 点の座標中の 2 点について、3 つ目のマグネットの位置の座標と近いものを選択し、それらの点を直線で繋ぐことで指定した。図 3 は削除範囲を指定した画像である。削除範囲はホワイトボードの色にできるだけ近い色で塗りつぶす処理を行った。

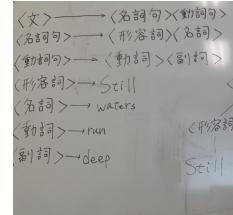


図 2: 保存した状態

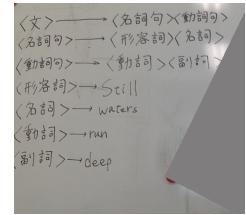


図 3: 削除した状態

4 削除範囲の指定の改善

3 章の削除範囲の指定方法では図 4 のように必要な部分も削除することがあった。そこで、マグネットを二つ使うことで、削除範囲を細かく指定できるようにした(図 5)。

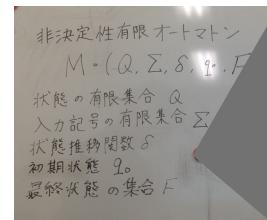


図 4: 単純な指定での状態

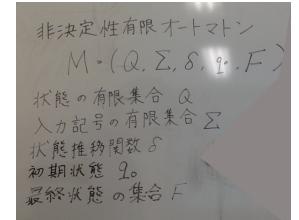


図 5: 正確に指定した状態

5 まとめ

板書内容の保存したい範囲の指定をマグネットを用いることで容易に行えるようになった。また保存する情報から必要なない部分の削除が可能になったことで、保存画像を使用するときに必要な内容だけを提示できるようになった。

参考文献

- [1] 松田智貴, "授業支援システムにおける板書内容の範囲指定の方法についての検討", 釧路高専卒業論文集, pp.26, 2018.

物体検出による野球ボールの軌道解析

松下晃矢 (情報工学分野)

指導教員 土江田織枝

1 はじめに

近年、スポーツの試合では審判の肉眼による競技の判定の難しい種目は機械による判定が取り入れられている。しかし、日本のプロ野球では数年前からビデオ判定が導入されたものの、機械による判定はほとんど進んでいない。野球中継中に紹介されるボールの軌跡の映像は、リアルタイムに行われているのではなく、映像を分析後に作成されたものである。野球中継の投球映像からボールを検出することで、ストライクとボールの判定を自動化でき、微妙な位置の投球に対しても正しい判定が行えるのではないかと考えた。本研究では物体検出アルゴリズムを使用して、野球ボールの軌道解析システムの開発を目的とする。

2 ボールの検出方法

2.1 SSD

SSD とは機械学習を用いた一般物体検知アルゴリズムのことで、学習時には画像のどこに何があるのかをアノテーションをつけて学習させる手法である。検出時には画像に 8723 個の大きさが異なる枠を設け、その枠の中に何があるのか、枠が物体からどれだけ離れており大きさが異なるのかを予測する。その結果から正しいと思われる枠を残し、重なり具合から枠が一つになるまで不要な枠を除去する。この作業を繰り返すことで最終的には、検出対象となる物体の枠だけが残ることになる。このアルゴリズムの利点は他の物体検知アルゴリズムより高速で比較的低画質でも高精度で検出を行える点である。これに対して欠点は、対象物体が小さいときの検出精度が低い点である。

2.2 学習データの検討

学習データでは、ボールの特徴がわかりやすいデータを用意する必要がある。改善前のデータでは中継映像をそのまま画像として使用していたため、ボールが小さいままうまく学習させることができなかった。そのため、ボールが大きく映る画像を使い学習データの改善を行った。

3 ボールの検出方法の検討

中継映像をそのまま解析するとボールが小さく、小さな物体の検出が苦手な SSD では検出することが困難であった。そのため、解析する画像を 16 分割した。分割したそれ

ぞれの画像で検出を行い、その後に画像を 1 つにまとめる方法 [1] で実験を行うことにした。更に、ロジンバッグなどを誤検出しないように、16 分割したときに上下の段の 8 枚を解析しないようにした。

4 実験結果と問題点

画像を 16 分割したデータを使い SSD によるボールの検出を行った(図 1(a))。その結果、正しくボールを検出しボールの軌道を描くことができた(図 2)。しかし、球場にある広告やユニフォームやシューズなどの他の物体もボールとして検出することがあった(図 1(b))。また、ユニフォームの背番号など白い場所とボールが重なるときに検出ができなかった。



(a) 正しく検出できた例 (b) 誤検出の例

図 1: SSD 実験結果



図 2: 軌道を描画した結果

おわりに

本研究では野球中継の投球映像を分割して SSD を用いてボールの検出を行い軌道の解析を行った。しかし、解析に 30FPS の一秒間の動画に対して 20 分程度の時間を要した。そのため、この方法ではリアルタイムでの検出は困難であることが分かった。今後の課題は、ボールの検出精度を上げるために学習データを増やすこと、そしてボールの検出処理を高速で行うためには SSD 以外のアルゴリズムを検討することが挙げられる。

参考文献

- [1] 物体検出プログラム (SSD) で小さい物体を検出するプログラムを作ったのでソースコードを解説します
”<https://tomomai.com/python-ssd-program/>”, (参照 2020-11-17)

セットプレーの練習補助を目的とした サッカーボールの軌跡抽出

相木宏太 (情報工学分野)

指導教員 林裕樹

1 はじめに

サッカーというスポーツにおいて、セットプレーは重要な役割を持っている。特に欧米や南米出身のような、背丈や体つきなどのフィジカル面で優位性を持つ選手が多い国やクラブチームにとって、得点に大きく繋がるチャンスになりやすい。逆に、アジア系のフィジカル面で大きく劣る選手が多い国やクラブチームにとって、セットプレーをいかに凌ぐかで失点を大きく減らすことが可能になる。このように、セットプレーは得点や失点に大きく関わってくる。野球やバスケットボールのように、点の取り合いになるような試合展開になることが少ないサッカーにおいて、1点の重みが他のスポーツより大きい。そのためセットプレーにおける情報を基にプレーを研究することは、サッカーの戦略的な面において重要である。そこで本研究では、固定カメラで撮影したサッカーの練習中の映像からボールを抽出し、その軌跡を表示することを目的とする。

2 ボールの抽出方法

映像中のボールを抽出するには、観測する画像と事前に取得しておいた背景の画像を比べ、背景の画像内に存在していないかった物体を抽出する背景差分で得た前景領域の輪郭を求め[1]、それらの輪郭の中でハフ変換によって円形の輪郭を持つものを選択する[2]。複数の円形の輪郭が選択された場合は、半径が最も小さい円を選択する。このように求めた円の中心を線で結ぶことで、ボールの軌跡を描画する。

3 ボールの軌跡抽出

正しくボールの軌跡を抽出できるかを確かめるため、ボールだけが映っている映像とボールを蹴る人物も一緒に映っている映像の2種類を用意した。



図 1: ボールだけが映った映像の軌跡抽出結果

図 1 はボールだけが映っている映像での実験結果であり、ボールは画面左から現れ、その後ゴールに向かっていく。図 1 では、ボールの軌跡がほぼ正しく抽出できているものの、ボールが奥の草むらに入ると軌跡が途切れてしまっている。これは、草むらに入った際に雑草にボールが隠蔽されることで、ボールの領域の形状を円として認識できなくなったためだと思われる。



図 2: 人物も映っている映像の軌跡抽出結果

図 2 は、画面下の中央にあるボールを、画面内の人物がゴールに蹴る様子に対する結果であり、ボールの軌跡が上手く抽出できていないのがわかる。これは、前景領域である人物付近に影や服の色の影響で円形とも判定されるような領域が発生し、それによりボール以外の領域を含める形で軌跡を抽出してしまったのが原因と考えられる。また、複数の円形の輪郭が選択された場合に半径が最も小さい円を選択するため、仮にボールの領域が抽出されていたとしても、ボール以外の円形領域が選択される可能性があるのも原因と考えられる。

おわりに

ボールのみが映った映像での軌跡抽出は成功したものの、ボールを蹴る人物も一緒に映っている映像での軌跡抽出は良好な結果を得られなかった。今後の課題として、円形の輪郭の新たな選別方法、ボール以外の円を棄却するためのノイズ除去や、ボールの色や模様の違いによる抽出結果のプレへの対応、ボールの検出が途切れてしまった場合の対処などが挙げられる。

参考文献

- [1] Hough 変換による画像からの直線や円の検出 <https://codezine.jp/article/detail/153>
- [2] 【Python/OpenCV】背景差分・フレーム差分で移動物体の検出 <https://algorithm.joho.info/programming/python/opencv-background-subtraction-py/>

OpenPoseを用いたトレーニング中の姿勢評価

石川 悠宇 (情報工学科)

指導教員 林 祐樹

1 はじめに

近年, 在宅時間の増加により, 自重トレーニングを行う人が増えている。自重トレーニングで高い効果を得るために、正しい姿勢を守ることが重要である一方、自分自身の姿勢を確認するのは難しく、他人に確認してもらう等の必要があり、他人との接触を避ける必要がある環境下ではハードルが高い。そこで本研究では、Zhe Cao らが発表した姿勢推定ライブラリである OpenPose[1] を用いて、自分で姿勢を確認できるシステムの開発を目的とする。

なお、このシステムでは対応姿勢をスクワットのみにし、評価試験を以下の 3 枚の画像に対して行った。



図 1: 基準



図 2: 深め



図 3: 浅め

2 姿勢の評価手法

OpenPose は、人物が写った画像を入力すると、解析結果の画像と、推定された 25 点の関節点の画像上の座標を出力する。このように画像から出力された関節の座標から姿勢を評価するためには、関節の角度が理想的な角度とどの程度異なるかを知る必要がある。そこで、下半身の沈み込みを見るためコサイン類似度を、膝の曲り具合を見るためベクトルのなす角を直接比較する方法を、それぞれ用いて姿勢を評価する。

コサイン類似度 [2] は、ベクトル同士の向きの近さを -1 から 1 で表し、 1 に近いほど類似していることになり、比較を行うベクトルをそれぞれ \vec{q}, \vec{d} とすると、

$$\cos(\vec{q}, \vec{d}) = \frac{\vec{q} \cdot \vec{d}}{\|\vec{q}\| \cdot \|\vec{d}\|} \quad (1)$$

で表される。本研究では画像左下から腰と左右の股関節、膝、足首の関節点までのベクトルを参考用と評価用の画像から求める。

膝の曲り具合については、画像の手前側にあり結果を確認しやすい右脚について、膝から股関節のベクトル \vec{X} と、膝から足首のベクトル \vec{Y} から右膝の角度を評価する。右膝の角度 θ は、内積の関係式より、

$$\theta = \arccos \frac{\vec{X} \cdot \vec{Y}}{\|\vec{X}\| \|\vec{Y}\|} \quad (2)$$

で求められる。

3 評価結果

表 1 に、コサイン類似度比較法と角度比較法の結果を示す。コサイン類似度は 7 点の関節点の平均を表し、角度比較法は基準画像での角度 (91°) との差を表している。これらの結果より、深めにスクワットを行った場合(図 2)の方が基準の姿勢(図 1)に近いという結果になった。

表 1: 評価の結果

姿勢	類似度	基準角度との差 [°]
深め	0.99909	17.21147
浅め	0.99896	-32.37018

4 おわりに

本研究では、自重トレーニングをより効果的に行うためのシステムを提案し、OpenPose による姿勢推定の結果から姿勢を評価できた。膝の角度を直接比較する方法は、画像の撮影位置に関係なく結果を求められるのに対し、コサイン類似度を用いる方法では、撮影位置によっては正しい結果が得られないという問題があるので、どう対処するかを今後の課題とした。

参考文献

- [1] Zhe Cao, et.al., OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17 July, 2019.
- [2] コサイン類似度について, "<https://www.cse.kyoto-su.ac.jp/g0846020/keywords/cosinSimilarity.html>"

撮影環境の変化に対して 頑健な点字ブロックの検出

小谷侑太朗 (情報工学科)

指導教員 林 裕樹

1 はじめに

視覚に障害がある歩行者の通行を補助するための点字ブロック部分を映像から検出するには、輝度を用いて二値化した画像から輪郭を抽出する。しかし、自然光の度合いや周辺環境、点字ブロック自体の劣化による色の変化によって検出できない場合もあり、高い精度での検出が求められるシステムでの利用は難しい。

そこで本研究では、機械学習を用いることで環境の変化に対して頑健で高い精度を持つ点字ブロック検出システムを開発することを目的とする。

2 画像の二値化と点字ブロックの判定

本システムは、まず動画をフレーム単位に分割し、各フレームを二値化することで点字ブロックの候補を抽出する。二値化処理では、自然光の変化によって点字ブロック部分の見え方も変化するため、安定して良好な結果を得ることが難しい。そこで本研究では複数のしきい値で画像を二値化し、すべての結果を候補とする。しきい値は、実際の画像にて良好な結果を示した四つの値を用いる。このように複数の値で二値化することで、従来よりも安定して分離することが可能となった。

二値化した画像から点字ブロック部分の候補の領域を切り出すために、OpenCV の `findContours()` を用いて連結領域の輪郭を求め、その領域を囲う最小の矩形を候補の領域とする [1][2]。このようにして求めた矩形群から点字ブロックだけを選択するため、機械学習によってこれらの領域が点字ブロックのものかを識別する。識別器は、実際の画像から得られた領域をラベル付けした画像データで学習する。

実験では、識別器に LinearSVC(LSVC) と RandomForestClassifier(RFC) の二つを用いて比較した。学習時は学習用のデータを 64×64 ピクセルにリサイズして正規化し、カラーの状態で一次元に展開した 12,288 次元ベクトルを用いた。表 1 に、学習データの 20% をランダムにテストデータとする交差検証を 10 回試行した際の平均の識別精度とシステム全体の実行時間を示す。

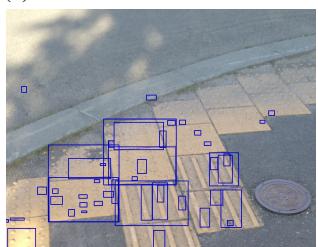
表 1: 交差検証での平均識別精度と平均実行時間

	識別精度 [%]	実行時間 [秒]
RFC	87.029	6.016
LSVC	75.652	1.795

表 1 より、識別精度は RFC の方が高かったものの、実行時間では LSVC の方が優れていた。また、LSVC で実際に画像から点字ブロックを検出してみると(図 1)，検出した領域は正しく点字ブロックになっていることから、速度が優先される場合は有用であると考えられる。



(a) 画像から輪郭を抽出したもの



(b) 点字ブロックと判断されたもの

図 1: 実行結果画像

おわりに

画像から輪郭として抽出された部分の中から、点字ブロックの部分を集中的に検出することが可能となった。しかし、まだすべての点字ブロックにかかる輪郭を正しいと判定できているわけではなく、取りこぼしてしまっている部分も多いため、学習データを更に強化していくとともに、RFC や LSVC だけではなく、ディープラーニングを使用し、さらなる精度の向上を目指したい。

参考文献

- [1] 北山直洋, Python で始める OpenCV4 プログラミング. 株式会社カットシステム, 2019.
- [2] クジラ飛行机, 杉山陽一, 遠藤俊輔, Python による AI・機械学習・深層学習アプリのつくり方. ソシム株式会社, 2020.

画像変形に対して頑健な電子透かしの開発

柴崎 俊伸 (情報工学科)

指導教員 林 裕樹

1 はじめに

電子透かしとは、著作権者の署名や ID といった情報を一見しても判別が難しい形で画像などに埋め込むものである。画像に対する電子透かしは、対象画像自体に情報を埋め込んでいるため、多くの情報量を埋め込むと大幅な画質の低下につながる。また、画像の圧縮や編集による透かし情報の喪失に対して耐性を持たせるためには、ある程度冗長な情報が必要になる。このように、電子透かしを埋め込んだ画像では、画質と電子透かしの耐性強度が互いにトレードオフの関係にあり、バランスを取ることが重要である。

そこで、本研究では、画質を考慮しつつ一定の耐性強度を持つ電子透かしを埋め込む手法を検討する。

2 電子透かしの埋め込み手法

透かしを埋め込む主要な手法として、ドットパターン法とビットプレーン法がある。ドットパターン法は、階調を特定のドットパターンに置き換えて 2 値化する際に、パターンを変調してテキストデータなどを埋め込む。しかし、画像の変形などに弱く、簡単に埋め込んだデータが消失してしまう。一方、ビットプレーン法は画素値の特定位置のビットからなるビットプレーンに透かし画像を混入する。しかし、特定のプレーンに挿入されたデータは容易に除去できてしまうので、本研究では一定のルールに従ってプレーンを変えながらデータを埋め込むことで、除去耐性を持たせる。

3 実験結果

図 2 の画像を透かし情報を埋め込んだ画像(図 1)と、図 1 を JPEG で圧縮した画像(図 3)、図 3 から透かし画像を抽出したもの(図 4)、図 1 の一部を黒く塗りつぶした画像(図 5)、図 5 から透かし画像を抽出したもの(図 6)を示す。

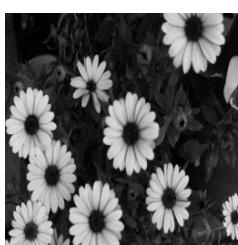


図 1: 透かしを埋め込んだ画像



図 2: 透かしとして埋め込んだ画像

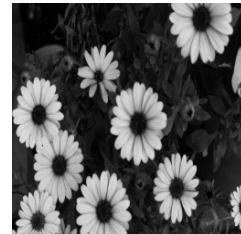


図 3: 圧縮した画像

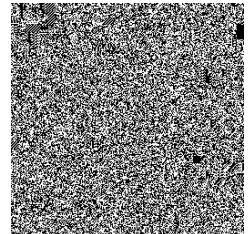


図 4: 抽出画像

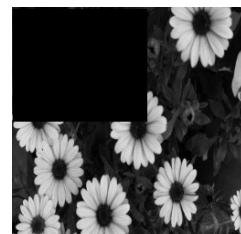


図 5: 一部を塗りつぶした画像

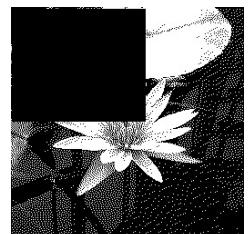


図 6: 抽出画像

4 今後の課題

本研究では、ビットプレーン法を用いることで、画質の低下もほぼ感じられず、塗りつぶし操作に対して一定の強度を有することがわかった。しかし、各画素の画素値に画像を埋め込んでいるので圧縮操作や拡大縮小などで座標が変化してしまう操作に対しての強度を有していないという課題がある。そこで、離散コサイン変換と量子化により情報を埋め込むという手法であれば、圧縮、拡大縮小操作に対しても耐性を持つことが可能になると考えられるので、今後の課題としたい。

参考文献

- [1] 柴田 且崇, 沢田 克敏, 中村栄治, DCT および VQ を用いた電子透かし. 愛知工業大学, 研究報告 第 47 号, 237–245, 2012.

ビックファイブ理論を用いた部活動レコメンデーションシステムの開発

石倉直樹 (情報工学分野)

指導教員 天元 宏

1 はじめに

現在、多くの学生が部や同好会に所属し、活動をしている。しかし、部や同好会に馴染めずやめてしまう学生も少なくない。また、新入生は部活体験期間に体験できる部や同好会の数には限界があり、自分にあった部活を探すのが困難な状況である。

そこで本研究では、ビッグファイブ理論を用いて、使用者の性格特徴に近い性格特徴を持つ部や同好会を推薦するシステムを開発する。

2 システムの概要

ビッグファイブ理論とは、1990年に Goldberg が提唱した、基本的な性格の次元が外向性、神経症傾向、勤勉性、協調性、開放性の 5 つの因子で構成されるという仮説である。

今回開発するシステムでは、日本語版 Ten Item Personality Inventory (以下、TIPI-J)[1] という 10 間の質問項目によって、ビッグファイブ理論の各因子の特徴量を 2 から 14 の範囲で求める。また、各因子の特徴量は正規分布に近い分布形状を示すため、各因子の特徴量の分散を考慮し、マハラノビス距離を使用する。

学習データは、TIPI-J の回答結果と、回答者の所属している部や同好会とする。入力は、TIPI-J の回答結果とする。出力は、部や同好会を性格特徴の近い順番に 10 個まで表示することとする。

システム全体の処理の流れは以下のようになる。

1. TIPI-J の回答結果から TIPI-J にある採点方法を用いて、各因子の特徴量を求める。各因子の特徴量をまとめて、性格特徴量と呼ぶ。
2. 各部や同好会ごとの、各因子の特徴量の平均と分散の値を求める。
3. システムの使用者の性格特徴量と各部や同好会の性格特徴量の平均と分散を使い、使用者の性格特徴と各部の性格特徴との近さを調べ、性格の近い順番に部や同好会を順位付けし、10 位まで出力する。

2.1 システムの検証方法と結果

このシステムの検証は、Leave-One-Out 法および使用者の評価によって検証した。

Leave-One-Out 法では、部や同好会に所属している釧路高専の学生に対して、アンケートを実施して集めた 265 個

の学習データを用いて、所属している部や同好会が何位にランクインするかを調べた。

表 1 より、約 42% は所属している部活動が 10 位以内に出力されることがわかる。

表 1: 順位に関する統計情報

部活動数	42	1 位	3.77%
平均値 (位)	14.68	5 位以内	21.89%
中央値 (位)	13	10 位以内	41.89%
最頻値 (位)	7	21 位以内	73.96%

使用者の評価では、釧路高専の学生 252 人に、開発したシステムを WEB アプリケーションで使用してもらい、「システムの出力結果に満足したかどうか」と、「出力された部や同好会に興味を持ったか」について回答してもらった。回答数は、それぞれ 228 通と 224 通である。

図 1 および 図 2 より、約半数は出力結果に満足し、興味がある部活動が出力されたことがわかる。

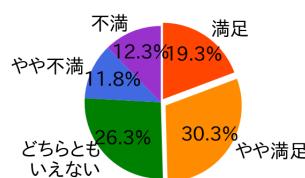


図 1: 出力結果に満足したか
出力されたか

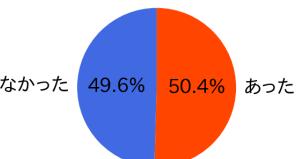


図 2: 興味がある部活動が
出力されたか

3 おわりに

本研究では、使用者の性格特徴に近い部活動を推薦するシステムを開発することができた。しかし、検証結果より、約半数しか満足できないシステムであったと考えられる。

したがって、今後の課題として、さらなる精度の向上が求められる。そのために、学習データを増やす、学習データが少ない部活動を推薦しない、ランキングアルゴリズムを使用する、などの工夫が必要である。

参考文献

- [1] 小塩真司、阿部晋吾、カトローニ ピノ、日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み. パーソナリティ研究, 第 21 卷, 第 1 号, 2012, 40-52.

CNNを利用した回帰の植物工場応用への挑戦

Yavang Phongkhang(情報工学分野)

指導教員 天元 宏

1 はじめに

画像認識の技術は、機械学習におけるディープラーニングが実用化されたことで、広く社会への応用が可能になった。それらの応用の中でも特にCNN(畳み込みニューラルネットワーク)を用いた識別は広く研究され、様々な対象に応用されている。しかし、CNNを用いた回帰についてはあまり研究されておらず、応用例も限られている。

そこで、本研究では、CNNを用いた回帰の可能性に着目する。特に本校に設置されている植物工場[1]で収穫された野菜の計測への応用を目的として、CNNを用いた回帰を可能にするシステムの構築及び評価実験を行う。

2 システムの概要

機械学習における識別とは学習データの分類モデルを構築することであり、未知パターンの所属するクラスの出力が可能となる。それに対し、回帰も教師あり学習データのモデルを構築することではあるが、未知パターンに対する連続値の予測が可能となる。

本研究ではそれを多層ニューラルネットワークからなるDNN(ディープニューラルネットワーク)を用いて実装する。更に、本研究では画像認識を対象とすることから、DNNの一部としてCNNを用いる。CNNは畳み込み層やプーリング層からなる画像処理用のニューラルネットワークで、学習により画像からの特徴抽出が可能となる。

本研究では、DNN構築のためのフレームワークとして、Python言語の深層学習ライブラリKerasを用いる。DNNの中間層は5層とし、それらの中間層はそれぞれ64個のユニットで構成する。

3 システムの検証方法

画像認識で回帰を行うためには、データとして画像と連続値の組が必要となる。本研究では植物工場で収穫されたリーフレタスの画像(図1)とサイズの50組に対して、 $k=10$ の k -分割交差法(k -Cross Validation Method)を用いて、実験を行った。

図1の様なスケールと対象物の関係性のモデルが構築できれば、様々な野菜にも応用可能と期待できる。



図1: リーフレタスの画像の例

4 実験結果

DNNの学習における誤差の変化を図2に示す。今回、平均絶対誤差(Mean Absolute Error)関数を利用して評価した。この値が0に近ければ高いほど精度が高いことを意味する。MAEの曲線は学習データに対する誤差を、VAL_MAEの曲線は検証データに対する誤差を表す。図2より、検証データに対する誤差も低下させることに成功していることが確認できる。

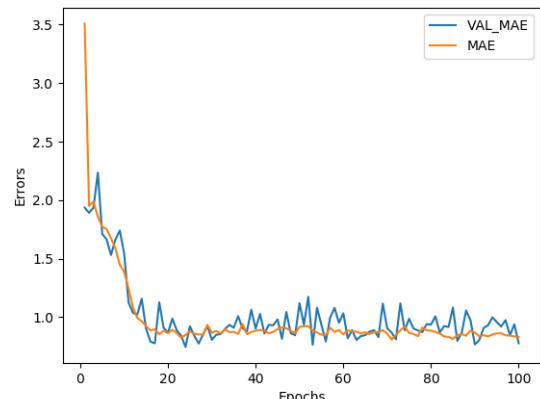


図2: k -分割交差法での誤差の変化

5 おわりに

DNNを利用して、 k -分割交差法(k -Cross Validation Method)で、画像を学習データとした回帰モデルを構築し、未知パターンの連続値を予測することができた。

今後はDNNのハイパーパラメータの最適化や転移学習の活用などを行い、実用的な画像回帰システムを目指す。

参考文献

- [1] 川村 淳浩, 天元 宏, 小林 豊, 植物工場向けエネルギー管理システム(FEMS)要素技術の開発. 第35回エネルギーシステム・経済・環境コンファレンス, 20–6, 2018.

リアルタイム情報を用いた満足度を最大にする観光ルート提供システムの開発

佐藤侑哉 (創造工学科 スマートメカニクスコース 情報工学分野)

指導教員 本間 宏利

1 はじめに

今日では、交通手段の発達により、観光がより身近なものとなっている。そのため、観光客に必要な情報を提供することのできる情報源の存在は、以前にも増して重要性を増しているといえる。そのため近年では、最適な観光ルートを提案するシステムが増加している。しかし、それらは予め用意されたルートを提案するのみであり、観光客の予算や所持時間、観光客の求める要素などに基づいて観光ルートを提供するシステムは、私の知る限りでは存在しない。また、悪天候や混雑などの旅行の際に起こる様々なアクシデントに適宜対応した観光ルートを提供するシステムはほぼ皆無である。

このような背景を基に本研究では、天気や混雑量などのリアルタイム情報を広範囲から取得することのできる通信網が整備された近未来を想定し、観光客の希望とリアルタイム情報を活用した最適な観光ルート提案システムの開発に挑戦した。

2 システムの概要

本研究では、時間帯や天候、混雑度などのリアルタイム情報を用いて、観光客が最も満足する観光ルートを推薦するシステムを開発した。今回は、観光地を接点とし、それらを結ぶ経路を辺として、節点ごとに「満足度」として定めた重みが最大となる経路を求める手法をとった。

各種「満足度」は、各観光地ごとのベースとなる魅力度の値と、条件ごとに定めた魅力度調整係数(天候、混雑度、観光客の趣味嗜好によって変動)を全て乗算したもの最終的な値とする。

本システムは以下に示したデータによって構成される。

- マップデータ(観光地間の移動時間)
- 各観光スポットの基本情報(ベースとなる魅力度や平均滞在時間、費用、魅力度調整係数)

システム使用時には、利用者が入力した条件とリアルタイム情報を基に、上で示した魅力度から今回の観光地ごとの満足度を求める。その後、決定した満足度を用いて、満足度が最も高い観光ルートとその到着時刻、費用を出力する。利用者が入力する条件を右上に示す。また、本システムが行うルート提供のイメージを、図1に示す。

- 出発地、目的地
- 出発時間、到着時間
- 希望する要素(自然、食、体験など)
- 人数、予算
- 訪れた観光地、エリア

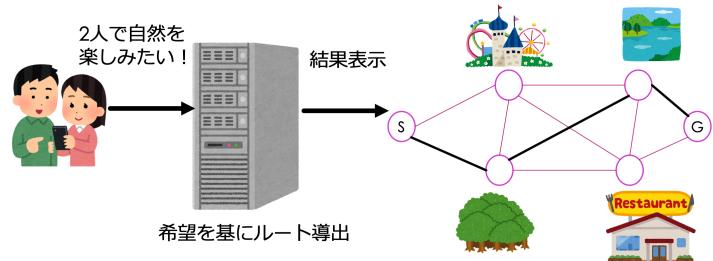


図1: ルート提供のイメージ

3 最適ルートの導出法

本研究では、最適なルートを導出するために、数理最適化と呼ばれる手法を用いた。これは、ある問題を数式群(数理モデル)として捉え、その中の目的関数の値を最大化する変数を求める手法である。今回のモデルの内容は、巡回セールスマン問題をモデルとしたものを基本とし、その上に追加の制約条件を加えたものとなった。モデルの実装にはpython3とそのモジュールであるpulpを、解の導出にはCBCソルバーを用いた[1]。

4 進歩、今後の課題

今回の実験により、リアルタイム情報と数理最適化を用いて、観光ルートを推薦することに成功した。

今後の課題としては、観光客に利用してもらうためのインターフェースの開発や、実際のリアルタイム情報の利用が挙げられる。最終的には、実際の観光地での運用が目標となる。

参考文献

- [1] 斎藤 努, データ分析ライブラリーを用いた最適化モデルの作り方, 近代科学社, 2018.

円弧グラフにおける迂回度最大要節点問題の効率的解法に関する研究

中島 哲 (情報工学科)

指導教員 本間 宏利

1 はじめに

グラフから節点 u を削除した時、最短経路長が増大する 2 節点が存在するとき、その節点 u を「要節点」と呼ぶ。コンピュータネットワークをグラフで表したとき、要節点に当たる端末が故障した場合、端末間の通信距離が延びることになる。そのため、要節点を見つけ、その頑健性を高めることはネットワークの安定性の向上に繋がる。

また、要節点が削除されたときに、最短距離が延びる度合を「迂回度」と呼ぶ。迂回度最大の要節点を導出することができれば、最も強度を高めるべき端末が分かり、低コストで安定したネットワークの構築に繋がる。^[1]

本研究の目的は、円弧グラフにおける迂回度最大要節点を導出する効率的なアルゴリズムを開発することである。

2 円弧モデル・グラフ

円弧モデルは、円上に並ぶ n 個の円弧の集合で表される。各円弧の左端、右端には昇順に整数値が割り振られる。また、各円弧は右端の大きさによって番号付けされる。図 1 は 12 個の円弧からなる円弧モデル M の例である。また、 M 上の円弧を節点とし、交差している円弧に対応する節点同士を辺で繋げたものを円弧グラフと呼ぶ。図 1 から導出された円弧グラフ G を図 2 に示す。図 2 のグラフにおいて、要節点は 1, 2, 5, 11 である。

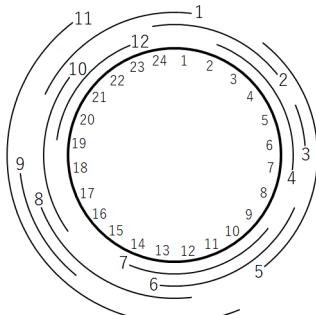


図 1: 円弧モデル M

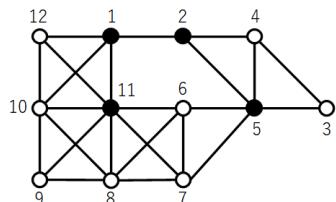


図 2: 円弧グラフ G

本研究では、円弧モデルを区間モデルに置き換えたモデルである拡張円弧モデル EM を利用する。拡張円弧モデルは、円の一部で展開し、区間モデルとしたものである。しかし、展開した部分では交差している情報が失われてしまうため、区間が左右にコピーされる。

3 提案アルゴリズム

Step 1: モデル M を拡張円弧モデル EM に変換

Step 2: 要節点 u が削除されたとき、最短経路長が増大する 2 節点の組を調べる。

Step 3: 最適なアルゴリズムを使用して [2], Step2 で求めた 2 節点の最短経路長を求める。その中で最大の最短経路数を迂回度とする

Step 4: Step2 から Step3 を全ての要節点に行う

Step 5: 迂回度最大要節点を導出する

このアルゴリズムを図 2 のグラフに用いると、表 1 の結果が得られる。よって、節点 5 が迂回度最大要節点となる。

表 1: アルゴリズムの実行結果

要節点	迂回度
1	2
2	2
5	3
11	1

4 結論

単純グラフで迂回度最大の要節点を求めるには $O(n^3)$ 時間掛かる。しかし、グラフを円弧グラフに限定し、上記のアルゴリズムを適用すると、 $O(n^2)$ 時間で計算することができる。

参考文献

- [1] An Algorithm for the Influential Hinge Vertex Problem on Interval Graphs, Hirotoshi Honma, Yoko Nakajima, Shigeru Masuyama, Journal of Information Processing, preprint, 2020.
- [2] F.R. Hsu et al. “Some optimal parallel algorithms on interval and circular-arc graphs,” J. Inf. Sci. Eng., vol.21, pp.627 – 642, 2005

日本語学習者のための漢字読解問題生成システムの開発

相坂 敏希(創造工学科 スマートメカニクスコース情報分野)

指導教員 本間 宏利

1 はじめに

近年、国内外の多くの日本語教育機関で日本語の学習支援が実施されている。多くの日本語学習者は漢字を学習するが、漢字は非常に複雑で初学者にとっては理解が難しいものである。

このような背景を基に、本研究では、日本語学習者にとって利便性が高い学習システムの開発を目指す。また、このシステムは教育者の問題作成による負担を軽減することが可能であるため、操作がより簡単なシステムの開発を目指す。

2 漢字読解問題とは

漢字読解問題とは、回答者が漢字の正しい「読み」や、ひらがなに対して適切な「漢字」を選択する問題である。漢字読み問題の例を図1、漢字書き問題の例を図2に示す。

問題:1
一日ごとに、(王子)は、お姫さまが好きになりました。
1: たまご
2: おうし
3: おうご
4: おうじ

図1: 漢字読み問題の例 (N4 レベル)

問題:1
そんな美しい声は、とても(ちじょう)の人間にはありません。
1: 池上
2: 地上
3: 地状
4: 地城

図2: 漢字書き問題の例 (N4 レベル)

図1、2のように、問題文に対して正解の選択肢が1つと不正解の選択肢が3つ用意される。このとき、不正解の選択肢を「錯乱肢」と呼ぶ。

3 問題の難易度について

図1、2は、日本語能力検定の出題内容に準じている。日本語能力検定の難易度は、高い順に N1, N2, N3, N4 とあ

る。本システムではユーザーが難易度選択し、問題に取り組むことを可能にした。また、図1、2の問題難易度は、それぞれ N4 に基づいたものであり、ユーザーが初級者レベルを選択した場合に生成された問題である。

難易度表は、表1のように設定した。

表1: 問題の難易度

ユーザーのレベル選択	問題の難易度
Hard Level (上級者)	N1
Normal Level (中級者)	N2, N3
Easy Level (初級者)	N4

4 処理過程と実装

問題生成の処理過程は以下の通りである。

1. 青空文庫から引用した、テキストデータを整形する。
2. 漢字、熟語データベースから、ユーザーが選択した難易度に応じた漢字をランダムに選出する。
3. 選出された漢字を含み、難易度に相応しくない漢字を含まない文章を抽出する。
4. 抽出された文章を相応しい形に整形し、問題文とする。
5. 作成した問題文に、正解の選択肢と錯乱肢を設けて表示する。

5 結論

ユーザーが選択した難易度に応じて、問題文を生成することが可能になった。この技術を活用することで、より多くの人間が手軽に日本語を学習できると推察する。本システムは、リアルタイムによる問題生成が可能になっている。また、生成された問題はストックとして保存されるため、現在 100 問以上のストックがある。最終目標は、本システムを WEB アプリケーション化することで、より多くの人間がこのシステムを用いて日本語を学習できるようになることである。

漢字読解問題のための錯乱肢生成

館下 涼 (情報工学科)

指導教員 中島 陽子

1 はじめに

今日、技術学習などを目的に外国人留学生が増加している。しかしながら、日本語を完全には理解できず、実際の技術学習に支障をきたしてしまう場合が見受けられる。

本研究はそれら外国人滞在者を対象にした漢字読解問題を生成する際、誤答となる選択肢（以下、錯乱肢と呼ぶ）を、回答者が正解と間違えやすい語群から抽出、問題の選択肢として設定する一連の動作をシステム化し、実装することを目的とする。

本研究では、生成する錯乱肢のレベルを対象者の言語レベルに合わせたものに設定することで、より効率的な言語学習を行うことを図る。言語学習を効率化させ、学習者のモチベーション維持を期待するものである。

2 研究の目的

外国人留学生など、日本語ではない母国語を持ち、日本語を理解する必要がある人々に、日本語習熟のために漢字の読み、書きを理解できるよう、漢字読解問題の錯乱肢作成システムを作成する。従来は教育担当者や出版社の担当員が問題一つ一つに選択肢を用意していた作業を解消したい。

3 システムの概要

錯乱肢の生成の際、回答者が間違えやすい語群から選択することが望ましい。そこで、錯乱肢の選定手法をパターンとして、正解にそれらの手法を当てはめ、錯乱肢を生成する。漢字の読みを答える問題の例では、正解と部首や意味が共通する漢字（読み1）、昼と夜、硬いと柔らかいといった特定の事象を表すことが共通する字（読み2）、あるいは、対象の漢字の違う読み（読み3）などである。読みから漢字を答える書き問題ならば、概形が似ている漢字（書き1）の他、意味として真逆か同じようなものを指す（書き2）、読みが同じだが形状、意味が全く違う漢字（書き3）が挙げられる。文中で示したパターンの一例を表1に示す。

表1: 選択パターンと錯乱肢の例

パターン	正解	錯乱肢
読み1	きり（霧）	つゆ（露）、しも（霜）
読み2	ひる（昼）	よる（夜）、あさ（朝）
読み3	なお（直す）	ただ（直ちに）、す（直ぐに）
書き1	受	愛、授
書き2	暑	熱、寒
書き3	発	髪、初

これらの錯乱肢を同一、あるいはランダムにパターンを用いて抽出、選択し、問題の錯乱肢に設定する。

実際に同一のパターンを使った問題を図1に、ランダムなパターンを使った問題を図2に示す。

釧路は年を通して霧が出る			
あられ	きり	つゆ	しづく

図1: 読み1の規則のみで作成した錯乱肢

例年通りのはつ詣は厳しいだろう			
初	髪	始	後

図2: 書き4、書き5の規則を併用した錯乱肢

また、問題作成時に学習対象者の言語レベルに応じて問題難易度をある程度操作する必要がある。例えば、日本語を全く理解できない場合は小学校低学年と同等に、日常会話がこなせるのならば、新聞記事や書籍に使われるような、高度であったり、難解な読みをする漢字を含ませるなど、利用者が滞りなく学習できる難易度を設定する必要がある。

4 まとめ

小学校6年間で学習する漢字を低学年、中学年、高学年に分ける形で難易度を調整し、漢字ごとにタグを付けることで、音訓の読み、部首から関連性のある漢字を抽出した。しかし、抽出した漢字の選択方法の指定などに課題が残っている他、漢字の候補も実装パターンが少なかったため、完成とは言い難い。

特許文書からの教師なし 数値データの抽出手法の開発

大橋英一郎 (情報工学分野)

指導教員 中島陽子

1 はじめに

近年の材料開発分野では、論文や特許文書などの技術文献を解析し、未知化合物の発見に役立つことが示されている。現在、マテリアルズ・インフォマティクスを活用したこの解析手法が注目されている。

技術文献中の各種実験パラメータや物性値に代表される数値データを活用することで、数値ベースの物性予測などが可能である [1][2]。また、技術文献から物理量および数値を取得することは、技術の重要度を図るなどの特許分析においても有用である。

特許文書は主に「課題を解決するための手段」、「発明の効果」、「実施例」、「要約」文から構成される。所望データは、実施例文中の発明の効果を表す語である「効果語」の周辺に存在する。そこで、効果語を用いた数値データの自動抽出手法に関する研究を行う。

本研究では、前段階として自然言語処理技術を用いて特許文書から「効果語」の自動抽出手法の開発を行う。

2 予備実験

数値データで評価される特許分野を調査する。具体的には、実施例文に「物理量、数値」 = 「%，98.4」が、効果語 = 「収率」を含む形かつ効果文にも同じ効果語が出現するか調査を行う。

調査の結果、冶金分野¹（国際特許分類コード:C23C）の特許分野に焦点を当てて研究を進める。

実施例文例：「合成及び蒸留精製を行ったところ目的物質 A の収率は **98.4%** で得られた」

発明の効果文例：「本手法で物質 A を高収率で得ることが可能である」

効果語例： 収率、原料コスト、容量増大、吸収量など

3 提案手法

はじめに、効果語候補となる複合語を発明の効果文と実施例文より形態素解析器 MeCab（辞書は Neologd を使用）を用いて抽出する。続いて、各文の複合語間で GENSIM ライブリ Word2vec を用いて分散表現にし、Word Mover's

Distance(WMD)²を用い類似度を求める [図 1]。学習モデルには Wikipedia の文書データを用いる。類似度の値が 0 に近いほど類似度の高いことを表す。類似度の値が 25 以下を効果語とする。

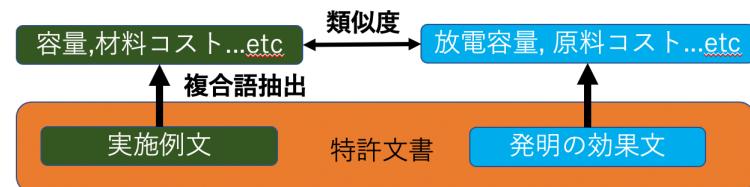


図 1: 特許文書からの効果語抽出手法

4 実験と結果

本実験は、冶金分野の特許文書 87 件を用いて行う。

実験の結果、実施例文と発明の効果文の効果語ペアの類似度上位 10 位以内を調べたところ、明らかに効果語と判断できる効果語ペアは 42.5% であった。明らかに効果語でないペアは材料名や手法名などの技術語であった。

おわりに

本研究では WMD の類似度を用いて、発明の効果文と実施例文で表現の異なる効果語も抽出を提案した。

効果語は抽出されていたが、技術語などのノイズも抽出されてしまっていた。そこで今後、ノイズとなる技術語を削除する事と学習モデルを適切なモデルに変更する事で精度の向上が期待できる。

また、効果語を抽出することができれば、効果語を用いた数値データの抽出が可能となる。

参考文献

- [1] Hirofumi Nonaka, Akio Kobayashi, Hiroki Sakaji, Yusuke Suzuki, Hiroyuki Sakai, Shigeru Masuyama, "Extraction of the Effect and the Technology Terms from a Patent Document," Journal of Japan Industrial Management Association, Vol.63, pp.105-111, 2012.
- [2] 作本猛、野中尋史、田中裕真、立花龍式、坂地泰紀、酒井浩之、小林暁雄、特許文書中の実験表現に関する属性定義と抽出方法の確立、NLP 若手の会 (Yans) 第 14 回シンポジウム, 2019.

¹冶金とは、鉱石から金属を精製したり、合金を製造することである。

²https://radimrehurek.com/gensim/auto_examples/tutorials/run_wmd.html

自然言語処理を用いた 介護コミュニケーション支援システムの開発

加藤 龍兵 (情報工学分野)

指導教員 中島 陽子

はじめに

現在、日本では高齢化が進行しており、介護施設や介護職員の数も増加しているが、介護職員の離職率は平均よりも高い傾向にある。

介護職員の離職原因の1つに、認知症に付随して発生する周辺症状への対応の困難さがある。本研究ではこれを解消するべく、介護事例に対して適切な対処を提示する介護コミュニケーション支援システムの構築を目指す。

研究内容

本研究の目的は、図1に示した介護コミュニケーション支援システムの構築である。これは、介護事例への対応方法を提示し、その評価をもとに性能を向上させる機能を持つ。そのため、様々な事例への対応が可能となる。

今年度は、介護事例の生データから学習データを作成し、これを用いて3つの手法でテキスト分類器を構築する。さらに、その性能を比較することによってテキスト分類器の構築手法の検討を行う。検討する手法は、Support Vector Machine (以下 SVM), Random Forest (以下 RF), BERT 言語モデル (以下 BERT) の3種である。

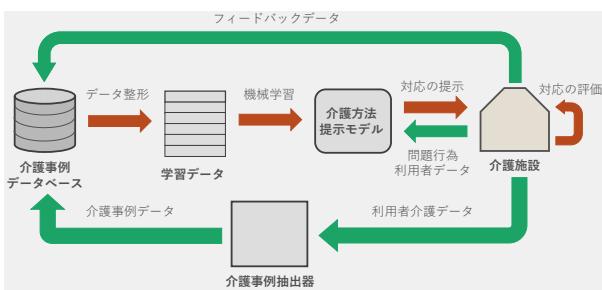


図1: 介護コミュニケーション支援システムの概略図

学習データの概要

学習データは、介護施設から提供いただいた介護事例データに対して、手動で要約とクラスラベル付与を行うことで作成した。なお、各クラスは徘徊や暴力といった

周辺症状の具体的な症状に相当する。また、データの総数は169個であり、クラスの総数は9個である。

構築手法の比較方法

構築方法の比較は、SVM, RF, BERT をそれぞれ用いてテキスト分類器を構築し、テストデータに対する正答率を求めることで行う。

テキスト分類器の構築は、SVM と RF では作成したモデルに k 分割交差検証を適用することで行い、BERT では東北大学の乾・鈴木研究室が公開している訓練済みモデルをファインチューニングすることで行う。また、SVM や RF を使用する際に必要となる文章の分散表現は n -gram と tf-idf を組み合わせて取得する。

実験結果と評価

各手法で構築したテキスト分類器のテストデータに対する正答率の平均値を表1に示す。なお、各手法ごとに試行は10回行い、結果は小数点第2位を四捨五入した。

この結果から、現時点でも最も優れたテキスト分類器の構築手法は SVM であることがわかった。しかし、SVMにおいても正答率は 58.4% に留まっている。これは、学習データの数が少なくクラスごとのデータ数の偏りが大きいという学習データの問題に起因すると考えられる。

表1: 各手法の正答率の平均値 [%]

SVM	RF	BERT
58.4	49.8	33.1

今後の課題

実験結果から、学習データ改善の必要性が浮き彫りとなった。これを達成するには、データの追加はもちろん、専門用語の辞書登録やストップワードの厳選といったデータの前処理の工夫も必要となると考えられる。

高次脳機能障害者向けスクリーニング支援 Web アプリケーションの開発

平澤慎 (情報工学科)

指導教員 中島陽子

1 研究の背景

高次脳機能障害とは、病気や事故による脳の損傷により、記憶、学習などの認知機能が損なわれている状態を指す。患者は、病院またはリハビリテーション施設において、患者の障害や回復の程度を判断するためのスクリーニング検査として、MMSE(Mini-Mental State Examination)やHDS-R(長谷川式簡易知能評価スケール)を利用して機能回復を目指す。双方に共通する「3 単語復唱テスト」では、一定期間後に 3 単語を復唱できるかを確認することで、短期記憶能力を測るものである。

リハビリテーションの効果を評価するため、検査は繰り返し実行される。しかし、検査を繰り返すことで 3 単語を長期記憶してしまい、検査の有効性が失われてしまう。しかし、検査の信頼性を欠くことを防ぐため、病院やセラピストは「3 単語復唱テスト」における 3 単語を変更することは許可されていない。よって、医療現場では、検査の有効性を担保した上で、3 単語の変更をサポートする支援ツールが求められている。

2 研究の目的

本研究は、自然言語処理技術を用いて、「3 単語復唱テスト」の問題点の克服をサポートするアプリケーションを開発することを目的とする。3 単語は、認知度が高いこと、音韻長が短いこと、単語間の連想度と類似度が低いことといった 3 つの規則を満たす必要がある。アプリケーションの機能として以下を提案する。

- 病院やセラピストの提案する 3 単語が、規則に従うかを判定する機能。
- 規則に従う 3 単語を自動的に生成し、病院やセラピストに提案する機能。

3 3 単語の選択規則とアプローチ

3 つの規則に従う 3 単語を判定・生成するため、自然言語処理の観点から以下のようなアプローチを行う。

認知度が高いこと

認知度の高い単語として、老若男女に知られる童話や童謡に登場する単語を用いる方法がある。しかし、先行研究によって、そのような手法では堅苦しい単語が多く抽出されることが問題点として挙げられる。よって、本研究では、小児向け図鑑に掲載されている単語を用いて、認知度の高い単語を取得することを試みる。

音韻長が短いこと

音韻長(モーラ数)とは、単語の拍数を意味する。例えば、「サル」はモーラ数 2、「チョコレート」はモーラ数 5 となる。捨て仮名(「や」「ゅ」「ょ」など)は、1 つ前の文字と合わせて 1 文字とカウントする。本研究では、規則を満たすモーラ数を 2 ~ 5 に定める。

単語間の関連性が低いこと

認知度や音韻長に問題がなくても、連想されやすい単語組が存在する。例えば、「冬」「こたつ」「猫」という単語組は、不思議と連想されやすいものである。これを避けるため、連想語辞典を用いて判定を行う。また、「船」「馬車」「飛行機」のような同じ分類に含まれるものは、類似度が高いと判断する。類似度の算出にはシソーラスを用いる。

4 研究の成果と今後の課題

研究の成果として、入力された 3 単語が「3 単語復唱テスト」に適切かどうかを判定する機能、「3 単語復唱テスト」に適切な 3 単語を生成する機能を持つ web アプリケーションを構築した。フロントエンドには Vue.JS を用い、各種入力のインターフェースを提供する。サーバーサイドでは Flask を用いて API を構築し、単語の判定・生成を行う。どちらも Docker を用いてコンテナ化し、Docker-compose で管理を行う。構築した web アプリケーションは、study.shindex.net:8080 で公開されている。

今後の課題として、連想語辞典の改良が挙げられる。現状では、連想されやすい 3 単語が生成されることがあり、検査の有効性を欠く恐れがある。より厳しい基準で作成された連想語辞典を用いて、更に信頼性の高いアプリケーションを構築する必要がある。

WebSocketを用いた リアルタイムウェブ環境の構築

坂上 光紀 (情報工学科)

指導教員 柳川和徳

1 はじめに

一台のパソコンをサーバ兼モニタとして、複数台のスマートフォンまたはタブレットをクライアント兼コントローラとして扱い、それらが協調動作するためのリアルタイムウェブ環境の構築について検討した。そして、アプリを手軽に開発するためのライブラリを作成した。

なお、クライアント／サーバ間の通信には WebSocket を利用する。

2 WebSocket とは

従来のウェブ通信では HTTP が多く利用されている。しかし HTTP では、通信を開いてから閉じるまでに一度のやり取りしかできない為、複数のデータを扱う際に手間(図 1 の青矢印の本数にあたる)が増える。

一方 WebSocket は、一度通信を開くと閉じるまで自由に複数のデータのやり取りが可能で、クライアント／サーバ間で双方向通信も可能である。そのため、WebSocket はよりリアルタイムな通信が可能となる。

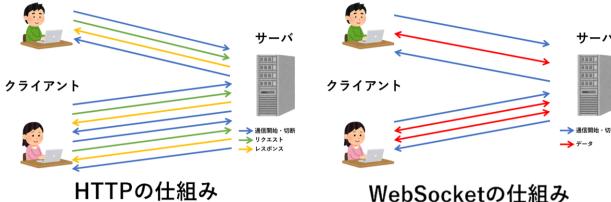


図 1: HTTP と WebSocket の違い

3 研究方法

はじめに、WebSocket を用いてリアルタイムウェブ環境で動くチャットアプリの作成を行い [1]、そのチャットアプリを基に、スマホ画面のボタンを押すと対応する文字列が PC に送られ、その文字列に応じて PC の画面を書き換えるような動作に変更した。また、これらの処理をライブラリ化し、ボタンを用いた三目並べと早押しゲームを作成した。

4 ファイル構成について

図 2 のように 3 つのファイルで構成されており、サーバファイルでサーバを立て、HTML ファイルで画面を構成し、スマホ画面のボタンが押されるなどの動作によりライブラリが呼び出され、対応する文字列を WebSocket を用いてサーバへ送り、サーバからアプリへ送られる。その後、

送られた文字列に応じてアプリ内で処理を行い、画面を書き換えるという流れになっている。

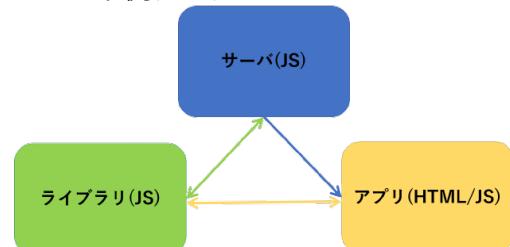


図 2: ファイル構成

5 研究成果

図 3 は作成した三目並べで、左から PC の画面、スマホの画面となっているが、ゲームの操作はほぼスマホのみで行うことができ、PC 画面が操作に応じてリアルタイムで変化するようになっている。なお、スマホは複数台接続が可能である。

作成したゲームは HTML ファイルのみを書き換えて作られているため、ライブラリやサーバ側のファイルは書き換える必要がほぼなく、HTML と JS を修得していれば簡単に別の作品に応用できる。

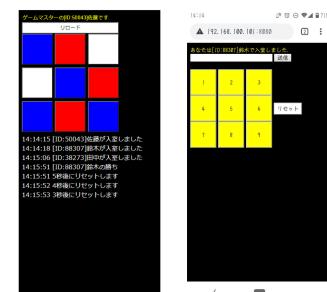


図 3: 三目並べの PC とスマホの画面

6 まとめ

リアルタイムウェブのライブラリを作成し、それを利用してチャットアプリと 2 つのゲームを作成した。

また、作成したゲームは HTML ファイルのみ書き換えることで、簡単に他の作品に応用が可能である。

参考文献

- [1] 斎藤 毅, “Node.js と Socket.IO でリアルタイムのチャットアプリを作ってみる”, <https://saitodev.co/article/Node.js と Socket.IO でリアルタイムのチャットアプリを作ってみる/>, 参照 Jul.2020.

板書映像の講師透明化システム

瀧上蒼人(情報工学科)

指導教員 柳川和徳

1 はじめに

今日多くの学校で感染対策での遠隔授業が行われている。遠隔授業の手段としてはウェブサイトを用いるほか、板書映像を用いる場合もある。

しかし、板書を用いた場合、講師が板書を遮ってしまう問題が発生する。対面の授業であれば学生が体を動かすなどして見ようとはできるが、カメラが固定となる遠隔授業ではそれは不可能である。見えないことを伝えるのも困難な場合があり、録画された映像ならばなおさらである。

そこで、板書の映像中の講師を透明化し、この問題を解決する。



図 1: 处理前



図 2: シルエット



図 3: 透明化

2 システムの概要

このシステムでは深度画像を背景用と最新の 2 枚使用して処理を行う。深度画像から時間的に最も距離が遠い部分だけを背景として登録し、背景よりも近い部分に透明化処理を行う。

また、既存の研究 [1] から講師を完全に透明化した映像よりもシルエット状態の映像のほうが情報が伝わりやすいという結果が出ているため、透明化部分の画素値を変更して色を付けて半透明にした。なお、本システムでは背景画像を更新していくため、最初に背景の撮影をする必要がない。さらに追加で無処理のカメラ映像とシルエットの映像と透明化した映像とを切り替えることができる機能を実装した。

3 研究成果

カメラから取得した映像(図 1)にシルエット処理を行うと図 2 のように人を半透明化し実際には遮蔽されても板書が見えるようになる。また図 3 のように完全な透明化も可能である。

4 まとめ

本研究では、板書映像での講師の遮蔽問題を解決するシステムの開発を行った。本システムを用いれば遠隔授業であっても講師による遮蔽を気にすることなく板書を見ることができるようになると期待される。

参考文献

- [1] 岡本隼, 山根恵和, 吉田光男, 岡部正幸, 梅村恭司 (2017) 講師のシルエットを透過表示した板書映像の生成とライブビューシステム. 日本教育工学会論文誌 41 (2) :177-186

POV-Rayにおけるヘアーの表現

竹田光 (情報工学科)

指導教員 柳川和徳

1 はじめに

POV-Ray とは、フリーのレンダリングソフトであり、ガラスなどの無機質な物体の質感の表現が得意である。しかし、動物の毛皮や植物の群生などの有機的な物体の表現は苦手としている。そこで本研究では、これらの有機的物体の一例としてヘアーの表現手法について模索する。特に、本研究では平面、円柱、球面でも均一にヘアーを群生させたためのマクロライブラリを作成した。

2 球面等分割

本研究にて力を入れたものは、動物の頭部などを模した球面上でムラなく植毛を行うことである。式(1)は三次元極座標を三次元直交座標で表現した関係式である。^[1] この式を利用した素朴な手法では、緯度 θ と経度 ϕ を等角度に分割し、ヘアーの群生位置を設定するため、球体の極で密に、反対に赤道では疎となりムラが発生する。これを抑えるべく、式(2)を用いて等間隔な分割を提案する。まず、式(2)の緯度方向の間隔 D_θ を経度方向の間隔の目標とする。しかし、緯度 θ の値が変わると、経度方向の円周の長さが変わってしまうので D_θ に完全に等しくは分割できないが、できるだけ揃えられるように、緯度の円周を D_θ で分割できる最大値 N_ϕ を用いて、経度方向の間隔 $D_\phi(\theta)$ を決められる。以上により、球体の等間隔分割を近似的に実現できると考えられる。

$$\begin{cases} x = r \sin \theta \sin \phi \\ y = r \cos \theta \\ z = r \sin \theta \cos \phi \end{cases} \quad (0 < \phi < 2\pi, 0 < \theta < \pi) \quad (1)$$

$$\begin{cases} D_\theta = \frac{\pi R}{N_\theta} \\ N_\phi = \text{int}\left(\frac{2\pi R \sin \theta}{D_\theta}\right) \\ D_\phi(\theta) = \frac{2\pi R}{N_\phi} \end{cases} \quad (2)$$

3 植毛マクロライブラリ

本手法の再利用性を確保するため、ヘアーの配置や生える方向を自動生成させるライブラリを作成した。この

ライブラリでは、群生させたい毛の位置や方向などの配置情報を規則的に自動生成し、配列に格納できる。そして、シーンファイルでは、自動生成された配置情報を利用して、多数の毛からなる毛皮モデルを簡単に定義することができます。図1は、植毛マクロライブラリを利用して、球体に等角度分割法でヘアーを群生させた画像である。

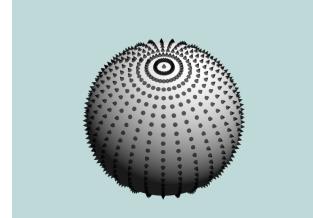


図1: 等角度分割法でヘアーを群生させた画像

4 実行結果

図2は、植毛マクロライブラリを利用して、球面に等角度分割でヘアーを群生させたものと、等間隔分割でヘアーを群生させたもので比較した画像である。

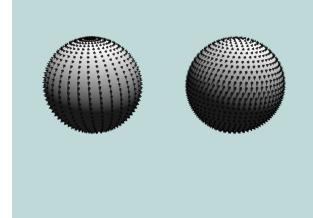


図2: 等角度分割法と等間隔分割法の比較画像

5 おわりに

現状として、球面上でのヘアーの群生のムラを抑えることができたよう見える。しかし、現地点では曲面の自由が効かないため、より有機的な物体を表現できる多角形ポリゴンでのヘアーの群生をしていきたい。

参考文献

- [1] 難波博之, 三次元極座標についての基本的な知識, 高校数学の美しい物語, <https://mathtrain.jp/rthetaphi>

ARを応用したヘッドトラッキング

渡部龍一(情報工学科)

指導教員 柳川和徳

1 はじめに

ヘッドトラッキングは運動立体視視 [1] による 3DCG を違和感なく再現するために重要な技術であるが、普通は特殊な機器を必要としている。そこで本研究では、これを手軽に実現するため ARToolKit [2] を応用する。

2 ヘッドトラッキングの実現方法

既存の研究 [3] ではカメラをモニタ側、マーカをユーザ側に置きヘッドトラッキングを実現していた。しかし、この配置では、カメラの視野外にユーザがいると、ユーザがモニタを見ることができたとしても立体視を実現することはできない。そこで本研究では既存の研究とは異なり、カメラをユーザ側、マーカをモニタ側に配置している。ARToolKit ではカメラから見たマーカの位置が分かるので、ユーザとカメラの移動に応じて 3DCG の仮想カメラを動かし、映像を違和感なく変えるプログラムを作成した。

3 実験結果

図 1 は機材の配置状況の写真である。図 2 と図 3 はこのプログラムを動かしたときの様子である。マーカを印刷しなくても使えるようにディスプレイに表示している。図 2 はモニタを上から見たときのカメラ映像で、ディスプレイの右上のマーカを認識しており、モニタにはしっかりと 3DCG を上から見た様子が映っている。図 3 はモニタを左から見たときのカメラ映像で、ディスプレイの左上のマーカと左下のマーカを認識しており、モニタにはしっかりと 3DCG を左から見た様子が映っている。このように、運動立体視は実現できている。しかし、画面にマーカを映したことで認識の安定性の精度が低くなってしまった。これは、ディスプレイに光が反射するためだと考えられる。手軽に使えるようにマーカを表示したが、マーカを印刷してディスプレイに貼る方が認識の精度が良かった。そこでマーカの表示/非表示を切り替え可能にした。

4 おわりに

ARToolKit を使いヘッドトラッキングはできた。今後の課題として認識精度の向上が必要である。

参考文献

- [1] Wikipedia, “立体視”, <https://ja.wikipedia.org/wiki/立体視#運動視差立体視>
- [2] 工学ナビ, “「攻殻機動隊」「電腦コイル」の世界を実現！ - ARToolKit を使った拡張現実感プログラミング”, <http://kougaku-navi.net/ARToolKit/>
- [3] 道後千尋ら, “視点追跡による立体視を用いたプロジェクト型 AR の提案”, FIT2015(第 14 回情報科学技術フォーラム), K-001, 2015



図 1: 機材の構成

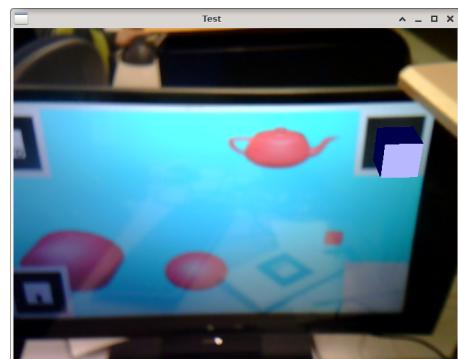


図 2: プログラムの実行例 1

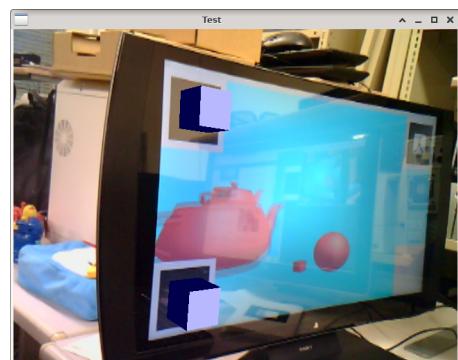


図 3: プログラムの実行例 2

シナリオ型チャットボットにおける 辞書の構築に関する研究

菊地 風汰 (情報工学科)

指導教員 高橋 晃

1 はじめに

現在チャットボットは、大きく分けシナリオ(ルールベース)型と機械学習型が存在する。

シナリオ型は、予め人の手によって一問一答の形式で構築されたシナリオに沿い対話をするものであり、構築は簡単なもの複雑な対応は難しいという欠点がある。

機械学習型は、予め様々なデータを学習させ、表現の多少の揺れを読み取ることもある程度可能になっているが、シナリオ型に比べ構築難度は高い。

そこで、本研究では構築が簡単なシナリオ型の長所に着目し、より柔軟な回答が可能とする辞書の開発を目的とする。

2 チャットボットについて

シナリオ型のチャットボットというのは、前述の通り予め一問一答のシナリオが用意されている必要がある。

質問を形態素解析し重要な単語を読み取る必要があるが、表現の揺れがあった場合に正しく判断できない場合があり、対応しようとした場合膨大な作業量となる。

そこで、表現の揺れに対応し意図を汲み取れるように別途辞書を作成することが、本研究の目的である。

2.1 チャットボット用辞書

従来の環境では、シナリオ型はチャットボット本体と応対のためのシナリオのみが存在した、しかしそれでは表現の揺れに対応できないため本研究では、シナリオとは別途の類義語等を示した辞書を用いる。

そのために日本語 WordNet[1] 用い類義語を検索し、別途辞書に登録しチャットボット使用時に類義語用辞書から検索し適用するためのシステムを作成する。

2.2 表現の揺れへの対応

表現の揺れについては例えば、「天気」、「空模様」など同じものを表しているにもかかわらずシナリオ型は機械学習を行わないため全く別物と判断し、意図したとおりに処理が行われない現象が発生する [2]。

他にも、「会議」、「ミーティング」等の日本語表記と英語表記の揺れなども存在する。

こういった現象を回避するために語彙とその単語に対し関連性を示した辞書の搭載は意味を持つ。

実際の辞書への記述は以下となる。

```
{  
    "word": "天気",  
    "synonyms": "空模様", "日和",  
    "similarity": 0.9, 0.89 ,  
}
```

このように、類義語及び類似度を json 形式で記述することによって、ある単語の入力があった時に柔軟な対応が可能となる。

3 成果及び今後の課題

辞書を利用するためのチャットボットの作成、シナリオの記述を行うシステムの作成が完了した。

例えば従来の辞書を用いない場合と用いた場合を比べると、

従来の例

入力 明日の空模様は?

出力 わかりません

辞書を利用した例

入力 明日の空模様は?

出力 明日の天気は晴れです

辞書がない場合、空模様と天気の表現の揺れに対応が不可能であったが、辞書を用いると意図した回答を出力することが可能となった。

しかし、表現の多少の揺れは対応可能なものの、一単語の枠に収まらない表現の揺れや、ユーザ独自の類義語の類似度の登録及び適用が不十分なため今後の課題としていきたい。

参考文献

[1] 日本語 WordNet, compling.hss.edu.sg/wnja/

[2] 表記揺れとは?11 の事例と対策, <https://xn-3kq3hlnz13dlw7bzic.jp/hyoki-yure-taisaku/>

耳を利用した非接触で行う認証システムの構築

志田優真(情報工学科)

指導教員 高橋晃

1 はじめに

生体認証とは、指紋や静脈など身体の一部やそれに準ずる要素を使って本人を特定する認証システムのことであり、スマートフォンなど様々な部分で使用されている。そんな生体認証の中に耳を用いる耳介認証というものがある。これは耳の穴の凹凸の個人差を利用して個人を識別するものである。しかし耳の画像を用いて識別を行うものは、過去の研究で認証率が高いという結果が出ているにもかかわらず耳画像による識別は一般的な技術として広まっているではない。コロナが世間を騒がせている今日、マスクを着用していても影響が少なく、認証率の高い耳画像認証に再び目を向けるべきだと感じた。また、研究当初よりデバイス等も進化しているため、より手軽に耳画像認証が行えるはずである。そのため本研究では、手軽に認証率の高い耳画像認証を実現するため Convolutional Neural Network(CNN) を用いた認証システムの構築を目指す。

2 研究概要

過去の研究では、耳の画像を2値化あるいは量子化し、その画像からモルフォロジースペクトルを抽出し、スペクトルの類似性を調べることで識別する手法などが用いられた[1]。しかし本研究では、より手軽に耳画像の識別を行うために CNN を用いる。CNN とは機械学習の一種であり、畳み込み層で画像の局所的な特徴を抽出し、続くブーリング層でその特徴をまとめることで入力画像の特徴を維持しつつ縮小処理が行える。すなわち抽象化が可能なため画像の分類に秀でている。

今回学習に用いた識別したい人の耳画像が図1である。まず、識別したい人の耳の画像を水増しさせたものと、その他2人の耳の画像を水増しさせたものをそれぞれ1000枚用意し、8割を学習に使用し残りの2割を検証に使用し実験を行った。与えられた学習用データを用いて学習し、検証データを識別する。これを指定したエポックの回数行うことでパラメータを学習させる。今回はエポック数を20に設定し実験を行った。実験の結果得られたモデルの正答率の推移を図2に示す。最終的には98.16%の確率で検証データを正しく識別できる学習モデルが完成した。

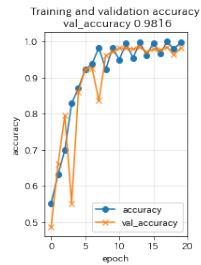


図1: 学習に用いた耳画像 図2: 正解率グラフ

次に、先ほど作成したモデルに学習させた3人の新たな耳画像を用意し、未知の画像でも正しく識別ができるかどうか検証を行った。結果を図3に示す。このように未知のデータを与えると98.81%というかなりの高確率で本人であるという結果を出すことができた。

```
./modelear/ear1.jpg -> [[2.1184766e-13 1.0000000e+00]] は 別人です  
./modelear/ear2.jpg -> [[9.9885988e-05 9.9990106e-01]] は 別人です  
./modelear/myear.jpg -> [[0.9881385 0.01186151]] は 本人です
```

図3: 未知データの識別

3 まとめ

サンプルとなる耳の画像を収集し識別したい本人とその他の人とで分類し、分類したそれらの画像を学習させ新たに未知の耳画像を与えると、その耳画像が本人かそれ以外かを識別するプログラムを作成。また、1枚の元となる画像に対して回転や拡大などの処理を施した画像を任意の枚数生成するプログラムも作成した。しかし現在は識別したい対象を1人しか選べないため、与えられた耳画像がAさんかBさんかそれ以外の人の耳かなどといった事例に対応できない。そのため、複数の人の耳データを学習させておき新しい画像を与えた際、その画像が学習している人のうちの誰なのかを判別できるようにすることを目指したい。

参考文献

- [1] 結城 義徳他, モルフォロジー演算に基づく耳介の特徴抽出および個人識別への応用 https://www.jstage.jst.go.jp/article/tvtr/18/27/18_KJ00001959626/_article/-char/ja/