

平成 30 年度
修士論文要旨集

高知大学大学院

総合人間自然科学研究科

理学専攻

応用理学コース 情報科学分野

清酒製造業における情報技術の活用に関する研究

応用理学コース

情報科学分野

江口大介

清酒製造業者にとって精度の高い清酒需要の予測，効率的な生産設備運用計画，市場に応じた清酒品種を選別し，酒造プロセスを管理することは，企業存続にとって不可欠である．より精度の高い清酒の需要予測ができれば，販売量の先取りができ，在庫費用の削減も可能になる．また生産設備，例えば酒造タンク，を100%稼働できる運用計画ができれば，生産効率を格段に上げることができる．さらに清酒品種の品質を醸造プロセスから判別して品種ごとの製造量を判定することや外的要因と醸造プロセスをコントロールする技術があれば，専門家による勘と経験に頼ることなく，より安定した製品提供が可能になる．

本研究において，1) 自己組織化マップ (SOM) を用いた清酒需要の予測技術，2) シミュレーテッドアニーリング (SA) 法を用いた酒造タンクのジョブスケジューリング技術，3) 自己組織化マップ (SOM) を用いた醸造プロセスからの清酒品種(品質)判定技術および 4) 人工ニューラルネット(ANN)を用いた醸造プロセスシミュレーション技術創出を試みる．

第1の清酒需要の予測法は，まず過去の月毎の需要動向を SOM で分類し，対象となる年度の数カ月の需要動向を同マップに分類することで残り月の需要を予測する方法である．過去8年間の需要動向を入力して評価したところ，月毎の予測平均誤差は0.114となり高精度に見積もれることがわかった．

第2の酒造タンクのジョブスケジューリング(JSP)最適化技術は，タンクの容量，用途などの属性と，使用スケジュール(順番)などをパラメータとして利用率を評価関数としてシミュレーテッドアニーリング(SA)法を用いるアルゴリズムを評価した．実験では，タンク数，ジョブ数が12以上でも，解の改善率が0.45以上となる良質な解を得ることが判明した．

第3の醸造プロセスから品質の推定法は，製品毎の発酵プロセスパラメータをパーセプトロンにより学習させ，結合の重み値を自己組織化マップ(SOM)で品質毎に分類させる方法を評価した．その結果，各品質(製品)の発酵プロセスパラメータの重み値から精度100%で推定できることが分かった．

第4の清酒品種(品質)のプロセスシミュレーション技術は，2層フィードフォワードニューラルネットワーク(FF)で清酒品種の発酵プロセスパラメータを学習させて発酵中のプロセスパラメータを再現する方法である．実験によれば推定精度が約70%となるシミュレーションが可能であることがわかった．

以上，本研究により，1) 清酒需要の予測，2)酒造タンクの利用効率化，3)醸造プロセスからの品質判定，さらに4)外部要因による醸造プロセスの劣化を早期に是正できる技術を提供できる．

本技術は，清酒製造業者が適切な販売計画を立案し，生産設備を効率的に運用し，需要に合う品種を確実に製造するために役立つものであり，情報化の遅れが目立つ清酒製造業に貢献すると確信している．

習慣化支援システムの評価実験コスト低減のためのモチベーション推移シミュレータ

応用理学コース 情報科学分野

中城 将太郎

本研究室では学習などの習慣化支援を目的としたシステムの開発を行っている。この開発中の習慣化支援システムは、複数人で1組のチームを作り、その中で互いにエールを送って励まし合うチーム機能を有し、そしてユーザにとってとっつきやすいものとするため、スマートフォンアプリとしての開発までが実装されている。しかし習慣化支援システムの評価実験を行うに当たっては、次の2つの理由から評価実験の規模が大きくなりすぎてしまうという問題がある。①ユーザが行動を変えてから定着に至ったとみなすには行動変容ステージモデルに基づく最低6ヶ月という長期的な実験期間を要することになる。②チーム機能の効果を比較により評価するためには大人数の被験者の協力が不可欠となる。さらに今後習慣化支援に関する仮説を立てるたびに、スマートフォンアプリ上で動作する支援機能として実装し、大規模な評価実験を繰り返すという検証方法は、非現実的で大変困難である。

そこで本研究では、評価実験を実施する前に仮説の検証を簡易的に行うためのシミュレータを提案し、設計・試作を行った。本シミュレータは、被験者らのモチベーションがどのように推移し、被験者らがどのように行動するかをシミュレーションするものである。シミュレータを用いて仮想的に実験を行えば、シミュレーションによって得られた結果から、被験者の行動の定着、または定着の失敗までの推移を観察し、予め原因の予測を立てることが可能である。それにより現実での評価実験を行うに当たって必要となる情報が整理でき、議論が不十分なまま打ち立てられた仮説を検証してしまうことを未然に防ぐことが期待できる。また、現実での検証を無視できるわけではないが、シミュレータによって評価実験における負担を減らすことができると思われる。そこで理想的な規模で行う評価実験を想定し、シミュレータを用いない場合と用いる場合の実験計画を立て、実験コストがどの程度低減できるかの見積もりを検討した。本論文ではその検討内容について報告する。

実際にシミュレータを実装するに当たって必要となるのは、被験者のモチベーション推移のモデル化である。モチベーションに関する理論研究は過去に数多く行われており、これらを大別すると、何によってモチベーションが動機付けられるかに着目した内容理論、モチベーションが動機付けられるまでの過程に着目した過程理論に分けられ、その2種類の中で様々な理論が展開されている。本研究では、その過程理論の中のある行動における動機付けの強さを定式化した期待理論に着目し、期待理論に基づいたモデルを用いてシミュレーションを行えるよう、シミュレータの設計を試みた。またこのシミュレータは、将来的に習慣化支援研究で何度も使われることが予想される。そのことから、ユーザビリティの高さや機能の拡張性が求められ、シミュレーションに使用するパラメータの追加やモデルの選択などが容易に行えるように設計する必要がある。この条件を満たすように設計したシミュレータの細微な仕様についても、本論文にて明記する。

シミュレータで用いるモチベーション推移モデルや習慣化支援に関する仮説のモデル化については、動作検証用の仮のモデルしか現状では作成できていない。従って、シミュレーション結果の妥当性の検証に関する議論は本論文の範疇を越えるため今後の課題である。

Greedy EM 法に基づく混合分布モデリングによる時空間オブジェクトの抽出と追跡

応用理学コース 情報科学分野

林 諒

近年のデータストレージの大容量化、計算速度の増加、インターネットなどの通信インフラストラクチャの普及によって、諸分野で大量のデータがアーカイブされるようになっている。また、深層学習をはじめとする機械学習や統計学習による大量データからの知識発見手法の研究の進展により、こうした手法が情報システムの処理の自動化や予測に活用されるようになっている。これらの大規模データには、センサネットワークなどの充実から、時間とともに変化する空間データも現れるようになってきている。特に、監視カメラや地球観測衛星の画像、気象レーダのデータなどはその典型例であり、その中には時間、空間ともに変化する対象物（オブジェクト、気象画像では雲塊にあたる）が含まれる。これらの時空間オブジェクトは時間とともに生成、融合、分裂などのイベントを経過して変化しながらやがて消滅する。こうしたデータはグリッド上のスカラー値として表現されることが多い。松永ら(2016)は連続的な分布を持つ2次元の観測点集合に対して、観測点のクラスターをオブジェクトとみなし、その集合を多変量正規分布の混合分布でモデル化して EM (Expectation and Maximization) 法と BIC による成分数決定によってパラメータ推定をすることによってオブジェクトの抽出を行なった。しかし、この手法をしきい値処理したグリッドスカラーデータに適用するとオブジェクト検出の際に不自然な融合や分裂が起こってしまった。また、追跡時には前の時間の解をラベルとともに簡易に引き継いで再推定していたため、識別されるイベントのタイプが限られ、その精度も悪いという問題があった。

本研究では、時間とともに変化する3次元グリッドスカラーデータも対象にいて、含まれるオブジェクトを自動的に抽出、追跡する過程の精度を向上させる手法を検討した。まず、グリッド上のスカラー値が混合分布の確率密度に比例するものと仮定し、この重みを取り入れるよう EM 法のパラメータ推定式を修正した。また Greedy EM 法(Vlassis and Likas, 2002)に基づく手法を導入してオブジェクトの追跡や発生、消滅の検知の精度を高める手法を開発した。Greedy EM 法は初期のオブジェクト数を1として、停止条件に到るまで逐次に増加させながらパラメータ導出を繰り返し、最適なオブジェクト数とモデルパラメータを求める手法である。本研究では、最初の時間フレームにおいては Greedy EM 法を高速化したものを使用し、その後の追跡過程では、この手法を変形して、前の時間の解を出発点として固定し、Greedy EM 法によって現時間でのオブジェクトの増加処理を、停止条件を満たすまで実施し、その後パラメータの更新を行い、逆のプロセスによる消滅成分の処理を実施するものとした。また移動、発生、分裂、結合、消滅といったイベントの詳細な判別を行うことで時間遷移中に発生する各イベントに対応したラベリングを実現した。

人工データを使用した実験より、この改良によってスカラー値のピークにオブジェクトの中心が正しくフィッティングすることで抽出される成分数とパラメータの精度が向上し、不自然な分裂や融合の発生を抑制できることを確認した。また導入した Greedy EM 法を用いるアプローチはオブジェクトの追跡において抽出、及び追跡の精度を向上させることを確認した。実データとしてこの手法をフェーズドアレイ気象レーダデータからの雨雲の抽出に適用し、その効果を確認した。

プライバシーを保護したビッグデータ再現法と その SOM 分析品質の研究

応用理学コース

情報科学分野

三戸理誠

近年、ビッグデータの分析は、ビジネスや新たな知見を得るために活用されている。一方、データ漏洩によるプライバシー侵害への懸念も大きい。そこで本研究は、個人情報を含まない統計データから、ビッグデータを再現することで安全にビッグデータ分析可能な環境を提供する方法を検討する。

本研究では、ビッグデータの統計データ（ヒストグラムと相関係数）から仮想的なビッグデータを再現する方法、およびそれらのビッグデータとしての品質、すなわち元のビッグデータ分析と同等の知見が得られるかどうかを研究する。

統計データからの再現手法として、まず、ヒストグラム（クラスごとの頻度）から再現する方法を考案した。まずヒストグラムの各クラスのクラス代表値を頻度数だけ一次元配列に並べ、最大範囲の乱数で一次元配列から値をランダムに取り出す（モンテカルロ法）。これを値とするデータ群によりヒストグラムを再現させる。もし1レコード1属性のデータならば、まずクラス頻度 K だけ K 匿名性をもち、さらにクラス幅分の曖昧性を持つ。もし1レコード M 属性ならば、さらに個人特定が困難になる。評価実験として、「とさでん交通バス乗客データ」を用いたところ、再現手法で平均値、最大、最小、標準偏差、分布関数（ヒストグラム）がほぼ元データと同等のデータを再現することがわかった。しかし、データ品質の評価として、ビッグデータ分析 (SOM) を比較したところ、両者が大きく異なる結果となり、ビッグデータとして品質が低いことがわかった。

ビッグデータの品質を向上するため、ヒストグラムに加え他の統計量（相関係数）も含む再現法を考案した。具体的に、前述の方法に加え、まず M 属性の 2 レコードの属性 a の値を交換し、次に属性間のすべての相関係数が元データの属性間の相関係数に近くなれば受理、遠ざかれば拒否（ペアワイズ交換法）する改善処理を一定回数繰り返すことで相関係数を再現させる方法を行う。評価実験として、「とさでん交通バス乗客データ」（レコード数約 3,600, 属性数約 9）を用いたところ、10 万回の改善処理で、相関係数が最大誤差 0.1 程度まで再現できることが分かった。さらにビッグデータ分析(SOM)比較では、ほぼ同等なマップが得られ、同様な知見も得られることが分かった。他のデータ評価として、過去分析した医療データのヒストグラムと相関係数(レコード数約 45,000, 属性数約 300)では、相関係数が最大誤差 0.05 で再現できることが分かった。また、ビッグデータ分析比較でも同等の知見が得られることが分かった。

本研究で提案する統計データのみから再現するビッグデータ再現法によるデータは、ビッグデータをプライバシー漏洩問題から解放し、安全に分析できる環境を提供すると思われる。本手法により多様なデータ分析法の研究推進や、ビッグデータ利用の活性化が期待できる。

手話トレーニングマシンの開発 – 表情判別システムの高精度化 –

応用理学コース

情報科学分野

藪中 奎介

本研究室では健聴者に対して手話学習の理想的な学習環境を提供することを目標とした、手話の学習を支援するシステムである「手話トレーニングシステム」の研究・開発を行っている。システムは学習者が行う指や手の動きなどの様々な情報を深度センサとグローブを用いて取得する。これをプログラムを用いて判別、学習者に対して正誤結果や動作に関するフィードバックを提供する。手話学習者はこのフィードバックによって正しい手話動作を効率的に学習することができる。手話トレーニングシステムは長らく手指動作のみを対象としたシステムであったが、2017年度の大栗氏により表情学習を行うための仕組みが追加された。これは機械学習を用いて学習者が行っている表情を顔の各部位の変化の程度に変換し、この程度の集合をファジィ推論により集約する。この集約された値を利用して学習者の表情の有無を調べ、学習者に対して適切なフィードバックを返すことを可能とした。表情は手話において非常に重要な要素の1つであり、これが追加されたことにより実用的な手話の学習が可能になった。

手指動作や非手指動作の程度は人により大きく異なり、この程度の違いを総称して個人差と呼ぶ。手話トレーニングシステムでは手話動作に適切なフィードバックを返す上で、個人差は誤認識を引き起こす大きな要因となってきた。手指動作については手指の動きを言葉により定義するファジィ推論を利用し曖昧さを許容することで、一定程度の個人差の吸収に成功している。しかし表情はいくつかの問題からまだ個人差の吸収には至っていない。表情において個人差は顔の各部位の動きの程度の違いとして表れるが、この程度は人により大きく異なる。現在の表情システムは機械学習とファジィ推論により構成されるが、もし機械学習のみでこの個人差を吸収するならば、教師データを作成する際にこれらの動きの程度全てを想定することが必要になる。このようなデータを得るためには多数の教師を必要とするため困難で、行ってもデータ作成作業が非常に煩雑となる。また現在ファジィ演算は顔の各部位の変化量を適切な優先度で統合することを目的に利用されており、そのままの構造では個人差の吸収には不向きである。

本研究はこれらの個人差により生じる学習のための表情判別システムに関する問題点を解決するため、機械学習により得られた顔の各部位の変化量を個人の動作に合わせて適切に補正する処理方法の導入を提案する。まずこの仕組みは、機械学習部、補正部、ファジィ推論部からなる。このうち、機械学習部とファジィ推論部は大栗氏の考案した表情判別システムを基礎としている。機械学習部は学習者の行った顔形状の変化から、表情に係る顔の各部位の形状変化量に変換する。補正部は機械学習により得られた顔の各部位の変化量に対する結果を、その個人から得た情報を元に機械学習の結果を適切に補正する箇所である。この箇所の導入により、個人差やセンサ精度の問題により生じた影響を吸収できることが見込まれる。ファジィ推論部は補正部によって補正された結果をファジィ推論を用いて統合し表情の程度に変換する。この時、表情におけるそれぞれの顔部品の影響度を加味して演算が行われる。

この処理方法の有用性を検証するための予備実験として、顔の眉間の変化量を補正しその結果精度がどのように変化するかを確認した。その結果、理想的に補正が行われた場合、有意に精度が向上することが分かった。この予備実験の結果を踏まえて、機械学習から得られた顔の各部位の程度の補正の計算を自動で行う仕組みを作成した。この提案手法について評価実験を行い、精度の変化を通して処理方法の有用性について検証を行った。