

令和 4 年 5 月 11 日現在

機関番号：33903

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K12073

研究課題名（和文）対話システムに名推理をさせるための誘導質問術に関する研究

研究課題名（英文）A study of a method to collect the properties of a person who to talk to from light conversation

研究代表者

菱田 隆彰（Hishida, Takaaki）

愛知工業大学・情報科学部・教授

研究者番号：30329627

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では大きく分けて2つのテーマについて着手した。一つは人工対話での雑談の内容から話し相手の属性を推定する技術である。対話内に含まれる歴史的な事象からその発生日時を変換するイベント履歴収集システムの提案、LINEやChatwork上などで使用可能な対話生成機構の構築を行った。もう一つは、コンピュータシステムと人間との対話において必要となる話し相手の周辺情報を入手する技術である。主に全天球カメラを2台利用した全周囲の距離推定を行うための基礎技術の検討を行った。

研究成果の学術的意義や社会的意義

本研究では、人工的な対話が可能なインターフェイスを試作し、対話を進めることで相手の属性の一部を推定することが可能であることを示した。また、話し相手の周辺の情報を収集するための距離推定手法の提案も行った。

これらの成果によって人工的な雑談型対話システムにおける機能性の向上が期待できる。意図的な質問をすることなく雑談を交わす中で話し相手の経歴や出自を含めた人間的な背景、現在の会話環境を知ることによって、様々なサービスにおいて相手をより満足させる提案が可能となる。

研究成果の概要（英文）：In this research, we started on two main themes. One is a technique for estimating the attributes of the talk partner from the contents of chats in artificial dialogue systems. We proposed an event history collection system that converts the date and time of occurrence from historical events contained in the chat dialogue, and constructed a chat generation mechanism that can be used on LINE and Chatwork. The other is a technique for obtaining peripheral information the talk partner, which is necessary for dialogue between a computer system and a human being. We mainly examined the basic technique for estimating the distance around the entire circumference using two omnidirectional cameras.

研究分野：教育工学

キーワード：雑談 人工対話 年齢推定 距離推定 人物分類 関係性分析 感情推定

1. 研究開始当初の背景

音声対話システムは最近様々な機器に装備されるようになった。例えば Apple の Siri や Softbank の pepper などから、サポートセンターなどの受付など様々な場面で遭遇するようになった。他方で人とのコミュニケーションを図るために雑談を行うシステムも徐々に実用的なものが現れ始めている。しかし、人工的な雑談対話システムを用いた対話は、実際に人と雑談をするようには上手く会話が進まないのが現状である。

その理由の一つは、対話システムが相手の興味や状況を察知して、適切な発話を行うことが未だ十分にできていないからである。我々はその問題を補う一つの方法として会話の流れを戦略的に捉え、相手の話す内容や話している状況から自然な形で話し相手の興味を推察できるような情報収集を進める手法の検討を行う。

2. 研究の目的

自然な雑談を続けるためには、話題に対する適切な受け答えが必要である。また雑談を長く続けるためには、様々な話題の提供ができなくてはならない。雑談を目的とするシステムとしては、多様な相手との対話が必要であるため、対話する相手によって好ましい話題を提供する必要がある。相手に合わせた話題を提供するためには、相手の興味を示すような様々な情報が必要になる。我々の研究の目的は「人工的な雑談対話型のシステムにおいて、話題をユーザごとにカスタマイズするための情報収集の困難さ」を解消することである。

3. 研究の方法

本研究では雑談型対話に有用な情報支援技術、対話戦術のモデル化手法、対話者情報の収集方法、対話エンジンとインターフェイスなどの確立とシステムへの実装を目指す。

大きく分けて2つのテーマについて行う。一つは人工対話での雑談の内容から対話者の属性を推定する技術である。対話内に含まれる歴史的な事象からその発生日時を変換するイベント履歴収集システムの提案、LINE や Chatwork 上などで使用可能な対話生成機構と人間関係図を生成システムの構築、対話中の文面から感情を推定するための辞書の拡張が可能な感情分析ライブラリの構築を行う。迂遠な対話の中に含まれる事象から対話者の年齢の推定やグループ内での人間関係の推定が可能であることを示す。

もう一つは、コンピュータシステムと人間との対話において必要となる対話者(人間)の周辺情報を入手する技術である。似た音声に対しての深層学習を用いた人物の識別の精度に関する研究、物体の位置と音声出力との協調に関する研究、全天球カメラを2台利用した全周囲の距離推定を行うための基礎技術の検討を行う。既存のカメラやマイク、スピーカを用いることでコンピュータの視覚や聴覚に高い認識力を持たせることが可能であることを示す。

4. 研究成果

(1) イベント履歴収集システムの構築

雑談型の対話システムにおいて、相手を楽しませながら自然な会話を行うためには、提供する話題を相手の興味を持つ話題を選ぶ必要がある。そのために、相手の個人的な情報を会話から手に入れる必要がある。既存研究によって会話の中に現れる特徴的なフレーズを個性として紐付けることで、相手の嗜好や状態を推測できることが示されているが、実際には個人を特徴付けるような情報を平易な会話の中から取得するのは困難な作業である。

そこで我々は、よりユーザ個人にあった話題を提供するために、雑談の内容に含まれる相手が体験した出来事の話から、その時期とその時の相手の年齢がわかれば自ずと現在の年齢が推定できることになることに着目した。問題点は様々な出来事からその実施時期を得ることは容易ではないことである。地域で起こった些細な出来事や特定の分野で有名な出来事の発生時期を知ることは難しい。

これらの問題を解決するために、大小様々な出来事について発生した時期を一元的に管理し、必要に応じて出来事のデータを提供できるようにするための一つの方法としてその発生時期を相互に変換することのできるイベント履歴収集システム、Historical information Acquisition System、略して HAS を提案し、システムの構築を行った。

構築した HAS の扱えるデータ量を増やす方法として、Wikipedia からデータを抽出して HAS で利用する方法の検討を行った。表1に実際に習得できたデータの分野と件数を示す。DBpedia で公開されている SPARQL のエンドポイントを用いて特定の出来事の欲しい情報の取得を行い、取得したデータに対して整形を行い HAS で提供可能な

表1 収集したデータ件数

分野	件数
事件	1,023
災害	305
映画	12,355
アニメ映画	1,420
ゲーム	8,649
音楽	16,343
漫画	1,343

データに変換する機構の実装を行った。結果として、事前に大量のデータを用意することなく、多くの出来事の情報を提供できるようになった。

(2) イベント履歴収集システムを用いた年齢推定

研究成果(1)で構築したイベント履歴収集システム HAS を利用した年齢推定可能な対話システムの構築を目指し、年齢推定のための LINE のチャットボットを作成し、HAS とそのテストデータの有効性を検証した。

チャットボットの基本的な対話の流れは3つのステップからなる。ステップ 1 では過去の体験について事件、災害、映画、アニメ映画、ゲーム、音楽、漫画のいずれかの体験についてランダムに問いかける。対話者の回答を受けて、HAS に問い合わせを行い、出来事の年代情報を取得することができたら、ステップ 2 としてその出来事を体験した年齢を問う発話を行いユーザの返答を待つ。ユーザが適切な年齢に関する返答をした時、STEP3 として得られた出来事と体験年齢を元に推定年齢の算出を行う。

検証は 1995 年度生まれ (23 又は 24 歳) の男性 10 人に対して、チャットボットと最大 20 回の対話を行ってもらい、ステップ 1 の問いかけの回答がテストデータと適合した回数と年齢推定の結果を集計する。表 2 は被験者の体験内容がテストデータに適合した回数とその比率を示している。問いかけの半数は年齢推定に利用できることがわかった。

本研究では、イベント履歴収集システム HAS を用い Wikipedia から収集したテストデータを利用して対話を行い、対話者の年齢を推定する LINE のチャットボットを構築した。対話者の推定年齢の算出が可能であることを示した。

表 2 テストデータとの適合率

分野	問合せ回数	適合回数	適合率
映画	28	17	60.7%
曲	41	22	53.7%
漫画	40	19	47.5%
ゲーム	52	30	57.7%
事件	13	3	23.1%
災害	11	5	45.5%
総件数	185	96	51.9%

(3) ユーザが変更可能な辞書を持つ感情分析ライブラリ

近年、UGC プラットフォームが様々な分野に拡大している。Web 小説はその一つである。コンテンツが飽和する中で、自身のコンテンツに注目させるためには目を引くタイトルを作成する必要がある。我々は小説などのタイトルについてユーザの興味を惹くタイトル生成方法の検討を行った。本研究では、目を引くタイトルの生成に必要な機構の一つとして文章の感情分析に着目し、ユーザによって適宜変更可能な辞書ベースの感情分析ライブラリ kamaboko を構築した。

既存の日本語を対象にした感情分析ライブラリには asari bert-base-japanese-sentiment, oseti などがある。あらかじめ positive・negative・neutral のラベルを付与した小説の一文でこれらのライブラリを評価した結果、いくつかの特徴が得られた。そのうち oseti は正解率がそれほど高くないもののルールの改良や辞書の拡張によって正解率の向上が期待できることが分かった。しかし oseti には二つの課題がある。一つ目は辞書の動的な拡張ができないことである。辞書ベースの感情分析では辞書の語彙量が精度に大きな影響を及ぼす。二つ目の課題は、否定語の適用能力の問題である。

我々は構築する感情分析ライブラリ kamaboko に対し分析に使用する辞書をユーザが独自に拡張できるよう実装し、形態素解析や依存構造解析、感情語彙のマッチング等をあらかじめ実施し、そのシーケンスに対する否定語の適用や並列関係の考慮を行っており、事前に特定した主語に依存する否定語を先に評価するアルゴリズムを導入し、これらの問題に対処した。また、構築した kamaboko の評価を行った。oseti に利用されている標準的な辞書に加え複数の拡張辞書を用意することで、感情分析の精度が向上することが分かった。kamaboko を利用することにより、既存のライブラリでは分析が難しい文章に対する感情分析が可能であることを示した。

(4) 携帯電話の電話音声画像を用いた兄弟間の判別について

近年、スマートフォンの普及により、携帯電話でのコミュニケーションが増えてきている。携帯電話の通話の音声は通信しやすいように加工されるため、似た声の人物の音声は、電話越しの聞き手にとって個人を判別することが難しい場合がある。

本研究では、携帯電話の通話時に筆者の声とその声に似た 2 人の弟の声を判別するため機械学習を活用した判別方法を検討する。電話音声画像を学習に用い、学習させるデータや学習済みモデルの出力を変化させて、精度比較を行った。精度比較の結果から、本研究の機械学習の分類の精度は1つのビットレートだけの学習データでは精度が悪くても、複数のビットレートの画像をまとめて学習に用いると精度が良くなることが分かった。また、今回の結果においては、モデルの最終出力よりも途中の出力を用いた学習の方が良い精度が得られた。モデルの層の多い VGG19 は良い結果が残ったが、他の条件によっては大きな差にならないことが分かった。

(5) 2台の全天球カメラを用いた距離推定手法

全天球カメラを有効に活用できる場面を増やすためには、画像以外にも情報が取得できた方が望ましい。例えば、カメラに映る被写体をリアルタイムで検出し、その被写体との間の距離を取得できれば、それを反映したVRコンテンツやシステムの製作が可能となる。しかし、現行の全天球カメラには被写体の検出や距離を測定する機能は搭載されていない。

本研究では、全天球カメラとしてRICOHの「THETA V」を2台用意し、それらのステレオ撮影により特定の被写体までの距離を推定する手法を提案し、その手法による距離推定の精度についての検証を行った。また、動画像によるリアルタイムでの距離推定についても検証した。

はじめに距離推定が実際に可能かどうか、実用に足る精度が出るのかどうかについて、1フレーム分の静止画を撮影して検証した。事前準備としてベースライン上にTHETA Vを2台、カメラ間の距離を1[m]として設置し、手作業でキャリブレーションを施した。画像の解像度は2K(1920×960)とした。その後、ベースラインから0.5、1.0、2.0、3.0、4.0[m]の各計測ライン上において、複数のポイントで被写体を撮影し、距離推定を試みた。距離推定の結果を表3に示す。どの計測ラインにおいても、誤差率が5[%]内に収まっていることが分かった。動作速度についても検証したが、安定した速度で動作させることが出来ず、こちらは有用性を示すには至らなかった。よって、提案手法は距離推定については有用であるが、リアルタイムでの安定した動作については未だ改善の余地があると言える。

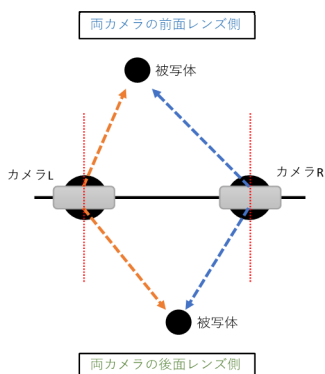


図1 距離推定の概要

表3 距離推定の精度

実際の距離[m]	計測値[m]	誤差率[%]
0.5	0.489	2.2
1.0	0.990	1.0
2.0	1.946	2.7
3.0	2.951	1.6
4.0	3.803	4.9

(6) ステレオ全天球カメラによる距離推定と「音のAR」への活用

本研究では研究成果(5)を元に提案システム(図2)に対して改良を行い、精度、速度の向上を目指した。また、ステレオ全天球カメラによって得られる被写体までの距離情報を活用する一例として、被写体の3次元位置を音声の定位によって知覚する機能をシステムに追加した。昨今では、視覚だけでなく聴覚から得る情報で現実を拡張する「音のAR」と呼ばれる仕組みが提唱されており、音のARを利用することを前提としたウェアラブルデバイスも登場し始めている。図3の「BOSE FRAMES」はサングラスにスピーカ、マイク、モーションセンサが内蔵されたスマートグラスであり、スマートフォンなどの外部デバイスと連携することで、音のARを利用することができる。本システムではカメラで得た視覚情報を聴覚情報に置き換えることで、音のARによる新たなユーザ体験を提供することができる。

具体的には、ステレオ全天球カメラで得られたベースラインから被写体までの距離、及びカメラから見た被写体の緯度、経度より、カメラと被写体を結ぶ3次元ベクトルの距離を求める。これらの情報を元に、カメラをユーザ、被写体を点音源とした時にユーザに聞こえる音声に生じる両耳間時間差や両耳間強度差を計算し、それを反映した音声データを生成する。また、3次元ベクトルの距離に応じて音量を調節し、音源の遠近を表現する。これらの処理をリアルタイム距離推定と並行して行い再生することで、カメラ内での被写体の動きを音声の定位の動きに置き換えてユーザに提供できる。構築したシステムによってBOSE FLAMESの装着者に対して実際に被写体の動きと音声の定位が連動することが確認できた。



図2 距離推定システム



図3 スマートグラス

(7) 全地球カメラによる深度マップ作成手法の検討

本研究では、研究成果(5),(6)を元に災害現場で収集される情報の中から距離情報の重要性に注目し、実用に足る精度を保ちつつ、高速に距離情報を収集し利用出来るシステムの構築を目指す。その端緒として、2台の全地球カメラで撮影した画像から深度マップを作成する手法について検討し、深度マップ生成システムの試作を行った。

構築したシステムは左カメラで撮影された画像全体を対象に特徴点抽出を行い、見つかった全ての特徴点において左右画像での特徴点マッチングと距離推定を行う。そして、各特徴点の周辺領域を距離に対応した色で塗りつぶすことで、簡易的な深度マップの生成を試みる。

結果として、概ね実際の距離変化に応じたグラデーションで表示されることが確認できた。(図4)また、最前列の机などレーザー距離計で距離が測定出来た構造物に関しては、近い距離が推定されていることも確認出来た。更にカメラの仰角を45度、90度に傾けて撮影した画像でも深度マップを作成したが、その場合も実際の距離変化に応じたグラデーションが確認できた。撮影から深度マップの生成まで10秒程度であり、速度の点でも当初の目的を達成出来たとと言える。

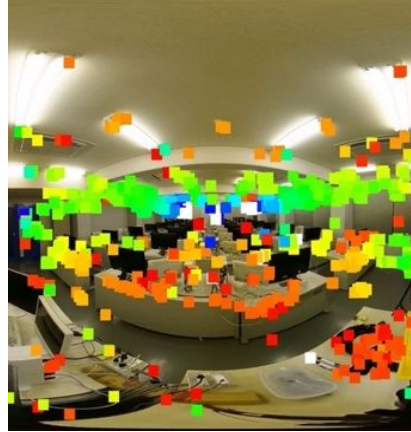


図4 深度マップ生成例

(8) グループワークの状況分析システムの構築

業務をはじめとしたグループワークの効率化には良好な人間関係や個々の作業状況の把握が求められるが、活動の途上でそれらの変化を細かく読み取ることは難しい。近年、業務にチャットツールによるメッセージのやり取りを採用する企業が増えている。

本研究では、グループワークの効率化を支援するシステムの構築を目的とし、チャットツールと連携したグループワークの活動状況を把握するシステムを構築した。本システムはチャットメッセージからメンバー間の対話状況やグループワークの状況を収集、分析し、人間関係図や活動の全体図の可視化を行った(図5)。

人間関係を分析するに、本研究では作業における人間関係を分析するためにチャットツールでやり取りされる本文中に「ありがとう」をはじめとしたキーワードが含まれている感謝のメッセージを情報源としメンバー間の人間関係を推定する。分析により取得した送り手、受け手、送信日時、感情スコアから時間経過を含めたメンバー間の人間関係を推定することが可能であることを示した。

タスクの偏りを把握するには過去、現在のタスクの存在やタスクの受け手と送り手、タスク完了の有無、タスクの発生日時や期限といったタスクに関する情報が必要である。タスクに関する情報取得にはChatworkのタスク機能を利用した。タスクに関する情報を取得することで過去のタスク状況や現在誰がどれほどのタスクを担い期限までにどれほどの猶予があるのか、グループ内の誰にタスクの依存をしているかといった情報を取得可能であることが分かった。

本研究にあたり実データを提供頂いた株式会社リオの方々から可視化された関係図を評価して頂いた所、業務上での人間関係やタスクの状態を把握することが可能であり、実際の状況に沿っているという評価を得た。一方、本システムでは取得不可能な情報やさらなる機能の追加といった改善点もあげられた。



図5 作業状況分析の概要

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計8件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 財満敦也, 菱田隆彰
2. 発表標題 携帯電話の電話音声画像を用いた兄弟間の判別について
3. 学会等名 情報処理学会第83回全国大会講演論文集
4. 発表年 2021年

1. 発表者名 池田輝政, 遠藤正隆, 中嶋裕一, 松井瑠偉人, 菱田隆彰
2. 発表標題 非属人性タスクを対象としたタスクマネジメントシステムの検討
3. 学会等名 情報処理学会第83回全国大会講演論文集
4. 発表年 2021年

1. 発表者名 池田輝政, 遠藤正隆, 中嶋裕一, 松井瑠偉人, 菱田隆彰
2. 発表標題 2台の全天球カメラを用いた距離推定手法
3. 学会等名 分散、協調とモバイル(DICOMO 2019)シンポジウム
4. 発表年 2019年

1. 発表者名 大竹栄一, 遠藤正隆, 中嶋裕一, 松井瑠偉人, 菱田隆彰
2. 発表標題 Historical information Acquisition System (HAS) の実装
3. 学会等名 分散、協調とモバイル(DICOMO 2019)シンポジウム
4. 発表年 2019年

1. 発表者名 鈴木克弥、遠藤正隆、中嶋裕一、松井瑠偉人、菱田隆彰
2. 発表標題 グループワークにおける作業状況の健全性分析システムの提案
3. 学会等名 第17回情報学ワークショップ(WiNF 2019)
4. 発表年 2019年

1. 発表者名 三十尾直也、遠藤正隆、中嶋裕一、松井瑠偉人、菱田隆彰
2. 発表標題 人為的ミスによる損害を減らすための業務向けdashcamの構築
3. 学会等名 第17回情報学ワークショップ(WiNF 2019)
4. 発表年 2019年

1. 発表者名 大竹栄一、遠藤正隆、中嶋裕一、松井瑠偉人、菱田隆彰
2. 発表標題 イベント履歴収集システムの試作と年齢推定への活用
3. 学会等名 情報処理学会第82回全国大会
4. 発表年 2020年

1. 発表者名 池田輝政、遠藤正隆、中嶋裕一、松井瑠偉人、菱田隆彰
2. 発表標題 ステレオ全天球カメラによる距離推定と「音の AR」への活用
3. 学会等名 情報処理学会シンポジウム インタラクシオン2020
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------