研究発表コンテンツ内の意味的対応関係を活用した研究理解の支援

太田 彩 *　脇田 建 †

概要. 遠隔開催の学会の増加に従い，学会運営組織から提供されるコンテンツの種類は以前に比べ多様になった．しかしコンテンツは個別に視聴・消費され，研究の統合的な理解の手助けにはなっていない．そこで本研究は各種コンテンツ間をハイパーリンクで結びつけた相互参照を提供することで新たな学会参加の形を提案する．具体的には自然言語処理・画像処理を用いて，コンテンツを超えて関連した内容を含む箇所の対応関係を自動で抽出する．加えて対応関係を相互参照の形でユーザーに提示することで，研究理解過程におけるコンテンツ間の移動を容易にする．以上によりコンテンツ間の結びつきを深め，論文の内容理解の効率向上・認知負荷の軽減を目指す．本論文では対応関係を自動で抽出する手法を提案し，相互参照をユーザーが容易に確認できるシステムのプロトタイプを作成した．加えて抽出手法の性能評価と有効性を検証するための実験を行った．

1　はじめに

近年の学会では口頭発表を補完する形で, 論文やシステムのデモ動画など多様なデータ種別のコンテンツが提供されている．加えて COVID-19 の蔓延により学会の遠隔視聴の機会が増加したことで, 発表動画や紹介動画などの追加の情報も取得できるようになった．その結果, 参加者は目的や関心に合わせ, 研究理解に使用するコンテンツの形態を選択することが可能になった．加えてコンテンツを組み合わせることで，研究への深い理解を得ることもできる．

しかし一方でコンテンツを横断して研究を理解するための体勢は整えられておらず，コンテンツは個別に消費されている．例えば研究の概要を発表動画で, 実験部分の詳細を論文で確認したい時，ユーザーはパソコンの各種ウインドウを使い分け目的を達成しようと考える．しかしウインドウの切り替えや特定箇所の探索はユーザーに認知負荷を与え，統合的な理解は阻害されてしまう [24]．このような状況からコンテンツは個別に消費され，各コンテンツの特性を活かした使い分けは行われていない．

本研究では現状を踏まえ，多様なコンテンツの統合的な理解の補助を目標とする．加えて，そのことによって論文の内容理解の効率向上，認知負荷の軽減を目指す．

本研究で注目したのは**相互参照**機能である．相互参照の利用は World Wide Web や Wikipedia のハイパーリンクに遡り，今日では PDF や Microsoft Word 文書でも使われ，有効性は幅広く認識されている．本研究では口頭発表動画・発表資料・論文内の要素の対応関係を自動で検出し，コンテンツ間にまたがる相互参照の形でユーザーに提示する．例として論文と口頭発表動画と類似した議論が展開されている論文中の箇所を示唆すること，選択されたスライドを用いて発表動画の箇所を提示することなどが挙げられる．コンテンツの切り替えを容易にすることで，データ種別を横断した内容の深い理解を支援することが目的である．

その上でコンテンツの要素を取り出し，要素間の対応関係を自動で抽出する手法を提案した．加えて対応関係を相互参照の形でユーザーに提示するシステムのプロトタイプを作成した．また対応関係の抽出手法の速度と性能を評価するための簡易的な評価実験を行った．

以上を踏まえ，論文の貢献は以下の通りである．まず相互参照に注目し，新しい学会視聴の形を提案する．加えて論文，発表動画から要素を抽出し，類似箇所を自動で検出する手法を提案する．さらに相互参照を確認するためのプロトタイプを作成し，自動検出手法の性能の予備調整を行った．

2　関連研究

2.1　アクティブ・リーディング

アクティブ・リーディングとは注釈をつけたり，メモをとるなどの行為をしながら読書を行うことである．学会参加者も論文を読み研究への理解を深めているため，本研究と関連深い分野である．これらの行為は教育機関や知的労働の場において大きな役割を果たしている．そのためアクティブ・リーディングの現状に関する調査 [6, 21, 17] や，それを支援するツールを提案する研究が数多く提案されてきた．

LiquidText は紙の読書の快適性を模倣するのではなく，流動的な文書表現と様々なインタラクションを提供することでアクティブリーディングを促進させた [25]．Head らは難解な論文に対し，専門用語や記号の定義を読者が最も必要とするときに提供

* 東京工業大学
† 東京工業大学

するインタフェースを提案した [12].

これらの研究は主に文書を読む読書の快適性と効率を向上することに重点をおいているが，本研究がめざすような関連する複数のデータ種別を横断的に理解する支援は提供しない.

2.2 オンライン学会への参加意識を高める研究

COVID-19 の蔓延により学会の遠隔視聴の機会が大幅に増加した．そのため近年遠隔方式の問題点を解決する形で，オンライン学会を支援する方法 [23, 13] が提案されてきた．加えて遠隔の学会や会議を後から振り返り理解するための研究 [10] も行われてきた.

Murali らはプレゼンテーションにおいてセンシング技術を利用し聴衆の反応を取得し発表者へ伝達した [18]．Aseniero らは参加者が会議中に自身の反応を発表者に匿名で送ることができ，あとから会議の様子を収集した反応を通し振り返ることのできるシステムを提案した [7].

El-Assady らは会議における話者の話すテーマの流れを表現することで，話者の行動パターンを分析するというアプローチを提案した [9]．阿辺川らは研究の映像と文書要素に対し，関係性などの付加情報を人間の手で注釈づけることで，情報を統合し読者の理解を支援した [5].

本研究は学会参加者による研究発表の深い理解を目的の中心に据え，必要なシステムの要件分析を行った．加えて論文と発表資料から図表や段落などの要素を自動で抽出し，要素間をまたいだ類似箇所を検出することで，ユーザーが短時間で必要な情報へアクセスできる体制を整えた.

3 提案手法

3.1 現在の学会参加の問題点

近年の学会は論文に発表動画，紹介動画など多様なコンテンツが提供されている．しかし複数コンテンツを統合的に理解するための体勢やシステムは提供されていない．そのためユーザーはコンテンツを個別に消費しており，各コンテンツの特性に応じた使い分けは行われていない．そこで本研究ではそれら問題を解決するために，様々なコンテンツの統合的な理解を促す新しい学会参加のコンセプトを提案する．そして論文の内容理解の効率向上，認知負荷の軽減を行うことを目標とする.

3.2 提案手法の概要

本研究では自然言語処理や画像処理技術を用いて，学会の提供コンテンツ内における関連した内容を含む箇所の対応関係を抽出する．加えて対応関係を相互参照の形でユーザーに提示し，研究理解の過程におけるコンテンツ間の移動を容易にする．こうして学会参加者の意識の中でのコンテンツ内の結びつきを深め，研究を様々な方面から理解することを可能にする.

PacificVis 2022[4] で提供されたデータをもとに，相互参照に用いるデータを決定した．具体的には論文のメタデータ，紹介動画，発表動画，論文を対象に研究を進める．ここでメタデータは研究の表題や著者を BibTeX などのフォーマットで保存したもの，紹介動画は学会の開催前に提供される研究の内容を 30 秒程度で簡潔に紹介した動画のことである．発表動画とは発表者が自身の研究をスライドウェアや開発したシステムを用いて紹介した研究紹介の録画動画である．本研究では特にスライドやシステムが動画の全面に映し出され，発表者の音声によって説明されたものを特に発表動画と定義づけた．これらコンテンツの中でも発表動画と論文が研究の内容を知る上で大きな役割を担っているため，この二つのコンテンツ間の結びつきを深めることを重視する.

3.3 要件分析

システムを開発するにあたり，研究発表の参加者の行動を想定し要件の分析を行った。加えて関連分野であるアクティブリーディングの研究を参考にした。O'Hara らは実験を通し人々のアクティブリーディングの様子を明らかにし [20]，Tashman らはアクティブリーディングの主要プロセスを注釈，内容抽出，ナビゲーション，レイアウトという 4 点とした [25]。本研究ではアクティブリーディングの主要プロセスを参考に以下の 5 つの要件を設定した。

(R1) 論文の概要を表示する機能

(R2) コンテンツ本体を視聴する機能

(R3) 関連箇所を示唆し，移動する機能

(R4) 関心のある箇所を選択し，自身の疑問や意見・感想を記述する機能

(R5) (R4) で記述した内容を他者に共有する機能

(R2) はナビゲーションに，(R3) はレイアウトに，(R4) は注釈と内容抽出に対応している。Tashman らの提案した Liquid Text[25] が注釈付システムであることに対し，本研究は論文の概要確認システムであるため，本システムではナビゲーションやレイアウトのプロセス中でもコンテンツ自体の配置を変更する機能は提供しない。加えて関心のある研究の発見と，他者との意見や情報の交換は研究理解の過程において重要な要素と認識されている [8, 23] ため，それらを支援するための機能 (R1)(R5) を追加した。

3.4 ユースケース

この節ではシステムの全体像を明確にするために，要件をもとにしたユースケースについて述べる.

花子さんはリアルタイムで見られなかった論文発表について，あとから発表動画を視聴することにした．まず研究の概要を知ろうと思い，発表動画から自動生成した発表資料を眺めた．このうち特に関心のある提案手法を説明するスライドに目をとめ，その発表資料に対応する動画の箇所付近を視聴した．動画を視聴する中，ユーザー実験の説明を見て，細かい条件設定が気になった．そこで動画のその部分から，内容に関連のある論文の段落を参照した．そうすると論文には実験の詳細が記述されており，疑問を解消することができた．

発表動画を視聴後，他の人にこの研究を詳しく紹介したいと思ったため，論文も最初から確認するようにした．そうすると文中に表を参照した箇所を発見したため，クリックするとシステム上で別のページにあった表を確認することができた．また読み進める中で似た話を論文の以前のページで既に読んだ気がしたため，類似段落をまとめたクラスターを確認した．その結果現在読んでいる段落と似た内容が他の節でも記述されていることを確認した．花子さんはシステムを使うことで効率的に研究内容を理解し，また論文の構造も深く理解することができた．

4 システム

本研究では要件分析とユースケースをもとにプロトタイプを作成した．その際に IEEE の 2 カラム国際会議用の様式に合わせて、要素と対応関係の抽出手法を考案した．本節では最初にシステムの概要について紹介を行い，その後開発に至るまでの工程について紹介を行う．

4.1 システム概要

システムは大きく論文選択画面と論文視聴画面の二つの画面に分けられる．論文選択画面では論文の概要と内容を表した WordCloud・紹介動画を確認しながら関心のある論文を選択することができる．論文視聴画面は図 1 のような構造になっており，6 つの部分に分けることができる．

発表動画・発表資料ビュー　このビューでは発表動画と，発表動画から自動生成された発表資料，また音声の文字起こしを閲覧できる．発表資料の一覧から発表の概観を眺め，円滑に自分の関心のある部分を選択することができる．加えて発表資料の 1 ページを選択すると，自動でそのページと内容が最も類似した論文の段落がハイライトされる．この画面は要件分析の (R2, R3) と対応している．

類以度ヒートマップ　このビューでは現在選択している発表資料のページの，論文の段落に対する類似度の一覧をヒートマップを通し確認することができる．ヒートマップが濃いほど段落との内容が類似していることを表しており，ヒートマップをタップすると，選択した段落を論文ビューで確認することができる．このビューは要件分析の (R3) と対応している．

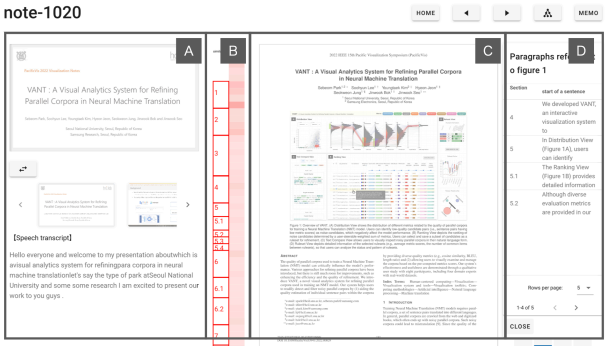

図 1. 現状のプロトタイプの論文視聴画面 (A) 発表動画・発表資料ビュー (B) 類以度ヒートマップ (C) 論文ビュー (D) 図表参照段落発見ビュー

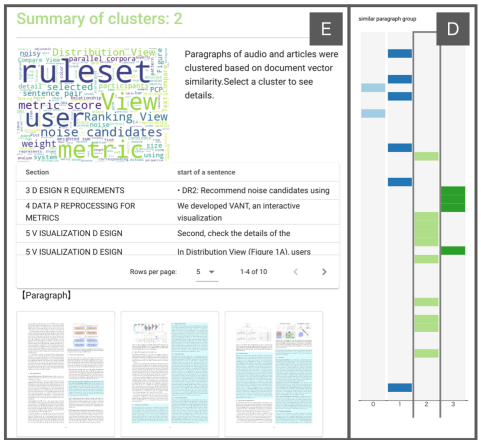

図 2. 論文視聴画面の一部分 (E) 論文グループビュー (F) 論文グループ詳細ビューの概要

論文ビュー　このビューでは論文を閲覧することができ，要件分析では (R2) と対応している．

図表参照段落発見ビュー　このビューは論文内の図表をクリックすることで閲覧することができる．ビュー内の表には選択した図表が参照されている論文の段落の一覧が表示されており，行を選択することで論文の段落の詳細を確認することができる．この画面は要件分析の (R3) と対応している．

論文グループビュー　このビューは論文の段落を内容の類似度に応じてクラスタリングした結果を表示する．クラスターを表す列をタップで選択すると，論文グループ詳細ビューでクラスターの詳細を確認できる．この画面は (R3) と対応している．

論文グループ詳細ビュー　このビューはクラスターを構成する段落の一覧を示すとともにクラスターの概要を Word Cloud で要約する（図 2）．また画面下部に表示される論文のハイライト表示で，クラスターを構成する段落を確認できる．このビューは要

件 (R2, R3) と対応する．

4.2 分析の流れ

ここでは開発したシステムについて主にデータの処理の流れを中心に紹介する．以下の説明はコンテンツ内の要素の抽出，要素間の関連の検出，ユーザーを補助する機能，システム全体の構成に分けられる．

4.2.1 コンテンツ内の要素の抽出

コンテンツ間の関連した箇所の検出を行うために，この段階では図 3 のように各コンテンツから発表資料や図表，文字起こしなどの要素を抽出する．発表動画については Amazon Transcribe[1] を使用し，発話内容の音声から文字起こしをする．加えて画像処理によって発表動画を一定秒数ごとにフレームに分割し，類似したフレームをまとめることで発表資料を作成する．参考となる発表動画を閲覧したところ一枚のスライドを 1 分程度発表に使用していることを確認できたため，本研究では 10 秒ごとにフレームを分割する．最後に発表資料の各スライドから図表を抽出し，文字認識ライブラリを用いて文章を抽出する．

論文については文字認識システムを使用し，テキスト化する．このままでは 2 段組の論文に適応できないため，正規表現を用いて各行が 2 段組か判定し，2 段組の場合は行を左右に分割し、左右の文章をそれぞれまとめることで文章の出現順を正しいものへと整える。そして調整した文章に対し正規表現を適用することで、節の表題と図表のキャプション，本文を段落ごとに取り出す．加えて論文の中の図表も画像処理ライブラリを用いて抽出する．これらの操作により相互参照を行うために必要な要素を一通り抽出することができた．

4.2.2 関連箇所の対応付け

この段階では自然言語処理と画像処理により，図 3 のように要素間の関連箇所を自動検出する．

まず論文の段落と発表動画の文字起こしを比較する．この際に各段落と文字起こしの文章をベクトル化し，コサイン類似度をもとに段落と音声間の類似度を計算した．ベクトル化手法は TF, TFIDF, word2vec[15, 16] の平均, doc2vec[14] の 4 手法を比較し、最も優秀な TFIDF を採用した.（詳細は 5.2 節）加えて論文の段落間の類似箇所を紐づけるために，TFIDF によって生成したベクトルを用いたクラスタリングを行った．論文内の類似した段落のグループ数は論文ごとに異なり，事前に不可知である．このため，クラスタリング手法としてはクラスタ数を事前に与えるものは避け，外れ値にも強い DBSCAN[11, 22] を採用した．また注目した段落と関連した段落の数を削減するため，近接性基準にあたる eps を 0.01 から 1.00 の範囲で 0.01 ごとに変化

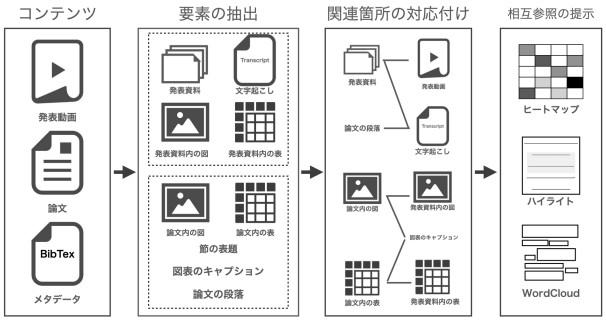

図 3. コンテンツの処理から相互参照を提示するまでの流れ

させ、小規模なクラスターを最大数作成できた eps の値を採用した．

続いて図表とキャプションの紐付けを行う．しかし既存の図抽出ライブラリの図の認識精度は不十分だった．本研究では論文内の空白を認識し，領域分割することで図表を検出している．しかし領域分割をすると図表とキャプションが分離されたり、図表自体も小さな領域に余分に分割される場合がある．そこで図表が論文の本文よりも上に配置される IEEE 様式の特性を活かし，キャプションの直上に位置する領域を対応する図表として扱い、キャプションとの対応関係を記録する。ただし、本稿図 4 のように複数の小さな図表が組み合わさっている場合は、それらを全て包含する領域を対応する図表とした。

その他，画像処理ライブラリを用いて論文内の図表と発表資料内の図の対応関係を検出した．加えてタイムスタンプを用いることで，発表動画のタイミング，文字起こし，発表資料の対応関係も導き出すことができた．

4.3 ユーザーを補助する機能の開発

この段階では対応関係を相互参照の形でユーザーに提示するための可視化と論文の内容を表すための可視化を作成する．発表資料と段落間の類似度を見るための可視化としてヒートマップを採用し，選択した論文の段落をハイライトを用いて指し示すことにした．加えて論文の内容を一目で理解できるように，出現頻度によって単語の大きさを変えながら並べる可視化である Word Cloud[19] を採用した．

4.4 開発に用いたツール群

前段階で作成した可視化を繋ぎ合わせ，システムを開発した．要素の抽出・要素間の対応付けには openCV[2], pyocr, pdftotext, `Pillow`[3] を使用している．システムのバックエンドは `Flask`, フロントエンドは Vue.js と WordCloud, 可視化は D3.js を使用している．

表 1. コンテンツ抽出の結果: 再現率 (適合率)

	論文１	論文２	論文３	論文４
発表資料の再構成	1.0(1.0)	0.89(0.96)	0.87(0.77)	0.84(0.89)
段落の抽出	0.90(0.83)	0.80(0.77)	0.83(0.79)	0.77(0.75)
段落の抽出 (分割ミスを除いた場合)	0.97(0.98)	1.0(1.0)	0.99(0.98)	0.96(0.98)
節の表題の抽出	1.0(1.0)	1.0(1.0)	0.94(1.0)	1.0(1.0)
図表キャプションの抽出	1.0(1.0)	1.0(1.0)	1.0(1.0)	1.0(1.0)
実行時間 (s)	87.55	174.44	262.37	169.65

5 実験

5.1 要素の抽出

IEEE の 2 カラム国際会議用の様式を採用した PacificVis 2022 の “Session 8: Text and Mixed Data” の 4 つの論文[1234]に対し，開発した要素抽出手法を適用し実行時間を計測した．加えて発表資料の再構成，段落・表題・図表キャプションの抽出という 4 つのタスクについて精度を評価した．発表資料の再構成では動画から 10 秒ずつ切り出したフレームを人間の目で確認し，重複したフレームを削除することで発表資料を再構成する．その際に再現率と適合率を用いて，作成したものと自動処理した結果の一致率を計測する．続いて段落・表題・キャプションの抽出においても，人間の目で論文の原本を確認しそれぞれ正解データを作成した．その後自動抽出した結果と正解データを比較し再現率を計算した．以上の実験の結果が表 1 である．

表題とキャプションは再現率・適合率ともに 100% に近い精度で抽出することができた．一方発表資料の再構成と段落の抽出では論文 2〜4 において精度が少し下がった．段落抽出の精度が下がった要因の多くは，段落を余分に分割してしまうなどの分割ミスである．実際に分割ミスを除いた場合の再現率は概ね 96%以上と高い値を出している．発表資料の再構成で精度が下がった原因としては，発表資料内のアニメーションの使用が挙げられる．口頭発表では注目箇所を強調するためにアニメーションが使用されるが，その際にどこまでが同じスライドだと感じるかは視聴者によってばらつきが存在する．今回特に精度が低かった論文 3 では，スライド内の重要箇所を強調するためにペンを使って書き込みをしており，アルゴリズムでは書き込みのあるスライドを別のスライドとして認識したため精度が低下した．

5.2 関連箇所の対応づけ

続いて 5.1 節の実験で用いた中で特に精度がよかった論文 1 を用いて，追加の実験を行う．その際に確認するタスクは発表資料と文字起こしの対応，発表動画と発表資料の対応，図表と図表番号の対応，段落と発表資料の対応，論文内の図表とスライド内の図表の対応の 5 つである．

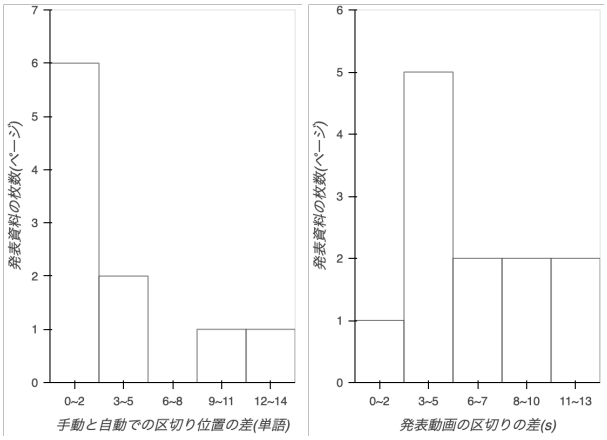

図 4. 一致箇所の検出結果 (左: 発表資料と文字起こしの対応， 右: 発表動画と発表資料の対応)

発表資料と文字起こしの対応　このタスクでは人間の手で発表資料の区切れに合わせて，文字起こしを分割し正解データを作成した．そして自動処理結果と正解データの分割位置が何単語ずれているのか比較した．その結果，図 4(左) のように平均のずれが 3.5 単語で， 6 割は 2 単語以内のずれとなった．全体が 943 単語で 11 枚のスライドに分割されることを考えると，ずれは非常に小さいと言える．

発表動画と発表資料の対応　このタスクでは発表動画の画面が変わったと判断した時刻が，人間とアルゴリズムによって何秒ずれるのか計測した．その結果が図 4(右) で，3〜5 秒ほどのずれが一番多いことがわかる．実験に用いた論文 1 の発表動画は 455 秒で 1 ページの平均秒数が 45 秒であることを考えると，差は十分に小さいと考えられる．

図表と図表番号の対応　5.1 節の実験で用いた論文 1〜4 の全ての論文を使用し，36 組の図表とキャプションの紐付けを試みた．その結果全ての図表を抽出し，正しい図表番号を付加できることを確認できた．

[1] (Mishra+ 2022)10.1109/PacificVis53943.2022.00022
[2] (Zheng+ 2022)10.1109/PacificVis53943.2022.00023
[3] (Wang+ 2022)10.1109/PacificVis53943.2022.00024
[4] (Park+ 2022)10.1109/PacificVis53943.2022.00029

論文内の図表とスライド内の図表の対応　唯一論文とスライドで同じ図表が使用されていた論文 4 に着目し，提案アルゴリズムによる対応関係の検出精度を確認した．その結果再現率 0.64，適合率 1.00 で対応関係を取り出すことができた．失敗した原因はスライド内でのアニメーションの使用や画像の再配置が挙げられる．今後は類似スライドでの対応関係を参照することで，精度の向上が期待される．

論文の段落と文字起こしの対応　このタスクでは手動で論文と文字起こしを確認し，内容に関連性のある箇所を紐付け，正解データを作成する．アルゴリズムでは各スライドに対し類似度の高い上位 5 段落を紐づける．正解データと比較し再現率を計測する．その結果 TFIDF の再現率が 0.68 と一番高く，他の手法に比べ 0.3 以上高いという結果になった．

6　議論

6.1　自動抽出手法の精度に関する考察

本研究では要素の抽出と関連箇所の検出を行う手法を考案し，5 章の実験を通し性能を測定した．その結果要素抽出タスク 4 つのうち 2 つのタスクで適合率 1.0 という非常に高い精度を確認することができた．加えて関連箇所の検出タスクにおいても，発表資料と発表動画・文字起こし間の対応関係を高い精度で検出することができた．そのため，提案手法を用いることでユーザーに対しある程度正確な対応関係を提供できることを確認することができた．一方で議論の余地のある以下の 3 点がある．

最初の点は発表動画からの発表資料の作成に伴う困難である．スライドウェアを用いた発表では頻繁にアニメーションが使用される．その際にどこまでを一枚のスライドウェアとして認識するかについては人によってばらつきがあり，明確な正解を設定することは困難である．開発アルゴリズムでは少しでも画面に変化があった場合では違うスライドとして認識したが，今後はユーザーのニーズに合わせ，枚数を減らしたりする機能が必要であると考えられる．

つぎに文字起こしと論文の段落に関する関連箇所の検出精度の低さである．これには複数の理由が存在する．まず 1 点目は音声認識の精度の限界である．論文の中で定義した語句を認識しないことで，文章同士の類似度の精度が下がることがある．今後は文字認識を併用して利用することで，語句の認識を改善したいと考える．2 点目の原因はスライドの内容が複数の段落の内容を包含していることである．その場合文字起こしと段落の内容が局所的に類似していても，全体の類似度は下がる傾向にある．今後は段落と文字起こしを文単位に分割し，文同士の類似度を計算することで性能を向上させたいと考える．最後は要素抽出にかかる計算時間の長さである．図 1 のように，長い場合では処理に 4 分以上の時間を要することがある．一見するとこの計算時間は長く感じる．しかし論文の読み込みや発表動画の視聴に要する時間と比較すれば遥かに短い．加えて実行時間の 50%以上は論文から図表やセクション・段落を抽出するための前処理時間である．論文は多くの場合発表当日よりも前に公開されるため、個々のユーザーに発表の前に論文の前処理を行ってもらうことで体感時間を減少させることができる．以上から現在の実行時間は十分に実用可能である．

6.2　要件に対する達成度

本研究では 3.3 節において，学会参加者を支援するための要件を (R1-R5) の 5 つ定義した．その上で現状のシステムの達成度について議論を行う．(R1) を達成するためにシステムでは論文選択画面内で論文の概要を表す WordCloud と論文概要の表示を行った．続いて (R2) を実現するために発表動画・発表資料ビューと論文ビューを導入し，コンテンツの本体を視聴できるようにした．そして (R3) を満たすために類似度ビューと論文グループビューと論文グループ詳細ビューを導入した．これらのビューではユーザーは関心のある箇所から関連のある箇所へと，コンテンツを超えて素早くアクセスすることができる．加えて論文グループ詳細ビューでは類似段落へのアクセスも容易にできるため，論文全体の構造も理解できる．一方，(R4, R5) を満たすための機能はまだ搭載していない．今後は (R4, R5) を満たすために新たなメモ機能とメモ共有機能を実装する予定である．その際にもメモと各コンテンツを紐づけることで学会参加者の作業量と認知負荷を軽減することを目指す．その事によってユーザーが研究を理解する過程において、関心のある論文の発見から研究を理解し、他者へと意見を共有するという一連のタスクを支援することが今後の目標である．

7　まとめ

本研究では学会で提供される様々なコンテンツの統合的な理解を目指し，相互参照に注目した新しい学会参加の支援の形を提案した．また支援を実現するために論文や発表動画から要素を取り出し，関連箇所を検出する手法を考案した．加えてユーザーが相互参照を目で見て確認するためのシステムのプロトタイプの作成を行った．そして，提案アルゴリズムの性能の予備調整を行った．一方今後の研究の課題は 2 点ある．1 点目はユーザー自身の考えを筆記し，それを他者へ共有することの補助である．2 点目は提案手法の有効性を検証するための本格的なユーザー実験の実施である．システムを使用した場合と使用しなかった場合の学会参加・研究理解を比較することで，提案手法がユーザーの論文の内容理解の効率と認知負荷へどのような影響を及ぼすのか定性的に評価する．

謝辞

研究室の皆さんには常に刺激的な議論を頂きました．特に同期の牧修平氏，Chen Xueyi 氏には精神的にも支えられました．心より感謝申し上げます．

参考文献

[1] Amazon Transcribe ホームページ. https://aws.amazon.com/jp/transcribe/.

[2] openCV ホームページ. https://opencv.org/.

[3] Pillow のホームページ. https://pillow.readthedocs.io/en/stable/.

[4] PacificVis 2022 のホームページ. https://pvis2022.github.io/pvis2022/, 2022.

[5] 阿辺川 武, 相澤 彰子. 内部構造解析機能と脚注表示機能を備えた論文閲覧システム. 人工知能学会第二種研究会資料, 2014(AM-07):03, 2014.

[6] A. Adler, A. Gujar, B. L. Harrison, K. O'Hara, and A. Sellen. A Diary Study of Work-Related Reading: Design Implications for Digital Reading Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '98, pp. 241–248, USA, 1998. ACM Press/Addison-Wesley Publishing Co.

[7] B. A. Aseniero, M. Constantinides, S. Joglekar, K. Zhou, and D. Quercia. MeetCues: Supporting online meetings experience. In *2020 IEEE Visualization Conference (VIS)*, pp. 236–240. IEEE, 2020.

[8] F. Beck, S. Koch, and D. Weiskopf. Visual Analysis and Dissemination of Scientific Literature Collections with SurVis. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):180–189, 2016.

[9] M. El-Assady, V. Gold, C. Acevedo, C. Collins, and D. Keim. ConToVi: Multi-party conversation exploration using topic-space views. In *Computer Graphics Forum*, Vol. 35, pp. 431–440. Wiley Online Library, 2016.

[10] M. El-Assady, R. Sevastjanova, B. Gipp, D. Keim, and C. Collins. NERExr: Named-Entity Relationship Exploration in Multi-Party Conversations. In *Computer Graphics Forum*, Vol. 36, pp. 213–225. Wiley Online Library, 2017.

[11] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, Vol. 96, pp. 226–231, 1996.

[12] A. Head, K. Lo, D. Kang, R. Fok, S. Skjonsberg, D. S. Weld, and M. A. Hearst. Augmenting Scientific Papers with Just-in-Time, Position-Sensitive Definitions of Terms and Symbols. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[13] E. Hu, M. A. R. Azim, and S. Heo. FluidMeet: Enabling Frictionless Transitions Between In-Group, Between-Group, and Private Conversations During Virtual Breakout Meetings. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.

[14] Q. Le and T. Mikolov. Distributed representations of sentences and documents. In *International conference on machine learning*, pp. 1188–1196. PMLR, 2014.

[15] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[16] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26, 2013.

[17] M. Morris, A. Brush, and B. Meyers. Reading Revisited: Evaluating the Usability of Digital Display Surfaces for Active Reading Tasks. pp. 79–86, 10 2007.

[18] P. Murali, J. Hernandez, D. McDuff, K. Rowan, J. Suh, and M. Czerwinski. AffectiveSpotlight: Facilitating the Communication of Affective Responses from Audience Members during Online Presentations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[19] L. Oesper, D. Merico, R. Isserlin, and G. D. Bader. WordCloud: a Cytoscape plugin to create a visual semantic summary of networks. *Source code for biology and medicine*, 6(1):7, 2011.

[20] K. O'Hara and A. Sellen. A Comparison of Reading Paper and On-Line Documents. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, p. 335–342, New York, NY, USA, 1997. Association for Computing Machinery.

[21] N. Roy, M. V. Torre, U. Gadiraju, D. Maxwell, and C. Hauff. Note the Highlight: Incorporating Active Reading Tools in a Search as Learning Environment. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*, CHIIR '21, pp. 229–238, New York, NY, USA, 2021. Association for Computing Machinery.

[22] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)*, 42(3):1–21, 2017.

[23] J. Song, C. Riedl, and T. W. Malone. Online Mingling: Supporting Ad Hoc, Private Conversations at Virtual Conferences. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, 2021.

[24] C. S. Tashman and W. K. Edwards. Active Reading and Its Discontents: The Situations, Problems and Ideas of Readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 2927–2936, New York, NY, USA, 2011. Association for Computing Machinery.

[25] C. S. Tashman and W. K. Edwards. LiquidText: A Flexible, Multitouch Environment to Support Active Reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 3285–3294, New York, NY, USA, 2011. Association for Computing Machinery.

未来ビジョン

本研究では様々なデータ種別のコンテンツを統合することによって，ユーザーによる研究への深い理解を支援することを目的とした．その上で本論文では手始めとして，学会発表が行われた後を想定しユースケースを設定し，システムのプロトタイプを作成した．加えて簡易的な実験を通し，開発したデータ処理アルゴリズムの性能を定量的に評価した．しかしシステムの要件として定義した,「研究に対する疑問や意見・感想を記述する機能」と「記述した内容を他者へと共有する機能」は現状実現できていない．今後はユーザーのメモも論文や発表動画と関連づけることによって，発表後のユーザーの行動をさらに手厚く支援したいと考える．

本研究の大きなビジョンでは，学会参加における全てのプロセスを支援したいと考えている．学会参加者が行う行動は多岐に及んでおり，発表を聴講する前の事前の準備から研究内容を理解しその内容を他者へと還元する部分まで様々なプロセスが存在する．その中で参加者は発表プログラムや著者の過去の研究論文，自身の記述したメモ，他者へ説明するために作成したスライドなど幅広いコンテンツに絶えず触れ続ける．そのためコンテンツの統合的な理解及び利用は学会参加における全てのプロセスにおいて必要不可欠であり，それを支援することで参加者の学会参加の方法も大きく変化すると考える．支援の一部としてプログラムから自動でカレンダーアプリへ予定を追加したり，発表中のメモを自動的に発表動画・発表資料と関連づけることなどが考えられる．参加者の負担となっている動作を相互参照によりサポートすることで，参加者はより深く研究の詳細を学び，発表者と接触する機会を有効に活用できるようになる．以上が本研究の描く未来ビジョンである．