

Information Retrieval of Compound Data by Linguistic Querying

鈴木, 釈規

<https://hdl.handle.net/2324/4475149>

出版情報 : 九州大学, 2020, 博士 (工学), 課程博士
バージョン :
権利関係 :



氏 名 : 鈴木 稔規

論 文 名 : Information Retrieval of Compound Data by Linguistic Querying
(言語的な問い合わせによる複合データ情報検索)

区 分 : 甲

論 文 内 容 の 要 旨

近年、ネット上で共有されるデータが多様化及び複雑化している。例えば、ソーシャルメディアサービスでやり取りされるデータは、テキストデータから動画や画像等に変化している。さらに投稿動画とそれに対するユーザのコメントがコンテンツとなっている。今後も新たなデータの出現による多様な様式からなる複合化したデータの増加が予想される。本論文では、こうした多様な形式から構成されるデータを複合データと呼称する。

言葉を検索質問とする複合データ検索は、検索質問の入力方法として簡便であり効率的な情報アクセスのために重要である。しかしながら、従来の検索において、複数の様式をまとめる複合データという観点は重視されてこなかった。複合データを対象にした検索であっても、検索質問と検索対象が同一であることが一般的である。例えば、画像とテキストからなる複合データに対して、画像のみやテキストのみを検索質問として受け付ける。この様な検索では検索者側が、検索に用いるデータを適切に選択する必要性があり、検索への習熟が要求される。また、多様な様式に対して横断的な検索ができなく、検索の質を低下させる。

従来単一の様式で行われていた検索を複合データとして扱うことにより、複合データの有効的な検索が可能になる。複合データ検索では、単一データ検索よりも豊富な情報を検索に動員できる。例えば、画像とテキストから構成される複合データ検索において、それらの片方のみの単一データよりも情報が多い。そのため複合データ検索では、単一データ検索では上手くいかなかったことが解決できると期待できる。

本研究では、複合データを言語による問い合わせで検索を行うことを目的とする。そうした検索を可能とする技術について具体的な研究課題を通して検討した。初めに、複合データの豊富な情報を活用する検索技術として、複合データへの参照キーワード付与について検討した。具体的には、1) 画像とテキスト情報を活用するキーワードの曖昧性解消及び、2) 画像データを複合データとして扱うことによるキーワード付与に取り組んだ。これらによってキーワードによる複合データへのアクセスを実現した。次に、言語による複合データ検索手法を具体的な検索課題について検討した。この検索課題として、3) 画像とそれに付加された多様なメタ情報からなる複合データ検索を対象として、情報要求文を用いて検索する手法を構築した。

1) 画像とキーワードからなるデータに対して、キーワードの曖昧性解消を検討した。この曖昧性の具体例として、ウェブ上の画像データとそれに付与される説明ラベル(タグ)が持つ曖昧性を対象とし、タグを百科事典の項目に対応付けすることで解消する。テキスト情報のみを利用する既存手法では、数単語からなるタグという限られたテキスト情報しか活用していなかったため、特に

多義語、例えば動物の名前（動物名、商品のブランド名やスポーツチーム名などに用いられる）のタグについて正確な対応付けが困難であった。この課題に対して、既存手法では考慮されていなかった画像の特徴を組み合わせた曖昧性解消手法を提案した。提案手法では、畳み込みニューラルネットワークを用いた画像分類器により、画像から分類ラベルをテキスト情報として取り出す。タグと画像情報を表す分類ラベル両者の類似度による順位づけによって、百科事典項目への対応づけを行う。評価実験では、単一データのみを扱う既存検索手法よりも高い性能を示し、6割以上の画像とタグの組に正確に対応づけすることが出来た。

2) 画像データを複合データとして扱うことによるキーワード付与手法について、画像アノテーション課題について検討した。画像アノテーションは、入力画像に対してそれを説明するキーワードを付与する課題である。画像アノテーション課題に対する既存手法の多くは、画像とキーワード群を訓練データとする教師有り学習手法である。訓練データに用いられる正解キーワードは人手によって付与されるため、付与されるキーワードの揺れが自然に発生する。例えば、人が映る画像に対するキーワードであっても、観光地で記念に撮られた画像では「観光客」、スポーツを楽しむ画像であれば「プレイヤー」などが付与される場合もあれば、より一般的な「女性」などが付与される場合もある。これは訓練データの品質ではなく、画像を説明するキーワードを付与するという人の主観性に依存する当課題ならではの性質に由来する。そのために訓練データのキーワードの一貫性が低くなり、その影響によって既存教師有り学習手法ではアノテーション性能が低くなる問題があった。この問題に対して、画像からテキスト情報を抽出する手法を提案した。画像とキーワードの百科事典文章の組みを教師データとするニューラルネットワークを定義し、変換器の訓練を行う。この変換器により、上述の「観光客」の画像から「女性」の画像と共通する文脈情報を持つ特徴量（例えば、「人に関係する」）を学習することが可能になる。このように変換したデータとキーワードの組みを訓練データとして、キーワード付与を学習する。評価実験では、画像アノテーション課題で、提案手法による変換データを用いた分類では既存手法と比較して、アノテーション性能を F 値で約 0.1 向上した。

3) テキストを検索質問として受け付けて、複合データを検索する手法を構築した。この検索の具体例として、ある人の生活を記録したログデータから特定のシーンを探し出すライフログデータ検索を対象とした。ライフログデータは、スマートデバイスによる利用者の生活に関する多種多様なレコードから構成されるデータである。例えば、スマートフォン、ウェアラブルカメラ、腕時計型端末などの機器によって、利用者の生活シーンの画像、歩いた歩数、脈拍数、GPS による位置情報などの記録情報が含まれる。ライフログデータ検索では、検索者は、例えば「家電売り場で買い物をしている」シーンを探し出したいといった漠然とした情報要求を持つ。しかしながら、ライフログデータは、上述の様に生活シーンの画像や位置情報などのレコードであり、「買い物をしている」と明示的に書かれているわけではなく、情報要求とデータの間には意味的な隔たりが存在する。この問題に対して、情報要求文の条件記述による絞り込み検索手法を提案した。提案手法ではライフログデータ中の画像データから、シーンや物について訓練された特定器によって、それらの単語（コンセプトと呼ぶ）を取得する。コンセプトと検索質問間の類似度を、単語のベクトル表現により算出し順位付けする。この順位付きレコードについて、情報要求文の条件記述から、規則によって構成したブーリアン型クエリにより絞り込む。例えば、情報要求文に「レストランで朝食を取っている時」であれば、「時間帯が午前」と「場所がレストラン」という条件を指定するクエリを構成する。評価実験では、既存評価用テストコレクションにおいて、既存手法と比較しても高い平均適合率 0.3 を超える検索結果を示した。