



HOKKAIDO UNIVERSITY

Title	Extracting Hierarchical Structure of Web Video Groups for Realizing Advanced Web Video Retrieval [an abstract of dissertation and a summary of dissertation review]
Author(s)	原川, 良介
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第12192号
Issue Date	2016-03-24
Doc URL	https://hdl.handle.net/2115/61753
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	doctoral thesis
File Information	Ryosuke_Harakawa_abstract.pdf, 論文内容の要旨



学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称 博士（情報科学） 氏名 原川 良介

学 位 論 文 題 名

Extracting Hierarchical Structure of Web Video Groups for Realizing Advanced Web Video Retrieval
(高度な Web 映像検索を実現する Web 映像集合の階層構造抽出に関する研究)

本論文は、高度な Web 映像検索を実現するために Web 映像集合の階層構造を抽出する研究を行い、その成果をまとめたものである。

映像共有サービスの普及によって、Web 上に存在する映像（以降、Web 映像）が増加しており、多数のユーザがそれを閲覧している。具体的に、世界最大の映像共有サービスである YouTube は 2015 年現在、十億人以上のユーザに利用されており、一日あたり数十億件もの Web 映像が閲覧されている。このような社会的背景から、ユーザが望む Web 映像を効果的に検索する技術の実現が急務となっている。一般に、Web 映像を検索する際、ユーザは望む Web 映像を特定するキーワードをクエリとして検索システムに入力する。検索システムは、入力されたキーワードに対応する Web 映像をランキング形式で提示する。したがって、適切なキーワードが入力されない場合、上位の検索結果に望む Web 映像が提示されず、下位の検索結果まで探索しなければ望む Web 映像が得られないという問題が存在する。

この問題を解決するため、類似したトピックを含む Web 映像の集合（以降、Web 映像集合）の階層構造を用いた検索手法が提案されている。これらの従来手法は、ユーザによって入力されたキーワードに関連する大量の Web 映像から Web 映像集合の階層構造を抽出し、ユーザに提示する。その結果、ユーザは階層構造に従って求めるトピックを含む Web 映像集合に辿り着くことが可能となり、先に記した問題を解決する検索が実現される。しかしながら、これらの従来手法は、階層構造を抽出する際、Web 映像の画像特徴もしくはタイトルや説明のテキスト特徴の一方のみを用いており、階層構造の抽出精度の向上に限界がある。Web 映像は画像、音響、テキスト特徴やリンク関係の特徴という異なる種類の特徴（以降、異種特徴）を有するため、これらの関連性を活用することで、問題の解決が期待できる。

そこで、本論文ではまず、単一の特徴のみを用いた際の精度の限界を打破するため、異種特徴間の関連性を活用して Web 映像集合の階層構造を抽出する手法を提案している。異種特徴間の関連性を用いることで、従来手法に比べ高度な Web 映像検索が実現される。提案手法で用いる異種特徴は、Web 映像の画像、音響およびテキスト特徴と Web 映像間のリンク関係の特徴とする。提案手法は、画像、音響およびテキスト特徴の相関に基づき潜在特徴を算出し、それを Web 映像間のリンク関係の特徴と協調的に用いることで、Web 映像集合の階層構造を抽出可能とする。さらに、本論文では、提案手法を大量の Web 映像に適用可能とする改良手法を提案している。改良手法では、事前のクラスタリングに基づき Web 映像の潜在特徴算出を効率化する。続いて、予め大量の Web 映像をグルーピングすることで、提案手法において必要であった Web 映像間のリンク関係の大局的な特徴算出の計算コストを削減し、階層構造抽出の高速化を可能とする。さらに、この改良手法の更なる高速化と高精度化を実現する手法を提案している。具体的には、計算コストが高い Web 映像間のリンク関係の大局的な特徴算出を行わず、その局所的な特徴のみを用いることで、階層構造抽出の高速化と高精

度化を実現する。これら二点の改良を経て、提案手法の実データへの適用可能性が示された。

本論文を構成する各章の概要は以下の通りである。第 1 章では、本論文の研究背景および目的について述べる。第 2 章では、本研究の関連研究として種々の Web 映像検索手法を紹介し、本論文で解決すべき問題を明らかにする。第 3 章では、異種特徴間の関連性を活用して Web 映像集合の階層構造を抽出する手法について説明を行う。提案手法では、まず、Web 映像の画像、音響およびテキスト特徴に対して正準相関分析を施し、その結果に基づき Web 映像の潜在特徴を算出する。得られた潜在特徴と Web 映像間のリンク関係の特徴の両者を用いることで、Web 映像の潜在特徴間の類似性をエッジの重みとして表現するグラフを構築する。このグラフから、強連結成分を抽出し、さらにモジュラリティを算出する。併せて、グラフのエッジに対して媒介中心性を算出することで、異種特徴間の関連性に注目したサブグラフの構造とその包含関係が明らかとなり、Web 映像集合の階層構造が抽出される。第 4 章では、第 3 章で説明した手法を高速化する手法について説明する。具体的には、クラスタリングに基づき選択された少数の特徴ベクトルに対して正準相関分析を施した後、複数の Web 映像から成るノードを持つ新たなグラフを構築する。これにより、大量の Web 映像が少数のノードで処理可能となり、Web 映像集合の階層構造が効率的に抽出される。さらに、第 5 章では、第 4 章で説明した手法を高速化かつ高精度化する手法について説明を行う。具体的には、Web 映像の潜在特徴間の類似性を表すグラフに対し、その局所的な構造に注目した解析アルゴリズムを適用することで、Web 映像集合の階層構造抽出の大幅な高速化が実現される。さらに、第 3 章および第 4 章で説明した手法は、大量の Web 映像を対象とした場合に事前のスクリーニングが必要なため、階層構造抽出の精度が劣化する問題が存在する。第 5 章の提案手法は、事前のスクリーニングを不要とすることで、この問題を解決し、階層構造抽出の高精度化についても実現できる。最後に、第 6 章において、本論文のまとめと今後の課題について述べる。

以上を要約すると、本論文は、高度な Web 映像検索を実現するために、Web 映像の画像、音響特徴、タイトルや説明のテキスト特徴、さらには Web 映像間のリンク関係の特徴という異種特徴間の関連性を活用し、Web 映像集合の階層構造を抽出する手法を提案している。本論文の貢献は、異種特徴の関連性を活用することで、Web 映像集合の階層構造を高精度に抽出可能とし、得られた階層構造を用いることで、適切なキーワードがクエリとして入力されない場合においても、望む Web 映像の高精度な検索を実現する点にある。また、YouTube から収集された Web 映像に対して実験を行うことで、提案手法の有効性が示されている。