



HOKKAIDO UNIVERSITY

Title	複数特徴を利用した映像意味解析とリランキングの導入による検索の高精度化に関する研究〔論文内容及び審査の要旨〕
Author(s)	吉田, 壮
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第12197号
Issue Date	2016-03-24
Doc URL	https://hdl.handle.net/2115/61838
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	doctoral thesis
File Information	Soh_Yoshida_review.pdf, 審査の要旨



学位論文審査の要旨

博士の専攻分野の名称 博士 (情報科学)

氏名 吉田 壮

審査担当者 主 査 教 授 長谷山 美紀
副 査 教 授 山本 強
副 査 教 授 坂本 雄児

学位論文題名

複数特徴を利用した映像意味解析とリランキングの導入による検索の高精度化に関する研究
(Accurate Video Retrieval Using Multiple Feature-Based Video Semantic Analysis and Reranking)

高速・大容量な通信回線の普及に伴い、インターネットによる映像配信サービスの利用が増加している。このような状況の中、ユーザが望む映像を効率的に検索し閲覧するための技術開発が急務とされている。一般に、映像検索の結果は、予め映像にメタデータを付与し、それとクエリとして与えるキーワードとの適合度に基づく順位付けに従い提示される。このとき、Web 上の映像は極めて膨大であることから、手動によるメタデータの付与は現実的でなく、自動化が必須である。現状の研究では、映像を同一意味内容のまとまりを持つ区間に分割した後、各区間に何が映っているのかを識別することでメタデータの付与を可能とする。これにより自動で得られたメタデータとクエリを比較して適合度を算出し順位付けすることで検索結果が提示される。しかしながら、メタデータの自動付与に誤りが生じる場合があり、そのような場合に、順位付けに影響を与え、検索精度の劣化を招き、ユーザが望む映像が下位に配置されてしまう。この問題を解決するための研究として、順位付けにおいて下位に存在する映像の中でユーザが望む映像の順位が高くなるように並べ替えるリランキング手法が存在する。

そこで、本論文ではまず、映像の複数特徴から得られるシーン境界の判定結果を協調的に用いるシーン分割手法の提案を行っている。提案手法では、各ジャンルの中でも映像構造が明確に存在する歌謡番組に焦点を当て、歌謡番組固有の構造を利用する。この構造に基づいてショットの階層クラスタリングを行い、その結果から取得される有向グラフのカット辺を探索する。画像特徴については、シーンが遷移するとき、同時に各シーンの中心となる出演者も変化することを考慮し、新たに出演者の変化を検出する特徴を顔認識法の結果から定義する。各々の特徴から得られる結果を統合する協調処理により歌謡番組固有の繰り返し構造を検出することで、シーン分割の高精度化を実現する。次に、撮像物体の自動選定処理を導入した識別器の提案を行っている。具体的に、Support Vector Machine (以降、SVM) を用いた識別超平面を算出し、特徴空間における特徴ベクトルと超平面との距離に基づいて、撮像物体の存在する領域を内包する矩形を自動検出する処理を提案する。提案手法では、背景領域に存在する特徴を除去することで、撮像物体の領域に限定した特徴抽出を可能とし、新たな識別超平面を算出する。これにより、背景や他の物体による悪影響を抑制した物体識別が実現される。最後に、映像特徴とタグ情報の両者の類似性に注目したリランキング手法の提案を行っている。提案手法では、映像のリンク関係を基に、映像とタグそれぞれの類似性を表現可能なネットワークを構築し、そのネットワーク上において、映像とタグ、各々の類似性を伝播させる。さらに、キーワード検索による順位が、類似性の高い映像ノード間で近くなるように目的関数を与え、その最小化問題を解くことで、ネットワークの整合性を保つようなリランキングが可能となり、望む映像が上位に配置されることが可能となる。

本論文では、第 2 章で従来手法について説明し、現状の問題点と本研究が解決すべき課題を明確にする。第 3 章では、映像特徴の協調利用による映像の構造に注目したシーン分割手法を提案する。第 4 章では、物体領域の自動選定処理を導入した SVM に基づく物体識別手法を提案する。第 5 章では、映像特徴とタグ情報の類似性の両者を考慮したネットワーク構造に基づくリランキング手法を提案し、有効性を実験により示す。最後に、第 6 章において、本研究の成果について要約し、論文全体のまとめとする。

これを要するに、著者は、複数特徴を利用した映像の構造に注目したシーン分割手法と物体領域の自動選定に基づく撮像物体の識別手法の提案による映像意味解析の改良、さらには、リランキング手法の新たな提案により映像検索の高精度化を実現している。また、本手法を実データに適用した実験を行うことにより、その有効性を示しており、本研究を通じて、情報科学分野の発展に貢献するところ大なるものがある。よって著者は、北海道大学博士 (情報科学) の学位を授与される資格あるものと認める。