

報告番号	※甲	第	号
------	----	---	---

## 主 論 文 の 要 旨

論文題目 Webコミュニティ活動に基づく映像アノテーションとその応用に関する研究  
氏 名 山本 大介

## 論 文 内 容 の 要 旨

本論文では、Web上で共有されている映像コンテンツをより高度に利用するために、映像コンテンツに対するアノテーション技術、とりわけ、Web上の一般ユーザによるアノテーション技術とそれに基づく応用技術について提案する。

映像コンテンツはテキストコンテンツなどとは異なり、テキスト検索のように意味内容を考慮した内容検索などの高度な処理を行うことが一般に困難である。映像共有サービスの普及とともに、Web上で扱われる映像コンテンツは氾濫しつつあり、いかに効率よくこれらのコンテンツの管理・検索・要約・宣伝などの応用を実現するかという問題が顕在化してきた。

従来、とりわけ、基礎研究の分野において、映像コンテンツの高度利用に対するアプローチは自動解析技術のみに基づく手法に傾倒してきた。映像解析技術は映像コンテンツの構造情報を理解することは可能であっても、映像コンテンツの意味を理解することは困難であり、将来にわたって信号処理技術のみでの実現は困難であると予想する。我々の提案手法は、人手によって意味内容を映像コンテンツと関連付けることを支援する仕組みである。いかにアノテーションのコストを低く、また、より正確なアノテーションを関連付ける仕組みを作るかということが問題になる。

そこで、本論文では、効率よく映像コンテンツに対するアノテーションを付与する手法について検討する。そのために、まず、これらの手法の基盤となる、アノテーションモデルについて検討する。具体的には、アノテーションに記述する内容の形式によって、様々なアノテーション型があることを示す。さらに、これらのモデルに基づき、効率よく映像コンテンツに対するアノテーションを付与する手法として以下の3つのアプローチを提案し、それぞれ具体的なアノテーションシステムを開発した。

1. VAE: オントロジーに基づく半自動アノテーション
2. iVAS: 閲覧者によるオンライン映像アノテーション
3. Synvie: Web コミュニティ活動に基づく映像アノテーション

オントロジーに基づく半自動アノテーションとは、映像コンテンツに対する2種類のアノテーションの作成を支援する仕組みである。1つは、映像に対する構造アノテーションであり、信号処理技術とその解析結果の修正を支援する仕組みを備える。もう1つは、映像に対する意味アノテーションであり、映像に関するオントロジーと映像の要素を関連付けることによって行うオントロジー型アノテーションの仕組みを備える。具体的なシステムとしてVAEを開発した。機械にできることは極力機械によって、人間にしかできないことは人間によって、協調的にアノテーションを作成する仕組みである。この方式のアノテーションは、専門家が商用コンテンツに対して詳細にアノテーションを付与することに適している一方、アノテーション作成コストが比較的高いのでWebで共有されているようなコストをかけることが困難な映像コンテンツに対して同様の仕組みを適用することは困難であるという欠点もある。

閲覧者によるオンライン映像アノテーションとは、映像の閲覧者にアノテーションの参加を促す仕組みである。そのために、映像シーンを話題とした掲示板型コミュニケーションを支援するユーザコメント型アノテーションの仕組みを備えるiVASというシステムを開発した。また、不特定多数の閲覧者がアノテーションを作成する方式であるためアノテーションの信頼性が一様ではない問題があるため、アノテーション信頼度と呼ばれる評価手法を導入し、それに基づくアノテーションの選別手法を提案した。さらに、閲覧者のアノテーション履歴から閲覧者の嗜好を評価するための仕組みとして、印象距離と呼ばれる指標を提案し、これらの評価を行った。この方式はアノテーション作成のためのコストが低いという特徴があり、より多くのコンテンツに対して適用可能である。

Web コミュニティ活動に基づく映像アノテーションとは、基本的には閲覧者によるアノテーション方式と同じであるものの、映像を話題としたコミュニティ活動は映像の閲覧時だけではなく、ブログなどでも映像を話題とした記事が書かれている点に着目し、それらの情報もアノテーションとして獲得した。具体的には、映像シーンを引用したブログエントリーの執筆を支援する、引用型アノテーションの仕組みを提案した。引用型アノテーションでは、映像に関する意味情報だけではなく、引用の共起関係に基づく構造情報の獲得が可能である。さらに、これらの仕組みを実装したシステムとしてSynvieを開発し、比較的大規模な公開実験を行った。2006年7月から一般公開し、284名の登録ユーザと、223個の映像コンテンツと、3534個のユーザコメントに基づくアノテーションを含む19182個のアノテ

ションの取得に成功した。これらの取得されたアノテーションの定量的な分析を行うことによって、引用型アノテーションの優位性を示した。これにより、iVASよりもより良いアノテーションの収集が可能になるという利点がある。

次に、Synvieによって獲得されたユーザアノテーションに基づいた応用について提案する。具体的には、Webコミュニティ活動に基づくアノテーションの仕組みによって獲得されたアノテーションに基づく検索システムを提案する。アノテーションのテキスト情報からタグ集合を生成することによって映像シーンに対する意味情報と、複数の映像シーンを同時引用することから得られる共起関係やアノテーションの参照頻度を用いた構造情報を利用する。これらのアノテーションは必ずしも信頼性が高いとはいえないという欠点がある。これらの欠点を補うために、タグクラウドの仕組みを利用した映像シーン検索システムを開発した。このシステムは、既存のタグクラウドと同様に、人間が提示されたタグを探索的に検索する仕組みであり、Synvieによって獲得されたアノテーションに適した仕組みである。本システムを利用した評価実験を行うことによって、必ずしも信頼性が高くないアノテーションを用いた応用システムにおいても、アノテーションの有用性を確認した。

Webユーザやコミュニティによるコミュニティ活動から、映像コンテンツに対するアノテーションを間接的に作成するというアプローチが本論文の最大の新規性にあたる部分である。また、実際に実証実験や応用システムを開発することまで行うことによって、本論文の有用性を検証した。それにより、Webコミュニティからアノテーションを獲得することは有用であることを示した。Weblogを書く、映像を閲覧しながら掲示板型のコミュニケーションをするなどといったWebコミュニティ活動は、ユーザによって自発的に行われる活動であり、実質的なコストを0として捉えることができる。そういった意味で、映像コンテンツの自動解析技術と同等に、人的なコストが掛からず意味情報の抽出が可能になる点が新しい。しかしながら、Webコミュニティから取得されたアノテーションのみで十分な精度が得られるとはいえない。将来的には、信号処理技術的なアプローチと、人間が介在したアノテーション的なアプローチとを融合させた仕組みを実現させることが重要である。これらの仕組みが実現することによって、コンテンツ作成者にとっても、コンテンツ閲覧者にとっても、コンテンツ権利者にとっても有益なシステムになることを期待する。