

平成25年度 情報工学コース卒業研究報告要旨

問瀬 研究室	氏 名	内 藤 貴
卒業研究題目	時空間特徴を用いた YouTube 上のビデオ内の暴力シーン検出	
<p>暴力的な表現は成長過程の子供に悪い影響を与え、現在では多くのメディアで暴力シーンは規制され、子供に暴力表現の影響のない環境が作られている。テレビでは各放送局が自主的に規制を行い、また映画やテレビゲームなどの分野には倫理委員会が設置された。しかし、近年では YouTube やニコニコ動画などインターネット上でも動画が視聴可能になり、非常に多くの動画コンテンツを見ることができるようになった。これらの動画サイトでは、暴力シーンの規制が投稿する個人の良心によるのでまだ不十分であり、膨大な数が一日に投稿されるため手動で検出することが難しい。そこで暴力シーンを自動で検出する必要がある。</p> <p>本研究では、web 上の映像から自動で暴力シーンを削除、規制することや、手動によるコストを削減するために、暴力シーンを検出することを目的とする。</p> <p>これまでも行動認識の分野における暴力シーン検出の研究は数多く存在する。多くの研究が映画やテレビ番組などの高画質で音声のズレやノイズのない映像から色ヒストグラムや音声を特徴量として扱い識別に用いた。本研究では、解像度が低く、ノイズが含まれ、あるいは音声がずれる動画も含む YouTube の動画に対して暴力シーンの検出を行う。</p> <p>そこで本研究では時空間特徴量を用いた暴力の検出法を提案する。提案手法は以下のようになる。まず、入力 YouTube 上の動画が与えられ、ショット検出によって 0.5 秒から 2 秒のショットに切り分ける。ショット検出では映像内の特徴としてフレーム間の色ヒストグラム距離と Edge Change Ratio を用いてショット検出を行う。次に、分割したショットに対して時空間特徴量を用いて暴力の抽出を行う。フレームごとに輝度の勾配から特徴点を取り出し、時間方向に追跡することで HOG 特徴量を算出する。取り出した HOG 特徴量は、Bag-of-Features 法を用いて、あらかじめクラスタ中心を求めた状態から 400 次元の特徴ベクトルを作成する。すべてのショットに対して特徴ベクトルを作成したのち、SVM にて 2 クラスの識別を行い、暴力と非暴力のショットを分類し出力する。</p> <p>提案手法の評価実験には、YouTube 上の Amazon の洋画アクション映画ランキング上位 40 位の映画の予告の動画を用いた。ショット検出で生成したショットに対して、殴打、爆発、銃撃、カーチェイス、流血、武器での攻撃の 6 種類の暴力のクラスとその他の非暴力のクラスに手動で分類し、データセットとして実験に用いた。合計 1280 個のショット、うち暴力ショット 250 個を準備した。対象クラスとそれ以外のクラスの 2 クラス分類器を各クラスごとに学習し各クラスごとに学習しテストを行った。そして、42 回の交差検定により評価した。評価は混同行列を用いて行った。</p> <p>実験の結果、全体で約 6 割の暴力を検出できた。しかし、タグごとの検出精度は訓練データの数が少なかったため低い精度を示した。今後は色ヒストグラムなどの時空間特徴量以外の特徴量をタグ別に組み合わせて精度の向上を目指す。本研究でも用いたアクションのみのデータセットに、SF、戦争、ホラーなども加え検証を行うことが課題に挙げられる。</p>		