

## 修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 電気通信 学研究科 情報工学 専攻 博士前期課程		
氏 名	野口顕嗣	学籍番号	0831031
論 文 題 目	Web 動画マイニングのための時空間特徴抽出手法の提案		
要 旨	<p>本研究では、Web 動画分類のための時空間特徴抽出手法の提案、その特徴を利用した Web 動画分類、及び動作認識を行う。現在行われている動作認識の研究のほとんどが、「カメラは動かない」、「動作している人間が一人」、「高い解像度」のような制限のある映像下における認識を行っており、実際に、低解像度、雑多な背景ノイズ、頻繁なカメラモーション、という特徴のある Web 動画における動作認識を行っている研究は少ない。そこで本研究では映像に一切の制限のない Web 動画の取り扱いを考える。このような動画に対する動作認識において重要なことは、以下の 4 つの事である。(1) 大量のデータを取り扱うので計算時間が短い。(2) 撮影位置の変化に頑健であること。(3) カメラモーションに対する対処。(4) 高い分類精度。本論文ではこれらを実現するため、新しい特徴を提案し、実際に分類を行うことでその実用性を検証する。</p> <p>本研究において提案する時空間特徴は、静止画のための局所特徴抽出手法である SURF によって抽出された局所特徴量に基づく点と、局所特徴が抽出された点の微小時間の動きを特徴化したものである。この手法は処理が単純で、計算時間が少ないが、高い分類精度が期待できる。また動きの回転に関しても頑健であり、Web 動画分類に適している。更にこの時空間特徴に視覚特徴、動き特徴を、Multiple Kernel Learning(MKL) で統合し、動作を認識する手法について提案する。</p> <p>実験として、他の手法との比較を行うために標準データセットである KTH データセットにおける動作認識を行った。その結果として時空間特徴のみで 91.0%、特徴を統合することで 94.0%の分類率を達成した。これは最新手法の結果と匹敵する値である。</p> <p>次に教師信号ありの Web 動画ショット分類を行った。"batting"など特定のキーワードで収集された動画をショット分割し、それぞれのショットから特徴を抽出する。その後全てのショットの中からランダムに選択をし、学習データを作成する。それらを support vectormachine(SVM) で学習、分類し、そのショットのランキング付けを行った。その結果として上位 200 位までの適合率が 57.2%という十分な結果を得ることが出来た。その後、教師信号なしの分類を試みた。全ショットを pLSA, k-means クラスタリングで分類し、その有効性について検証をした。</p> <p>最後に Web 動画ショットにおける動作認識を行った。データセットとして本研究で独自に構築した Our Youtube データセットと、公開されている Web 動画データセットの二種類のデータセットを使用した。Our Youtube データセットは六種類の動作を、公開データセットは 11 種類の動作を含んでいる。結果として、Our Youtube データセットでは 83.3%、公開データセットでは 68.7%であった。最新手法における公開データセットの分類率は 71.1%で、これに匹敵する結果であるといえる。</p>		