

論文審査の結果の要旨および担当者

報告番号	※ 甲 第	号
------	-------	---

氏 名 吉 穎

論 文 題 目

Subjective Video Attribute Recognition: Labeling, Training and Explaining

(映像における主観的属性の認識：ラベリング、学習、および解釈)

論文審査担当者

主 査 名古屋大学大学院情報学研究科 教授 森 健策

委 員 名古屋大学大学院情報学研究科 教授 井手 一郎

委 員 名古屋大学大学院情報学研究科准教授 小田 昌宏

委 員 中京大学大学院工学研究科 教授 目加田慶人

委 員 立命館大学大学院情報理工学研究科 教授 加藤ジェーン

吉穎君提出の論文「Subjective Video Attribute Recognition: Labeling, Training and Explaining」は、ビデオ画像の主観的属性の認識に関して、データセットの作成、属性認識、認識システムの説明性に関して述べており、6章から構成される。

第1章では、本論文における研究の背景、目的、ならびに位置づけについて述べている。ビデオ属性の認識は、客観的属性認識と主観的属性認識の2つに大分される。客観的属性認識では、ビデオ内の動作、人・物体などが客観的な基準に基づき認識される。一方、主観的属性認識は人間の知覚認知によって定義され、個人の経験や感情によって影響を受けることから困難なタスクであることを示している。

第2章では、ビデオ属性の客観的属性認識と主観的属性認識に関して、データセット作成、主観的属性認識手法、説明可能な機械学習手法について関連研究を述べている。

第3章では、暴力認識に特化したビデオ映像の主観的属性認識のためのデータセット作成について述べている。一般的に、データアノテーションにおいては、アノテータは客観的な事実に基づきラベル情報を与え、複数のアノテータが与えるラベル情報の多数決によって最終的なラベルが与えられる。しかし、主観的属性認識においては、正確なラベルを与えることが難しい。そこで、本手法では、ペアワイズ比較法の導入によりこの問題を解決している。本章では、暴力シーンを例にとり、ビデオクリップ間の相対的な評価を繰り返すことで、主観的属性認識のためのデータセットを作成する手法を示している。

第4章では、主観的属性を認識するための効率的な方法を紹介している。一般的にビデオ映像認識においては、End-to-Endのニューラルネットワークが利用される。しかしながら、このネットワークから抽出される特徴ベクトルは属性認識を行うには十分ではない。そこで、ビデオ映像から社会関係を認識するための最も重要な領域を抽出し、得られた領域に関する情報を活用することで、主観的な社会的関係を認識するフレームワークを提案している。このフレームワークの学習処理を行うため、主観的な社会的関係属性と客観的な社会的関係属性の両方がアノテーションされたデータセットも作成している。本章で示す方法により主観的属性の認識において有用な客観的情報を取り入れることが可能となり、主観的属性の認識精度が向上している。

第5章では、空間時間概念に基づいたニューラルネットワークの判断に関する説明方法について述べている。現状、ビデオ映像を解析するニューラルネットワークの判断の説明方法においてはピクセルレベルの解釈手法が主流であるが、この方法ではより高いレベルの説明を生成することは難しい。本論文では、空間-時間概念に基づく説明フレームワークを提案し、人間が理解できる空間時間概念を導入することでニューラルネットワークの判断の説明を行っている。本フレームワークでは、ビデオから抽出されたスーパーボクセルによって空間-時間概念をグループ化し、各概念の重要性スコアを評価している。高いスコアはネットワークがより注意を払っていることを表しており、ネットワークの主観的属性の認識過程を理解することに役立つ。

第6章では、本論文を総括し、今後の課題と展望について述べている。

以上のように、本論文は、コンピュータによるビデオ映像の主観的属性の認識手法について述べたものである。これらは、ビデオ映像の自動認識技術の向上やその利用を通じた暴力シーンを含む映像の自動レーティングにもつながる。以上から、吉穎君提出の論文は学術ならびに産業の向上に寄与すると考え、博士（情報学）の学位論文に相応しいものと判断する。