

論文の内容の要旨

論文題目 物理モデルと事例の確率的統合による画像認識と映像生成

氏名 島野美保子

人間の目に直接は見えないものを見る技術が実現するならば、我々は何を見たいか。見るターゲットは、人それぞれの様々な事象になるだろう。

この「見えないものを見る技術」は、まさにコンピュータビジョン研究の、カメラで撮影した「見たまま」の画像や映像データから、被写体である対象世界がどうなっているのかという、例えば、形状、反射特性、照明条件等のそのままでは「見えない」有用な情報を明らかにする技術と言える。つまり、コンピュータビジョンは、ある画像から外界の3次元世界を推定する情報処理であり、物理法則の成り立つ3次元実世界において光学に従って結像される画像と逆問題の関係にある。従って、コンピュータビジョン問題を、物理法則に従う現象を表現した物理モデルを立て解析的に解く手法が多く提案されてきた。

最近では、より一般的な自然画像や映像、複雑な現象を対象にすることが求められるようになり、撮影画像から、より現実的で複雑な対象の性質を精度良く推定する手法の確立が必要とされている。しかし、物理モデルに基づいた手法のみでは、対象世界を正しく表現しきれないため、あるいは未知な項目が多数あるため、不良設定問題となり解けない場合が多く存在する。

本論文は、このように物理モデルのみでは不良設定問題となってしまう、解析的なアプローチのみでの求解が難しいコンピュータビジョン問題に対して、事例に基づいた補完情報を利用することで、尤もらしい解を得るアプローチを研究する。具体的には、顔画像認識と高時間分解能映像の生成という、2つの応用分野にこのアプローチを適用した以下の4テーマに取り組む。

昨今では、各人物について多くの登録画像を用意することは手間がかかるため、各人物につき登録画像1枚の任意照明下における顔認識が求められるようになってきた。もし、1枚の登録画像から様々な照明下の顔画像を合成できるならば、これらをあたかもたくさん登録画像があるとみなし、顔認識の学習用画像に用いることで実現できる。しかし、このたった1枚の登録画像から新たな照明下の見えを推測することは、形状・反射係数・照明を同時推定しなければならず、制約不足の問題である。第一のテーマでは、まず、このような制約不足の問題に対して、物理モデルとしての反射モデルに加えて、その事例による補完情報として、別人の顔画像集合から得られる統計量を利用するというアプローチを提案し、その有効性を検証する。人の顔の形や反射特性は類似しているため、

顔のクラスを考えることによって、制約不足の部分の補完情報として利用し、様々な見えを推測することが可能になる。

更に、同じく顔認識への応用例として、各アプリケーションにより人為的に定められたカテゴリが線形分離不能な場合であっても、正しくカテゴリに分類する方法を提案する。本手法は、Fisherの判別基準と赤池情報量規準を統合し、与えられるカテゴリに合わせて線形分離可能なクラス構成を生成して、特徴空間を最適化する。これによって、モデルと事例から得られる補完情報について、確率分布が不一致であったため線形分離不能になっていたところを、一致するように統合を図ることができ、分類能力の向上が可能になる。

続いて、「高時間分解能映像の生成」の課題に取り組む。近年、動画のストリーミング配信等も広まり、映像がより身近なものとなってきたため、遠く離れたあんな場所のこんな状況を見たいという要望に加え、「もっとぶれの少ない映像を見たい」、「もっと詳細に見たい」、「もっと明るい映像を見たい」等、ユーザの所望する映像生成も多種多様になってきた。このようなユーザの所望する映像生成として「映像の高画質化」に注目し、「動きぶれを含む映像から、動きぶれを低減した高時間分解能映像の生成」を目指す。露光時間を長くとった連続露光で撮影された動きぶれを含む映像では、1つのフレームを分解するような高時間分解能化が必要である。ある高時間分解能のフレームから低時間分解能フレームへの生成は一意に決まるが、その逆の低時間分解能フレームから高時間分解能フレームに分解することは、いろいろな動き方の可能性があり、1つには定まらない。つまり、このような高時間分解能化は、フレームの分割の仕方が無数に存在している不良設定問題である。こうしたフレーム分割の不定性の問題を、異なる時間分解能の映像間に存在する自己相似な事例データを利用することにより、取り得る複数の解の中から、尤もらしいフレームの分割方法を選択することで解決する。加えて、物理モデルとしては、生成した映像と入力映像との整合性がとれるよう、各フレームの露光時間は連続的であるという映像復元モデルを導入し、最大事後確率推定によって、高時間分解能映像を復元する。

前述の高時間分解能化を活用するにあたり、映像符号化技術の利用も期待される。そこで、高時間分解能映像の生成手法の応用先として、ユーザに元映像よりも高画質の映像を提供するという映像符号化技術を提案する。伝送量には制限があるため、画質を改善しつつ高効率に伝送できるよう、トレードオフの関係にある符号量と画質を最適化する必要がある。このため、自己相似性による高時間分解能化の技術を利用することにより、フレーム毎に時間分解能と量子化パラメタの最適な組み合わせを最短経路探索問題として求める。更に、時間分解能等のコスト単調性に基づいた高速な探索、および符号化パラメタの最適化も可能にする。

以上のように、本論文では、2つの応用分野において、従来ならば不良設定として解けない問題についても尤もらしい解を得られる具体的な手法を提案する。更に、その根

幹を成す，物理モデルと事例に基づいた補完情報を確率的に統合するというアプローチは，コンピュータビジョン分野における重要な要素技術を成すと考える．本研究により得られた成果は，より一般的な自然画像や映像，複雑な現象を対象を対象にする上でも，幅広い分野への活用が期待される．