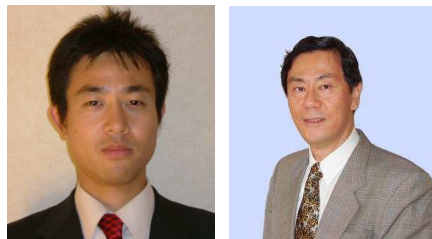


複数運動分離の教師なし学習による多段階最適化 (“Multi-stage optimization for multi-body motion segmentation”, Vol. E87-D, No. 7, pp. 1935–1942, July 2004)



菅谷保之 金谷健一
岡山大学

研究の目的

この研究の目的は、ビデオ画像中の異なる物体を識別することです。M フレームに渡って運動をする特徴点の軌跡は、その x, y 座標を並べた $2M$ 次元空間の1点として表せますが、異なる運動の軌跡はその $2M$ 次元空間の異なる空間に含まれるという事実が識別の原理です。

実ビデオ画像の謎

我々は以前からこれを理論的に解析して、分離アルゴリズムを構築し、シミュレーションによって予想通りの性能を確認していました。そこでこれを実ビデオ画像で実践しようとしたのが本論文の発端です。

ところが実際に画像を撮影して実験してみると、どうしても望ましい分離ができません。シミュレーションでうまくいくのに、実画像できないというのは不思議でした。

シミュレーションの落とし穴

ようやく気がついたのは、シミュレーションでは状況をなるべく一般的になるように設定していたことでした。

しかし、その「一般的な状況」をビデオカメラで撮影することが不可能であることに気がつきました。人の動きも車の動きも、実際の物体は平面的な並進が主で(映画「マトリックス」のような)複雑な運動をするものは見つかりません。

常識的には単純な運動は解析が簡単で、複雑な運動は解析が困難のように思えますが、運動の分離に関しては逆で、複雑な運動ほど識別しやすく、単純な運動ほど区別しにくいことがわかりました。

モデルの選択

結局、実ビデオ画像を解析するには、一般の運動が「退化」した実際に生じやすい単純な運動を

モデル化しなければならないという結論に至りました。例えば平面上の並進であるというようにモデルを特定すれば、そのモデルに対して教師なし学習(EMアルゴリズム)を適用すれば精度の高い分離ができるかと期待されます。

しかし、事前にはどのような退化が起きているのか不明です。最初は何らかのしきい値やモデル選択規準を用いて退化の種類を特定しようと試みましたが、どれも望ましい結果が得られませんでした。

多段階最適化

さまざまな試行錯誤を重ねた結果、最終的に、非常に単純な運動から、次第に一般化した運動へと階層的なモデルを仮定し、最も単純なモデルを仮定して分離を行ない、その分離を初期値として次のモデルを仮定する分離を行ない、さらに次のモデルを適用するという多段階最適化に到達しました。

こうすると単純な運動は複雑な運動の特殊な場合なので、もしあるモデルが成立していれば、そこで適切な分離が行なわれ、それをさらに一般的なモデルに当てはめてももはや解が変化しないこととなります。こうして、どういう退化が生じているかを特定することなく、あらゆる場合に分離が可能となる方法が実現されました。

今回の受賞を励みにして、今後も地道に研究活動を続けていきたいと思えます。

*提案手法のプログラムと用いたデータを公開しています。
<http://www.suri.it.okayama-u.ac.jp/>