KCF Trackerによる高速物体追跡*

内海 佑麻[†] (61602452)

2019/07/23

概 要

モダンな物体追跡手法として，KCF Tracker (Kernelized Correlation Filter Tracker) を紹介し，実装をおこなった．KCF Tracker は，データ行列に巡回性を持たせることにより，DCT による対角化が可能となり，処理にかかるストレージおよび計算量を大幅に削減することができる．さらに，カーネルトリックにより

目 次

1 はじめに

1.1 Optical flow

異なる時点の 2 画像間で，各画素値の移動量を表現したベクトルを Optical Flow という．Optical Flow により特徴点を検出することで，たとえば動画内の特定の物体を追跡することができる．まず，一般論として，勾配法 (Gradient-based method) による Optical Flow の求解手順を定式化する．勾配法では，Taylor 展開による近似を用いるため,「連続する 2 画像での対象物の移動が微小であること」を前提としている．

*慶應義塾大学 理工学研究科 2019 春学期 ”ComputerVision 特論” 最終レポート
[†]情報工学科 4 年，Email: uchiumi@ailab.ics.keio.ac.jp

勾配法による Optical Flow の球解　画像 I における画素 (x, y) の時刻 t における画素値を $I(x, y, t)$，時間 δt に対する画素値の移動量を $(\delta x, \delta y)$ とおく．移動前後で対象画素の画素値が不変であると仮定すると，

$$I(x, y, t) = I(x + \delta x, y + \delta y, \delta t) \tag{1}$$

が成り立つ．さらに，画素値の変化が滑らかであると仮定すると，1 次までの Taylor 展開により，以下の近似式を得る．

$$I(x, y, t) \approx I(x, y, t) + \frac{\partial I}{\partial x}\delta x + \frac{\partial I}{\partial y}\delta y + \frac{\partial I}{\partial t}\delta t \tag{2}$$

両辺を δt で割って，整理すると，

$$\frac{\partial I}{\partial x}\frac{\delta x}{\delta t} + \frac{\partial I}{\partial y}\frac{\delta y}{\delta t} + \frac{\partial I}{\partial t} = 0 \tag{3}$$

$\delta t \to 0$ とすれば，

$$\frac{\partial I}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial I}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial I}{\partial t} = 0 \tag{4}$$

ここで，画素 (x, y) の Optical Flow: $(u, v) = \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}\right)$ に着目する．$I_x = \frac{\partial I}{\partial x}$，$I_y = \frac{\partial I}{\partial y}$，$I_t = \frac{\partial I}{\partial t}$，とおけば，$(u, v)$ がみたすべき方程式[1]は，次式のようになる．

$$I_x u + I_y v + I_t = 0 \tag{5}$$

しかし，上式は 2 つの未知数 u, v に対して方程式は 1 つであり，解が定まらない[2]．そのため，上式に加えて，いくつかの仮定 (制約条件) を導入して Optical Flow: (u, v) を球解する手法が提案されている．古典的で重要な手法として，Lucas–Kanade 法や Horn–Schunck 法がある．

- Lucas–Kanade 法
  「各画素の近傍では，移動方向が相関する」という仮定をおく．画像中の特定の点 (領域) に絞って追跡を行うような用途に適しているため，sparse 型と呼ばれる．

- Horn–Schunck 法
  「各画素は，変化率が最小となる方向へ移動する」という仮定をおく．画像中の画素全体の動きを解析するような用途に適しているため，dense 型と呼ばれる．

2　理論

2.1　Ridge 回帰

SVM などの洗練されたモデルに近い性能をもち，かつ単純な閉形式解[3]をもつ Ridge 回帰を用いる．訓練データ $(\mathbf{x}_i, y_i)_{i=1\cdots N}$ に対する Ridge 回帰は，以下のように定式化される．

$$\max_{\mathbf{w}} \sum_i (f(\mathbf{x}_i) - y_i)^2 + \lambda\|\mathbf{w}\|^2, \quad f(\mathbf{x}_i) = \mathbf{w}^{\mathrm{T}}\mathbf{x}_i \tag{6}$$

[1] これを，OpticalFlow の拘束式と呼ぶ．
[2] これを，Aperture Problem(窓問題) という．
[3] 四則演算と初等関数の合成関数によって表せる解を，閉形式解 (closed-form solution) という．

Ridge 回帰の閉形式解は，以下のようになる．

$$\mathbf{w} = (X^{\mathrm{T}}X + \lambda I)^{-1} X^{\mathrm{T}}\mathbf{y} \tag{7}$$

さらに，転置行列 X^{T} を，エルミート転置 $X^{\mathrm{H}} := (X^*)^{\mathrm{T}}$ によって，拡張すると，

$$\mathbf{w} = (X^{\mathrm{T}}X + \lambda I)^{-1} X^{\mathrm{H}}\mathbf{y} \tag{8}$$

となる．[4]

2.2 巡回行列と DFT

$$P = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 1 \end{bmatrix} \tag{9}$$

巡回行列を定義する．

$$C(\mathbf{x}) = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & \cdots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \cdots & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{bmatrix} \tag{10}$$

2.3 カーネルトリック

3 実装

4 考察

5 結論

- 連続状態空間の制御タスクにおける好奇心駆動の探索戦略: VIME を提案．
- 状態遷移の前後で，KL 情報量が最大化するように報酬関数を設計．
- モデルパラメータ θ の事後分布を推定する際，変分推論 (VI) を用いる．
- モデルパラメータ θ は，探索環境のダイナミクスを表現している．
- VIME は，ヒューリスティックを上回る実験結果を示した．

[4] A^* は，行列 A の複素共役を表す．

参考文献