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

*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

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．

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．

[1] これを，OpticalFlow の拘束式と呼ぶ．
[2] これを，Aperture Problem(窓問題) という．

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．n 個の訓練データ $(\mathbf{x}_i, y_i)_{i=1\cdots n}$ に対する Ridge 回帰は，以下のように定式化される．

$$\min_{\mathbf{w}} \sum_{i=1}^{n} (f(\mathbf{x}_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2, \quad f(\mathbf{x}_i) = \mathbf{w}^{\mathrm{T}} \mathbf{x}_i \tag{6}$$

データ行列

$$X = \begin{bmatrix} x_1^{\mathrm{T}} \\ x_2^{\mathrm{T}} \\ x_{3^{\mathrm{T}}} \\ \vdots \\ x_N^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{3n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \cdots & x_{nn} \end{bmatrix} \tag{7}$$

を用いると，Ridge 回帰の閉形式解は，以下のようになる．

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

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，

$$\mathbf{w} = (X^{\mathrm{H}} X + \lambda I)^{-1} X^{\mathrm{H}} \mathbf{y} \tag{9}$$

となる．[4]

2.2 巡回行列と DFT

注目点の移動パスをベクトル $\mathbf{x} = (x_1, \cdots x_n)^{\mathrm{T}} \in \mathbb{R}^n$ とおく．このとき，cyclic shift operator として，$n \times n$ 行列

$$P = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 1 \end{bmatrix} \tag{10}$$

を定義すると，ベクトル $\mathbf{x}$ の巡回置換は，

$$\{P^u \mathbf{x} \mid u = 0, \cdots n-1\} \tag{11}$$

[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．

[4]A^* は，行列 A の複素共役を表す．

と得られる．すなわち，あるベクトル $\mathbf{x}$ から，n 個の仮想サンプルを作成することができる．そこで，ある信号 $\mathbf{x}$ を，P を用いて拡張すると，

$$C(\mathbf{x}) = \begin{bmatrix} (P^0\mathbf{x})^{\mathrm{T}} \\ (P^1\mathbf{x})^{\mathrm{T}} \\ (P^2\mathbf{x})^{\mathrm{T}} \\ \vdots \\ (P^{n-1}\mathbf{x})^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & \cdots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \cdots & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{bmatrix} \tag{12}$$

を得られる．ここで，$C(\mathbf{x})$ は巡回行列 (Circulant matrix) だから，これは DFT(Discrete Fourier Transform) によって対角化可能である．すなわち，

$$\forall \mathbf{z} \in \mathbb{R}^n, \quad F(\mathbf{z}) = \sqrt{n}F\mathbf{z} \tag{13}$$

をみたす，ある定まった DFT 行列 F と，$\mathbf{x}$ の周波数成分

$$\hat{\mathbf{x}} = F(\mathbf{x}) = \sqrt{n}F\mathbf{x} \tag{14}$$

を用いると，$C(\mathbf{x})$ は次式のように固有値分解できる．

$$C(\mathbf{x}) = F\mathrm{diag}(\hat{\mathbf{x}})F^H \tag{15}$$

2.3 巡回行列を使った Ridge 回帰

$C(\mathbf{x})$ をデータ行列として，Ridge 回帰を行う，中心化していない共分散行列は，DCT 行列 F を用いて次式のように表せる．

$$C(\mathbf{x})^{\mathrm{H}}C(\mathbf{x}) = F\mathrm{diag}(\hat{\mathbf{x}}^*)F^HF\mathrm{diag}(\hat{\mathbf{x}})F^H \tag{16}$$

さらに，要素積 $\odot$ を用いて

$$F^HF = (F^*)^{\mathrm{T}}F = I \tag{17}$$

$$\mathrm{diag}(\hat{\mathbf{x}}^*) \cdot \mathrm{diag}(\hat{\mathbf{x}}) = \mathrm{diag}(\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}}) \tag{18}$$

が成り立つから，

$$C(\mathbf{x})^{\mathrm{H}}C(\mathbf{x}) = F\mathrm{diag}(\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}})F^H \tag{19}$$

となる．ここで，$\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}}$ は，信号 $\mathbf{x}$ の自己相関 (auto correlation) またはパワースペクトル (power spectrum) と呼ばれる量である．すなわち，P によって空間方向にシフトさせた信号どうしの共分散を表している．以上から，データ行列として $C(\mathbf{x})$ を用いた Ridge 回帰の閉形式解は，次のようになる．

$$\begin{aligned} \hat{\mathbf{w}} &= \left(C(\mathbf{x})^{\mathrm{H}}C(\mathbf{x}) + \lambda I\right)^{-1}C(\mathbf{x})^{\mathrm{H}}\hat{\mathbf{y}} && (20) \\ &= \mathrm{diag}\left(\frac{\hat{\mathbf{x}}^*}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda}\right)\hat{\mathbf{y}} && (21) \\ &= \frac{\hat{\mathbf{x}}^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda} && (22) \\ & && (23) \end{aligned}$$

よって，逆 DFTF^{-1} を用いて，w が求められる．

$$\mathbf{w} = F^{-1}(\hat{\mathbf{w}}) \tag{24}$$

2.4 Kernel Ridge 回帰

任意の正定値関数 $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^{\mathrm{T}}\phi(\mathbf{x}_j)$ に対して，

$$\mathbf{w} = \sum_{i=1}^{n} \alpha_i \phi(\mathbf{x}_i) \tag{25}$$

とおけば，グラム行列

$$K = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & k(\mathbf{x}_1, \mathbf{x}_2) & k(\mathbf{x}_1, \mathbf{x}_3) & \cdots & k(\mathbf{x}_1, \mathbf{x}_n) \\ k(\mathbf{x}_2, \mathbf{x}_1) & k(\mathbf{x}_2, \mathbf{x}_2) & k(\mathbf{x}_2, \mathbf{x}_3) & \cdots & k(\mathbf{x}_2, \mathbf{x}_n) \\ k(\mathbf{x}_3, \mathbf{x}_1) & k(\mathbf{x}_3, \mathbf{x}_2) & k(\mathbf{x}_3, \mathbf{x}_3) & \cdots & k(\mathbf{x}_3, \mathbf{x}_n) \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ k(\mathbf{x}_n, \mathbf{x}_1) & k(\mathbf{x}_n, \mathbf{x}_2) & k(\mathbf{x}_n, \mathbf{x}_3) & \cdots & k(\mathbf{x}_n, \mathbf{x}_n) \end{bmatrix} \tag{26}$$

を用いて，Representer 定理より，

$$f(\mathbf{x}_i) = \mathbf{w}^{\mathrm{T}}\mathbf{x}_i = \sum_{j=1}^{n} \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) \tag{27}$$

$$\|\mathbf{w}\|^2 = \alpha^{\mathrm{T}} K \alpha \tag{28}$$

が成り立つ．すなわち，n 個の訓練データ $(\mathbf{x}_i, y_i)_{i=1\cdots n}$ に対する Kernel Ridge 回帰は，以下のように定式化される．

$$\min_{\alpha} \sum_{i=1}^{n} (f(\mathbf{x}_i) - y_i)^2 + \lambda \alpha^{\mathrm{T}} K \alpha, \quad f(\mathbf{x}_i) = \sum_{j=1}^{n} \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) \tag{29}$$

Kernel Ridge 回帰の閉形式解は，以下のようになる．

$$\alpha = (K + \lambda I)^{-1}\mathbf{y} \tag{30}$$

2.5 巡回行列を使った Kernel Ridge 回帰

Ridge 回帰と同様に，ベクトル $\mathbf{x}$ とその巡回シフトに対するカーネルは，P と $k(\cdot, \cdot)$ を用いて

$$\{k(\mathbf{x}, P^u\mathbf{x}) \mid u = 0, \cdots n-1\} \tag{31}$$

となる．よって，ベクトル $\mathbf{k}^{\mathbf{xx}}$ を，

$$\mathbf{k}^{\mathbf{xx}}_{\mathbf{i}} = k(\mathbf{x}, P^{i-1}\mathbf{x}) \quad (i = 1, \cdots n) \tag{32}$$

と定義し，これを用いた巡回行列

$$C(\mathbf{k}^{\mathbf{xx}}) = \begin{bmatrix} (P^0\mathbf{k}^{\mathbf{xx}})^{\mathrm{T}} \\ (P^1\mathbf{k}^{\mathbf{xx}})^{\mathrm{T}} \\ \vdots \\ (P^{n-1}\mathbf{k}^{\mathbf{xx}})^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} k(\mathbf{x}, P^0\mathbf{x}) & k(\mathbf{x}, P^1\mathbf{x}) & \cdots & k(\mathbf{x}, P^{n-1}\mathbf{x}) \\ k(\mathbf{x}, P^{n-1}\mathbf{x}) & k(\mathbf{x}, P^0\mathbf{x}) & \cdots & k(\mathbf{x}, P^{n-2}\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{x}, P^1\mathbf{x}) & k(\mathbf{x}, P^2\mathbf{x}) & \cdots & k(\mathbf{x}, P^0\mathbf{x}) \end{bmatrix} \tag{33}$$

を得られる．さらに，中心化していない共分散行列は，DCT 行列 F を用いて次式のように表せる．

$$C(\mathbf{k}^{\mathbf{xx}})^{\mathrm{H}} C(\mathbf{k}^{\mathbf{xx}}) = F\mathrm{diag}\left(\hat{\mathbf{k}^*}^{\mathbf{xx}} \odot \hat{\mathbf{k}}^{\mathbf{xx}}\right) F^{\mathrm{H}} \tag{34}$$

以上から，データ行列として $C(\mathbf{k}^{\mathbf{xx}})$ を用いた Kernel Ridge 回帰の閉形式解は，次のようになる．

$$\hat{\alpha} = (C(\mathbf{k}^{\mathbf{xx}}) + \lambda I)^{-1}\mathbf{y} \tag{35}$$

$$= \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}^{\mathbf{xx}} + \lambda} \tag{36}$$

よって，逆 DFTF^{-1} を用いて，w が求められる．

$$\alpha = F^{-1}(\hat{\alpha}) \tag{37}$$

3 実装

RBF kernel を用いる．

$$\mathbf{k}^{\mathbf{xx}'} = \exp\left\{-\frac{1}{\sigma^2}\left(\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 - 2F^{-1}(\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}}')\right)\right\} \tag{38}$$

4 考察

5 結論

- foo
- bar
- buz

参考文献