Generative models for human activity recognition

ROY team: Ilya Zharikov, Roman Isachenko, Artem Bochkarev

Skolkovo Institute of Science and Technology
Machine Learning course

March 20, 2017.

Project goal

Aim

Classification model for complex-structured objects.

Problem

Initial object has no appropriate feature description.

Applications:

- image processing;
- signal classification;
- topic modelling;
- *time series analysis*.

Related work

1. Wang W. et al. Human activity recognition using smart phone embedded sensors // *International Joint Conference on Neural Networks*. 2014.
2. Kwapisz, Jennifer R., Gary M. Weiss, and Samuel A. Moore. Activity recognition using cell phone accelerometers. // *ACM SigKDD Explorations Newsletter*. 12(2): 74-82. 2011.
3. Kuznetsov M. P., and Ivkin N. P. Time series classification algorithm using combined feature description. // *Journal of Machine Learning and Data Analysis*. 2015.

Problem Statement

- $s \in \mathcal{S}$ - complex structured object
- $y \in Y$ - class label

Task

Given: $\mathfrak{D} = \{(s_i, y_i)\}_{i=1}^{m}$ recover

$$y = f(s) \quad \forall s \in \mathcal{S}.$$

Approach

Suppose $f = g \circ h$, where

- $h(s) : \mathcal{S} \to H \in \mathbb{R}^n$, $\mathbf{h}_i = h(s_i)$;
- $g(\mathbf{h}, \theta)$ — parametric map from H into Y (classification model).

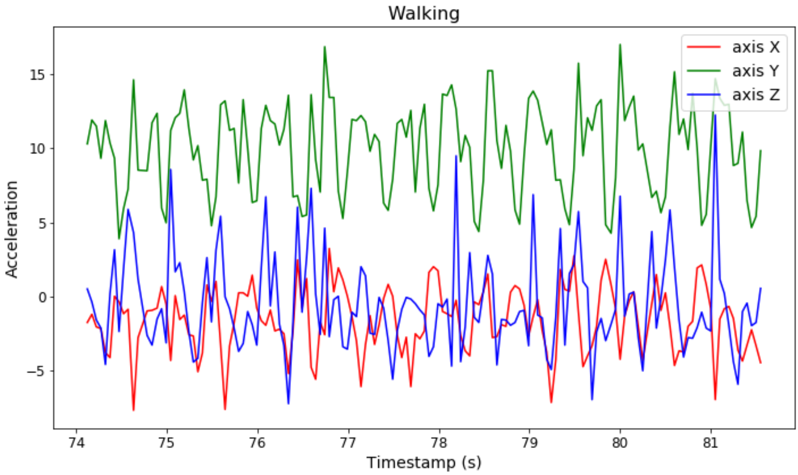
Optimal parameters

1. Choice of feature map $h(s)$ by
   - prior (expert) knowledge;
   - minimizing error functional.
2. Classification for $\{(\mathbf{h}_i, y_i)\}_{i=1}^m$

$$\theta^* = \arg\min_\theta L(g, \theta, \mathfrak{D}).$$

E.g. $g(\mathbf{h}, \theta)$ - classification model,
θ - model parameters,
$L(g, \theta, \mathfrak{D})$ - classification error function.

Time series example



Expert functions

Prior knowledge about the objects could allow to choose the features.

Feature description

$\mathbf{h}_i = h(s_i) \in \mathbb{R}^{40}$ — different statistics:

- average acceleration;
- standard deviation;
- mean absolute deviation;
- ...

Autoregressive model

Data generation hypothesis

Time series $s = (x_1, \dots, x_T)$ is generated by autoregressive model

$$\hat{x}_t = w_0 + \sum_{j=1}^{n} w_j x_{t-j}.$$

Feature description

$$h(s_i) = \mathbf{w}^* = \arg \min_{\mathbf{w} \in \mathbb{R}^{n+1}} \sum_{j=n+1}^{T} \|x_j - \hat{x}_j\|^2.$$

Singular Spectrum Analysis

Trajectory matrix for time series $s = (x_1, \dots, x_T)$:

$$\mathbf{X} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ x_2 & x_3 & \dots & x_{n+1} \\ \dots & \dots & \dots & \dots \\ x_{T-n+1} & x_{T-n+2} & \dots & x_T \end{pmatrix}.$$

Feature description

$$h(s_i) = (\lambda_1, \dots, \lambda_n),$$

where $\{\lambda_i\}_{i=1}^n$ are eigenvalues of the matrix $\mathbf{X}^\top\mathbf{X}$, obtained by SVD decomposition

$$\mathbf{X}^\top\mathbf{X} = \mathbf{V} \cdot \mathrm{diag}(\lambda_1, \dots, \lambda_n) \cdot \mathbf{V}^\top.$$

(?!) Splines

Data

WISDM[1]

- # of objects: 4321
- # of classes: 6

USC-HAD[2]

- # of objects: 13620
- # of classes: 12

[1] http://www.cis.fordham.edu/wisdm/
[2] http://sipi.usc.edu/HAD/

Results

Results

Feature union

Conclusion