Neurocomputing 128 (2014) 273–284

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Extreme learning machine towards dynamic model hypothesis in fish ethology research

Rui Nian [a], Bo He [a,*], Bing Zheng [a], Mark van Heeswijk [b], Qi Yu [b], Yoan Miche [b], Amaury Lendasse [b,c,d,e]

[a] College of Information Science and Engineering, Ocean University of China, 266003 Qingdao, China
[b] Department of Information and Computer Science, Aalto University, FI-00076 Aalto, Finland
[c] IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain
[d] Computational Intelligence Group, Computer Science Faculty, University of the Basque Country, Paseo Manuel Lardizabal 1, Donostia/San Sebastián, Spain
[e] Arcada University of Applied Sciences, 00550, Helsinki, Finland

ARTICLE INFO

Article history:
Received 29 August 2012
Received in revised form
21 February 2013
Accepted 20 March 2013
Available online 6 November 2013

Keywords:
Extreme learning machine
Fish ethology
Video surveillance system
Dynamic state space
Color distribution
Object recognition

ABSTRACT

In this paper, we present one dynamic model hypothesis to perform fish trajectory tracking in the fish ethology research and develop the relevant mathematical criterion on the basis of the Extreme Learning Machine (ELM). It is shown that the proposed scheme can conduct the non-linear and non Gaussian tracking process by multiple historical cues and current predictions – the state vector motion, the color distribution and the appearance recognition, all of which can be extracted from the single-hidden layer feedforward neural network (SLFN) at diverse levels with ELM. The strategy of the hierarchical hybrid ELM ensemble then combines the individual SLFN of the tracking cues for the performance improvements. The simulation results have shown the excellent performance in both robustness and accuracy of the developed approach.

1. Introduction

The 21 century is an Era of the 'Ocean'. There has been a growing trend towards the deployment of the ocean blueprint all over the world. Fish ethology, an emerging discipline to explore the inherent nature of the movement, behaviors and activities for either wild or cultured fish, has shown great prospects in the aquaculture, fisheries, and other marine related surveys and applications [1–4].

Among a mass of fish behavior descriptions, the fish trajectory tracking is essential and fundamental. Tracking the fish can be extremely complex due to the random fish movements, all kinds of shape variations, the non rigid or articulated nature, the partial and full occlusions, the scene illumination changes, the multiple viewpoints, the poor image quality, the projection of the 3D world on 2D images, the real-time processing requirements, and so on.

So far, object tracking in literature mainly focus on classical approaches such as the background subtraction, the inter-frame difference, the optical flow computation, the Kalman filtering, the particular filtering, the mean-shift algorithms, etc., and the primary differences come from the type of the object representation, the feature extraction, the motion modeling, the shape and appearance, and the context that the tracking is performed [5–8].

In practice, the fish activities often correspond to a complicated, nonlinear, and non-Gaussian dynamic system, and the Bayesian estimation theory can be seen to be a philosophically optimal solution [9–11]. In cases that there is no prior knowledge available for the overall functional form of probability distribution beforehand, the scheme of the dynamic model approximation from the observations is the central concern for the fish ethology research.

In the context of the machine learning, Artificial Neural Networks (ANN) has been playing the dominant roles due to benefits on generalization, flexibility, nonlinearity, fault tolerance, self organization, adaptive learning, and computation in parallel, while the bottlenecks such as the overfitting, local minima, time consuming etc., can probably restrict the scalability in the conventional implementations [12–18]. Recently, the Extreme learning machine (ELM) has made a great breakthrough in the single-hidden layer feedforward neural network (SLFN) instead of the classical gradient-based algorithms [19,20]. The achievements of ELM tend to provide better generalization performance than the traditional approaches, and seek straightforward solutions mathematically with inspiring abilities such that the hidden node parameters can be randomly chosen and the output weights can be analytically determined at extremely fast learning speed and

* Corresponding author.
E-mail address: bhe@ouc.edu.cn (B. He).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.03.054

with least human intervene. By far, ELM has not only developed for the conventional SLFN [21,22], but also extended to the generalized SLFN that need not be the neuron alike [23,24]. There are a great many ELM variations that have been proposed and have led to the state-of-the-art results in many applications both for the regression problem and the pattern recognition problem [25–32]. Random hidden layer feature mapping based ELM improves the stability in the calculation of the output weights according to the ridge regression theory [25–27]. The Kernel based ELM makes use of the corresponding kernel instead of the hidden layer feature mapping itself, and the dimensionality of the hidden layer feature space needs not be specified either [25,28,29]. The fully complex ELM can use the fully complex activation function directly with the universal approximation capability [20]. The incremental ELM (I-ELM) shows an efficient and practical way to construct the incremental feed-forward network with a wide type of activation functions, where the hidden nodes can be added one by one [23,24]. The online sequential ELM (OS-ELM) can learn the training data sequentially not only one-by-one but also chunk by chunk and discard the observations as soon as the learning procedure has already been done [20]. The optimally-pruned ELM (OP-ELM) starts with a large network and then eliminates the hidden nodes that have low relevance to the learning [33,34]. ELM ensembles are widely used to improve single network's performance with a plurality consensus scheme [30–32].

In this paper, we develop a scheme of dynamic model hypothesis by means of ELM learning algorithm for fish ethology research. The rest of the paper is organized as follows: In Section 2, the background of the Bayesian estimation theory will be briefly introduced. In Section 3, the basics of the ELM are outlined. In Section 4, the dynamic model with ELM is developed in detail. In Section 5, the simulation and discussion will be stated in support of the developed scheme. Section 6 comes to the conclusions.

2. Bayesian estimation

In principle, for the fish trajectory tracking, the Bayesian sequential estimation can seek an optimal model [9–11]. The general dynamic model can be considered as the state transition and the state measurement,

$$X_t = f(X_{t-1}, U_{t-1}), \quad Y_t = h(X_t, R_t) \tag{1}$$

where t is the time index, X_t refers to the state variable of the fish propagated by the possibly nonlinear process model f over time, such as the position, velocity, etc., h is the observation model mapping the state variable X_t to the corresponding observation variable Y_t, U_t and R_t are respectively the process noise and the measurement noise that are roughly supposed as white Gaussian noise. The state prediction function is formulated as

$$p(X_t|Y_{1:t-1}) = \int p(X_t|X_{t-1})p(X_{t-1}|Y_{1:t-1})dX_{t-1} \tag{2}$$

and the state variable can be updated by the posterior density $p(X_t|Y_{1:t})$ inferred from the prior density $p(X_t|Y_{1:t-1})$,

$$p(X_t|Y_{1:t}) = \frac{p(Y_t|X_t)p(X_t|Y_{1:t-1})}{p(Y_t|Y_{1:t-1})} \tag{3}$$

where $Y_{1:t} = \{Y_1, Y_2, \ldots, Y_t\}$ constitutes the complete solution to the sequential estimation problem, and the normalizing constant is

$$p(Y_t|Y_{1:t-1}) = \int p(Y_t|X_t)p(X_t|Y_{1:t-1})dX_t \tag{4}$$

In most cases, the above analytic solution can not be well determined in a direct way. Therefore, the particle filtering by the Monte Carlo simulation is usually taken to approximate the optimal Bayesian estimation recursively.

Let the posterior density function be characterized by N random samples $\left\{X_t^i, \omega_t^i\right\}_{i=1}^N$,

$$p(X_t|Y_{1:t}) \approx \sum_{i=1}^{N} \omega_t^i \delta(X_t - X_t^i) \tag{5}$$

where $\{X_t^i, i=0,\ldots,N\}$ is a set of support points with the associated weights $\{\omega_t^i, i=1,\ldots,N\}$.

In the sequential importance sampling, the recursive estimate for the importance weights of the particle i can be derived by

$$\omega_t^i = \omega_{t-1}^i \frac{p(Y_t|X_t^i)p(X_t^i|X_{t-1}^i)}{q(X_t^i|X_{0:t-1}^i, Y_{1:t})} \tag{6}$$

where $q(X_{0:t}|Y_{1:t})$ is an easy to sample and proposal density, $X_{0:t}$ is the historical state variable and $Y_{1:t}$ is the corresponding observation, and the estimated state can be approximated by $\hat{X}_t \approx \sum_{i=1}^N \omega_t^i X_t^i$.

The optimal importance density function will minimize the variance of the true weights, $q(X_t|X_{t-1}^i, Y_t) = p(X_t|X_{t-1}^i, Y_t)$. In practice, it is often convenient to choose the importance density to be the prior,

$$q(X_t|X_{t-1}^i, Y_t) = p(X_t|X_{t-1}^i) \tag{7}$$

Then the importance weight is updated as

$$\omega_t^i = \omega_{t-1}^i p(Y_t|X_t^i) \tag{8}$$

The fish swims as a quite complicated, nonlinear, and non-Gaussian dynamic system in the sea. Theoretically, it is possible to learn the fish behaviors of any complexity if the training database is quite adequate, while in case that the size of samples is in fact far from an optimum [35], we need to offer an efficient and consistent approximation of the dynamic model to estimate the posterior probability density function. The classic particle filtering has been developed as one of the most common and powerful technique for such a system [7,8], while the learning scale, the static reference model, the degeneracy problem, the sample impoverishment, the space dimensionality etc., may still hinder our implementations.

3. The basics of ELM

So far, ELM learning has attracted more and more attention in machine learning since proposed, which announces a novel learning framework that significantly improves generalization performance at surprisingly fast speed, needless of mathematically predetermined internal knowledge [19,20].

Suppose that there are Q arbitrary distinct training samples $\{I_q, O_q\}_{q=1}^Q$, with the input $I_q = [I_{q1}, I_{q2}, \ldots, I_{qd}]' \in R^d$ and the expected output $O_q = [O_{q1}, O_{q2}, \ldots, O_{qk}]' \in R^N$. In general, a standard SLFN can be modeled as the following matrix format,

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{O}$$
$$\mathbf{H}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{\tilde{Q}}, b_1, \ldots, b_{\tilde{Q}}, \boldsymbol{I}_1, \ldots, \boldsymbol{I}_Q) = \begin{bmatrix} g(\boldsymbol{a}_1 \cdot \boldsymbol{I}_1 + b_1) & \cdots & g(\boldsymbol{a}_{\tilde{Q}} \cdot \boldsymbol{I}_1 + b_{\tilde{Q}}) \\ \vdots & g(\boldsymbol{a}_i \cdot \boldsymbol{I}_q + b_i) & \vdots \\ g(\boldsymbol{a}_1 \cdot \boldsymbol{I}_Q + b_1) & \cdots & g(\boldsymbol{a}_{\tilde{Q}} \cdot \boldsymbol{I}_Q + b_{\tilde{Q}}) \end{bmatrix}_{Q \times \tilde{Q}}$$
$$\boldsymbol{\beta} = [\boldsymbol{\beta}_1 \ \cdots \ \boldsymbol{\beta}_i \ \cdots \ \boldsymbol{\beta}_{\tilde{Q}}]'_{k \times \tilde{Q}}, \quad \mathbf{O} = [\boldsymbol{O}_1 \ \cdots \ \boldsymbol{O}_q \ \cdots \ \boldsymbol{O}_Q]'_{k \times Q} \tag{9}$$

where $\mathbf{H}$ is defined as the hidden layer output matrix, $\tilde{Q}$ is the number of hidden nodes, $g(x)$ stands for the activation function, $\boldsymbol{a}_i = [a_{i1}, a_{i2}, \ldots, a_{id}]'$ is the weight vector connecting the ith hidden node and the input nodes, b_i is the threshold of the ith hidden