Robot Learning from Demonstration and the Problem of Target Defense by Team of Unmanned Surface Vehicles

Vladimir Popov

Department of Intelligent Systems and Robotics, Ural Federal University, 620083 Ekaterinburg, Russia

Copyright © 2014 Vladimir Popov. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

In this paper, we consider some approaches to the task-level robot learning from demonstration for the target defense by a team of unmanned water surface vehicles. We assume that the problem of target defense by a team of unmanned water surface vehicles represented as a problem of learning of rhythmic motor primitives. We consider for the problem neural networks as oscillators to learn rhythmic motor tasks. Also, we use the approximate period problem and introduce the approximate period problem for a set of strings with a set of restrictions. In this paper, we present experimental results for different synthetic test data set for stationary and moving targets.

Keywords: unmanned water surface vehicle, task-level robot learning from demonstration, rhythmic motor primitive, approximate period problem

Unmanned water surface vehicles (USV) have been extensively studied over the last thirty years [1]. It is clear that USVs can be used to solve many practical tasks. In particular, we can mention sea-bed mapping and ocean sampling tasks [2], environmental monitoring [3], cooperative surveillance [4], search and rescue [5], and patrolling and protecting different areas [6]. It should be noted that USVs can significantly increase the capability of other surface or
underwater vehicles by continuously following them in marine applications. In particular, USVs can be used for guarding of sensitive areas in naval missions [6].

Investigation of various problems of robot learning and self-learning forms a fundamental part of contemporary computer science (see e.g. [7] – [9]). In particular, we can mention the task-level robot learning from demonstration (see e.g. [10] – [12]). It should be noted that different motor skills of robots can be represented by motor primitives. Problems of motor primitives have received great attention recently (see e.g. [11, 12]). In particular, motor primitives can be considered as tasks for demonstrations. In this paper, we consider the task-level robot learning from demonstration for the target defense by a team of unmanned water surface vehicles. Motor primitives can be divided into two major groups: discrete motor primitives and rhythmic motor primitives (see e.g. [12]). There are a large number of rhythmic motor tasks that could be learned [11]. In particular, the problem of target defense by a team of unmanned water surface vehicles can be modeled as a problem of learning of rhythmic motor primitives.

Now we define basic notions and consider the problem formulation. Let

\[ \{V[1], V[2], \ldots, V[n]\} \]

be the set of USVs that used for the target defense. The learning of the task of the target defense can be considered as learning of a set of rhythmic motor tasks

\[ \mathcal{T} = \{T[1], T[2], \ldots, T[n]\} \]

where \( T[i] \) is the rhythmic motor task that should be performed by \( V[i] \) for all \( i \in \{1, 2, \ldots, n\} \). Since we consider missions of target defense as rhythmic motor tasks, we can assume that \( T[i] \) is periodically repeating sequence of motor primitives for all \( i \in \{1, 2, \ldots, n\} \). Therefore, without loss of generality, we can assume that

\[ T[i] = t[i, 1]t[i, 2]\ldots t[i, p] \]

for some \( p \) and set of motor primitives \( \{t[i, j] \mid 1 \leq i \leq n, 1 \leq j \leq p\} \). In this case, we can consider \( j \) in \( t[i, j] \) as point in time. Therefore, for any \( j \), some subsequence of

\[ t[1, j]t[2, j]\ldots t[n, j] \]

should represent a line of defense. It is easy to see that training data can be represented as a set \( \mathcal{L} \) of lines of defense and a set \( \mathcal{U} = \{U[i] \mid 1 \leq i \leq n\} \) of demonstrations, where

\[ U[i] = u[i, 1]u[i, 2]\ldots u[i, m_i] \]

is a sequence of actions that represents the demonstration for \( T[i] \).
It is clear that we can use some traditional approach to the task-level robot learning from demonstration. In particular, we consider neural networks as oscillators to learn rhythmic motor tasks $T[i]$ (see e.g. [13]). From other hand, we can use the approximate period problem AP (see [14]) to learn rhythmic motor tasks $T[i]$ (see [7]). In this paper, we extend the approach of [7] to the case of set of tasks.

Throughout the paper, for a fixed alphabet $\Sigma = \{a_1, a_2, \ldots, a_r\}$, we assume that $\Gamma$ is augmented alphabet $\Gamma = \Sigma \cup \{a_0\}$ where $a_0$ is a special symbol. Symbol $a_0$ is called an indel. We assume that $a_0$ represents the insertion or deletion of a particular symbol in one string relative to another. Note that the alignment notation can be used to illustrate a comparison between strings. Let $X = \{X_1, X_2, \ldots, X_k\}$ be a set of strings over $\Sigma$. A multiple alignment of $X$ is a set $A = \{A_1, A_2, \ldots, A_k\}$ of strings over $\Gamma$ such that $|A_i| = n$, $A_i$ is a copy of $X_i$ into which $n - |X_i|$ copies of $a_0$ have been inserted, for all $1 \leq i \leq k$. We assume that $\delta$ is a distance function that specified by a penalty matrix. In this case, a penalty matrix $M$ specifies the substitution cost for each pair of letters. Also, $M$ specifies the insertion and deletion cost for each letter. The weighted edit distance between $S$ and $T$ is the minimum cost to convert $S$ to $T$ using a penalty matrix. For given two strings $X$, $P$ and distance function $\delta$, we can define approximate periods as follows (see [14]). If there exists a partition of $X$ into disjoint blocks of substrings, $X = P_1 \ldots P_q$, $P_i \neq \epsilon$, $r > 2$, such that $\delta(P, P_i) < K$ for $1 \leq i < q$, and $\delta(P', P_q) < K$ where $P'$ is some prefix of $P$, we say that $P$ is a $K$-approximate period of $X$ (see [14]).

We consider the following problem.

**THE APPROXIMATE PERIOD PROBLEM FOR A SET OF STRINGS WITH A SET OF RESTRICTIONS (APSR):**

**INSTANCE:** A finite alphabet $\Gamma$, sets of strings $L$ and $U$ over $\Gamma^*$, a penalty matrix $M$, and a positive integer $K$.

**QUESTION:** Is there a set of strings $T \in \Gamma^*$ such that

- $|T[i]| = |T[j]|$, $1 \leq i < j \leq n$;
- $T[i]$ is a $K$-approximate period of $U[i]$, $1 \leq i \leq n$;
- for all $1 \leq j \leq p$, there exist $L \in \mathcal{L}$ such that $L$ is a subsequence of $t[1,j]t[2,j] \ldots t[n,j]$?

In our computational experiments we consider two types of synthetic test data sets, with stationary target and moving target. Let $ST[n,g,d]$ be the synthetic test data set for the stationary target where $n$ is the number of USVs, $g$ is the number of potential intruders, $d$ is the average proportion of obstacles in the general area of the environment. Similarly, let $MT[n,g,d]$ be the synthetic test data set for the moving target where $n$ is the number of USVs, $g$ is the number of potential intruders, $d$ is the average proportion
of obstacles in the general area of the environment. In our computational experiments we consider neural networks as oscillators to learn rhythmic motor tasks, the problem AP, and the problem APSR. Let $D_{NN}[S]$ be the average number of demonstration that needed for successful learning of the task of target defense by neural networks for test set $S$. Similarly, we can define $D_{AP}[S]$ and $D_{APSR}[S]$ for problems AP and APSR, respectively. Let

$$N_{AP}[S] = \frac{D_{AP}[S]}{D_{NN}[S]}$$

and

$$N_{APSR}[S] = \frac{D_{APSR}[S]}{D_{NN}[S]}.$$ 

It is clear that we can consider $N_X[S]$ as a measure of the performance of $X$. Selected experimental results for test sets $T[1]$ and $T[2]$ are given in Tables 1 – 4.

<table>
<thead>
<tr>
<th>$d$</th>
<th>$0%$</th>
<th>$10%$</th>
<th>$20%$</th>
<th>$40%$</th>
<th>$70%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = g = 1$</td>
<td>20.42 %</td>
<td>17.18 %</td>
<td>11.34 %</td>
<td>3.74 %</td>
<td>9.88 %</td>
</tr>
<tr>
<td>$n = 2$, $g = 1$</td>
<td>16.21 %</td>
<td>15.87 %</td>
<td>11.29 %</td>
<td>3.51 %</td>
<td>7.43 %</td>
</tr>
<tr>
<td>$n = 3$, $g = 1$</td>
<td>13.11 %</td>
<td>11.98 %</td>
<td>11.23 %</td>
<td>3.46 %</td>
<td>3.08 %</td>
</tr>
<tr>
<td>$n = 6$, $g = 1$</td>
<td>7.38 %</td>
<td>4.41 %</td>
<td>3.14 %</td>
<td>2.57 %</td>
<td>2.22 %</td>
</tr>
<tr>
<td>$n = 9$, $g = 1$</td>
<td>2.82 %</td>
<td>1.06 %</td>
<td>0.95 %</td>
<td>0.92 %</td>
<td>0.89 %</td>
</tr>
<tr>
<td>$n = 3$, $g = 3$</td>
<td>0.99 %</td>
<td>0.86 %</td>
<td>0.85 %</td>
<td>0.84 %</td>
<td>0.83 %</td>
</tr>
<tr>
<td>$n = 6$, $g = 3$</td>
<td>0.92 %</td>
<td>0.84 %</td>
<td>0.83 %</td>
<td>0.83 %</td>
<td>0.82 %</td>
</tr>
<tr>
<td>$n = 9$, $g = 3$</td>
<td>0.87 %</td>
<td>0.83 %</td>
<td>0.81 %</td>
<td>0.78 %</td>
<td>0.77 %</td>
</tr>
</tbody>
</table>

Table 1: Experimental results for $N_{AP}[ST[n,g,d]]$.

<table>
<thead>
<tr>
<th>$d$</th>
<th>$0%$</th>
<th>$10%$</th>
<th>$20%$</th>
<th>$40%$</th>
<th>$70%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = g = 1$</td>
<td>7.33 %</td>
<td>6.84 %</td>
<td>6.67 %</td>
<td>6.41 %</td>
<td>6.29 %</td>
</tr>
<tr>
<td>$n = 2$, $g = 1$</td>
<td>4.78 %</td>
<td>4.71 %</td>
<td>4.08 %</td>
<td>1.12 %</td>
<td>0.95 %</td>
</tr>
<tr>
<td>$n = 3$, $g = 1$</td>
<td>0.91 %</td>
<td>0.89 %</td>
<td>0.89 %</td>
<td>0.89 %</td>
<td>0.88 %</td>
</tr>
<tr>
<td>$n = 6$, $g = 1$</td>
<td>0.76 %</td>
<td>0.75 %</td>
<td>0.74 %</td>
<td>0.73 %</td>
<td>0.73 %</td>
</tr>
<tr>
<td>$n = 9$, $g = 1$</td>
<td>0.72 %</td>
<td>0.72 %</td>
<td>0.72 %</td>
<td>0.72 %</td>
<td>0.71 %</td>
</tr>
<tr>
<td>$n = 3$, $g = 3$</td>
<td>0.28 %</td>
<td>0.26 %</td>
<td>0.25 %</td>
<td>0.23 %</td>
<td>0.23 %</td>
</tr>
<tr>
<td>$n = 6$, $g = 3$</td>
<td>0.21 %</td>
<td>0.21 %</td>
<td>0.21 %</td>
<td>0.21 %</td>
<td>0.21 %</td>
</tr>
<tr>
<td>$n = 9$, $g = 3$</td>
<td>0.19 %</td>
<td>0.18 %</td>
<td>0.18 %</td>
<td>0.17 %</td>
<td>0.17 %</td>
</tr>
</tbody>
</table>

Table 2: Experimental results for $N_{AP}[MT[n,g,d]]$. 
<table>
<thead>
<tr>
<th>$d$</th>
<th>0 %</th>
<th>10 %</th>
<th>20 %</th>
<th>40 %</th>
<th>70 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = g = 1$</td>
<td>0.112 %</td>
<td>0.097 %</td>
<td>0.086 %</td>
<td>0.064 %</td>
<td>0.058 %</td>
</tr>
<tr>
<td>$n = 2$, $g = 1$</td>
<td>0.108 %</td>
<td>0.092 %</td>
<td>0.079 %</td>
<td>0.069 %</td>
<td>0.073 %</td>
</tr>
<tr>
<td>$n = 3$, $g = 1$</td>
<td>0.105 %</td>
<td>0.088 %</td>
<td>0.074 %</td>
<td>0.063 %</td>
<td>0.068 %</td>
</tr>
<tr>
<td>$n = 6$, $g = 1$</td>
<td>0.101 %</td>
<td>0.085 %</td>
<td>0.072 %</td>
<td>0.060 %</td>
<td>0.062 %</td>
</tr>
<tr>
<td>$n = 9$, $g = 1$</td>
<td>0.097 %</td>
<td>0.076 %</td>
<td>0.054 %</td>
<td>0.042 %</td>
<td>0.045 %</td>
</tr>
<tr>
<td>$n = 3$, $g = 3$</td>
<td>0.075 %</td>
<td>0.062 %</td>
<td>0.051 %</td>
<td>0.038 %</td>
<td>0.033 %</td>
</tr>
<tr>
<td>$n = 6$, $g = 3$</td>
<td>0.072 %</td>
<td>0.041 %</td>
<td>0.032 %</td>
<td>0.026 %</td>
<td>0.022 %</td>
</tr>
<tr>
<td>$n = 9$, $g = 3$</td>
<td>0.057 %</td>
<td>0.038 %</td>
<td>0.021 %</td>
<td>0.018 %</td>
<td>0.019 %</td>
</tr>
</tbody>
</table>

Table 3: Experimental results for $N_{APSR}[ST[n, g, d]]$.

<table>
<thead>
<tr>
<th>$d$</th>
<th>0 %</th>
<th>10 %</th>
<th>20 %</th>
<th>40 %</th>
<th>70 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = g = 1$</td>
<td>0.032 %</td>
<td>0.028 %</td>
<td>0.026 %</td>
<td>0.015 %</td>
<td>0.014 %</td>
</tr>
<tr>
<td>$n = 2$, $g = 1$</td>
<td>0.027 %</td>
<td>0.012 %</td>
<td>0.011 %</td>
<td>0.009 %</td>
<td>0.008 %</td>
</tr>
<tr>
<td>$n = 3$, $g = 1$</td>
<td>0.015 %</td>
<td>0.007 %</td>
<td>0.006 %</td>
<td>0.003 %</td>
<td>0.0034 %</td>
</tr>
<tr>
<td>$n = 6$, $g = 1$</td>
<td>0.0052 %</td>
<td>0.0023 %</td>
<td>0.0012 %</td>
<td>0.0009 %</td>
<td>0.0008 %</td>
</tr>
<tr>
<td>$n = 9$, $g = 1$</td>
<td>0.0038 %</td>
<td>0.0021 %</td>
<td>0.0009 %</td>
<td>0.00007 %</td>
<td>0.00007 %</td>
</tr>
<tr>
<td>$n = 3$, $g = 3$</td>
<td>0.0009 %</td>
<td>0.0007 %</td>
<td>0.0007 %</td>
<td>0.0005 %</td>
<td>0.0005 %</td>
</tr>
<tr>
<td>$n = 6$, $g = 3$</td>
<td>0.0008 %</td>
<td>0.0004 %</td>
<td>0.0004 %</td>
<td>0.0004 %</td>
<td>0.0004 %</td>
</tr>
<tr>
<td>$n = 9$, $g = 3$</td>
<td>0.0003 %</td>
<td>0.0002 %</td>
<td>0.0001 %</td>
<td>0.000028 %</td>
<td>0.00002 %</td>
</tr>
</tbody>
</table>

Table 4: Experimental results for $N_{APSR}[MT[n, g, d]]$.

**References**


Received: July 3, 2014