

# 认知雷达中基于 Q 学习的自适应波形选择算法

王彬<sup>1,2</sup>, 汪晋宽<sup>2</sup>, 宋昕<sup>2</sup>, 韩英华<sup>2</sup>

(1. 东北大学信息科学与工程学院, 辽宁 沈阳 110004;

2. 东北大学秦皇岛分校工程优化与智能天线研究所, 河北 秦皇岛 066004)

**摘要:** 自适应波形选择器是认知雷达中智能发射器的重要组成部分。有效的波形选择能够在不同的环境下选择发射最优的波形序列, 从而以更高的精度追踪目标。针对雷达目标转移概率未知这一特点, 把自适应波形选择问题建模为随机动态规划模型, 提出应用 Q 学习的方法来解决这个问题。仿真结果说明, 该算法接近于最优波形选择方案, 并且状态估计的不确定性低于固定波形。

**关键词:** 波形选择; Q 学习; 空间划分; 认知雷达

**中图分类号:** TN 951

**文献标志码:** A

**DOI:**10.3969/j.issn.1001-506X.2011.05.09

## Adaptive waveform selection algorithm based on Q-learning in cognitive radar

WANG Bin<sup>1,2</sup>, WANG Jin-kuan<sup>2</sup>, SONG Xin<sup>2</sup>, HAN Ying-hua<sup>2</sup>

(1. College of Information Science and Engineering, Northeastern University, Shenyang 110004, China)

2. Engineering Optimization and Smart Antenna Institute, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China)

**Abstract:** The adaptive waveform selector is an important part of intelligent transmitters in cognitive radar. Effective waveform selection can transmit an optimal waveform sequence in different environments so as to track targets with higher accuracy. The problem of adaptive waveform selection is modeled as a stochastic dynamic model, and a Q-learning method is proposed to solve this problem under the fact that the transition probabilities of radar targets are unknown. The simulation results demonstrate that the proposed algorithm approaches the optimal waveform selection scheme and has a lower uncertainty of state estimation compared with the fixed waveform.

**Keywords:** waveform selection; Q-learning; space division; cognitive radar

## 0 引言

认知雷达是加拿大的 Simon Haykin 教授于 2006 年提出的一种智能化的新体制的雷达系统, 它具有许多优点。然而, 认知雷达只是智能雷达的一种理想化结构, 还有很多问题尚待解决。自适应波形选择是认知雷达中的一个重要问题, 其目的是根据不同的环境选择最优的波形从而以更高的精度追踪目标。文献[1]首次给出了认知雷达的概念, 同时在认知雷达的整体框架中提出了针对弱目标的 Bayesian 跟踪方案。文献[2]给出了认知跟踪雷达的结构, 它的实现包括接收端和发射端两个功能模块, 以及从接收端到发射端的反馈回路。文献[3]提出了容积分卡尔曼滤波的方法, 这种方法成功地解决了在保存二阶信息情况下贝叶斯滤波的最优近似问题。文献[4]尝试将不同的波形进行组合构成发射波形, 提高系统检测跟踪性能, 避免单一波形存在的缺陷。文献[5]研究了多天系统波形优化问题, 根据目标和杂波的统计特性, 通过最大化检测器输出端的信噪比调整发射波形, 提高了检测性能。文献[6]将序列假设检验应用于自适

应波形设计, 提高雷达的目标检测性能。根据目标与波束之间的交互信息, 文献[7]给出了一种用于多目标的波形设计算法, 该算法可以同时估计和跟踪多个扩展目标参数。基于海杂波统计特性的在线估计, 文献[8]根据波形的自相关函数设计相位调制波形, 以提高信号与杂波的功率比。文献[9]构建了一个频率调制的 Chirp 信号库, 根据跟踪器的最小均方误差来选择相应波形, 可以提高跟踪的性能。文献[10]使用复合高斯杂波模型和雷达阵列的测量值, 利用最大似然方法估计目标和杂波的参数。波形选择问题可以被看作一个传感器调度问题, 每一种波形提供了不同测量环境的方法, 相关的工作在文献[11]中。文献[12]考虑了追踪加速目标的波形选择算法。文献[13]利用自相关系数和模糊函数, 使用基因算法来执行波形选择的任务。文献[14]使用了 Incremental Pruning 的方法解决了雷达追踪中的波形选择问题。

本文把认知雷达中的自适应波形选择问题建模为随机动态规划模型, 针对雷达目标转移概率未知这一特点, 提出应用 Q 学习的方法解决此问题。仿真结果证明了本文所提算法的有效性。

收稿日期: 2009-09-26; 修回日期: 2010-11-19。

基金项目: 国家自然科学基金(60874108), 中央高校基本科研业务费专项资金(N090604006)资助课题

作者简介: 王彬(1982-), 男, 博士, 主要研究方向为雷达信号处理。E-mail: wangbin\_neu@yahoo.com.cn

### 1 距离-多普勒空间划分

对于雷达来说,测量目标最重要的参数有 4 个,即距离、多普勒频移和两个正交的空间角度。雷达对目标测量是通过对波束覆盖范围进行分辨单元划分,对每一个分辨单元进行目标测量来进行的。假设处于某目标临近分辨单元内的目标不会对它的测量造成干扰,只要每一个目标占据一个分辨单元并且各个分辨单元互不相交,那么雷达就可以在不受其他目标干扰的情况下进行目标测量。

对于一个单一的雷达脉冲,可以通过分辨单元在距离、多普勒频移和两个空间角的范围内进行划分。在这种情况下考虑处于雷达波束中的一个目标处在一定距离处,并且以特定的速度运动。与该目标相匹配的雷达匹配滤波器会对该目标的延时特性以及多普勒频移特性产生一个最大响应,并进行测量。现在来考虑另外一个目标。如果此目标在一个或两个空间角方向上处于天线方向图的波束宽度之外,它就不会产生任何接收机响应,并且在角度上与第一个目标区分开来。如果两目标都处于天线方向图主波束之内,且第二个目标所处位置距离远大于压缩脉冲长度,那么就可以在距离上区分这两个目标。类似的,假设两目标都在主波束内,且处于相同的距离上,但是以不同的速度运动,如果速度差大于发送波形长度的倒数,那么可以通过滤波器的多普勒响应来区分这两个目标。

如果一个雷达需要同时测量已经在多普勒频率上区分的多个目标,那么它就可以并行运行多个匹配滤波器,当滤波器与某一目标的多普勒频率相匹配的时候,它的响应就可以作为测量值;在距离上划分的目标可以根据测量值和距离门限来分辨。因此,雷达可以同时测量多个目标的距离或多普勒频率,甚至可以二者兼具。这种测量需要多个主波束或是一个主波束内的所有目标轮流扫描。

从上面的讨论可以看到,角度分辨力独立于距离和多普勒分辨力。因此,雷达角度分辨力特性与距离和多普勒频率分辨力特性相独立。在本文的研究中只将雷达波束覆盖区域划分为距离-多普勒空间。一种可行方法就是将距离-多普勒空间划分为多个互不重叠的平行四边形分辨单元,如图 1 所示。图 2 为距离和多普勒示意图。

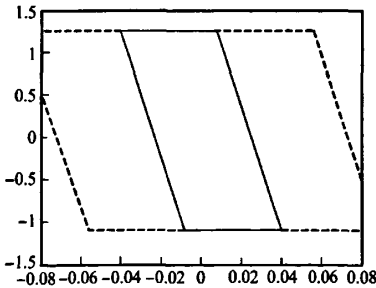


图 1 分辨单元和对应的平行四边形

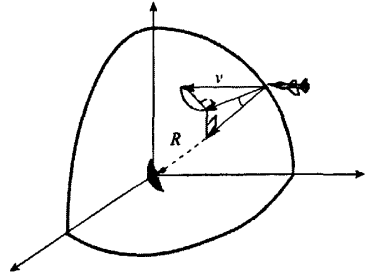


图 2 距离和多普勒示意图

雷达系统通常被设计为在最小距离  $R_{min}$  和最大距离  $R_{max}$  之间工作,用  $\Delta R$  表示距离分辨力,目标距离在  $\Delta R$  以上能够被完全区分开,距离可分辨单元数目定义为

$$N = \frac{R_{max} - R_{min}}{\Delta R} \tag{1}$$

雷达利用多普勒频移来测量目标的径向速度,同时来区分运动和静止的目标或物体。多普勒现象描述了附带波形中心频率的频移。

一般来说,一个波形可以被设计成或达到好的多普勒分辨力或好的距离分辨力,但是两者不能同时达到。所以需要考虑自适应波形选择问题。自适应波形选择一个最基本的方案是定义一个关于特定目标和选择波形的代价函数,目的就是在特定位置选择最优的波形使得代价函数最大。

对可能存在的目标数目不做任何假设,把雷达波束覆盖的范围划分为距离-多普勒单元,在这个单元中距离用  $\tau=1, \dots, N$  表示,多普勒用  $\nu=1, \dots, M$  表示。可能不存在目标,也可能存在 1 个目标或是  $NM$  个目标,所以所有可能存在的情况为

$$C_{NM}^0 + C_{NM}^1 + C_{NM}^2 + \dots + C_{NM}^{NM-1} + C_{NM}^{NM} = 2^{NM} \tag{2}$$

关于雷达的环境各种可能存在的情况为  $2^{NM}$ ,假设这个空间用  $\chi$  代表。模型的状态为  $X_t = x(x \in \chi)$ 。令  $Y_t$  为测量变量, $u_t$  为控制变量, $u_t$  表示在时刻  $t$  选择发射一个波形从而得到测量值  $Y_{t+1}(u_t \in U)$ 。得到一个特定测量值  $X_t = x$  依赖于真实的状态以及所选择的波形。

定义  $a_{x'x}$  为状态转移概率,其具体形式为

$$a_{x'x} = P(x_{t+1} = x' | x_t = x) \tag{3}$$

定义  $b_{x'x}$  为测量概率,其具体形式为

$$b_{x'x}(u_t) = P(Y_{t+1} = x' | X_t = x, u_t) \tag{4}$$

假设发射的基带信号为  $s(t)$ ,接收的基带信号为  $r(t)$ 。根据匹配滤波器理论,匹配滤波器具有冲激响应  $h(t) = s^*(-t)$ ,所以匹配滤波器的输出为

$$x(t) = \int s^*(\lambda - t)r(\lambda) d\lambda \tag{5}$$

由于回波信号是具有多普勒频移的,因此对于一个有  $\nu_0$  频移的回波信号,其匹配滤波器具有冲激响应

$$h(t) = s^*(-t)e^{2j\nu_0 t} \tag{6}$$

输出为

$$x(t) = \int s^*(\lambda - t) e^{-2\pi v_0(\lambda - t)} r(\lambda) d\lambda \quad (7)$$

式中,  $v_0$  为期望频移。

把基带接收信号建模为 Swerling 目标回波的形式

$$r(t) = As(t - \tau) e^{2\pi v_d t} I + n(t) \quad (8)$$

式中,  $s(t - \tau) e^{2\pi v_d t}$  是时延为  $\tau$ 、多普勒频移为  $v_d$  的发射基带信号  $s(t)$  的副本信号;  $I$  是目标指示器;  $A$  是接近于均值为 0、方差为  $2\sigma_A^2$  的复高斯随机变量;  $n(t)$  是独立于  $A$  的复高斯白噪声, 均值为 0, 方差为  $2N_0$ 。

在  $t$  时刻, 匹配于没有延时和多普勒频移的滤波器输出的模值的平方为

$$|x(t)|^2 = \left| \int_0^t r(\lambda) s^*(\lambda - t) d\lambda \right|^2 \quad (9)$$

下面分两种情况进行讨论。

当没有目标存在的时候

$$r(t) = v(t) \quad (10)$$

因此

$$x(\tau_0) = \int_0^{\tau_0} n(\lambda) s^*(\lambda - \tau_0) d\lambda \quad (11)$$

$x(\tau_0)$  是一个复高斯随机变量, 均值为 0, 方差为

$$\sigma_0^2 = E\{x(\tau_0)x^*(\tau_0)\} = 2N_0\xi \quad (12)$$

式中,  $\xi$  为发射脉冲的能量。

当目标存在时

$$r(t) = As(t - \tau) e^{2\pi v_d t} I + n(t) \quad (13)$$

$$x(\tau_0) = \int_0^{\tau_0} [As(\lambda - \tau) e^{2\pi v_d \lambda} + n(\lambda)] s^*(\lambda - \tau_0) d\lambda \quad (14)$$

这个随机变量的均值仍然是 0, 方差为

$$\begin{aligned} \sigma_1^2 &= E\{x(\tau_0)x^*(\tau_0)\} = \\ &\sigma_0^2 \left( 1 + \frac{2\sigma_A^2\xi^2}{\sigma_0^2} A(\tau_0 - \tau, v_0 - v) \right) \end{aligned} \quad (15)$$

$A(\tau, v)$  是模糊函数, 为

$$A(\tau, v) = \frac{1}{\left( \int |s(\lambda)|^2 d\lambda \right)^2} \left| \int s(\lambda) s^*(\lambda - \tau) e^{2\pi v \lambda} d\lambda \right|^2 \quad (16)$$

复高斯随机变量  $x$  幅值的平方为指数分布, 令  $y = x^2$ , 则

$$y = x^2 \sim \frac{1}{2\sigma_1^2} e^{-\frac{y}{2\sigma_1^2}} \quad (17)$$

因此虚警概率为

$$P_f = \int_D \frac{1}{2\sigma_0^2} e^{-\frac{x}{2\sigma_0^2}} dx = e^{-\frac{D}{2\sigma_0^2}} \quad (18)$$

检测概率  $P_d$  为

$$P_d = \int_D \frac{1}{2\sigma_1^2} e^{-\frac{x}{2\sigma_1^2}} dx = e^{-\frac{D}{2\sigma_1^2} \left( 1 + \frac{2\sigma_A^2\xi^2}{\sigma_0^2} A(\tau_0 - \tau, v_0 - v) \right)} \quad (19)$$

目标出现在  $(\tau, v)$  单元的时候, 假设它在这个单元的实际位置是均匀分布的, 可以得到检测概率为

$$P_d = \frac{1}{|A|} \iint_{(\tau, v) \in A} e^{-\frac{D}{2\sigma_1^2} \left( 1 + \frac{2\sigma_A^2\xi^2}{\sigma_0^2} A(\tau_0 - \tau, v_0 - v) \right)} d\tau_d v_d \quad (20)$$

式中,  $A$  是以  $(\tau, v)$  为中心的分辨单元。

## 2 基于 Q 学习的波形选择算法

定义  $\pi = \{u_0, u_1, \dots, u_T\}$ , 对于一个给定的波束  $T+1$  代表能够检测和确定目标最大的波形数目。  $\pi$  就是这个判决过程所使用的波形序列。在认知雷达中可以根据不同的环境获得不同的  $\pi$ 。

令

$$V_i(X_i) = E \left[ \sum_{t=0}^T \gamma^t R_t(X_i, u_t) \right] \quad (21)$$

式中,  $R_t(X_i, u_t)$  为处于  $X_i$  位置使用  $u_t$  波形所获得的报酬,  $\gamma$  为折扣因子。那么最终的目标就是寻找这样一个  $\pi^*$  满足

$$V^*(X_i) = \max_{\pi} E \left[ \sum_{t=0}^T \gamma^t R_t(X_i, u_t) \right] \quad (22)$$

然而, 在雷达中, 目标真实的状态一般是不知道的。使用文献[15]中介绍的方法, 可以通过求解式(23)来获得最优控制策略

$$V^*(p_0) = \max_{\pi} E \left[ \sum_{t=0}^T \gamma^t R_t(p_t, u_t) \right] \quad (23)$$

式中,  $p_t$  为给定测量值和波形情况下的条件概率密度;  $p_0$  为雷达环境的先验概率密度。  $p$  是状态  $X_t$  的充分统计量。所以需要解决式(24)的问题。

$$\max_{\pi} E \left[ \sum_{t=0}^T \gamma^t R(p_t, u_t) \right] \quad (24)$$

其中,  $p_t$  的更新公式为

$$p_{t+1} = \frac{B A p_t}{I' L A p_t} \quad (25)$$

式中,  $B$  为对角矩阵, 对角线上的元素为  $(b_{x,x}(u_t))$ ;  $I$  是列向量;  $A$  是状态转移矩阵。

$t$  时刻之后使用波形序列  $\pi$  所获得的收益为

$$G_t(p_t) = E \left\{ \sum_{\tau=t}^{T-1} R_{\tau}(p_{\tau}, u_{\tau}) + C_T(p_T) \mid p_t \right\} \quad (26)$$

其中, 在波形序列  $\pi$  下,  $u_t = U^*(p_t)$ 。

更自然的, 可以使用递归的方法计算  $V_t$ , 计算方法为

$$V_t(p_t) = R_t(p_t, u_t) + E\{V_{t+1}(p_{t+1}) \mid p_t\} \quad (27)$$

很明显,  $G_T(p_T) = V_T(p_T) = R_T(p_T)$ , 假设对于  $t+1, t+2, \dots, T$  也是成立的, 那么  $V_t(p_t)$  可以写成

$$V_t(p_t) = R_t(p_t, u_t) +$$

$$E \left\{ E \left[ \sum_{\tau=t+1}^{T-1} R_{\tau}(p_{\tau}, u_{\tau}) + R_T(p_T) \mid p_{t+1} \right] \mid p_t \right\} \quad (28)$$

由于随机变量是离散和有限的, 可以得到

$$\begin{aligned} E\{E[G_t \mid p_{t+1}] \mid p_t\} = \\ \sum_{p_{t+1} \in P} \sum_{g \in G} g P(G_t = g \mid p_{t+1}, p_t) P(p_{t+1} = p_{t+1} \mid p_t) = \\ \sum_{g \in G} g \sum_{p_{t+1} \in P} P(G_t = g, p_{t+1} \mid p_t) = E[G_t \mid p_t] \end{aligned} \quad (29)$$

因此

$$\begin{aligned} V_t(p_t) = \\ R_t(p_t, u_t) + E \left\{ E \left[ \sum_{\tau=t+1}^{T-1} R_{\tau}(p_{\tau}, u_{\tau}) + R_T(p_T) \mid p_{t+1} \right] \mid p_t \right\} = \\ R_t(p_t, u_t) + E \left[ \sum_{\tau=t+1}^{T-1} R_{\tau}(p_{\tau}, u_{\tau}) + R_T(p_T) \mid p_t \right] = \end{aligned}$$

$$E\left[R_i(p_i, u_i) + \sum_{t=i+1}^{T-1} R_t(p_t, u_t) + R_T(p_T) \mid p_i\right] = E\left[\sum_{t=i}^{T-1} R_t(p_t, u_t) + R_T(p_T) \mid p_i\right] = G_i(p_i) \quad (30)$$

如果应用传统的动态规划的方法求解这个问题,可以使用式(31)得到价值函数。

$$V_i(p_i) = \max_{u_i} (R_i(p_i, u_i) + \gamma E[V_{i+1}(p_{i+1}) \mid p_i]) \quad (31)$$

写成概率的形式为

$$V_i(p_i) = \max_{u_i} \left( R_i(p_i, u_i) + \gamma \sum_{p' \in P} P(p' \mid p_i, u_i) V_{i+1}(p') \right) \quad (32)$$

先来考虑简单一些的时间差分近似方法。

假设  $v$  是价值在状态  $p_i$  和策略  $\pi$  情况下的无偏估计。 $v$  的定义为

$$v_i^* = C_i(p_i^*, u_i^*) + C_{i+1}(p_{i+1}^*, u_{i+1}^*) + \dots + C_T(p_T^*, u_T^*) \quad (33)$$

使用标准的随机梯度算法来估计状态  $p_i$  的值为

$$V_i^*(p_i) = V_i^{-1}(p_i) - \alpha_n [V_i^{-1}(p_i) - v_i^*] \quad (34)$$

则时间差分为

$$D_i = C_i(p_i, u_i) + V_{i+1}^{-1}(p_{i+1}) - V_i^{-1}(p_i) \quad (35)$$

因此

$$v_i^* = V_i^{-1}(p_i) + \sum_{t=i}^{T-1} D_t \quad (36)$$

把式(31)代入式(29),可以得到

$$V_i^*(p_i) = V_i^{-1}(p_i) - \alpha_{n-1} \sum_{t=i}^{T-1} D_t \quad (37)$$

时间差分是在状态  $p_i$  估计价值函数的误差。这些误差是最小化估计误差问题的随机梯度,其折扣形式为

$$V_i^*(p_i) = V_i^{-1}(p_i) - \alpha_{n-1} \sum_{t=i}^{T-1} (\gamma\lambda)^{t-i} D_t \quad (38)$$

通过式(38),就可以更新  $V$  值。

事实上,这个方法近似了  $V$  值,但是仍然不太适合于雷达的情况。因为雷达目标转移的概率通常是不知道的。 $Q$  学习方法是一类特殊的时间差分方法,也是近似动态规划方法的一种。 $Q$  学习正好可以解决在状态转移概率未知情况下的最优求解问题,这很适合于雷达的情况。因此使用  $Q$  学习的方法解决此问题。

对于一个“状态-动作”对  $(p_i, u_i)$ , 定一个  $Q$  因子为

$$Q(p_i, u_i) = \sum_{p' \in P} P(p' \mid p_i, u_i) [R_i(p' \mid p_i, u_i) + \gamma V_{i+1}] \quad (39)$$

根据式(32)和式(39),可以得到

$$V_i^* = \max_{u_i} Q(p_i, u_i) \quad (40)$$

上面建立起来了状态价值函数和与状态相关  $Q$  因子两者之间的关系。如果  $Q$  因子是知道的,给定状态的价值函数就可以通过式(40)求得。

由此可得 Bellman 等式的  $Q$  因子形式为

$$Q(p_i, u_i) = \sum_{p' \in P} P(p' \mid p_i, u_i) [R_i(p' \mid p_i, u_i) + \gamma \max_{u_{i+1}} Q(p_{i+1}, u_{i+1})] \quad (41)$$

用  $s^i$  表示随机变量  $X$  的第  $i$  次独立抽样值,  $X^n$  表示  $X$

在第  $n$  次迭代的估计值,则期望为

$$E(X) = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n s^i}{n} \quad (42)$$

$$X^n = \frac{\sum_{i=1}^n s^i}{n} \quad (43)$$

可以得到

$$X^{n+1} = (1 - \alpha^{n+1})X^n + \alpha^{n+1} s^{n+1} \quad (44)$$

其中

$$\alpha^{n+1} = \frac{1}{n+1} \quad (45)$$

于是

$$Q(p_i, u_i) = E[R_i(p' \mid p_i, u_i) + \gamma \max_{u_{i+1}} Q(p_{i+1}, u_{i+1})] \quad (46)$$

可以采用式(47)进行迭代,即

$$Q^{n+1}(p_i, u_i) \leftarrow (1 - \alpha^{n+1})Q^n(p_i, u_i) + \alpha^{n+1} [R_i(p' \mid p_i, u_i) + \gamma \max_{u_{i+1}} Q^n(p_{i+1}, u_{i+1})] \quad (47)$$

下面给出波形选择的  $Q$  学习算法步骤。

- 步骤 1 初始化  $Q$  因子为 0, 设置  $n=1$ ;
- 步骤 2 对于  $t=0, 1, \dots, T$ , 执行步骤 3~步骤 6;
- 步骤 3 令现在的状态为  $p_i$ , 下一个状态为  $p_{i+1}$ ;
- 步骤 4 使用现在的  $Q$  因子找到判决量

$$u_i = \operatorname{argmax}_{u_i} Q_i^{-1}(p_i', u_i) \quad (48)$$

步骤 5 使用式(47)更新  $Q(p_i, u_i)$ ;

步骤 6 找到下一个状态

$$p_{i+1} = \frac{BAp_i}{I'BAp_i} \quad (49)$$

步骤 7 增加  $n$ , 如果  $n < N$ , 返回步骤 2;

步骤 8 对于任意的  $p_i \in P$ , 选择

$$d(p_i) \in \operatorname{argmax}_{u_i} Q(p_{i+1}, u_{i+1}) \quad (50)$$

这种算法生成的最优策略用  $\hat{d}$  表示。

### 3 仿真实验

在这一节,进行了 4 个实验。为了说明波形选择的必要性,做出了 3 种波形的测量概率随信噪比变化曲线和对于不同的目标测量概率随信噪比变化曲线。状态估计不确定性曲线说明了所提算法的有效性。

调频或调相波形可用来实现更宽的工作带宽,线性调频(linear frequency modulation, LFM)信号是经常用到的一种信号。LFM 信号表达形式为

$$s(t) = \operatorname{rect}\left(\frac{t}{T}\right) e^{j2\pi\left(f_c t + \frac{K}{2}t^2\right)} \quad (51)$$

式中,  $f_c$  为载频,  $\operatorname{rect}$  为矩形信号, 即

$$\operatorname{rect}\left(\frac{t}{T}\right) = \begin{cases} 1, & |t| \leq T/2 \\ 0, & \text{其他} \end{cases} \quad (52)$$

图 3 所示为 LFM 信号实部。

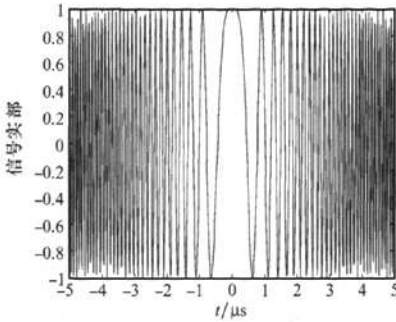


图 3 LFM 信号实部

步进频率波形 (stepped frequency waveform, SFW) 信号产生合成高距离分辨率的目标轮廓, 目标距离轮廓是通过真实目标距离轮廓的频率样本进行逆傅里叶变换而计算出来的。

线性调频波形和步进频率波形都可以实现高距离分辨率, 但是 SFW 信号需要更复杂的硬件实现, 同时对雷达工作带宽的要求限制较少。在这里采用 LFM 信号作为雷达发射信号。

在仿真中, 为了验证所提算法的有效性, 假定每一种波形以及目标不同位置的测量概率、折扣因子和状态转移矩阵均为已知。当然这些参数是由环境认知后可以得到的, 实际中如何简化和合理地根据模型选择这些参数还需要进一步研究。状态空间是  $4 \times 4$ 。考虑使用五种不同的波形。对于每一种波形  $u$  以及目标位于  $x$  和  $x'$  的时候检测概率的分布情况如表 1 所示。折扣因子  $\gamma = 0.85$ , 状态转移矩阵  $A$  为

$$A = \begin{bmatrix} 0.96 & 0.02 & 0.01 & 0.01 \\ 0.01 & 0.93 & 0.03 & 0.02 \\ 0.02 & 0.03 & 0.95 & 0.02 \\ 0.01 & 0.02 & 0.01 & 0.95 \end{bmatrix} \quad (53)$$

表 1 例子中的测量概率

	$x=1$ $x'=1,2,3,4$	$x=2$ $x'=1,2,3,4$	$x=3$ $x'=1,2,3,4$	$x=4$ $x'=1,2,3,4$
$U=1$	0.97,0.01,0.01,0.01	0.96,0.01,0.01,0.02	0.01,0.01,0.96,0.02	0.01,0.95,0.02,0.02
$U=2$	0.96,0.01,0.02,0.01	0.02,0.95,0.01,0.02	0.01,0.01,0.01,0.97	0.02,0.96,0.01,0.01
$U=3$	0.02,0.95,0.02,0.01	0.02,0.02,0.01,0.95	0.02,0.96,0.01,0.01	0.01,0.02,0.02,0.95
$U=4$	0.96,0.01,0.01,0.02	0.01,0.96,0.02,0.01	0.97,0.01,0.01,0.01	0.03,0.95,0.01,0.01
$U=5$	0.01,0.02,0.04,0.03	0.01,0.97,0.01,0.01	0.02,0.01,0.96,0.01	0.04,0.94,0.01,0.01

根据文献[16]中所提供的方法, 采用报酬函数的线性形式为

$$R(p, u) = p'p - 1 \quad (54)$$

图 4 为 3 种波形的测量概率随信噪比变化曲线。波形 1~3 的脉宽是依次减小的。可以看到, 在相同信噪比的条件下, 不同的波形对应于不同的测量概率。一般来说, 发射波形的脉宽越窄, 对于同一目标的测量概率越大。但是窄脉冲降低了平均发射功率, 这会妨碍雷达的正常工作模式。这是一对矛盾, 需要在二者中做一个折衷。因此, 对于不同的环境, 要合理地选择不同的波形。

是一样的。当信噪比大于 12 时, 这种波形对于测量 Swerling V 更有效。当信噪比小于 12 时, 对测量 Swerling I 更有效。因此对于同一个目标, 需要根据环境的不同来变换波形, 从而提高目标的检测概率。

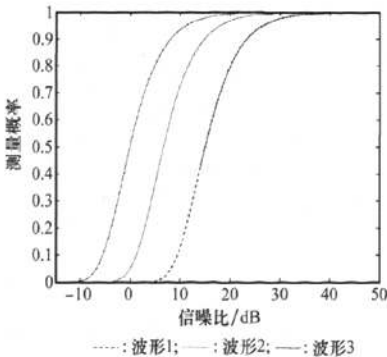


图 4 3 种波形的测量概率随信噪比变化曲线

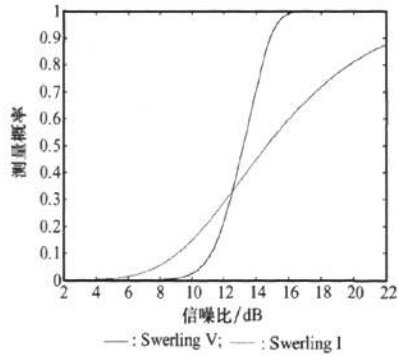


图 5 不同目标测量概率随信噪比变化曲线

图 5 为不同目标下测量概率随信噪比变化曲线。可以看到同一种波形对于不同目标的测量概率是不一样的。当信噪比为 12 时, 对于 Swerling I 和 Swerling V 的测量概率

图 6 为状态估计的不确定性曲线。对于所有的情况, 不管状态随着测量次数如何变化, 状态估计的不确定性随着测量次数的增加而降低, 和固定波形相比, 所提出的 Q 学习算法降低了状态估计的不确定性, 即降低了定位目标的不确定性。如果选用最优的波形调度方案, 这种不确定性能够进一步降低。同时, 所提出的 Q 学习算法已经非常接近最优调度方案。使用 Q 学习的好处在于可以不知道状态转移概率的精确知识, 这对于雷达是非常适用的。尤其在状态空间和波形空间很大的情况下, 每次迭代求解难

以实现,精确的后向动态规划方法已经失效。这意味着只能采用近似求解方法来代替精确的求解方法。Q 学习的方法很接近于最优波形选择方案,是一个不错的选择。

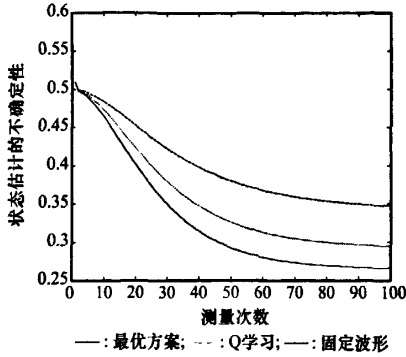


图 6 状态估计的不确定性曲线

图 7 比较了时间差分学习和 Q 学习两种方法。可以看到时间差分学习的性能稍优于 Q 学习,这是因为时间差分学习的使用条件要比 Q 学习苛刻一些,每一步的计算也更为精确,同时两者的性能都很接近于最优调度方案。但是在雷达场景下,目标的状态转移概率通常是未知的,因此 Q 学习更适合于这种环境,而且 Q 学习计算也更为简便。

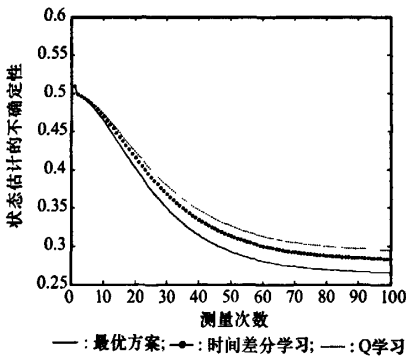


图 7 时间差分学习和 Q 学习的比较

图 8 为“Q 值-状态-波形”空间曲线,不同“状态-波形”对对应的 Q 值可以从图中得到。

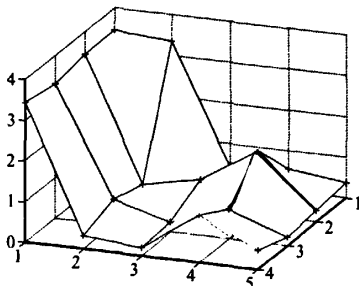


图 8 “Q 值-状态-波形”空间曲线

### 4 结论

自适应波形选择是认知雷达中的一个重要问题。认知

雷达中的自适应波形选择问题可以建模为随机动态规划模型。本文提出了基于 Q 学习的自适应波形选择算法,从仿真中可以看到该算法的优点。Q 学习算法跟固定波形相比,能够减少状态估计的不确定性,同时接近于最优调度波形,这是在不知道精确转态转移概率的情况下所得到的,十分适用于雷达环境。下一步需要研究的是如何研究高效的算法接近最优调度波形以及减少算法的计算量,这对于实际中大空间的情况是非常有意义的。

### 参考文献:

- [1] Haykin S. Cognitive radar: a way of the future[J]. *IEEE Signal Processing Magazine*, 2006, 23(1): 30 - 40.
- [2] Haykin S. Cognition is the key to the next generation of radar systems[C]// *Proc. of Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop*, 2009: 463 - 467.
- [3] Arasaratnam I, Haykin S. Cubature Kalman filters[J]. *IEEE Trans. on Automatic Control*, 2009, 54(6): 1254 - 1269.
- [4] Rago C, Willett P, Bar-Shalom Y. Detecting-tracking performance with combined waveforms[J]. *IEEE Trans. on Aerospace and Electronic Systems*, 1998, 34(2): 612 - 624.
- [5] Friedlander B. Adaptive waveform design for a multi-antenna radar system[C]// *Proc. of the 40th Asilomar Conference on Signals, Systems and Computer*, 2006: 735 - 739.
- [6] Goodman N A, Venkata P R, Neifeld M A. Adaptive waveform design and sequential hypothesis testing for target recognition with active sensors[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2007, 1(1): 105 - 113.
- [7] Leshem A, Naparstek O, Nehorai A. Information theoretic adaptive radar waveform design for multiple extended targets[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2007, 1(1): 42 - 55.
- [8] Sira S P, Cochran D. Adaptive waveform design for improved detection of low-RCS targets in heavy sea clutter[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2007, 1(1): 56 - 66.
- [9] Sira S P, Papandreou-Suppappola A, Morrell D. Dynamic configuration of time-varying waveforms for agile sensing and tracking in clutter[J]. *IEEE Trans. on Signal Processing*, 2007, 55(7): 3207 - 3217.
- [10] Wang J, Dogandzic A, Nehorai A. Maximum likelihood estimation of compound-gaussian clutter and target parameters[J]. *IEEE Trans. on Signal Processing*, 2006, 54(10): 3884 - 3898.
- [11] He Y, Chong E K P. Sensor scheduling for target tracking in sensor networks[C]// *Proc. of the 43rd IEEE Conference on Decision and Control*, 2004: 743 - 748.
- [12] Savage C O, Moran B. Waveform selection for maneuvering targets within an IMM framework[J]. *IEEE Trans. on Aerospace and Electronic Systems*, 2007, 43(3): 1205 - 1214.
- [13] Capraro C T, Bradaric I, Capraro G T. Using genetic algorithms for radar selection[C]// *Proc. of the IEEE Radar Conference*, 2008: 1 - 6.
- [14] Scala B L, Rezaeian M, Moran B. Optimal adaptive waveform selection for target tracking[C]// *Proc. of the International Conference on Information Fusion*, 2005: 552 - 557.
- [15] Bertsekas D. *Dynamic programming and optimal control* [M]. Nashua: Athena Scientific, 2001: 125 - 143.
- [16] Krishnamurthy V. Algorithms for optimal scheduling of hidden Markov model sensors[J]. *IEEE Trans. on Signal Processing*, 2002, 50(6): 1382 - 1397.

作者: [王彬](#), [汪晋宽](#), [宋昕](#), [韩英华](#), [WANG Bin](#), [WANG Jin-kuan](#), [SONG Xin](#), [HAN Ying-hua](#)

作者单位: [王彬, WANG Bin\(东北大学信息科学与工程学院, 辽宁, 沈阳, 110004; 东北大学秦皇岛分校工程优化与智能天线研究所, 河北, 秦皇岛, 066004\)](#), [汪晋宽, 宋昕, 韩英华, WANG Jin-kuan, SONG Xin, HAN Ying-hua\(东北大学秦皇岛分校工程优化与智能天线研究所, 河北, 秦皇岛, 066004\)](#)

刊名: [系统工程与电子技术](#) **ISTIC EI PKU**

英文刊名: [SYSTEMS ENGINEERING AND ELECTRONICS](#)

年, 卷(期): 2011, 33(5)

## 参考文献(16条)

1. [Krishnamurthy V Algorithms for optimal scheduling of hidden Markov model sensors](#)[外文期刊] 2002(06)
2. [Bertsekas D Dynamic programming and optimal control](#) 2001
3. [Scala B L;Rezaeian M;Moran B Optimal adaptive waveform selection for target tracking](#) 2005
4. [Capraro C T;Bradaric I;Capraro G T Using genetic algorithms for radar selection](#) 2008
5. [Savage C O;Moran B Waveform selection for maneuvering targets within an IMM framework](#) 2007(03)
6. [He Y;Chong E K P Sensor scheduling for target tracking in sensor networks](#) 2004
7. [Wang J;Dogandzic A;Nehorai A Maximum likelihood estimation of compound-gaussian clutter and target parameters](#) 2006(10)
8. [Sira S P;Papandreou-Suppappola A;Morrell D Dynamic configuration of time-varying waveforms for agile sensing and tracking in clutter](#)[外文期刊] 2007(07)
9. [Sira S P;Cochran D Adaptive waveform design for improved detection of low-RCS targets in heavy sea clutter](#) 2007(01)
10. [Leshem A;Naparstek O;Nehorai A Information theoretic adaptive radar waveform design for multiple extended targets](#)[外文期刊] 2007(01)
11. [Goodman N A;Venkata P R;Neifeld M A Adaptive waveform design and sequential hypothesis testing for target recognition with active sensors](#)[外文期刊] 2007(01)
12. [Friedlander B Adaptive waveform design for a multi-antenna radar system](#) 2006
13. [Rago C;Willett P;BarShalom Y Detecting-tracking performance with combined waveforms](#)[外文期刊] 1998(02)
14. [Arasaratnam I;Haykin S Cubature Kalman filters](#)[外文期刊] 2009(06)
15. [Haykin S Cognition is the key to the next generation of radar systems](#) 2009
16. [Haykin S Cognitive radar: a way of the future](#)[外文期刊] 2006(01)

## 本文读者也读过(10条)

1. [夏洪恩 基于目标跟踪遥波形自适应选择技术](#)[学位论文]2010
2. [王彬. 汪晋宽. 宋昕. 韩英华 一种基于互信息的雷达波形设计模型](#)[会议论文]-2010
3. [闫冬梅. 汪晋宽. 刘丽. 王彬. 许鹏 基于重叠分簇的目标协作跟踪算法](#)[会议论文]-2010
4. [李成铁. 汪晋宽 基于Web的三维虚拟现实远程监控研究和应用](#)[会议论文]-2010
5. [刘福来. 孙长银. 汪晋宽 一种有效的基于半定规划的零陷控制方法](#)[会议论文]-2010
6. [刘志刚. 汪晋宽 传感器网络环境下协作跟踪综述](#)[会议论文]-2010

7. [蔡凌, 汪晋宽, 王翠荣](#) [基于自动选择及资源分配的虚拟网络设计](#)[会议论文]-2010
8. [解志斌, 汪晋宽, 王赞, 高静, XIE Zhi-bin, WANG Jin-kuan, WANG Yun, GAO Jing](#) [多天线系统分集复用模式快速选择算法](#)[期刊论文]-[信息与控制](#)2009, 38(1)
9. [杨红兵, 周建江, 汪飞, 刘伟强, YANG Hong-bing, ZHOU Jian-jiang, WANG Fei, LIU Wei-qiang](#) [STLFMCW雷达信号波形设计与射频隐身特性分析](#)[期刊论文]-[现代雷达](#)2011, 33(4)
10. [刘丽, 汪晋宽, 闫冬梅, 杜瑞燕](#) [一种改进的MIMO系统串行干扰消除检测算法](#)[会议论文]-2010

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_xtgcydzjs201105009.aspx](http://d.g.wanfangdata.com.cn/Periodical_xtgcydzjs201105009.aspx)