

基于多任务复数因子分析模型的雷达高分辨距离像识别方法

和 华 杜 兰* 徐丹蕾 刘宏伟

(西安电子科技大学雷达信号处理国家重点实验室 西安 710071)

摘要: 传统的高分辨距离像(HRRP)统计识别方法大部分只使用雷达目标高分辨回波的幅值信息且需要大量的训练样本保证统计模型参数学习的精度。为了充分利用高分辨回波的相位信息,在雷达采样率有限、训练样本数不足条件下保证统计识别的性能,该文提出一种多任务学习(MTL)复数因子分析(CFA)模型,将数据描述推广到复数域,将每个方位帧训练样本的统计建模视为单一的学习任务,各学习任务共享加载矩阵,利用贝塔伯努利(Beta-Bernoulli)稀疏先验自适应地选择各任务需要的因子,完成多任务的共同学习。基于实测数据的识别实验显示,与传统的单任务学习(STL)因子分析模型相比,该文提出的多任务因子分析模型具有更低的模型复杂度且在小样本条件下可以显著提高识别性能。

关键词: 雷达自动目标识别; 高分辨距离像; 多任务学习; 因子分析

中图分类号: TN958

文献标识码: A

文章编号: 1009-5896(2015)10-2307-07

DOI: 10.11999/JEIT141591

Radar HRRP Target Recognition Method Based on Multi-task Learning and Complex Factor Analysis

He Hua Du Lan Xu Dan-lei Liu Hong-wei

(National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China)

Abstract: Most traditional recognition methods for High Resolution Range Profile (HRRP) only utilize the amplitude information and need large number of training samples to obtain better estimation precision of model parameters. To utilize the phase information contained in the complex echoes and obtain better recognition performance with small training data and low sampling rate, a statistical model based on Multi-Task Learning (MTL) and Complex Factor Analysis (CFA), referred to as MTL-CFA, is proposed in this paper. The MTL-CFA model directly describes the complex HRRP data. The statistical modeling of each training aspect-frame is considered as a single task, and all tasks share a common loading matrix. The factor number of each task is automatically determined via the Beta-Bernoulli sparse prior. Experimental results based on measured data show that the proposed model MTL-CFA can not only describe the observed data with lower order of model complexity, but also obtain satisfactory recognition accuracy with small training data, compared with the traditional Single-Task Learning (STL) based on FA models.

Key words: Radar Automatic Target Recognition (RATR); High Resolution Range Profile (HRRP); Multi-Task Learning (MTL); Factor Analysis (FA)

1 引言

高分辨距离像(HRRP)是宽带雷达目标散射点子回波沿雷达视线方向投影的叠加,它反映了目标散射体的散射截面积沿雷达视线方向的分布情况,包含了目标的尺寸大小和散射点分布等丰富的结构

信息^[1-3]。作为雷达目标的一种重要的特征,高分辨距离像易于获取和处理,已经成为雷达目标识别领域的研究热点^[1-5]。在雷达目标识别中,充分利用训练样本的信息,构建合适复杂度的模型描述数据对于提高识别性能相当关键。

根据散射点模型理论,高分辨回波是由同一距离单元的多个散射点子回波相关叠加而成,包含幅度和相位。目前大多数关于距离像识别的文献都是对实距离像进行统计建模和识别^[6,7],忽略了相位信息。文献[7]假设实距离像服从高斯分布,但取模后的雷达回波幅值与高斯分布显然是失配的。文献[8]将高分辨回波的距离单元分为两种类型,采用与高分辨回波更加匹配的 Gamma, Gaussian Mixture 双

收稿日期: 2014-12-11; 改回日期: 2015-06-29; 网络出版: 2015-07-17

*通信作者: 杜兰 dulan@mail.xidian.edu.cn

基金项目: 国家自然科学基金(61271024, 61201296, 61322103), 高等学校博士学科点专项科研基金(20130203110013)和陕西省自然科学基金基础研究计划(2015JZ016)

Foundation Items: The National Natural Science Foundation of China (61271024, 61201296, 61322103); The Foundation for Doctoral Supervisor of China (20130203110013); The Natural Science Foundation of Shaanxi Province (2015JZ016)

分布模型,提高了识别性能。文献[7]利用因子分析(FA)模型描述实距离单元的相关性,进一步提升了实距离像的识别性能。文献[9]的实验表明,复距离像的相位信息有助于改善识别性能,不过该文献提取的特征并不适合于统计建模。为了充分利用高分辨回波的相位信息,可以考虑对复距离像建模。相比于实距离像,直接对复距离像建模的主要问题在于其初相敏感性。文献[10]分析证明了在复高斯假设条件下初相的存在并不影响复距离像的统计特性。文献[11]采用复数因子分析(CFA)对复距离像建模,提高了识别性能,但该方法中因子个数需要提前设定,因子个数过大会产生过匹配,因子个数过小则难以较好地描述数据。

由于高分辨距离像的方位敏感性,一般需要按照一定角域(不发生越距离单元走动)对距离像进行分帧,对于宽带雷达,目标不发生越距离单元的角度较小,每类目标分帧后的帧数较多,对每一帧分别独立建模,模型复杂度和运算量相对较大,而同一目标各帧的距离像具有一定的相似性,多帧模型的学习也即多个学习任务之间具有很强的相关性。根据多任务学习^[12](MTL)的思想,不同方位帧的学习任务可以共享某些模型参数,通过减少待学习模型参数的个数,降低对训练样本数的需求,因此在相同训练样本数条件下,多任务学习可以提高模型参数的估计精度。文献[13]采用多任务隐马尔可夫模型对实距离像的谱图建模,在训练样本较小时仍能获得较好的识别性能。

本文将文献[11]的方法推广到多任务学习,采用多任务复数因子分析(MTL-CFA)模型对复距离像建模,在不同方位帧共享FA模型中的加载矩阵,并利用贝塔伯努利(Beta-Bernoulli)稀疏先验自适应地选择各方位帧需要的因子,解决了文献[11]中因子个数的选择问题,在利用不同学习任务相关性的同时,充分利用了高分辨回波的相位信息,进一步提高了识别性能。基于实测数据的识别实验验证了提出方法的有效性。

2 基于 MTL-CFA 的高分辨距离像的统计建模

在介绍 MTL-CFA 模型之前,先简单介绍一下传统的 FA 模型。

2.1 传统的 FA 模型与因子个数选择

FA 模型主要用于描述分布于低维线性子空间的高维数据。高维数据 $\mathbf{X} \in \mathbb{R}^{P \times N}$ 的每个样本可以表示为多个相互独立的因子与噪声的线性组合:

$$\mathbf{x}_i = \mathbf{A}\mathbf{w}_i + \mathbf{u} + \boldsymbol{\varepsilon}_i \quad (1)$$

其中, $\mathbf{x}_i \in \mathbb{R}^{P \times 1}$ 表示 \mathbf{X} 的第 i 个样本, $i = 1, 2, \dots, N$; $\mathbf{A} \in \mathbb{R}^{P \times K}$ 表示加载矩阵; $\mathbf{w}_i \in \mathbb{R}^{K \times 1}$ 表示第 i 个样本对应的因子,一般假设因子相互独立,服从零均值的高斯分布, $\mathbf{w}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_K)$, K 表示因子个数,一般 $K < P$; $\boldsymbol{\varepsilon}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda})$ 。经推导可知 \mathbf{x}_i 服从高斯分布,即 $\mathbf{x}_i \sim \mathcal{N}(\mathbf{u}, \boldsymbol{\Lambda} + \mathbf{A}\mathbf{A}^T)$ 。

FA 模型参数的求解属于不完全数据问题,当因子个数 K 给定时,可以采用期望最大化(EM)算法求得模型参数的最大似然估计,一般可以采用经典的贝叶斯信息准则^[14](BIC)确定因子个数 K 。文献[8]将 FA 模型用于实距离像识别,将每帧距离像的建模视为一个单任务,对每帧实距离像用一个 FA 描述,本文将这种建模方式称为单任务学习实数因子分析(Single-Task Learning and Real Factor Analysis),记为 STL-RFA。

为了能够充分利用相位信息,文献[11]采用复数因子分析在复数域对高分辨回波建模。假设目标有 C 类,每一类目标的距离像样本按角域均匀分帧,第 c 类分为 M_c 帧, $c = 1, 2, \dots, C$,一般认为复距离像的均值为零,将式(1)中的均值项 \mathbf{u} 去掉,然后将因子和噪声的分布从实高斯推广到复高斯得到 CFA 模型:

$$\mathbf{x}_i^{c,m} = \mathbf{A}^{c,m}\mathbf{w}_i^{c,m} + \boldsymbol{\varepsilon}_i^{c,m} \quad (2)$$

其中 $\mathbf{x}_i^{c,m}$ 表示第 c 类目标第 m 帧第 i 个复距离像,相比于式(1),式(2)中的加载矩阵 $\mathbf{A}^{c,m}$ 为 $P \times K_m^c$ 的复矩阵;因子和噪声服从零均值的复高斯分布 $\mathbf{w}_i^{c,m} \sim \text{CN}(\mathbf{0}, \mathbf{I}_{K_m^c})$, $\boldsymbol{\varepsilon}_i^{c,m} \sim \text{CN}(\mathbf{0}, \boldsymbol{\Lambda}^{c,m})$, $\text{CN}(\bullet)$ 表示复高斯分布。本文将这种建模方式称为单任务学习复数因子分析,记为 STL-CFA。

2.2 基于 MTL-CFA 模型的高分辨距离像的统计建模

下面将 CFA 模型推广到多任务学习,介绍基于 MTL-CFA 模型的高分辨距离像的统计建模。

将式(2)中的加载矩阵 $\mathbf{A}^{c,m}$ 换为各帧共享的加载矩阵 \mathbf{A} ,得到 MTL-CFA 模型:

$$\mathbf{x}_i^{c,m} = \mathbf{A}\mathbf{w}_i^{c,m} + \boldsymbol{\varepsilon}_i^{c,m} \quad (3)$$

其中, $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_K]$ 是 $P \times K$ 的复矩阵, \mathbf{A}_k 表示 \mathbf{A} 的第 k 列,一般 K 设得比较大,以保证某些任务间相似性较弱时,模型的适应能力。为了使模型更灵活,模型中的参数不再像 CFA 模型中是确定参数,而假定其是服从某种分布的随机参数,某些参数的先验分布的参数设有超先验。将文献[15]中加载矩阵的先验分布推广到复数域,服从零均值的复高斯分布: $\mathbf{A}_k \sim \text{CN}(\mathbf{A}_k | \mathbf{0}, P^{-1}\mathbf{I}_P)$ 与实高斯分布类似,复高斯分布属于指数族分布,具有共轭先验,后验分布推导比较方便。

为了能够同时自动确定每帧的因子个数 K_m^c ，将因子 $\mathbf{w}_i^{c,m}$ 分解为权值 $\tilde{\mathbf{w}}_i^{c,m}$ 和稀疏选择因子 $\mathbf{z}_i^{c,m}$ ：

$$\mathbf{w}_i^{c,m} = \tilde{\mathbf{w}}_i^{c,m} \odot \mathbf{z}_i^{c,m} \quad (4)$$

\odot 表示两个向量的 Hadamard 积，与文献[11]中 CFA 假设一致，权值 $\tilde{\mathbf{w}}_i^{c,m}$ 的先验分布服从零均值的复高斯分布：

$$p(\tilde{\mathbf{w}}_i^{c,m} | \alpha^{c,m}) = \frac{1}{\pi^P |1/\alpha^{c,m} \mathbf{I}_P|} \cdot \exp\left\{-\left(\tilde{\mathbf{w}}_i^{c,m}\right)^H (1/\alpha^{c,m} \mathbf{I}_P)^{-1} \tilde{\mathbf{w}}_i^{c,m}\right\} \quad (5)$$

给 $\alpha^{c,m}$ 以伽马共轭超先验：

$$p(\alpha^{c,m} | g_0, h_0) = h_0^{g_0} (\alpha^{c,m})^{g_0-1} \exp(-h_0 \delta^{c,m}) / \Gamma(g_0) \quad (6)$$

$\Gamma(\bullet)$ 表示伽马函数， g_0, h_0 表示超参数，这里设 $g_0 = h_0 = 10^{-6}$ ；稀疏选择因子的 $\mathbf{z}_i^{c,m}$ 中的元素取 0 或 1，当第 c 类第 m 帧的稀疏选择因子 $\mathbf{z}_i^{c,m}$ 的第 k 行元素全为零，则 \mathbf{A}_k 没有用来表示信号 $\mathbf{x}^{c,m}$ ；若给 $\mathbf{z}_i^{c,m}$ 稀疏先验，则可以选取 \mathbf{A} 中较少的列向量重构复距离像样本，剔除 \mathbf{A} 中没有用于表示第 c 类第 m 帧样本所对应的列，得到第 c 类第 m 帧对应的子加载矩阵 $\mathbf{A}^{c,m}$ 。同文献[11]，假设其服从伯努利分布， $\mathbf{z}_{i,k}^{c,m} \sim \text{Ber}(\mathbf{z}_{i,k}^{c,m} | u_k^{c,m})$ ，也即

$$p(\mathbf{z}_{i,k}^{c,m} | u_k^{c,m}) = (u_k^{c,m})^{\mathbf{z}_{i,k}^{c,m}} (1 - u_k^{c,m})^{1 - \mathbf{z}_{i,k}^{c,m}} \quad (7)$$

因为同一帧样本选择的因子高度相同，设同一帧的 $\mathbf{z}_{i,k}^{c,m}$ 取 1 的概率相同，均为 $u_k^{c,m}$ ，同样采用共轭先验假设 $u_k^{c,m} \sim \text{Beta}(u_k^{c,m} | e_0, f_0)$ ， $\text{Beta}(\bullet)$ 表示贝塔分布， $u_k^{c,m}$ 的先验均值 $\langle u_k^{c,m} \rangle = e_0 / (e_0 + f_0)$ ，为了使 $\mathbf{z}_i^{c,m}$ 稀疏，也即其元素 $\mathbf{z}_{i,k}^{c,m}$ 以较大概率取 0，要求 $\langle u_k^{c,m} \rangle$ 近似为 0，因为 K 通常比较大，设 $e_0 = 1/K$ ； $f_0 = (K - 1)/K$ 。

一般假设复距离像中的噪声为复高斯白噪声，服从零均值，各维相互独立的复高斯分布：

$$p(\boldsymbol{\epsilon}_i^{c,m} | \delta^{c,m}) = \frac{1}{\pi^P |1/\delta^{c,m} \mathbf{I}_P|} \cdot \exp\left\{-\left(\boldsymbol{\epsilon}_i^{c,m}\right)^H (1/\delta^{c,m} \mathbf{I}_P)^{-1} \boldsymbol{\epsilon}_i^{c,m}\right\} \quad (8)$$

同样给 $\delta^{c,m}$ 以共轭先验，假设其服从伽马分布， $\delta^{c,m} \sim \text{Gamma}(\delta^{c,m} | a_0, b_0)$ ，设 $a_0 = b_0 = 10^{-6}$ 。

与 CFA 相比，基于 MTL-CFA 模型的复距离像建模既共享了加载矩阵，利用了不同方位帧之间的相关性，又利用 Beta-Bernoulli 稀疏先验，自动选择每一帧的因子个数，其贝叶斯图模型如图 1 所示。

$$\left. \begin{aligned} q(\mathbf{A}_k) &= \text{CN}(\mathbf{A}_k | \boldsymbol{\xi}_k, \mathbf{B}_k) \\ \mathbf{B}_k &= \left[\sum_{c=1}^C \sum_{m=1}^{M_c} \sum_{i=1}^N \langle \delta^{c,m} \rangle \langle (w_{i,k}^{c,m})^H w_{i,k}^{c,m} \rangle \langle \mathbf{z}_{i,k}^{c,m} \rangle + P \right]^{-1} \mathbf{I}_P \\ \boldsymbol{\xi}_k &= \mathbf{B}_k \sum_{c=1}^C \sum_{m=1}^{M_c} \langle \delta^{c,m} \rangle \sum_{i=1}^N \left\langle (w_{i,k}^{c,m})^H \mathbf{z}_{i,k}^{c,m} \right\rangle \left(x_i^{c,m} - \left\langle \sum_{j=1, j \neq k}^K \mathbf{A}_j w_{i,j}^{c,m} \mathbf{z}_{i,k}^{c,m} \right\rangle \right) \end{aligned} \right\} \quad (10)$$

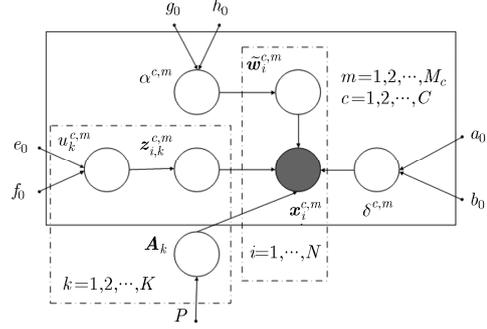


图 1 MTL-CFA 的贝叶斯图模型

按照贝叶斯理论^[16]，隐变量和参数的后验分布的推断方法主要有基于马尔可夫蒙特卡罗(MCMC)的吉布斯(Gibbs)采样方法^[16]和变分近似^[16]两种，考虑到计算精度和收敛速度，本文采用变分近似推断隐变量和参数的后验分布，下面介绍基于变分贝叶斯(VB)的后验分布的求解过程。

2.3 基于变分近似的隐变量后验分布推断

根据变分贝叶斯理论，完全数据由观测数据和隐变量构成，隐变量和参数的分布通过最大边缘似然准则^[14]求解。在 MTL-CFA 模型中，各目标各方位帧的所有的训练样本组成观测数据 \mathbf{X} ，隐变量与参数合起来记为 \mathbf{Z} ，参数先验分布中的超参数记为 θ 。若记隐变量和参数的联合概率密度函数为 $q(\mathbf{Z})$ ，那么观测数据的对数边缘似然可以表示为

$$\ln p(\mathbf{X} | \theta) = \int q(\mathbf{Z}) \ln p(\mathbf{X} | \mathbf{Z}, \theta) d\mathbf{Z} \quad (9)$$

一般来说，隐变量和参数之间具有一定的依赖关系，其联合分布的推导比较复杂，若假设其相互独立，则根据变分贝叶斯理论^[16]，每个隐变量或参数的分布正比于全数据的联合似然函数 $p(\mathbf{X}, \mathbf{Z} | \theta)$ 关于其他参数与隐变量的数学期望。变分贝叶斯推断就是更新隐变量和参数的分布 $q(\mathbf{Z})$ ，使得下界 LB 不断增大的过程， $q(\mathbf{Z})$ 的更新步骤如下：

(1)初始化：初始化隐变量和参数的分布，设定门限 $\kappa = 10^{-6}$ ，下界变化量 $\Delta = 1$ ，迭代计数 $t = 1$ ，下界 $LB(1) = \int q(\mathbf{Z}) \ln(p(\mathbf{X}, \mathbf{Z} | \theta)/q(\mathbf{Z})) d\mathbf{Z}$ ；

(2)更新 \mathbf{A}_k 的分布：对全数据联合似然关于除 \mathbf{A}_k 的隐变量和参数求数学期望，得到 \mathbf{A}_k 的分布：

其中, $\langle \bullet \rangle$ 表示求数学期望。

(3)更新 $\tilde{\mathbf{w}}_i^{c,m}$ 的分布: 对全数据联合似然关于除 $\tilde{\mathbf{w}}_i^{c,m}$ 的隐变量和参数求数学期望, 得到 $\tilde{\mathbf{w}}_i^{c,m}$ 的分布:

$$\left. \begin{aligned} q(\tilde{\mathbf{w}}_i^{c,m}) &= \text{CN}(\tilde{\mathbf{w}}_i^{c,m} | \boldsymbol{\omega}_i^{c,m}, \boldsymbol{\Sigma}_i^{c,m}) \\ \boldsymbol{\Sigma}_i^{c,m} &= \left\langle \alpha^{c,m} \right\rangle \mathbf{I}_K + \left\langle \delta^{c,m} \right\rangle \left\langle \mathbf{z}_i^{c,m} (\mathbf{z}_i^{c,m})^H \right\rangle \odot \left\langle \mathbf{A}^H \mathbf{A} \right\rangle^{-1} \\ \boldsymbol{\omega}_i^{c,m} &= \left\langle \delta^{c,m} \right\rangle \boldsymbol{\Sigma}_i^{c,m} \text{diag} \left\langle (\mathbf{z}_i^{c,m})^H \right\rangle \left\langle \mathbf{A}^H \right\rangle \mathbf{x}_i^{c,m} \end{aligned} \right\} \quad (11)$$

(4)更新 $\alpha^{c,m}$ 的分布: 对全数据联合似然关于除 $\alpha^{c,m}$ 的隐变量和参数求数学期望, 得到 $\alpha^{c,m}$ 的分布

$$\left. \begin{aligned} q(\alpha^{c,m}) &= \text{Gamma}(\alpha^{c,m} | g^{c,m}, h^{c,m}) \\ g^{c,m} &= g_0 + NK, \quad h^{c,m} = h_0 + \sum_{i=1}^N \left\langle (\tilde{\mathbf{w}}_i^{c,m})^H \tilde{\mathbf{w}}_i^{c,m} \right\rangle \end{aligned} \right\} \quad (12)$$

(5)更新 $u_k^{c,m}$ 的分布: 对全数据联合似然关于除 $u_k^{c,m}$ 的隐变量和参数求数学期望, 得到 $u_k^{c,m}$ 的分布

$$\left. \begin{aligned} q(u_k^{c,m}) &= \text{Beta}(u_k^{c,m} | e^{c,m}, f^{c,m}) \\ e_k^{c,m} &= e_0 + \sum_{i=1}^N \left\langle \mathbf{z}_{ik}^{c,m} \right\rangle, \quad f_k^{c,m} = f_0 + N - \sum_{i=1}^N \left\langle \mathbf{z}_{ik}^{c,m} \right\rangle \end{aligned} \right\} \quad (13)$$

(6)更新 $\mathbf{z}_{i,k}^{c,m}$ 的分布: 对全数据联合似然关于除 $\mathbf{z}_{i,k}^{c,m}$ 的隐变量和参数求数学期望, 得到

$$\left. \begin{aligned} q(\mathbf{z}_{i,k}^{c,m}) &= \text{Ber} \left(\mathbf{z}_{i,k}^{c,m} \mid \frac{Ep_{i,k}^{c,m}}{Ep_{i,k}^{c,m} + Eb_{i,k}^{c,m}} \right), \quad Eb_{i,k}^{c,m} \propto \exp \left\langle \ln(1 - u_k^{c,m}) \right\rangle \\ Ep_{i,k}^{c,m} &\propto \exp \left\langle \ln(u_k^{c,m}) \right\rangle - \left\langle \delta^{c,m} \left(\mathbf{x}_i^{c,m} - \sum_{j=1}^K \mathbf{A}_j \mathbf{z}_{i,j}^{c,m} \mathbf{w}_{i,j}^{c,m} \right)^H \left(\mathbf{x}_i^{c,m} - \sum_{j=1}^K \mathbf{A}_j \mathbf{z}_{i,j}^{c,m} \mathbf{w}_{i,j}^{c,m} \right) \right\rangle \end{aligned} \right\} \quad (14)$$

(7)更新 $\delta^{c,m}$ 的分布: 对全数据联合似然关于除 $\delta^{c,m}$ 的隐变量和参数求数学期望, 得到 $\delta^{c,m}$ 的分布

$$\left. \begin{aligned} q(\delta^{c,m}) &= \text{Gamma}(\delta^{c,m} | a^{c,m}, b^{c,m}), \quad a^{c,m} = a_0 + NP \\ b^{c,m} &= b_0 + \sum_{i=1}^N \left\langle \left(\mathbf{x}_i^{c,m} - \mathbf{A} \text{diag}(\mathbf{z}_i^{c,m}) \mathbf{w}_i^{c,m} \right)^H \left(\mathbf{x}_i^{c,m} - \mathbf{A} \text{diag}(\mathbf{z}_i^{c,m}) \mathbf{w}_i^{c,m} \right) \right\rangle \end{aligned} \right\} \quad (15)$$

(8)终止条件: 计算下界 $LB(t+1)$ 以及下界变化率 Δ , 同时令 $t = t + 1$, 当下界变化率 Δ 小于门限 κ 时, 迭代截止, 否则转至步骤(2)循环更新。

3 基于 MTL-CFA 的高分辨距离像的统计识别

本文采用统计识别方法, 根据最大后验准则的贝叶斯分类器判定测试样本的类别。首先将训练样本分帧, 然后采用 MTL-CFA 得到各帧样本的概率密度函数, 根据贝叶斯公式, 测试样本 \mathbf{x}^* 属于第 c 类第 m 帧的后验概率正比于先验分布和帧概率密度函数的乘积:

$$p(c, m | \mathbf{x}^*) \propto p(c, m) p(\mathbf{x}^* | c, m) \quad (16)$$

一般情况下, 假设先验分布 $p(c, m)$ 为均匀分布, 而帧概率密度函数可以通过式(17)得到

$$p(\mathbf{x} | c, m) = \iiint \text{CN}(\mathbf{x} | \mathbf{A}(\mathbf{w}^{c,m}), (\delta^{c,m})^{-1} \mathbf{I}_P) \cdot q(\delta^{c,m}) q(\mathbf{w}^{c,m}) q(\mathbf{A}) d\delta^{c,m} d\mathbf{w}^{c,m} d\mathbf{A} \quad (17)$$

其中隐变量和参数的分布 $q(\delta^{c,m})$, $q(\mathbf{w}^{c,m})$ 和 $q(\mathbf{A})$ 由 2.3 节得到。实际中, 式(17)的积分是很难计算的,

即使像本文中采用共轭先验根据最大边缘似然得到隐变量和参数分布具体形式的情况下。常用的解决办法是, 对难以积分隐变量和参数, 例如 \mathbf{A} 和 $\delta^{c,m}$, 将其后验均值代入式(17), 对式(17)关于剩余的隐变量和参数进行积分:

$$p(\mathbf{x} | c, m) = \int \text{CN}(\mathbf{x} | \langle \mathbf{A} \rangle (\mathbf{w}^{c,m}), \langle \delta^{c,m} \rangle^{-1} \mathbf{I}_P) \cdot q(\mathbf{w}^{c,m}) d\mathbf{w}^{c,m} \quad (18)$$

由于样本 \mathbf{x} 和噪声均服从零均值的复高斯分布, 可以合理假设 $\mathbf{w}^{c,m}$ 服从零均值的复高斯分布, 即 $q(\mathbf{w}^{c,m}) = \text{CN}(\mathbf{w}^{c,m} | \mathbf{0}, \boldsymbol{\Omega}^{c,m})$, 其中

$$\boldsymbol{\Omega}^{c,m} = \frac{1}{N} \sum_{i=1}^N \left\langle (\tilde{\mathbf{w}}_i^{c,m} \odot \mathbf{z}_i^{c,m}) (\tilde{\mathbf{w}}_i^{c,m} \odot \mathbf{z}_i^{c,m})^H \right\rangle \quad (19)$$

于是, 帧概率密度函数可表示为

$$p(\mathbf{x} | c, m) = \text{CN}(\mathbf{x} | \langle \mathbf{A} \rangle \boldsymbol{\Omega}^{c,m} \langle \mathbf{A}^H \rangle + \langle \delta^{c,m} \rangle^{-1} \mathbf{I}_P) \quad (20)$$

根据贝叶斯公式, 对于测试样本 \mathbf{x}^* , 将其判为后验概率最大的一帧, 该帧对应的类别即为测试样本的类别:

$$\text{label}(\mathbf{x}^*) = \arg \max_c p(c, m | \mathbf{x}^*) \quad (21)$$

4 实验结果和讨论

4.1 数据及实验设置介绍

为了验证本文提出的模型的识别性能，我们采用同文献[17]等相同的 3 类飞机的实测数据，雷达和飞机目标的详细参数如表 1 和表 2 所示。3 类飞机目标的航迹在地面的投影见图 2 所示，雷达位于坐标(0,0)处。每类目标数据分成多段，在本实验中，雅克-42 的第 2,5 段，奖状飞机的第 6,7 段，安-26 飞机的第 5,6 段样本用于训练，估计帧概率密度函数，其余段样本用于测试，验证模型的识别性能。

为了克服姿态敏感性，采用同文献[17]一致的做法对训练样本按角域均匀分帧，奖状和安-26 各 50 帧，雅克-42 有 35 帧，每帧含有 1024 个样本，每个

表 1 实验雷达参数

信号类型	Chirp 信号
中心频率(MHz)	5520
带宽(MHz)	400
脉冲重复频率(Hz)	400
Dechirp 后采样频率(MHz)	10

表 2 飞机参数

飞机类别	机长(m)	翼展(m)	机高(m)
安-26	23.80	29.20	9.83
奖状	14.40	15.90	4.57
雅克-42	36.38	34.88	9.83

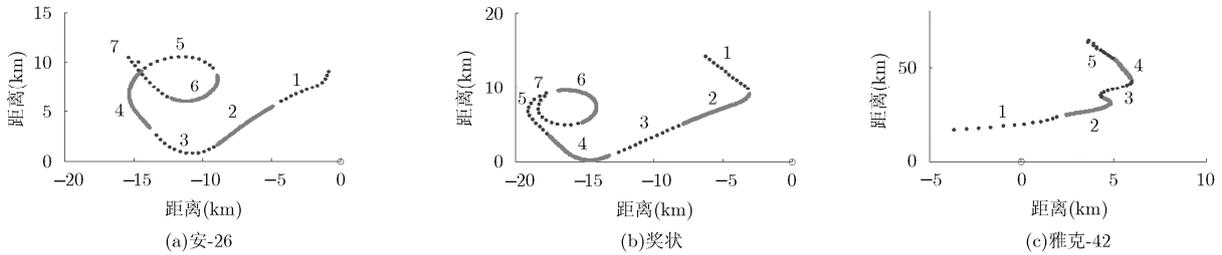


图 2 3 类目标飞行航迹

样本有 256 个距离单元，即 $C = 3$, $M_1 = 35$, $M_2 = M_3 = 50$, $N = 1024$, $P = 256$ ；同样，为了克服强度敏感性，对每个训练样本和测试样本采用 L2 范数归一化。经过上述预处理之后，数据就可以用于模型的训练和识别实验的测试。

4.2 模型选择性能分析

在 MTL-CFA 模型中，我们假设每一帧的子加载矩阵 $A^{c,m}$ 由共享加载矩阵 A 的某些列组成，具有稀疏性。为了验证各帧子加载矩阵原子的共享性和稀疏性，图 3 给出在每帧样本数为 70，共享矩阵初始列 $K = 256$ 时，3 类目标各取 5 帧的 $u_k^{c,m}$ 的后验均值分布图， $u_k^{c,m}$ 大于 0.5 的取 1，用黑色表示，小于 0.5 的取 0，用白色表示。图中横坐标表示帧标号，可以看出，3 类目标各帧的 $\langle u_k^{c,m} \rangle$ 的大部分值是小于 0.5 的，说明各帧子加载矩阵确实在共享加载矩阵上具有稀疏性；而且同一类不同帧 $\langle u_k^{c,m} \rangle$ 大于 0.5 对应的行比较相似，其对应的原子有一定的重叠，具有共享性，而不同类目标之间共享的原子又有所差别，这对识别是有利的。

4.3 相关模型比较

下面我们将模型复杂度、识别性能等两个方面比较本文提出的 MTL-CFA 与 STL-CFA^[11]及 2.1

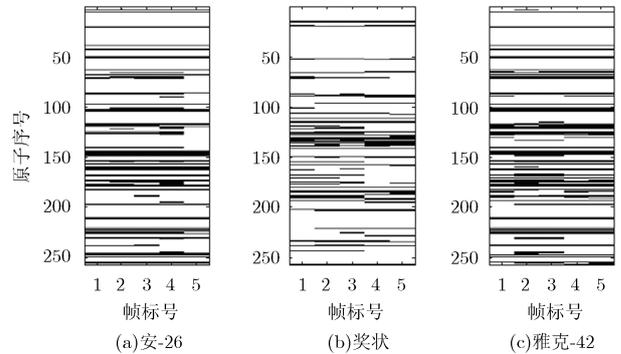


图 3 3 类目标 $u_k^{c,m}$ 的后验分布图

节提到的 STL-RFA^[8]的优劣，为了比较的公平性，STL-CFA 与 STL-RFA 中的因子个数均经过 BIC 模型选择确定。

4.3.1 模型复杂度比较 一般来说，模型参数越多，复杂度越高，所需的训练样本就越多，各模型用于 3 类飞机数据的实参数个数见表 3。

图 4 给出各个模型的总实参数个数随样本数的变化图。由于 MTL-CFA 学习过程中是多帧模型同时学习，对内存的需求量较大，本文在 8 核 i7-2.80 GHz 处理器，32 G 内存的移动工作站上进行实验，当每帧样本个数大于 130 时，对内存的需求大于 32

表 3 不同模型中实参数个数

方法	参数个数
STL-RFA	$\sum_{c=1}^C \sum_{m=1}^{M_c} PK_m^c + 1 - \frac{1}{2} K_m^c (K_m^c - 1),$ $K_m^c \text{ 表示第 } c \text{ 类第 } m \text{ 帧的因子个数}$
STL-CFA	$\sum_{c=1}^C \sum_{m=1}^{M_c} 2PK_m^c + 1 - (K_m^c)^2$
MTL-CFA	$2PK_s - K_s^2 + \sum_{c=1}^M M_c, K_s \text{ 表示共享因子个数, } M_c \text{ 表示第 } c \text{ 类的帧数}$

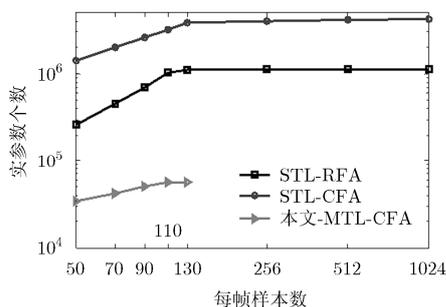


图 4 总参数随样本数的变化图

G, 所以本文中 MTL-CFA 的实验只做到每帧 130 个样本。从图中可以看出, 各模型估计的参数个数随着样本数的增加呈上升趋势, 这是由于文献[8,11]采用 BIC 自动选择因子个数, MTL-CFA 也通过稀疏先验自动确定因子个数, 各模型能够选择适合数据样本个数的复杂度。另外可以看出在不同样本数下, MTL-CFA 需要估计的参数个数明显少于 STL-RFA 和 STL-CFA, 其中, STL-CFA 需要估计的参数最多, 因此其复杂度最高。

4.3.2 识别性能比较 由于不同模型在不同样本数下模型参数个数不同, 其识别性能也有差异。各模型识别率随样本数的变化如图 5 所示。与 4.3.1 节复杂度比较实验类似, 由于内存的限制, MTL-CFA 的识别实验也只做到每帧 130 个样本。从图 5 可以看出, 各个模型下的识别率均随着样本数的增加在上升。当每帧样本数在 70~130 之间时, MTL-CFA 的识别率明显高于 STL-RFA 和 STL-CFA, STL-CFA 的识别率最低。若以识别率高于 0.85 为基准确定样本个数, MTL-CFA 只需要 70 个样本, 而 STL-RFA 和 STL-CFA 需要约 160 和 240 个样本, 约为 MTL-CFA 所需样本数的 2 倍和 3 倍。因此, 在相同识别性能时, MTL-CFA 所需的样本数更少, 更适用于小样本问题。随着样本数增多, STL-CFA 和 STL-RFA 的识别性能趋于稳定。当每帧样本数等于或大于 512 时, STL-CFA 的识别性能优于 STL-

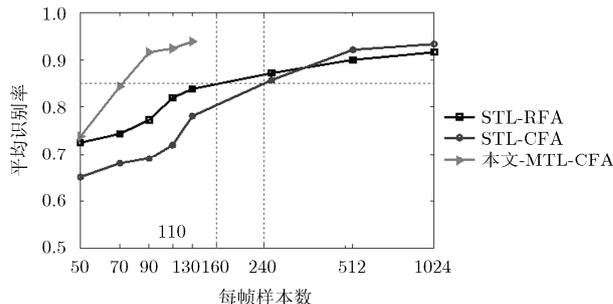


图 5 识别率随样本数的变化曲线

RFA, 说明复距离像的相位信息有助于改善识别性能。图 5 中 MTL-CFA 每帧 130 个样本时的平均识别率是 93.78%, STL-CFA 每帧 1024 个样本时的平均识别率是 93.22%, STL-RFA 每帧 1024 个样本时的平均识别率是 91.49%, 因此 MTL-CFA 的识别性能是最优的。

5 结论

本文提出了一种基于 MTL-CFA 模型, 在多类目标帧概率密度函数学习过程中, 利用不同子任务之间的相关性, 共享加载矩阵, 各帧自动选择合适复杂度的因子个数。该模型不仅减小了统计模型复杂度和对训练样本数需求, 而且充分利用了复距离像的相位信息。基于实测数据的识别实验表明, 在样本较少的情况下 MTL-CFA 可以获得更好的识别性能, 适用于小样本统计识别问题。

参考文献

- [1] 张瑞, 牛威, 寇鹏. 基于样本紧密度的雷达高分辨距离像识别方法研究[J]. 电子与信息学报, 2014, 36(3): 529-536.
Zhang Rui, Niu Wei, and Kou Peng. Radar high resolution range profiles recognition based on the affinity[J]. *Journal of Electronics & Information Technology*, 2014, 36(3): 529-536.
- [2] 张学峰, 王鹏辉, 冯博, 等. 基于多分类器融合的雷达高分辨距离像目标识别与拒判新方法[J]. 自动化学报, 2014, 40(2): 348-356.
Zhang Xue-feng, Wang Peng-hui, Feng Bo, et al. A new method to improve radar HRRP recognition and outlier rejection performances based on classifier combination[J]. *Acta Automatica Sinica*, 2014, 40(2): 348-356.
- [3] 张玉玺, 王晓丹, 姚旭, 等. 基于复数全极化 HRRP 的雷达目标识别[J]. 系统工程与电子技术, 2014, 36(2): 260-265.
Zhang Yu-xi, Wang Xiao-dan, Yao Xu, et al. Radar target recognition based on complex fully polarimetric HRRP[J]. *Systems Engineering and Electronics*, 2014, 36(2): 260-265.
- [4] 潘勉, 王鹏辉, 杜兰, 等. 基于 TSB-HMM 模型的雷达高分辨距离像目标识别方法[J]. 电子与信息学报, 2013, 35(7): 1547-1554.

- Pan Mian, Wang Peng-hui, Du Lan, *et al.*. Radar HRRP target recognition based on truncated stick-breaking hidden markov model[J], *Journal of Electronics & Information Technology*, 2013, 35(7): 1547-1554.
- [5] 姚莉娜, 吴艳敏, 崔光照. 基于随机森林的雷达高分辨距离像目标识别新方法[J]. 郑州大学学报(工学版), 2014, 35(4): 105-108.
- Yao Li-na, Wu Yan-min, and Cui Guang-zhao. New method of radar high resolution range image target recognition based on random forest[J]. *Journal of Zhengzhou University (Engineering Science)*, 2014, 35(4): 105-108.
- [6] Jacobs S P and O'Sullivan J A. Automatic target recognition using sequences of high resolution radar range profiles[J]. *IEEE Transactions on Aerospace Electronic Systems*, 2000, 36(2): 364-380.
- [7] Du L, Liu H, and Bao Z. Radar HRRP statistical recognition: parametric model and model selection[J]. *IEEE Transactions on Signal Processing*, 2008, 56(5): 1931-1944.
- [8] Du L, Liu H, and Bao Z. A two-distribution compounded statistical model for radar HRRP target recognition[J]. *IEEE Transactions on Signal Processing*, 2006, 54(6): 2226-2238.
- [9] Du L, Liu H, and Bao Z. A novel feature vector using complex HRRP for radar target recognition[J]. *Lecture Notes in Computer Science*, 2007, 4491: 1303-1309.
- [10] 王鹏辉, 杜兰, 刘宏伟. 基于复高斯模型的雷达高分辨距离像目标识别新方法[J]. 光学学报, 2014, 34(2): 275-284.
- Wang Peng-hui, Du Lan, and Liu Hong-wei. A new method based on complex Gaussian models for radar high resolution range profile target recognition[J]. *Acta Optica Sinica*, 2014, 34(2): 275-284.
- [11] 王鹏辉. 基于统计建模的雷达高分辨距离像目标识别方法研究[D]. [博士论文], 西安电子科技大学, 2012.
- Wang Peng-hui. Study of radar high resolution range profile target recognition based on statistical modeling[D]. [Ph.D dissertation], Xidian University, 2012.
- [12] Xue Y, Liao X, and Carin L. Multi-task learning for classification with Dirichlet process priors[J]. *The Journal of Machine Learning Research*, 2007, 8: 35-63.
- [13] Pan M, Du L, and Wang P. Multi-task hidden Markov modeling of spectrogram feature from radar high-resolution range profiles[J]. *EURASIP Journal on Advances in Signal Processing*, 2012(1): 1-17.
- [14] Watanabe S. A widely applicable Bayesian information criterion[J]. *The Journal of Machine Learning Research*, 2013, 14(1): 867-897.
- [15] Paisley J and Carin L. Nonparametric factor analysis with beta process priors[C]. Proceedings of the 26th ACM Annual International Conference on Machine Learning, 2009: 777-784.
- [16] Ando T. Bayesian Model Selection and Statistical Modeling [M]. CRC Press, 2010, 43-96.
- [17] Du L, Liu H, and Wang P. Noise robust radar HRRP target recognition based on multitask factor analysis with small training data size[J]. *IEEE Transactions on Signal Processing*, 2012, 60(7): 3546-3559.
- 和 华: 男, 1988 年生, 博士生, 研究方向为雷达目标识别.
- 杜 兰: 女, 1980 年生, 教授, 博士生导师, 研究方向为统计信号处理、雷达信号处理、机器学习及其在雷达目标检测与识别方面的应用.
- 刘宏伟: 男, 1971 年生, 教授, 博士生导师, 研究方向为雷达信号处理、MIMO 雷达、雷达目标识别、自适应信号处理、认知雷达等.