

孤立字语音识别

基于Python实现

时域特征提取与识别

数据收集

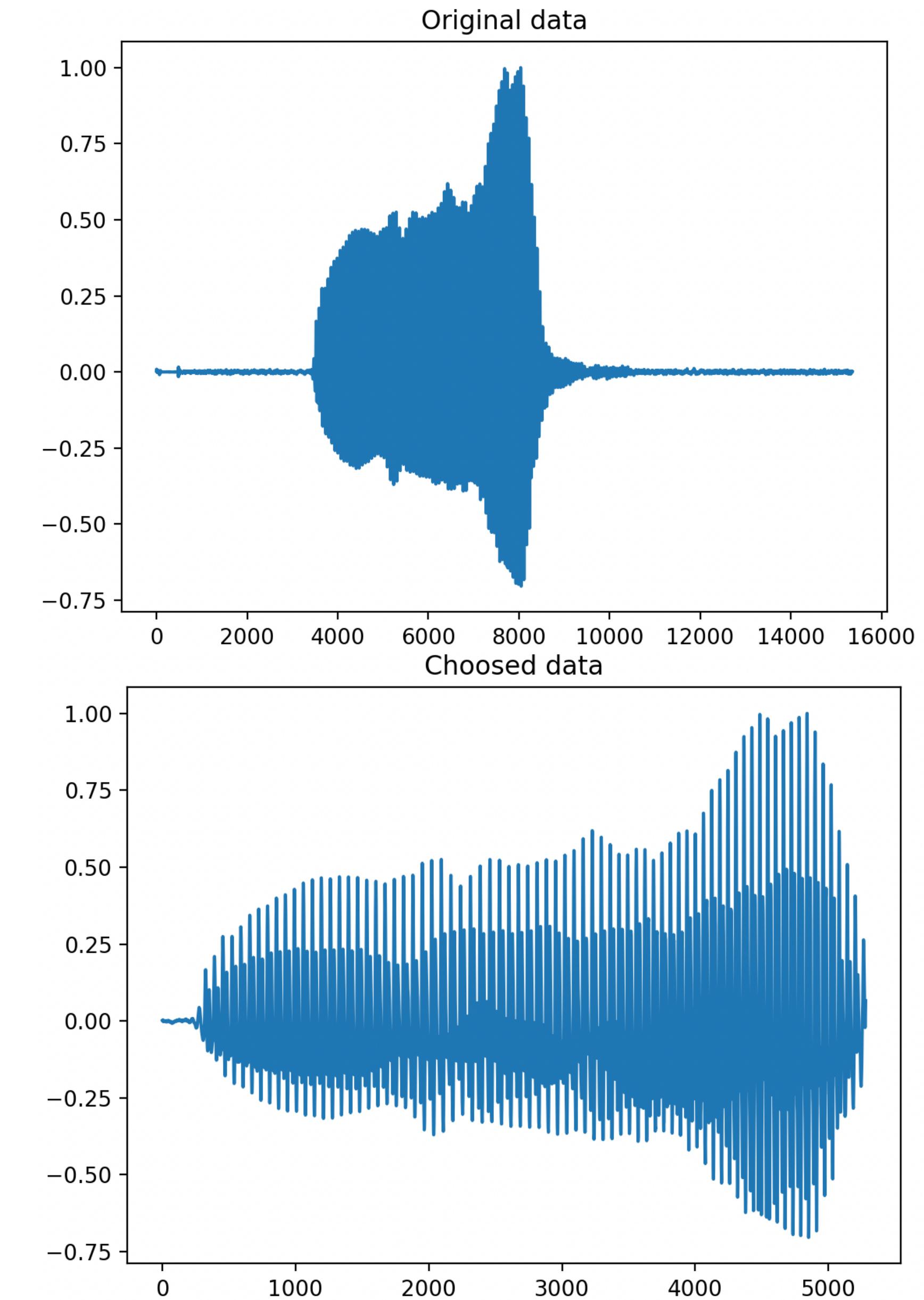
基于PyAudio实现

- 选取16位宽存储数据
- 采样率8000Hz
- 可连续采样
- 可采样不同数字
- 存储文件格式: number_times.wav, 每个文件存储一个数字

```
In platform: posix
Input the number: 0
Begin to collect data, input any to begin, "c" to change number, "q" to quit
1
In collecting number=0, times=0
* recording
100%|██████████| 15/15 [00:01<00:00, 7.86it/s]
* done recording
```

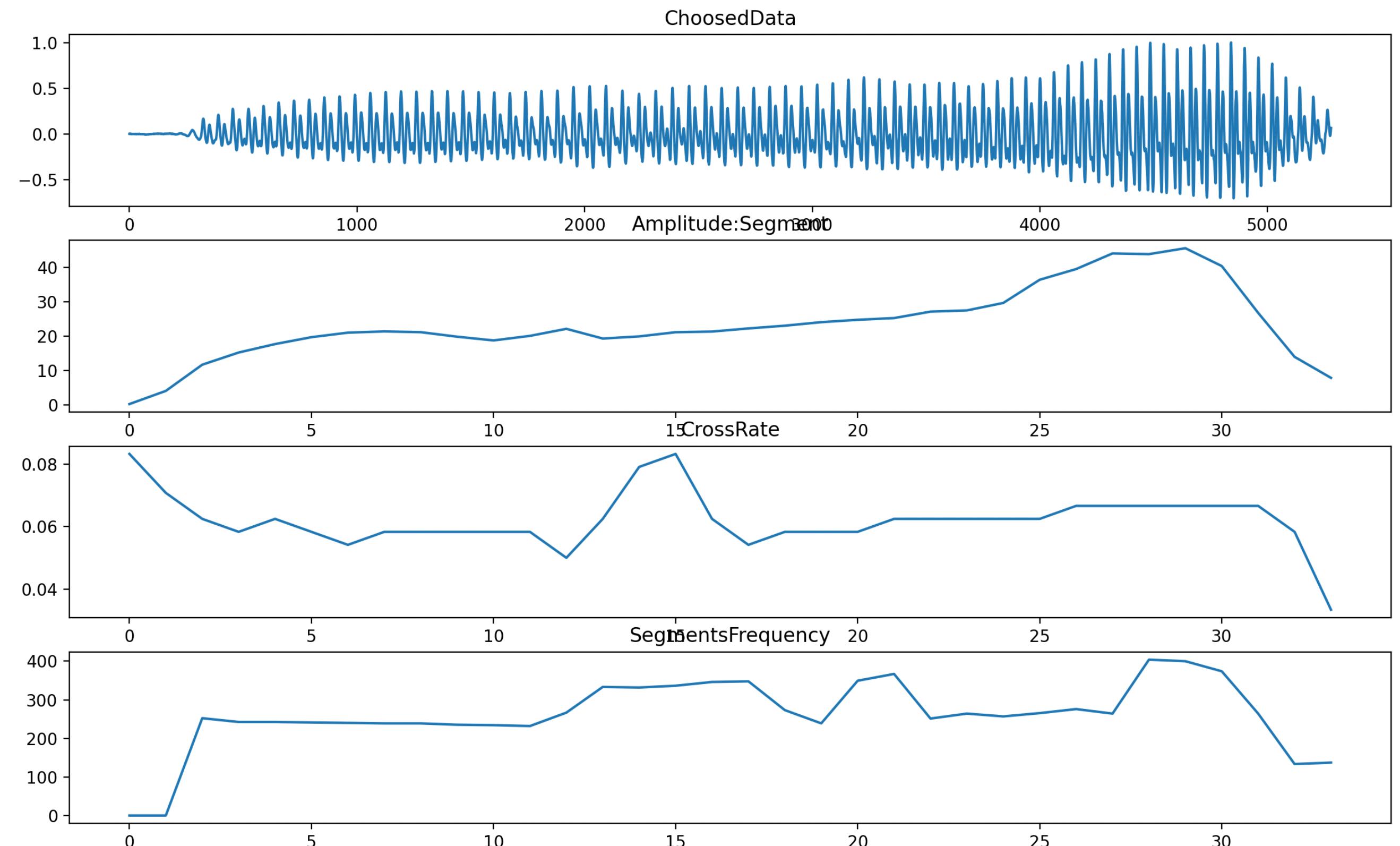
语音信号预处理

- 幅度归一化: 找出一段语音中最大的幅度差作为1进行归一化
- 语音分帧: 以30ms为一帧, 对原始语音信号分帧
- 加窗: 使用numpy.hanning, numpy.hamming生成汉宁窗与海明窗, numpy.ones生成矩形窗, 并对每一帧相乘
- 端点检测: 对分帧后每一帧信号计算能量和过零率, 先规定能量阈值并选取在能量范围内的帧作为第一次选取序列, 在此基础上分别向前向后搜索过零率大于阈值的帧作为最终选取序列
- 自动处理: 循环读入文件按照对应阈值处理



时域特征提取

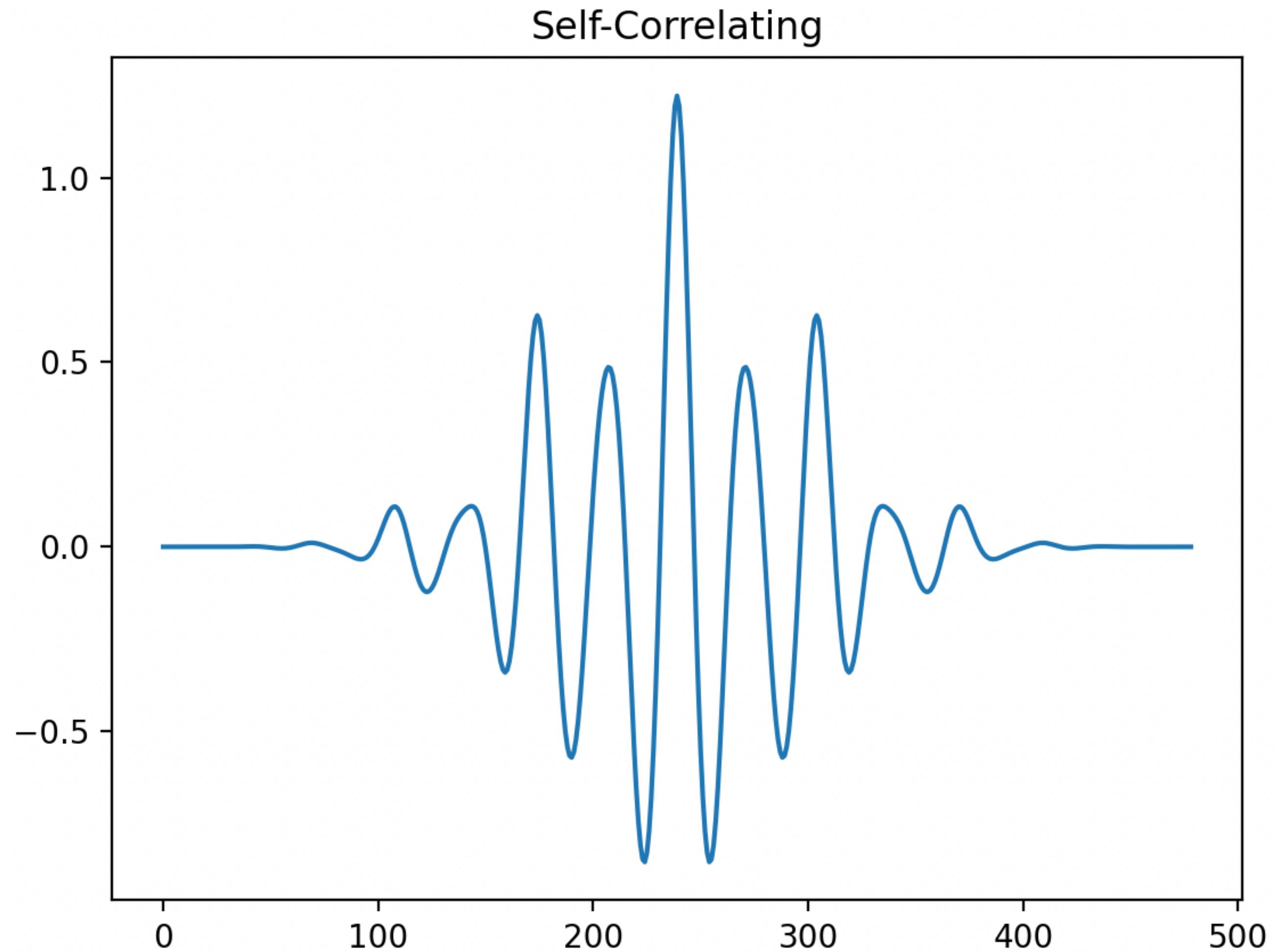
- 能量: 以每一帧的能量作为一个数据维度
- 过零率: 以每一帧的过零率作为一个数据维度
- 基音频率: 通过自相关运算以及峰值搜索算法获得浊音基音频率



时域特征提取

基音频率

- 浊音有一定的基音频率
- 清音不具备显著的基音频率
- 通过自相关运算得到幅值的类周期
- 自相关算法时间复杂度 $O(n^2)$, 故使用较低的采样率减少数据量
- 峰值搜索算法采用自动多尺度峰值查找算法(AMPD)^[1]



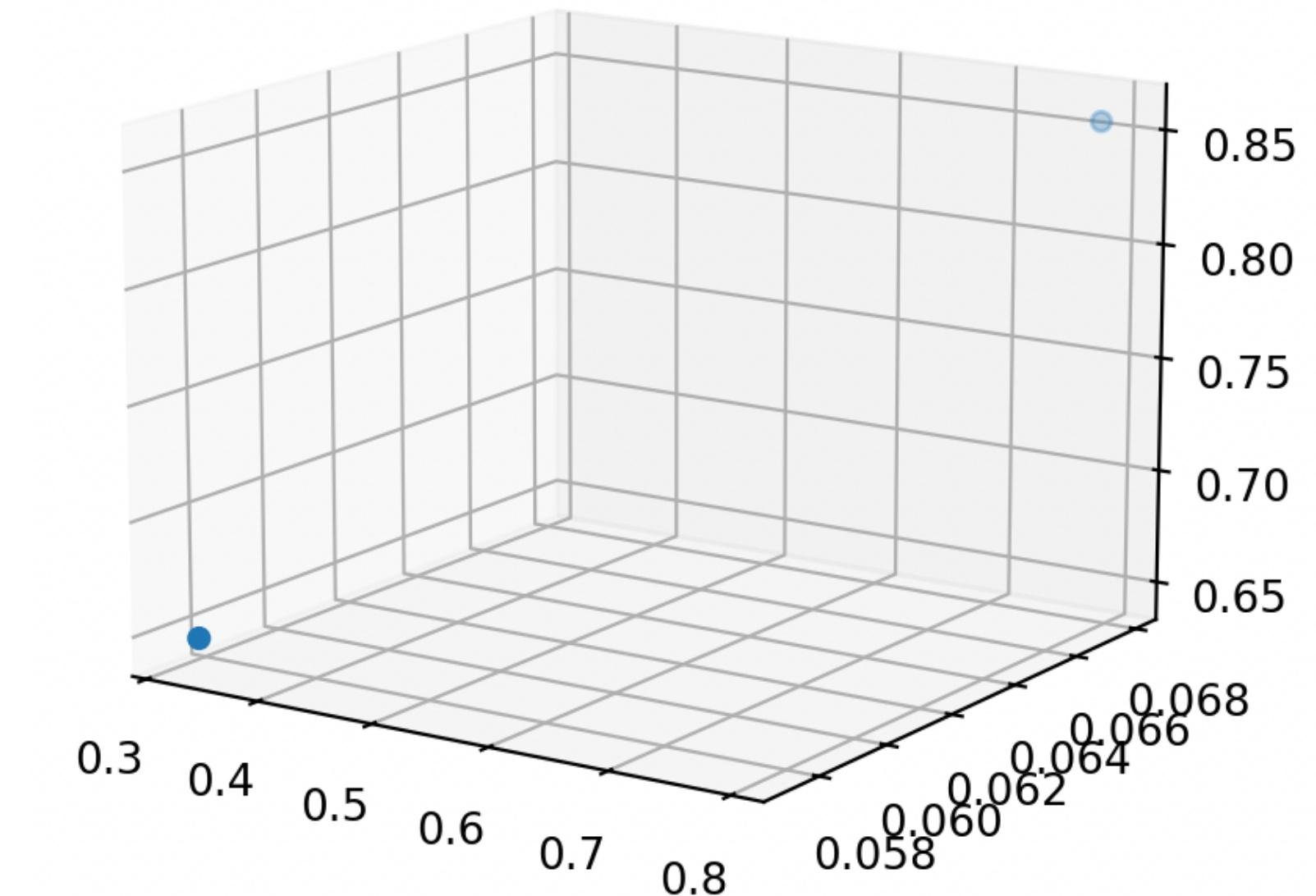
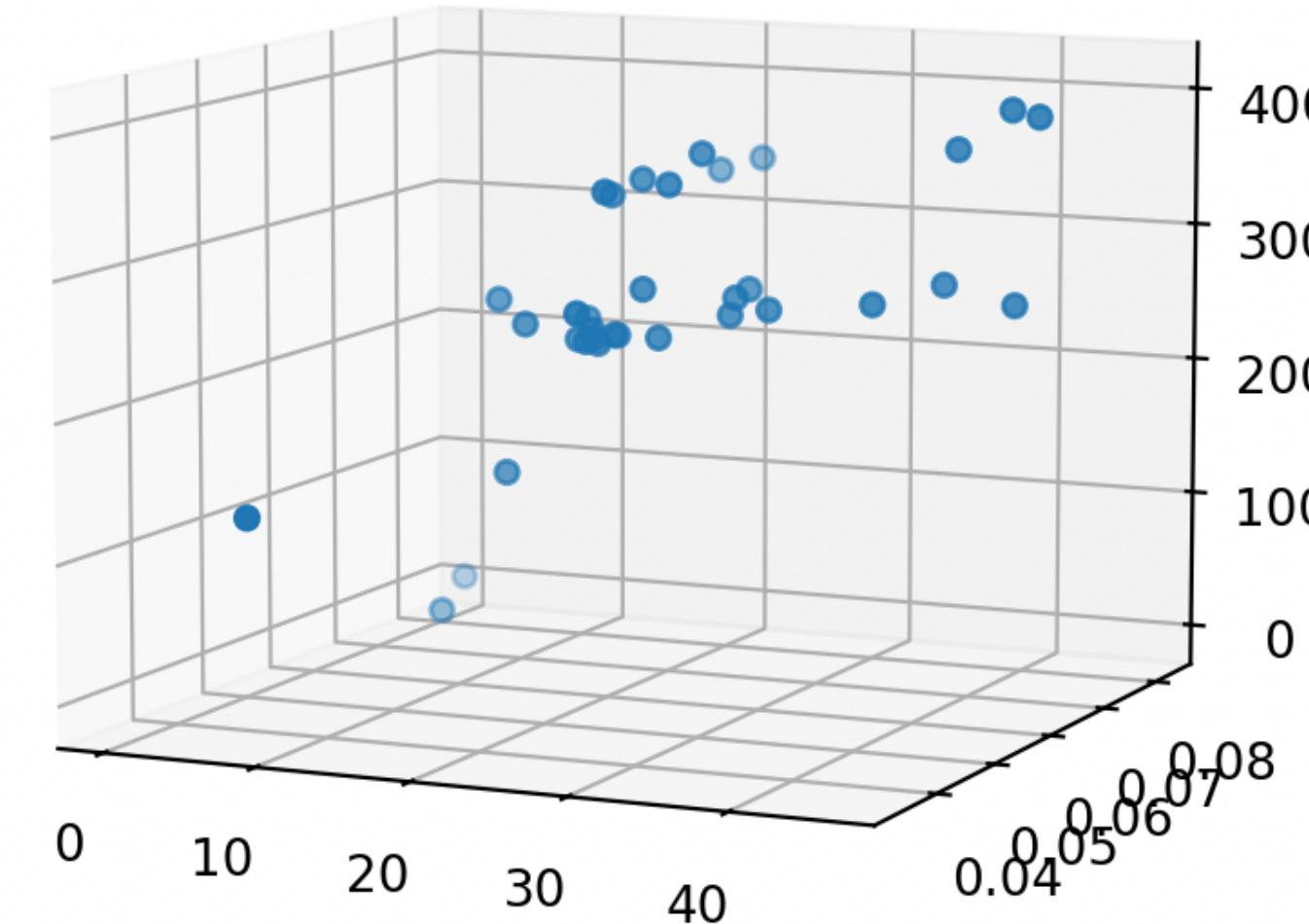
[1] Scholkmann, F.; Boss, J.; Wolf, M. An Efficient Algorithm for Automatic Peak Detection in Noisy Periodic and Quasi-Periodic Signals. *Algorithms* **2012**, *5*, 588–603.

<https://doi.org/10.3390/a5040588>

时域特征分析

聚类分析

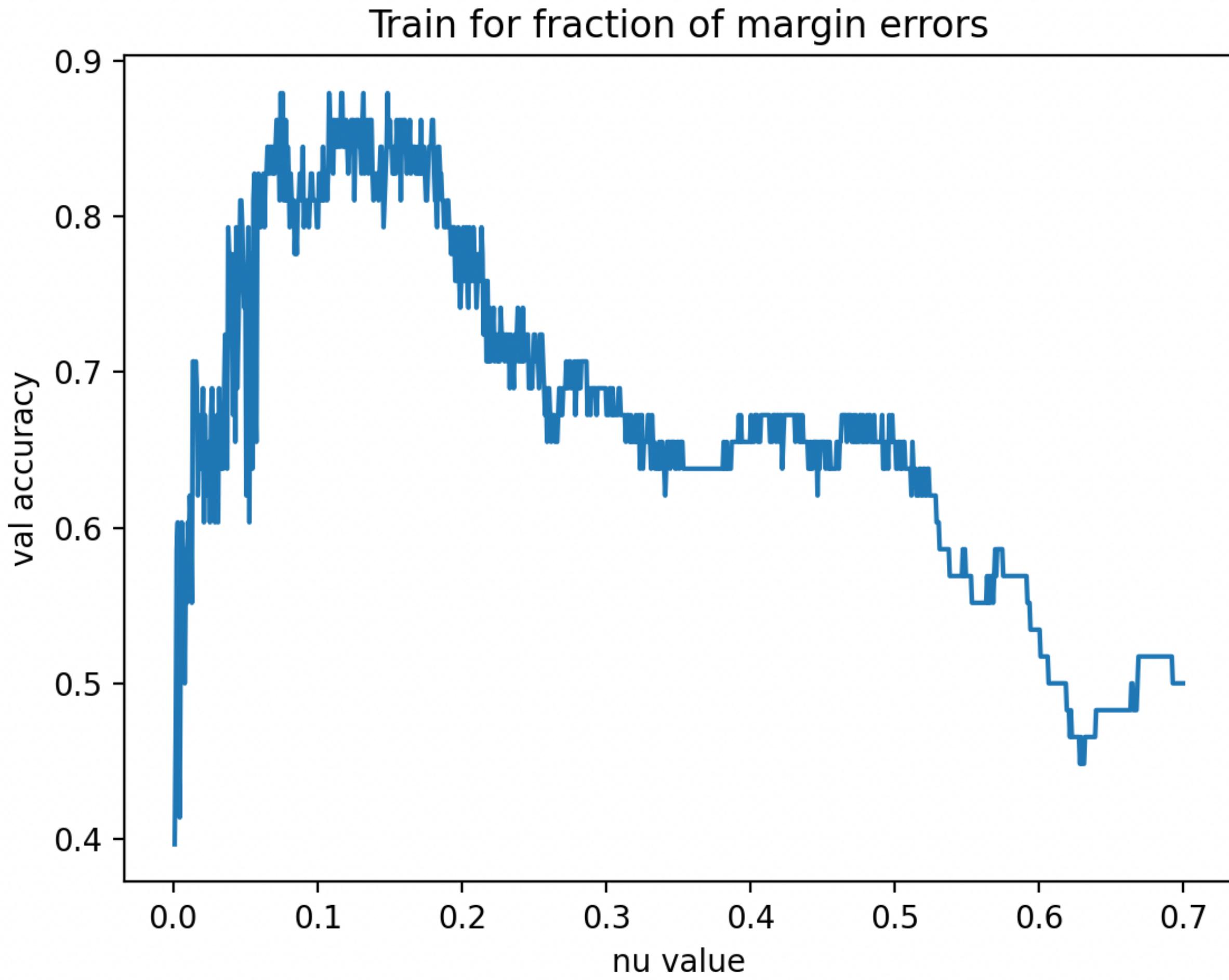
- 将每一帧信号的三个特征作为三个维度
- 清音浊音以是否有基音频率作为判别依据
- 设置基音频率阈值，当超过该阈值的作为浊音保存当前频率为一个特征
- 未超过该阈值的作为清音将基音频率置为0保存
- 归一化坐标轴，为每一个特征赋予同样的权重
- 聚类为清音点和浊音点



时域特征识别

使用sklearn.svm.NuSVC训练

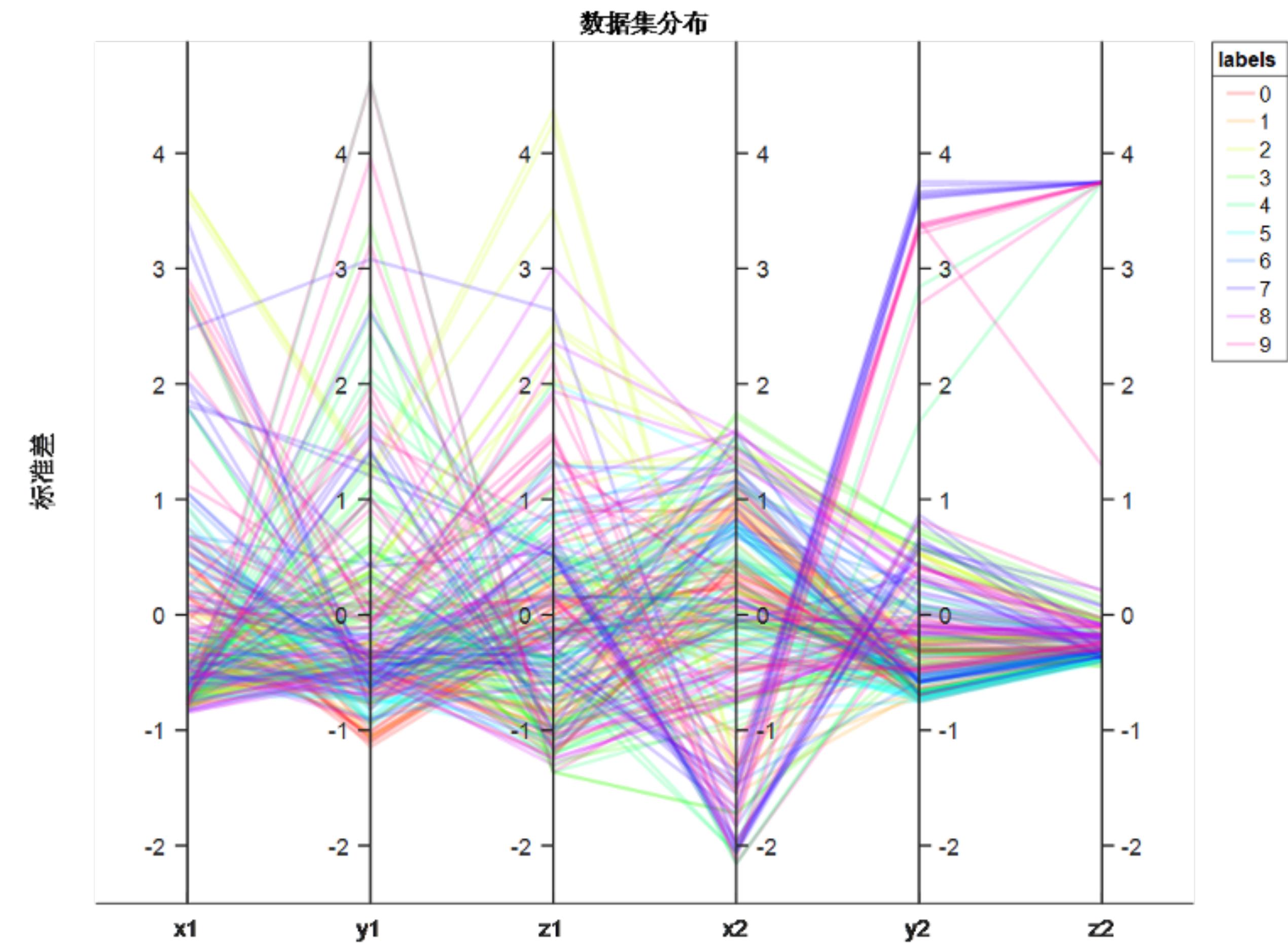
- 采用留出法(hold-out)随机选取1/5作为验证集, 剩余为训练集进行训练
- 使用在验证集中的准确率作为衡量标准
- 对NuSVM方法使用高斯核, 迭代优化边际误差因数(nu)
- 最终得到在测试集中当nu=0.07时, 达到最大验证集准确率86.21%



交融实验分析

基于MATLAB分析数据

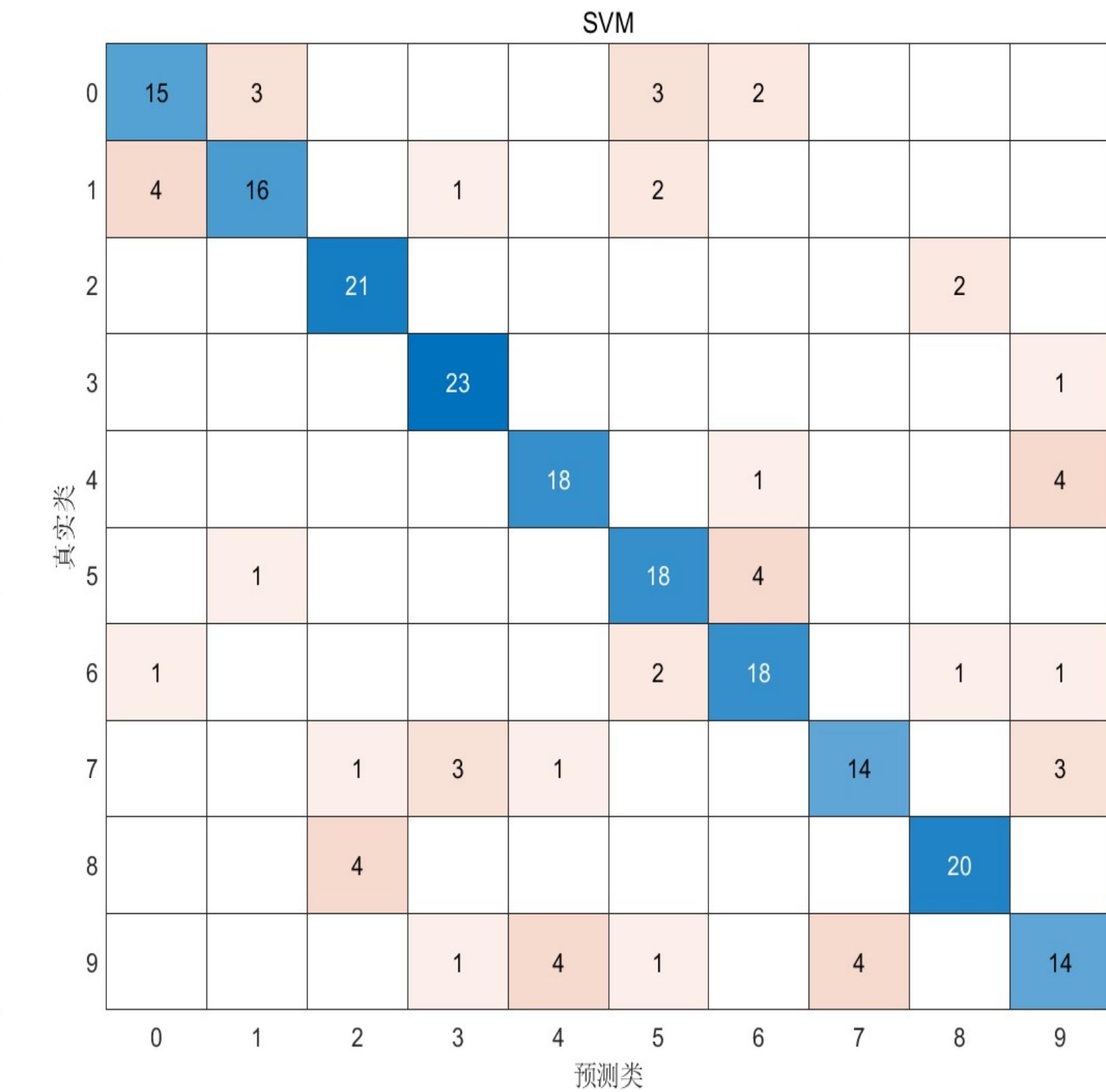
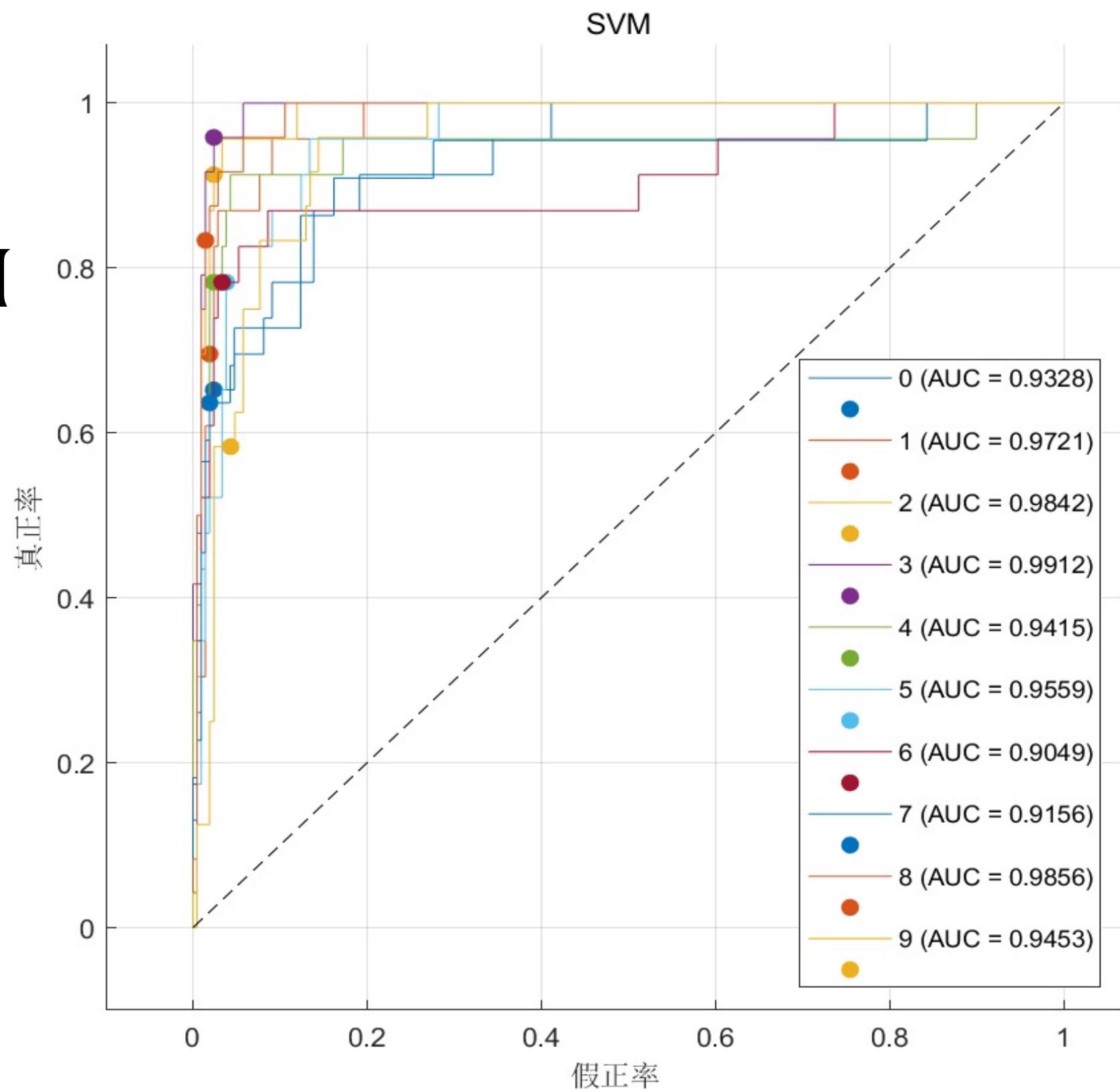
- 数据集分析



交融实验分析

基于MATLAB分析数据

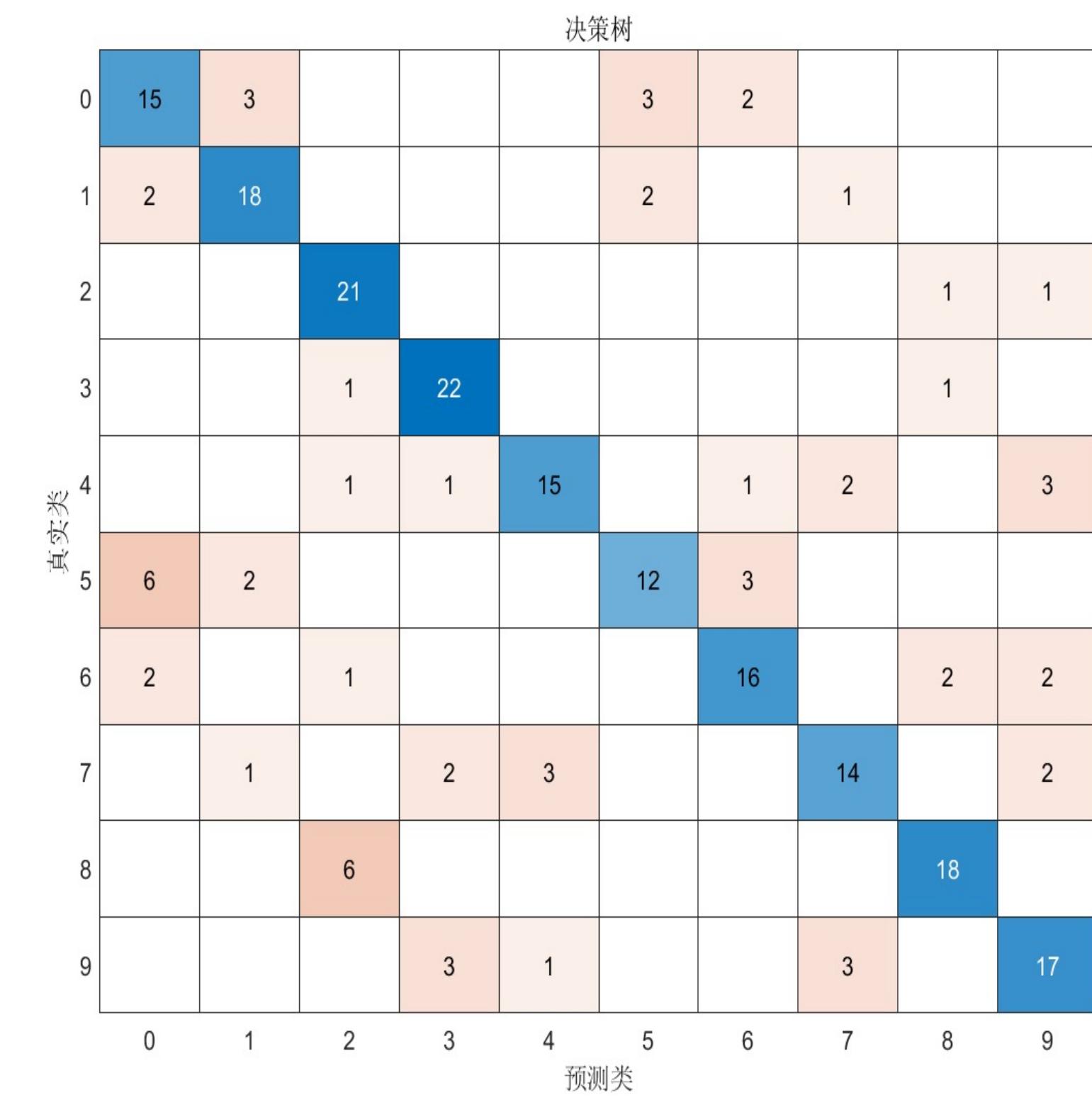
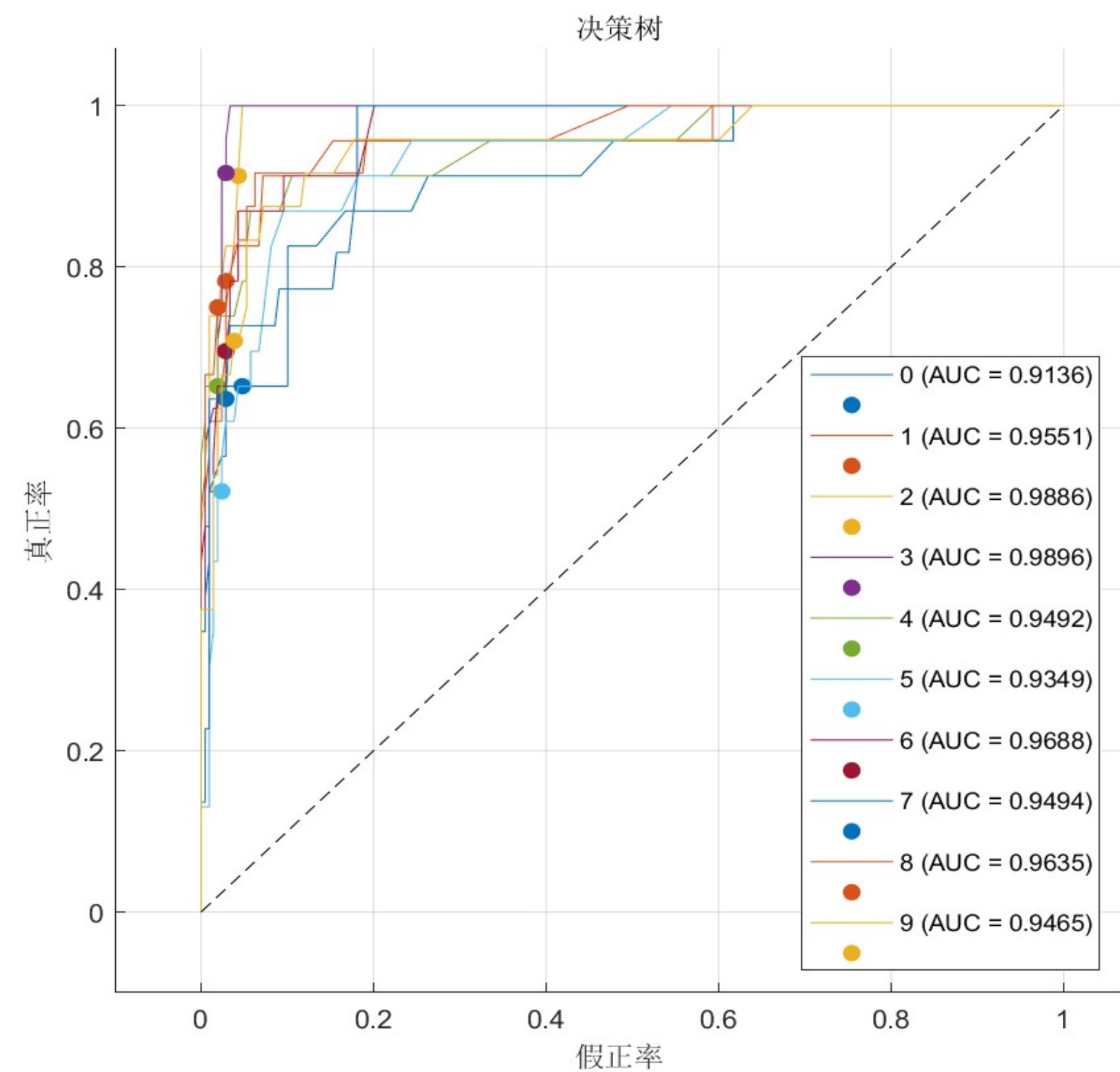
- 数据集分析
- 不同分类器效果分析-SVM
- 采用矩形窗



交融实验分析

基于MATLAB分析数据

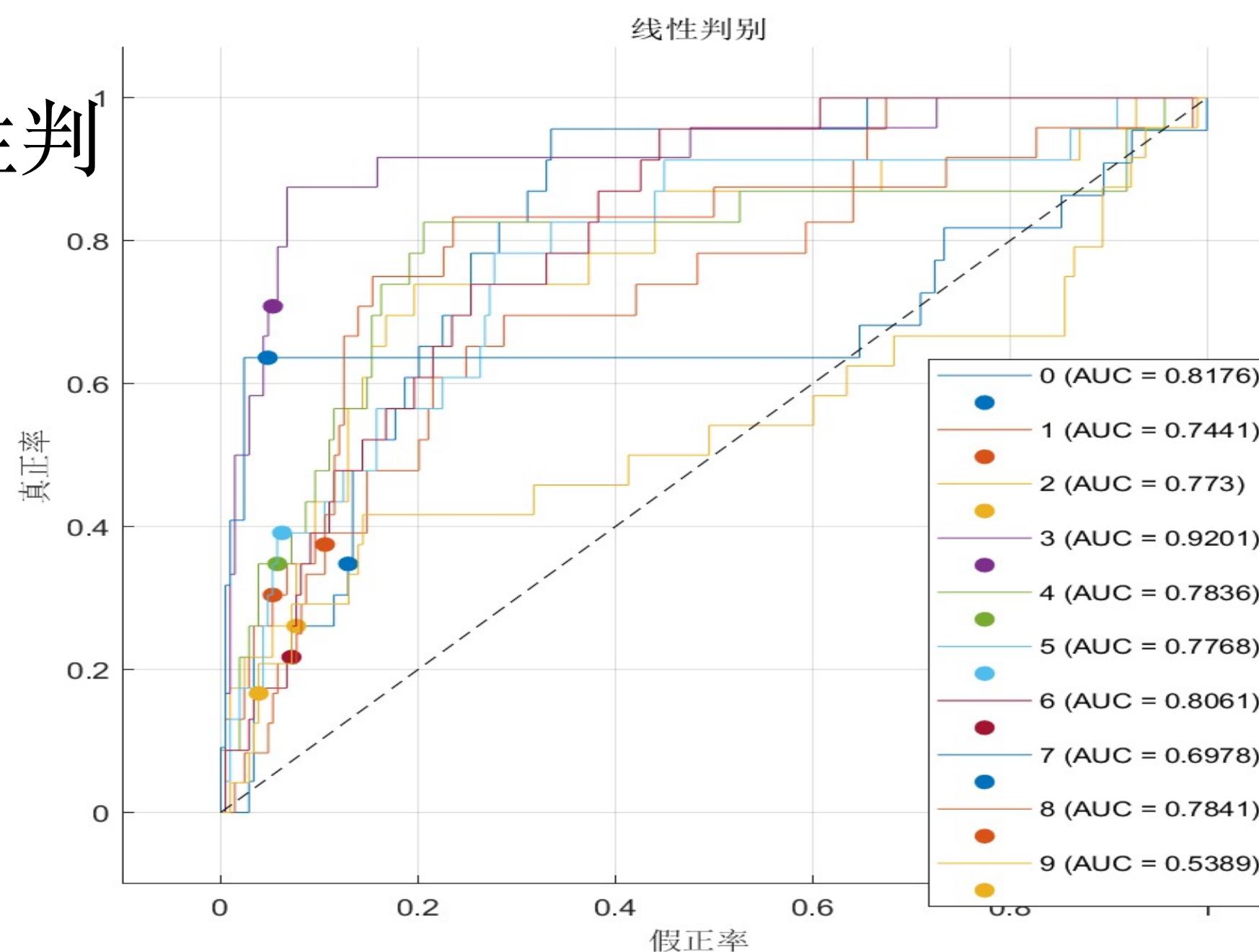
- 数据集分析
- 不同分类器效果分析-决策树
- 采用矩形窗



交融实验分析

基于MATLAB分析数据

- 数据集分析
- 不同分类器效果分析-线性判别分类器
- 采用矩形窗



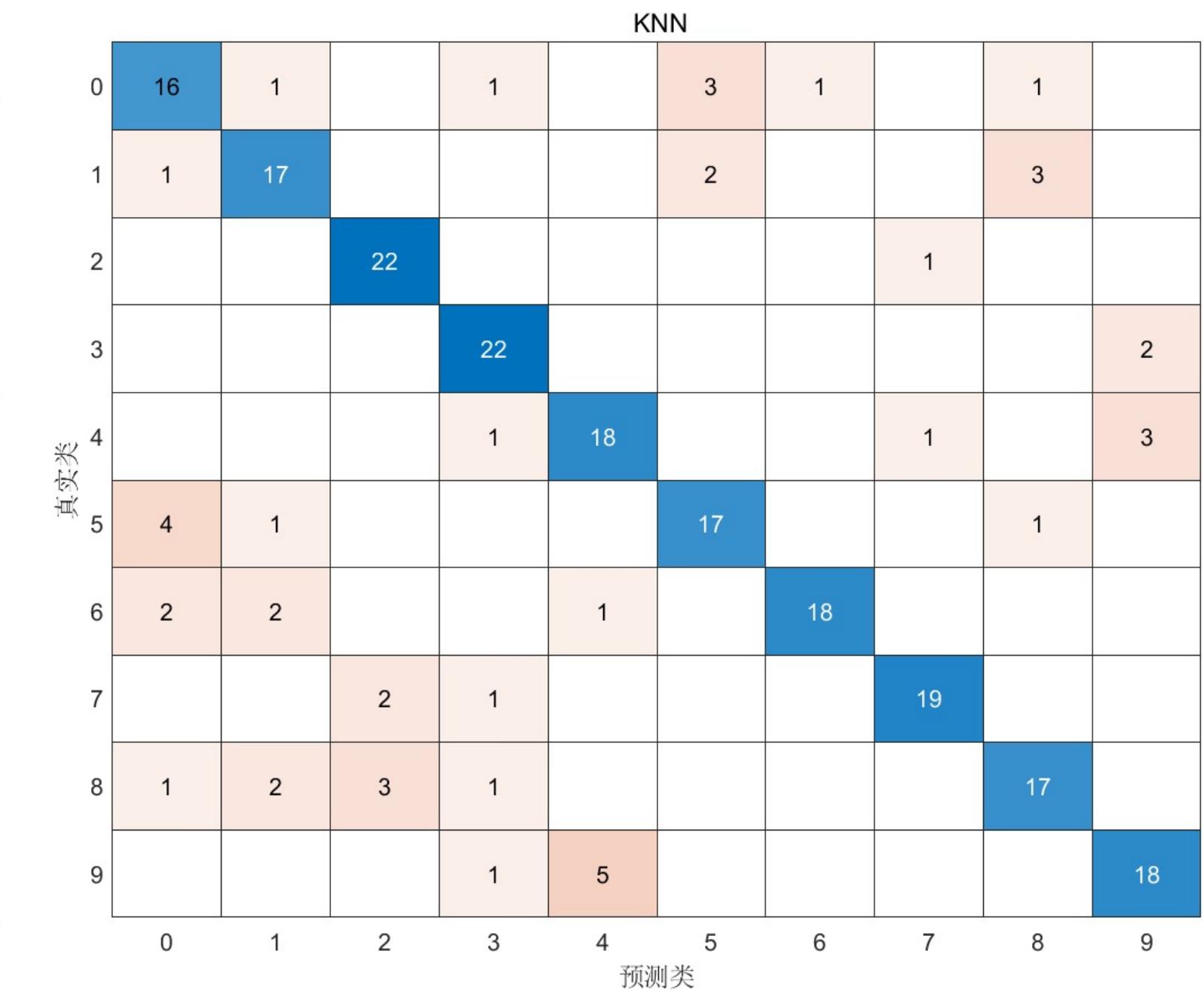
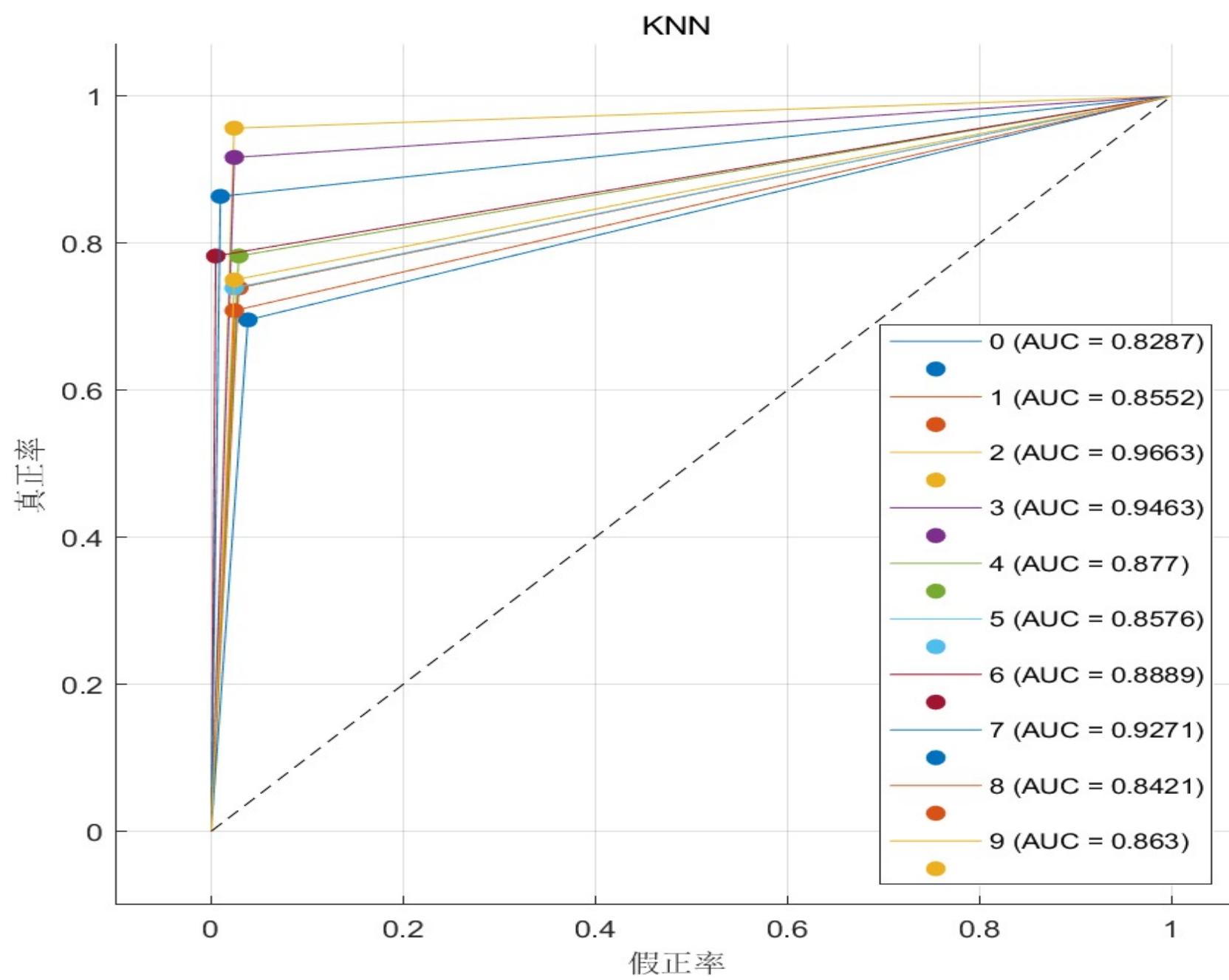
线性判别

真实类	预测类									
	0	1	2	3	4	5	6	7	8	9
0	8	5					6	3		1
1	3	7		1			4	3		5
2	1		6						13	3
3	1			17	2				3	1
4	1		4	1	8	2	3			4
5	8	2	1			9	3			
6	10	4	1			1	5		2	
7				5	2			14		1
8	3		6	3			3		9	
9			4	1	8		7			4

交融实验分析

基于MATLAB分析数据

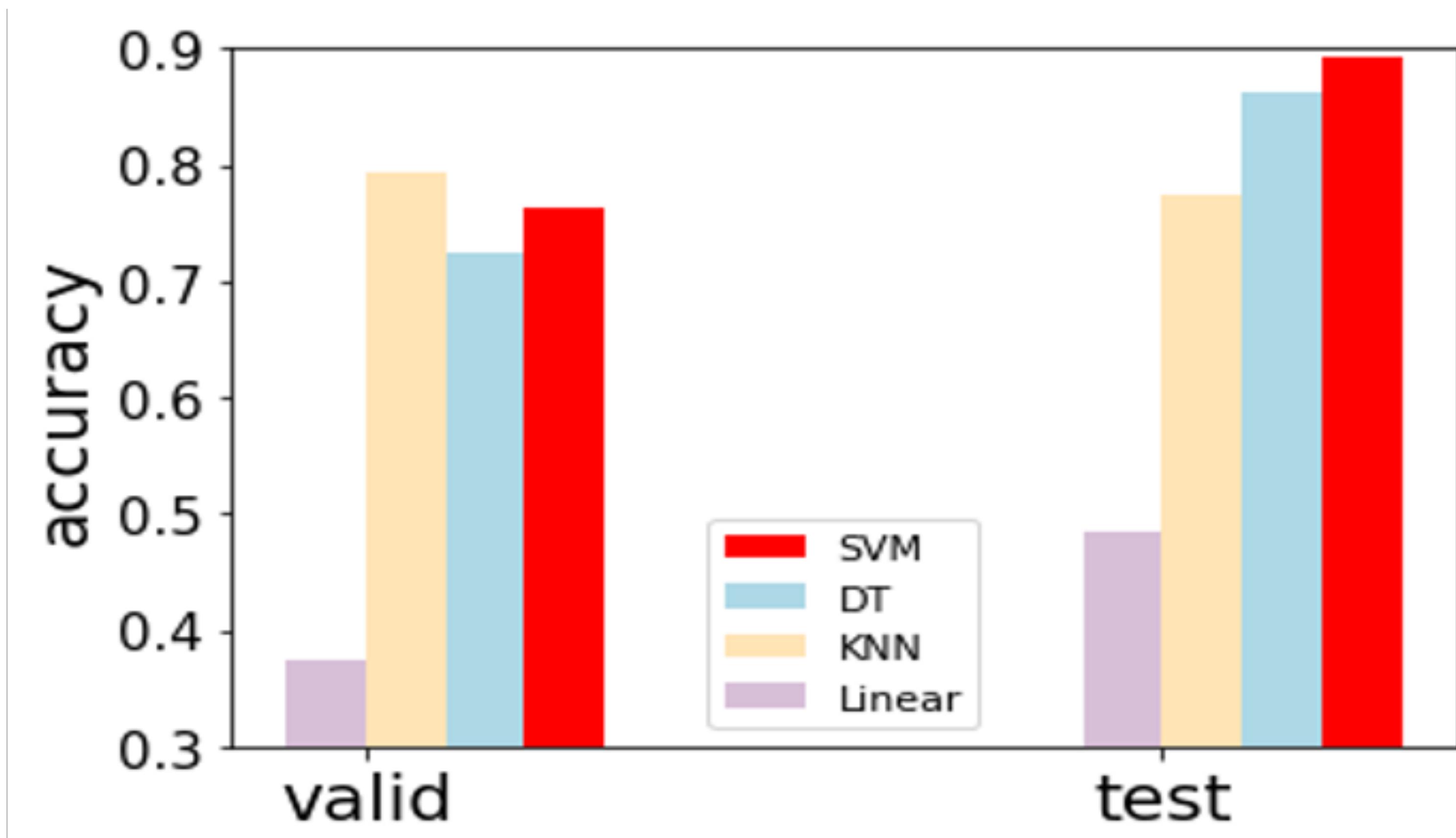
- 数据集分析
- 不同分类器效果分析-KNN
- 采用矩形窗



交融实验分析

基于MATLAB分析数据

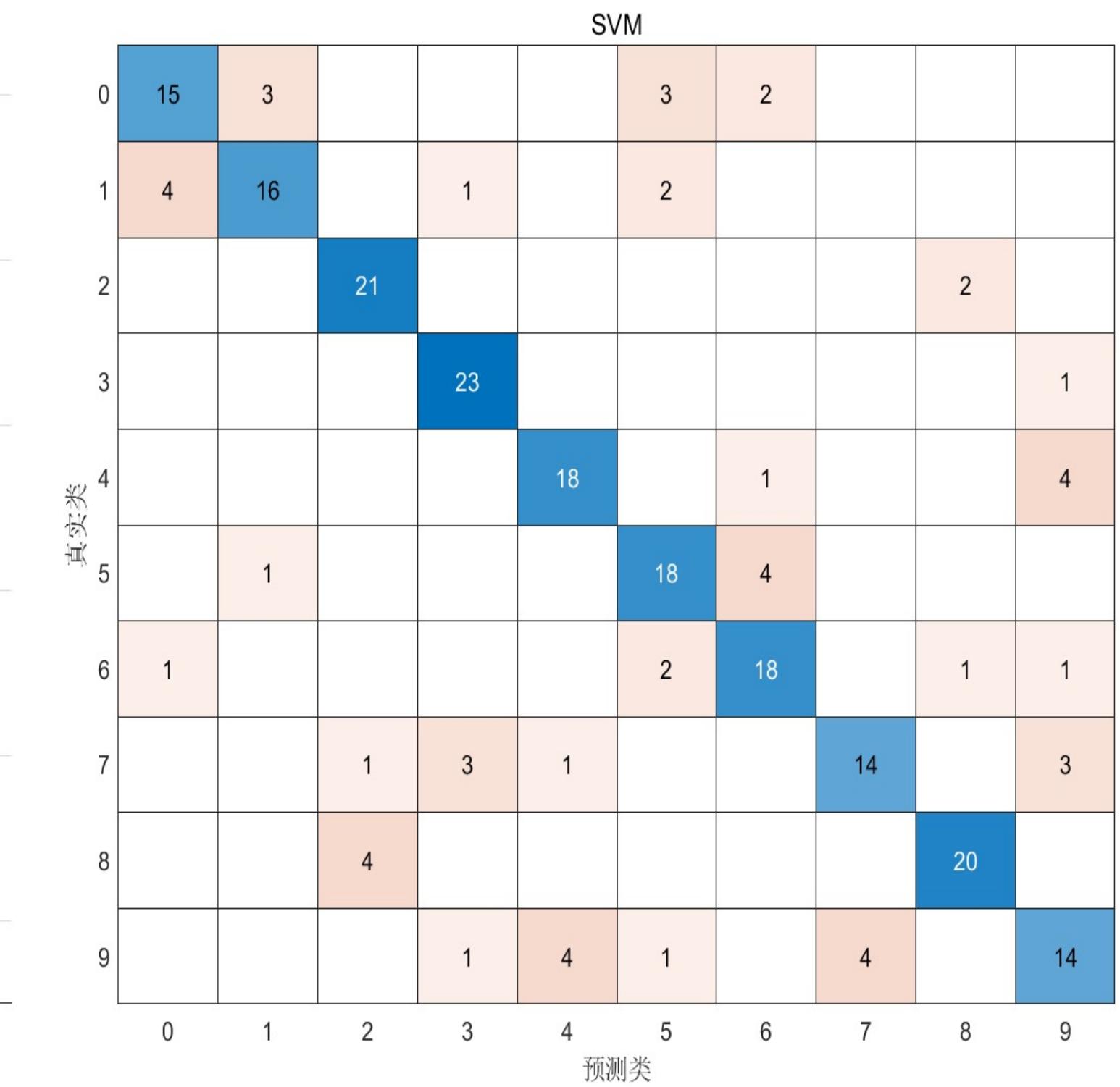
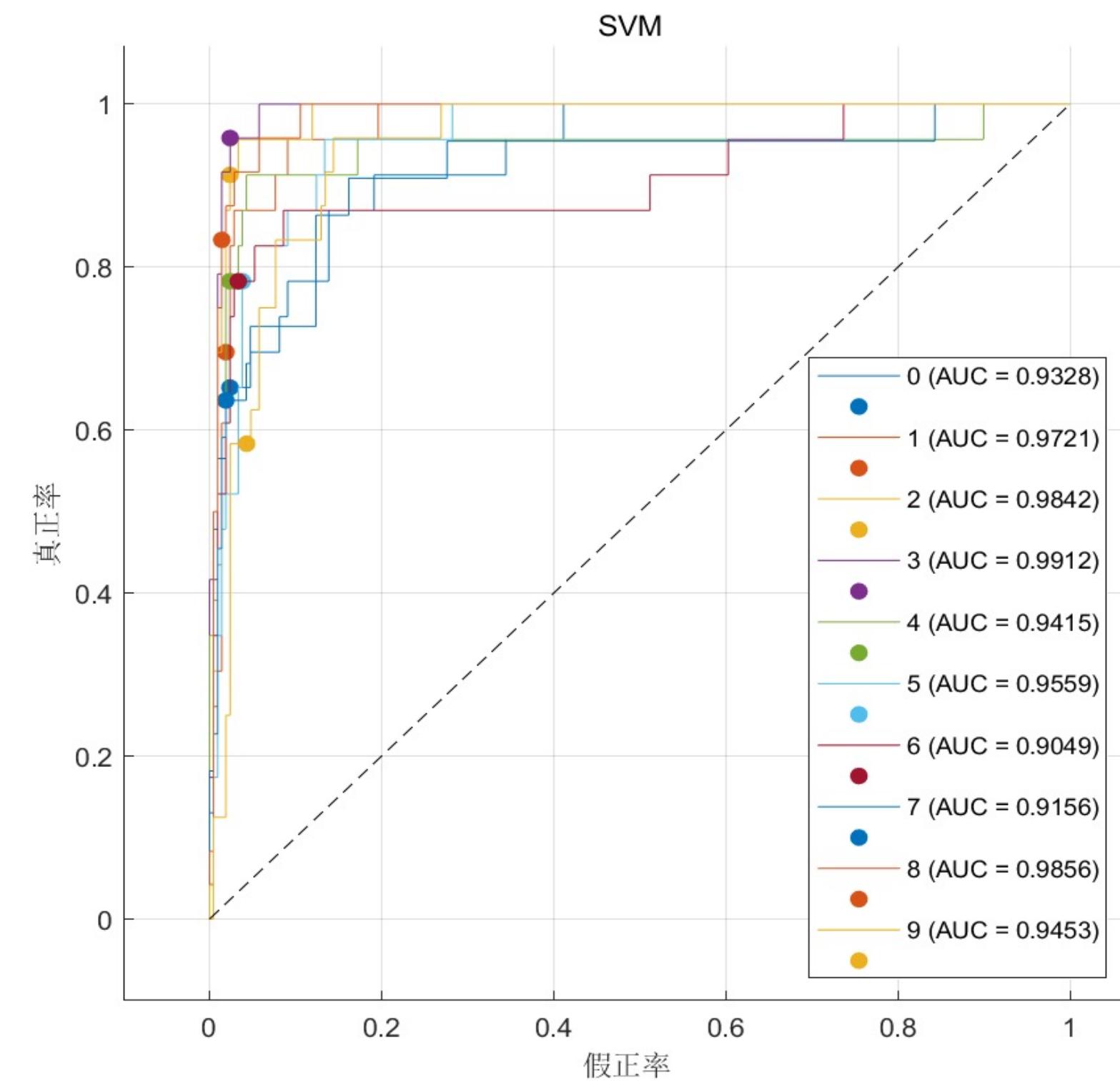
- 数据集分析
- 不同分类器效果分析-综合分析
- 采用矩形窗



交融实验分析

基于MATLAB分析数据

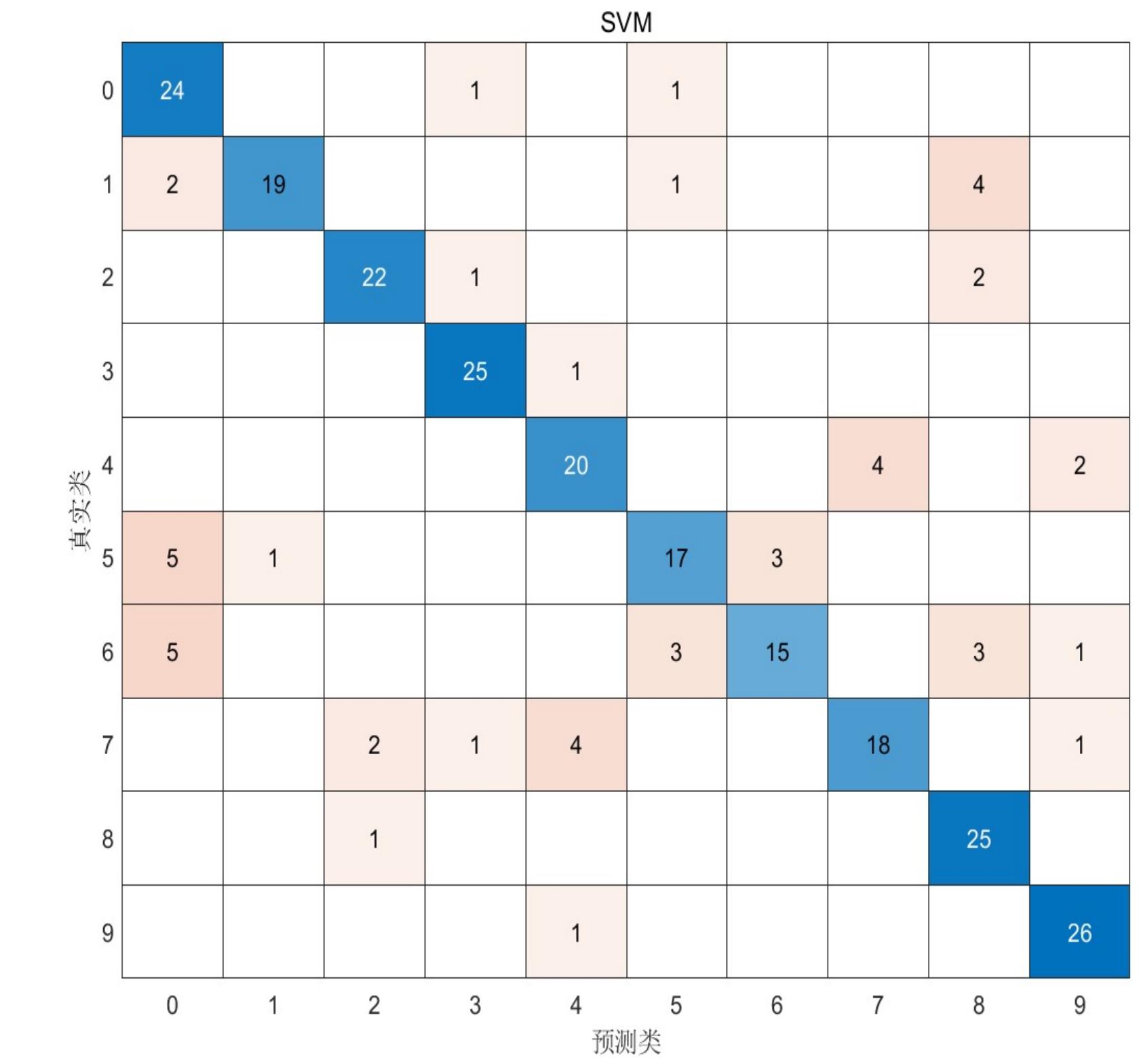
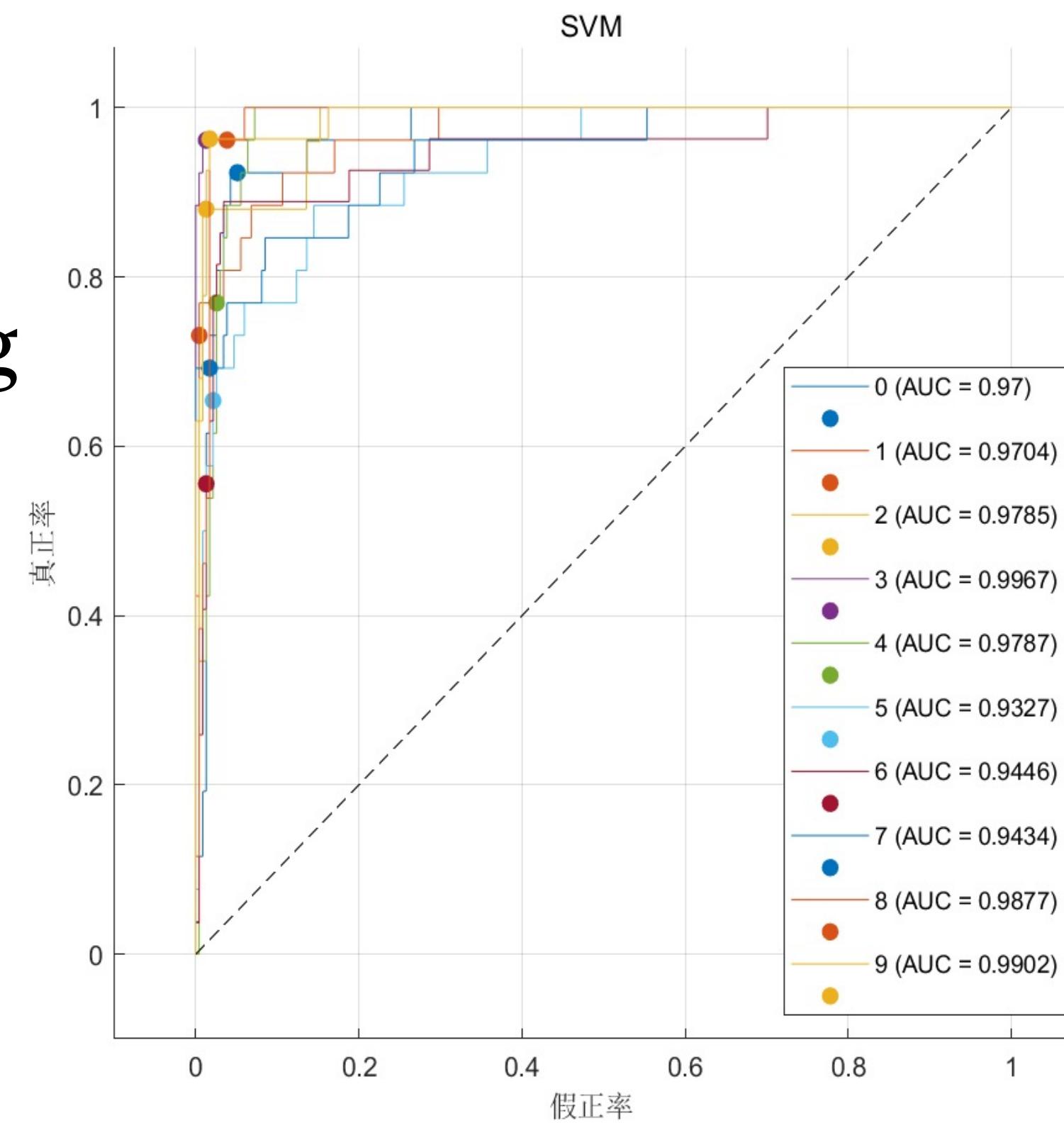
- 数据集分析
- 选取SVM作为分类器
- 不同窗函数分析- Boxcar



交融实验分析

基于MATLAB分析数据

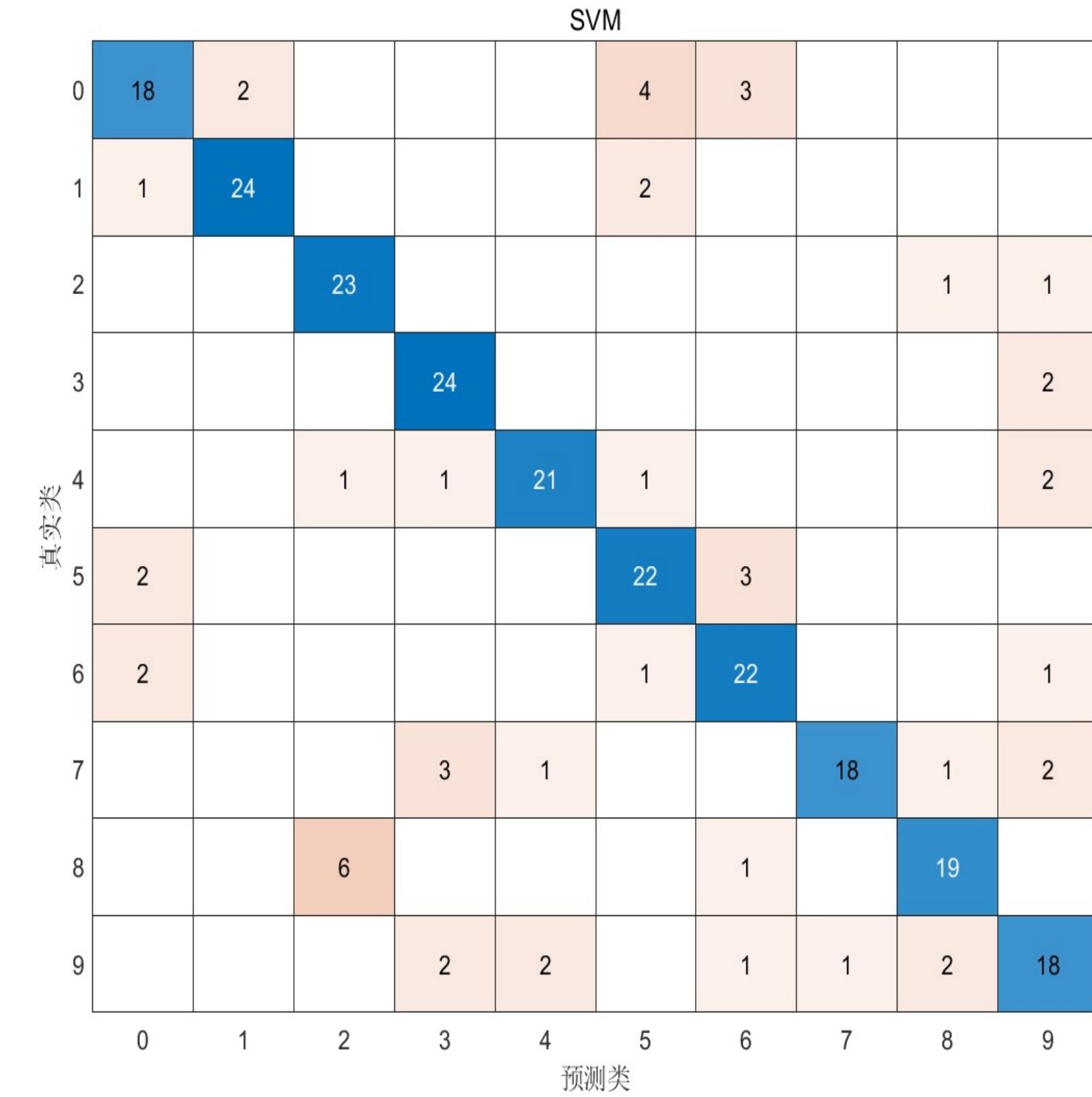
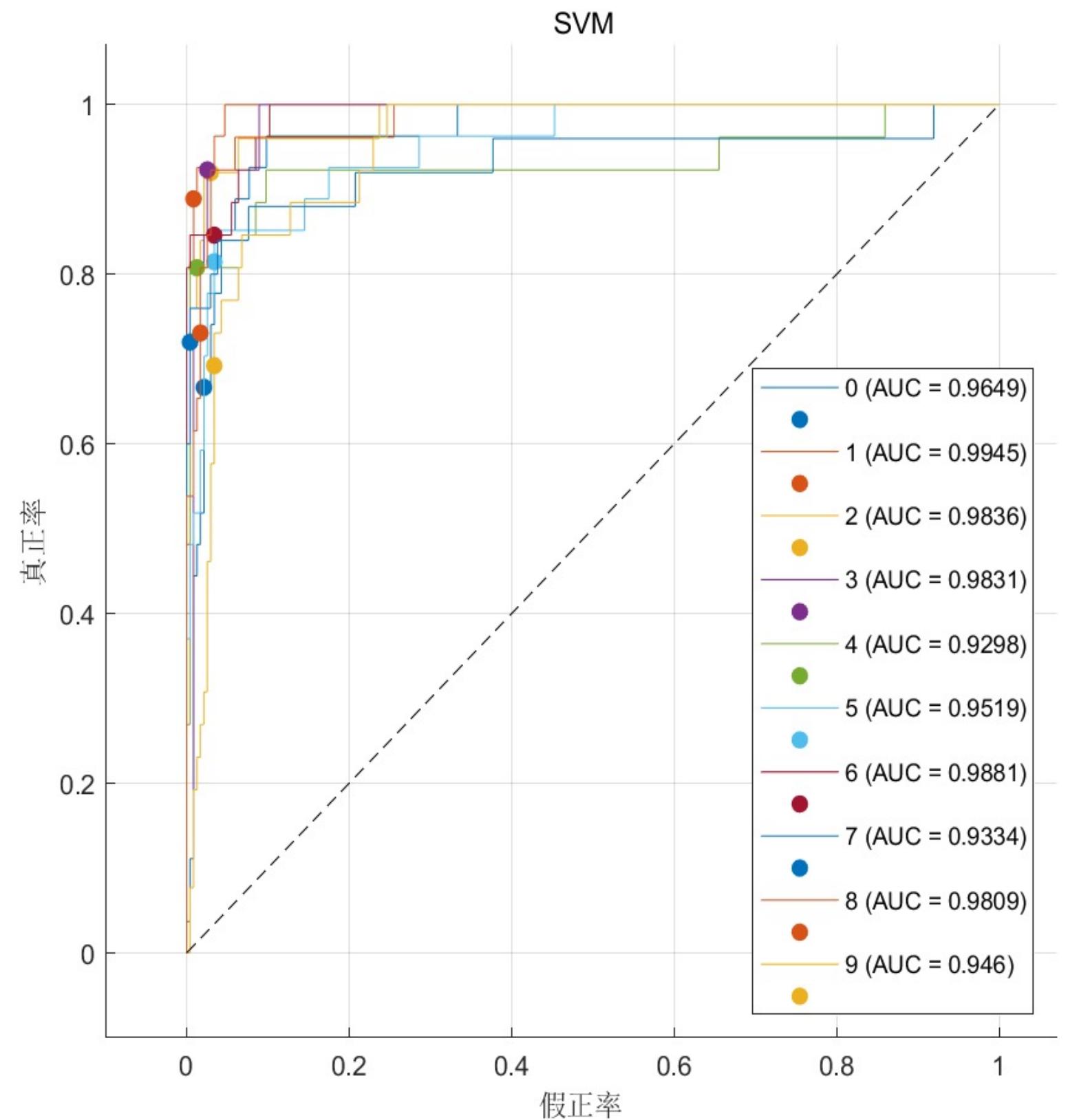
- 数据集分析
- 选取SVM作为分类器
- 不同窗函数分析- Hanning



交融实验分析

基于MATLAB分析数据

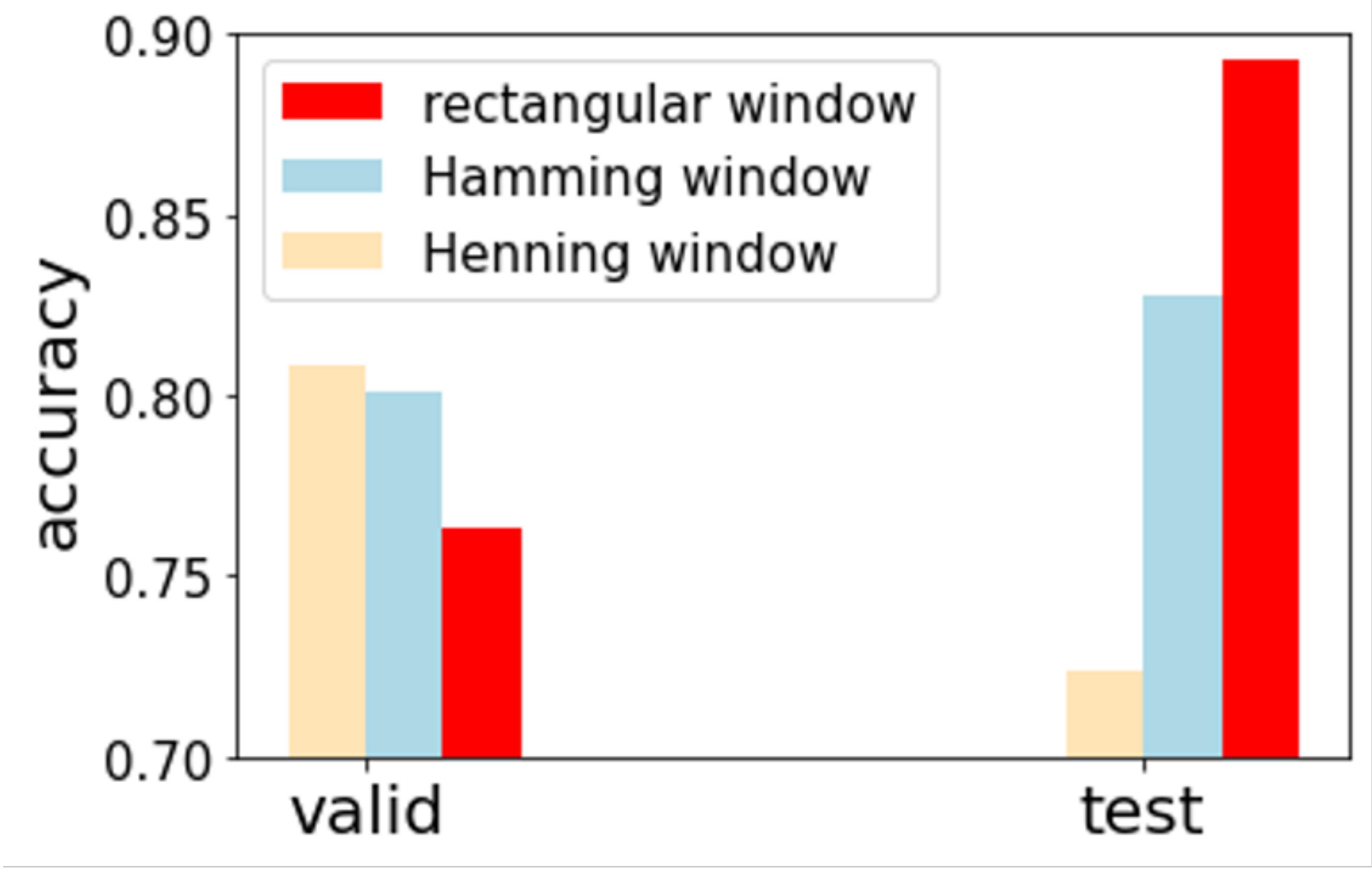
- 数据集分析
- 选取SVM作为分类器
- 不同窗函数分析- Hamming



交融实验分析

基于MATLAB分析数据

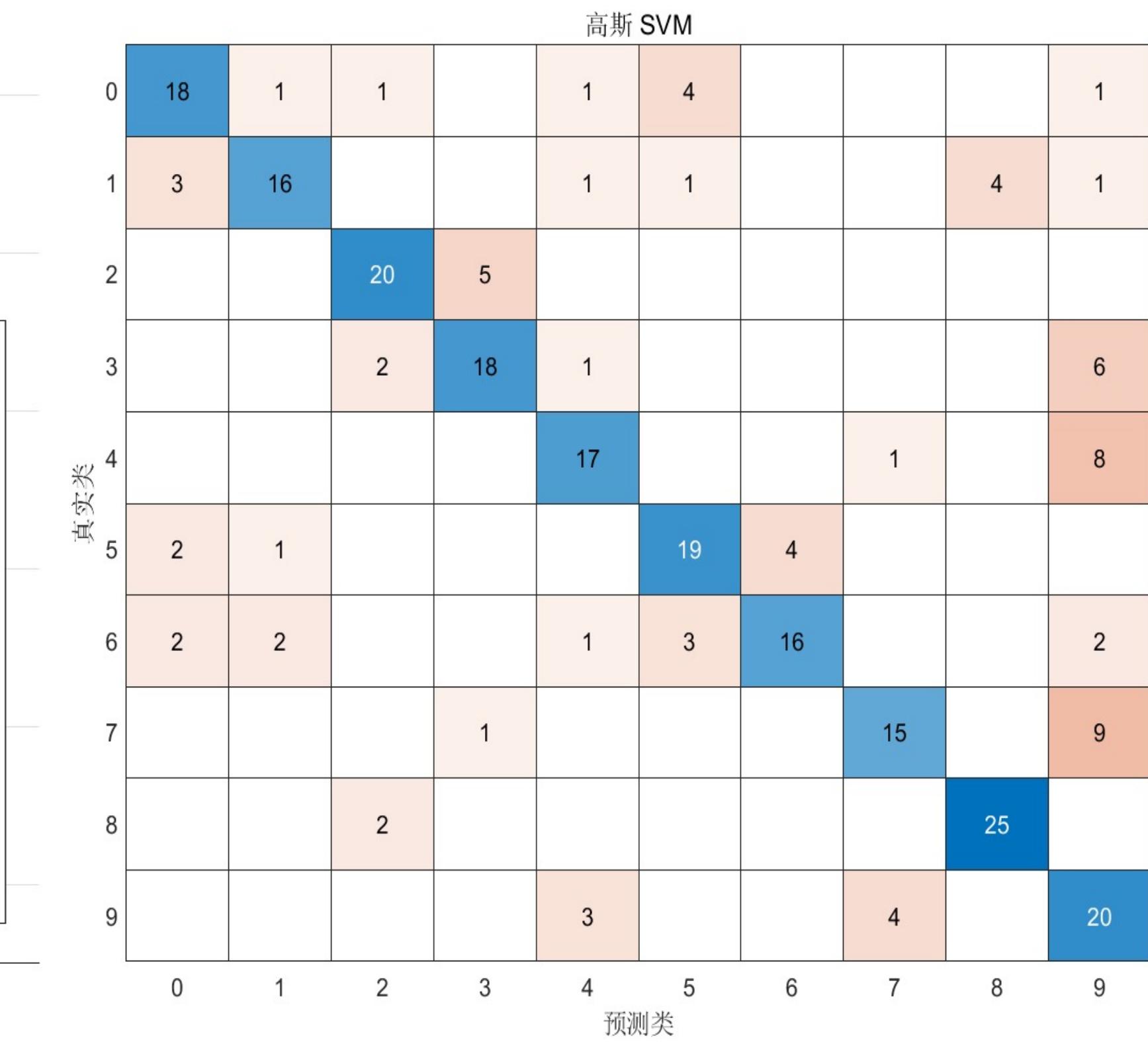
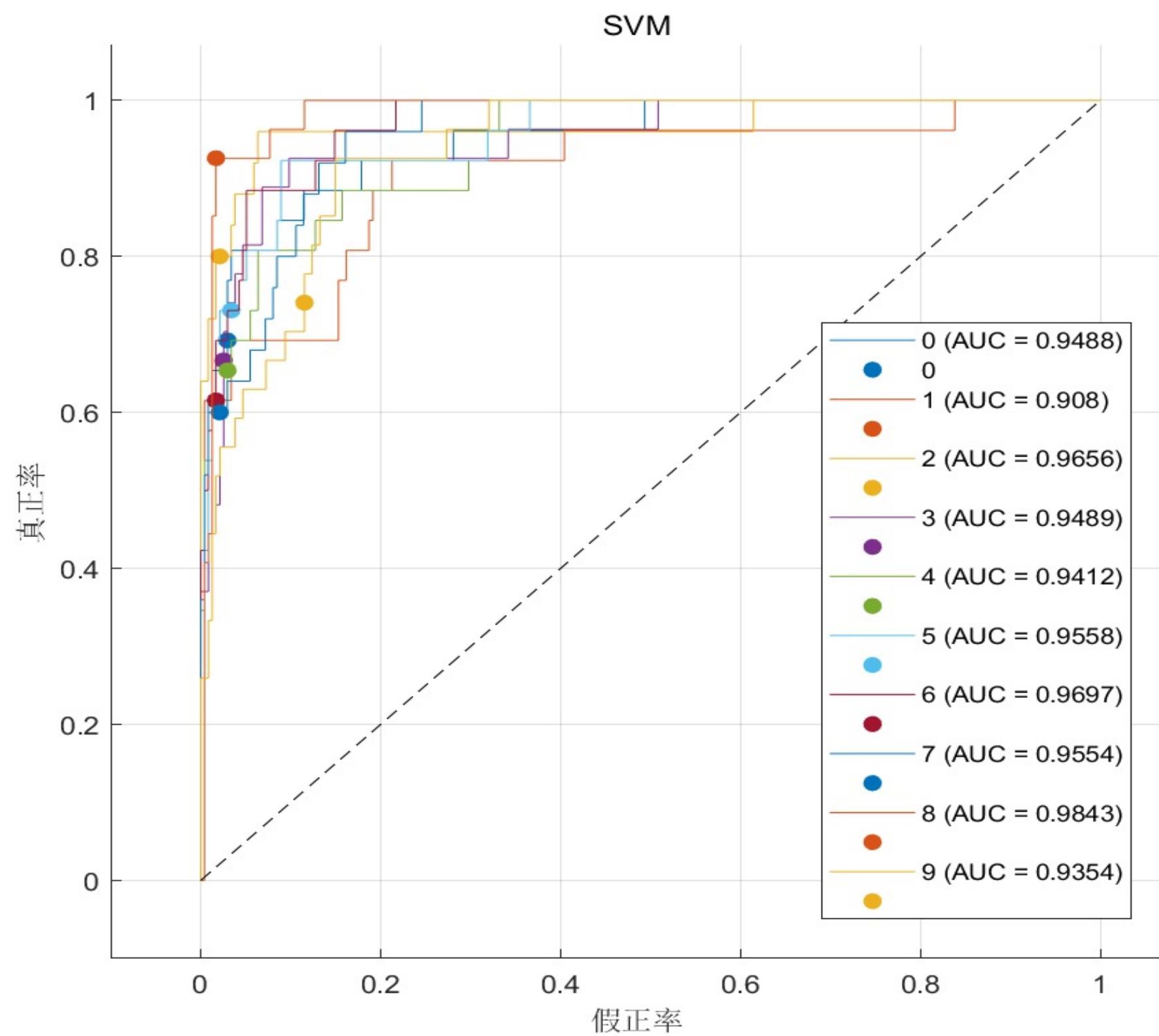
- 数据集分析
- 选取SVM作为分类器
- 不同窗函数分析- 综合分析



交融实验分析

基于MATLAB分析数据

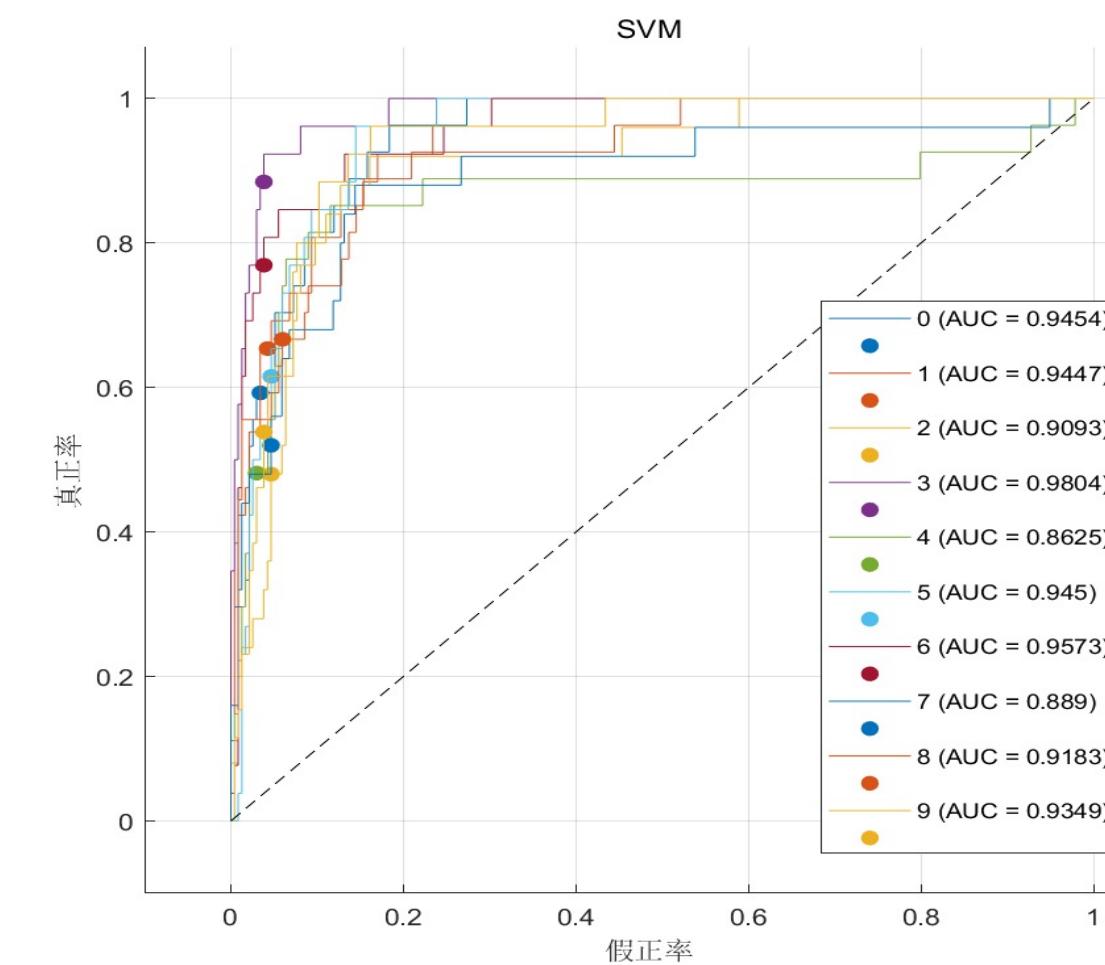
- 数据集分析
- 选取SVM作为分类器
- 选取矩形窗
- 特征分析-缺少能量



交融实验分析

基于MATLAB分析数据

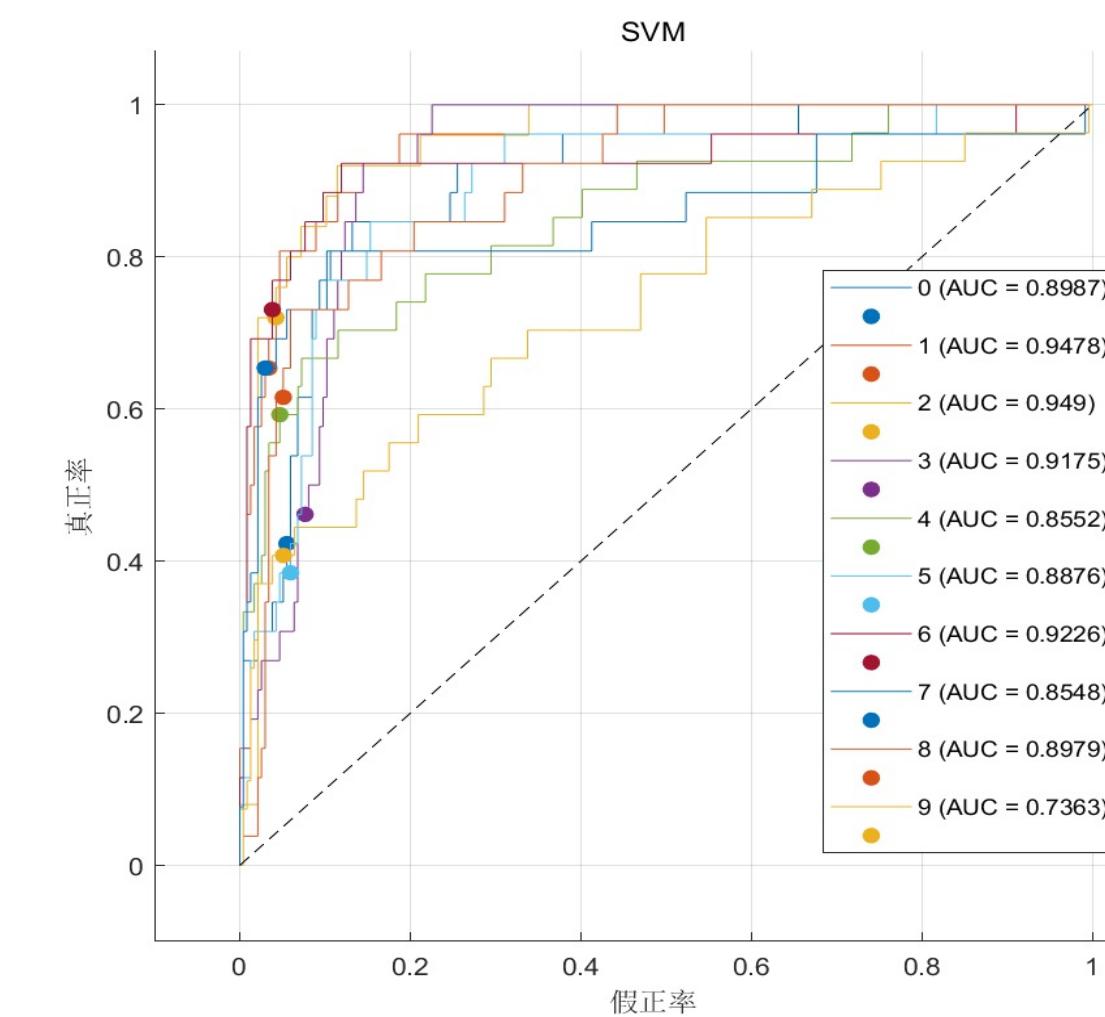
- 数据集分析
- 选取SVM作为分类器
- 选取矩形窗
- 特征分析-缺少过零率与缺少基音频率
- 综合对比, 基音频率影响最大



SVM

该图是一个混淆矩阵，展示了SVM模型在10个类别上的分类结果。矩阵的行代表“真实类”，列代表“预测类”。矩阵中的数字表示将某个真实类别的样本误认为另一个类别的数量。

真实类	0	1	2	3	4	5	6	7	8	9
0	16	6				4	1			
1	4	17		1		3	1			
2			12	2					7	4
3				23	1			1	1	
4			1	1	13		2	9		1
5	3	3				16	3			1
6	1	1			2	20		2		
7			1	3	4		1	13		3
8			8				1		18	
9			1	2	2		1	4	14	



SVM

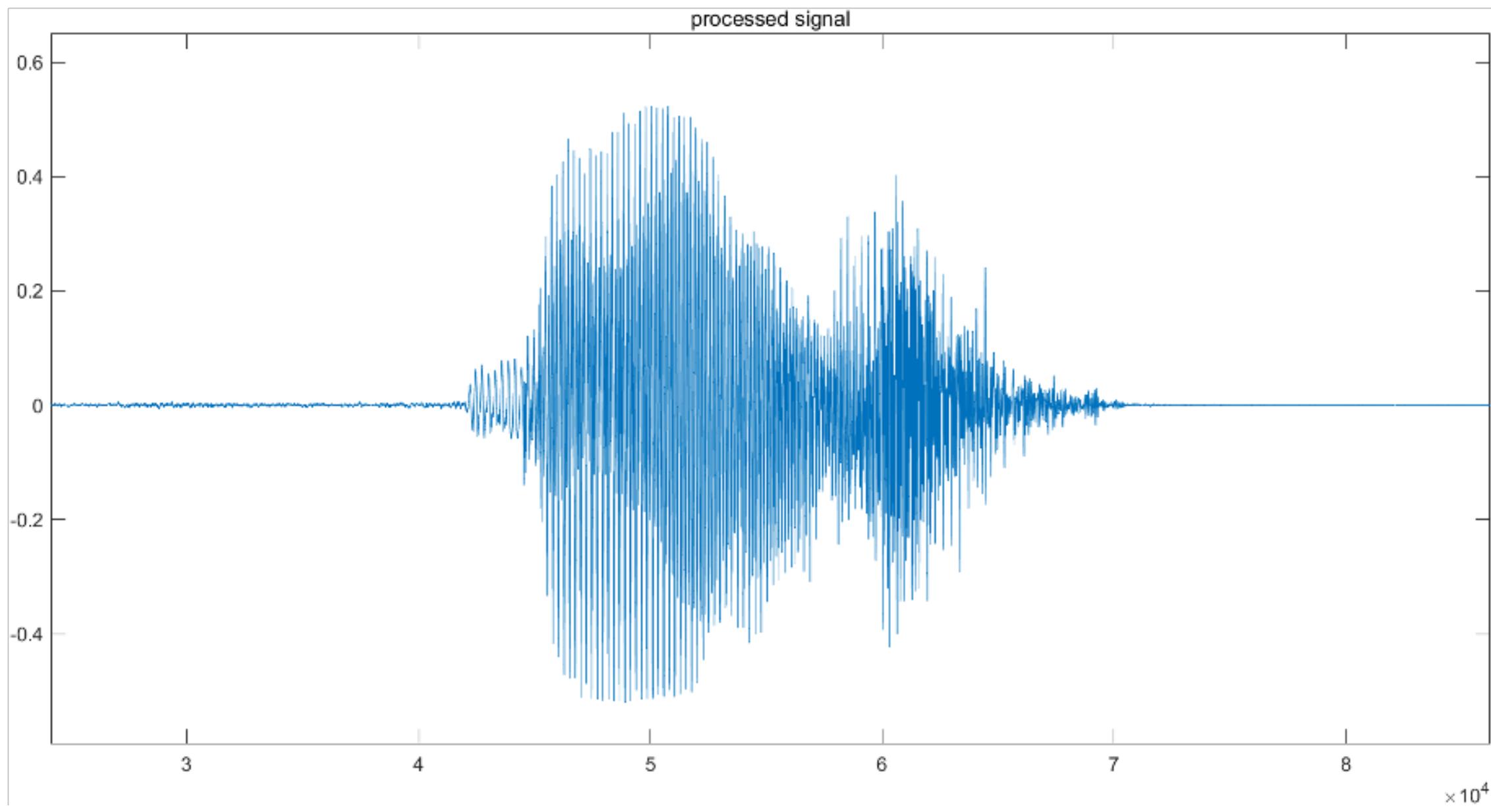
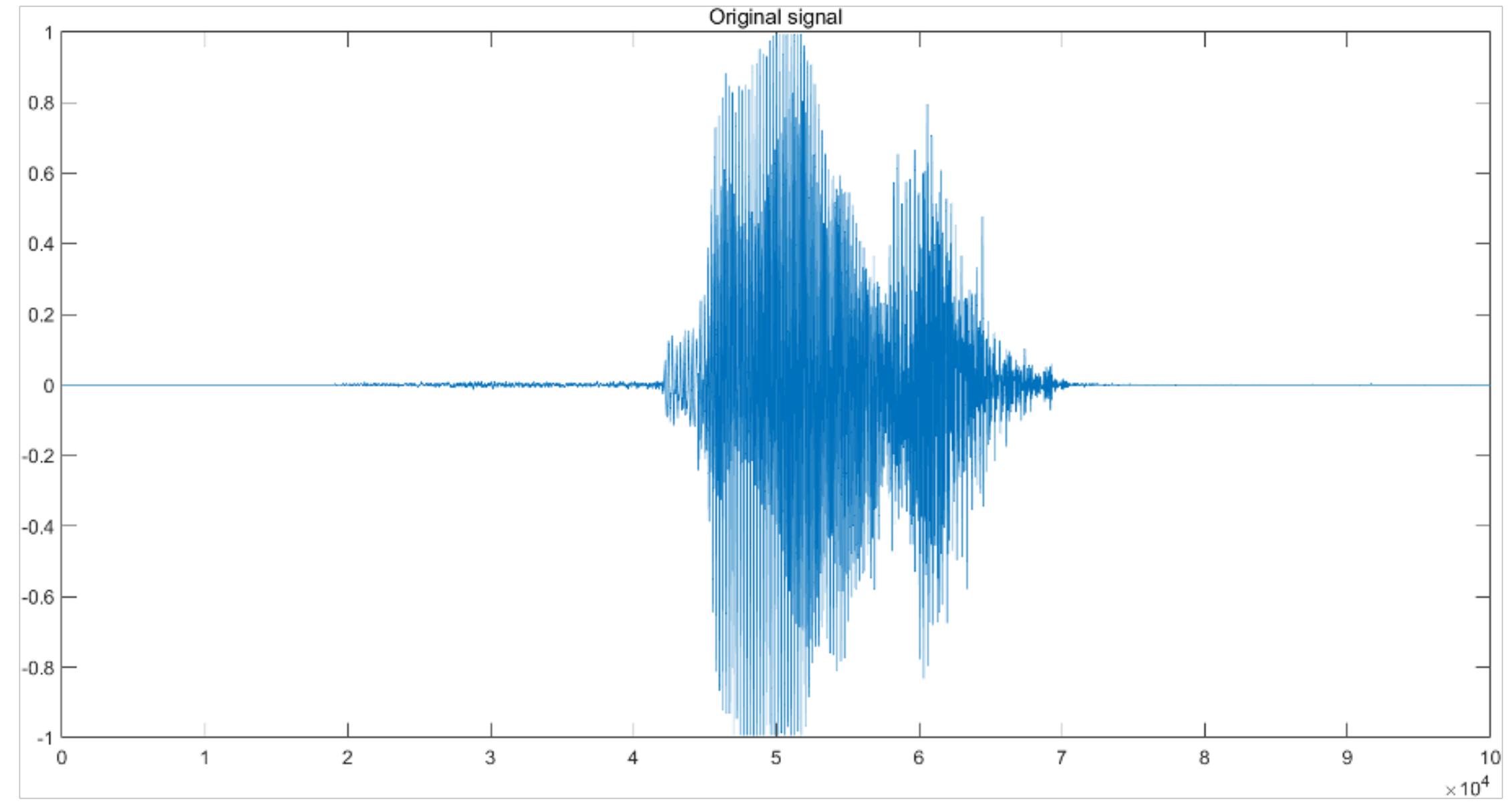
该图是一个混淆矩阵，展示了SVM模型在10个类别上的分类结果。矩阵的行代表“真实类”，列代表“预测类”。矩阵中的数字表示将某个真实类别的样本误认为另一个类别的数量。

真实类	0	1	2	3	4	5	6	7	8	9
0	11	3		1	2	5	4			
1	3	17		2	1	1		2		
2			18	1					6	
3	1	1		12	2	3		1	3	3
4	1		2	3	16	1	1		1	2
5	5	2	1	3	1	10	3			1
6	1			3	1	19	1			1
7	1	2		1	1		17		4	
8			6	3				16	1	
9	1		1	5	1	2	1	3	2	11

频域特征提取与分析

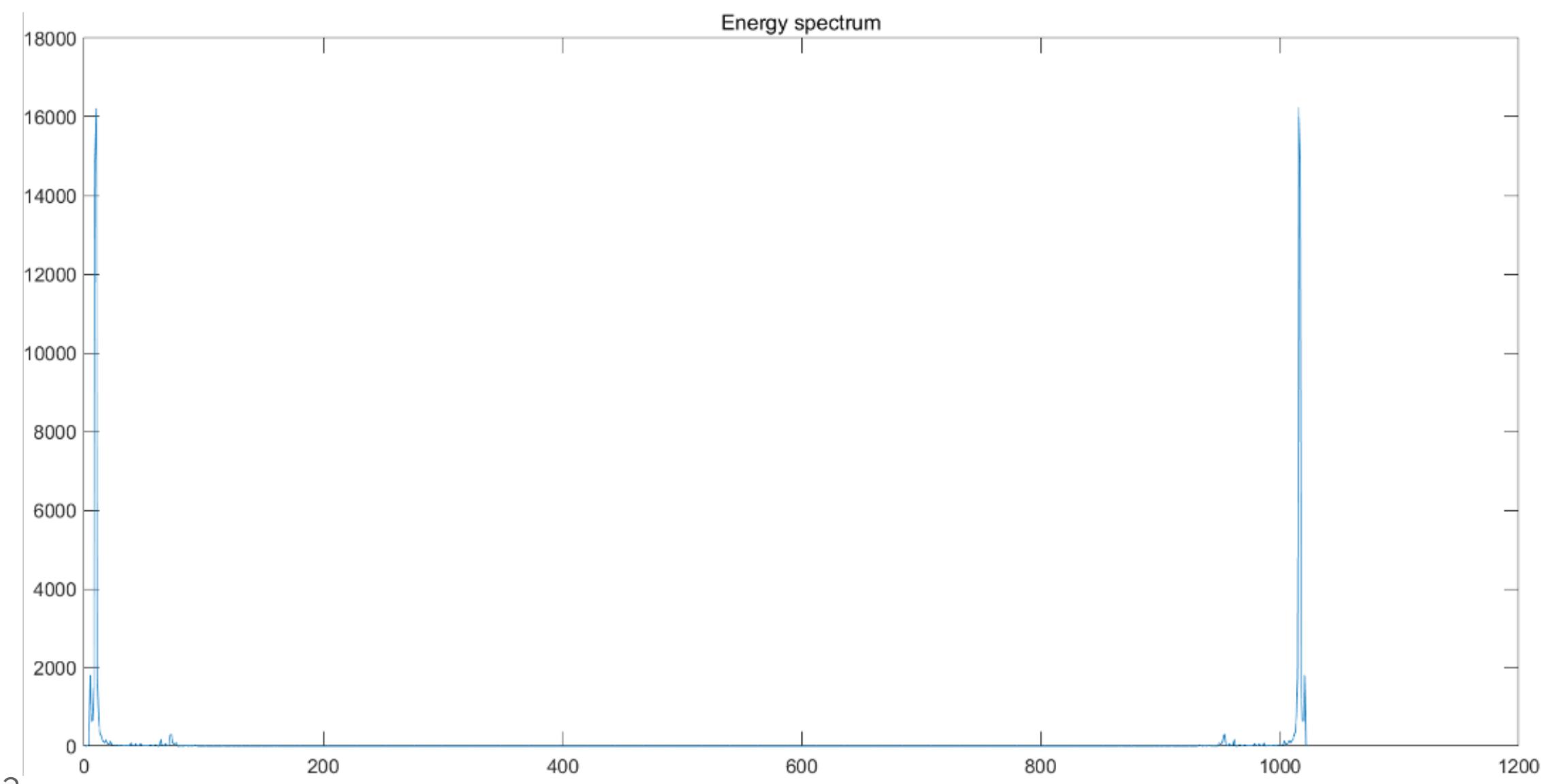
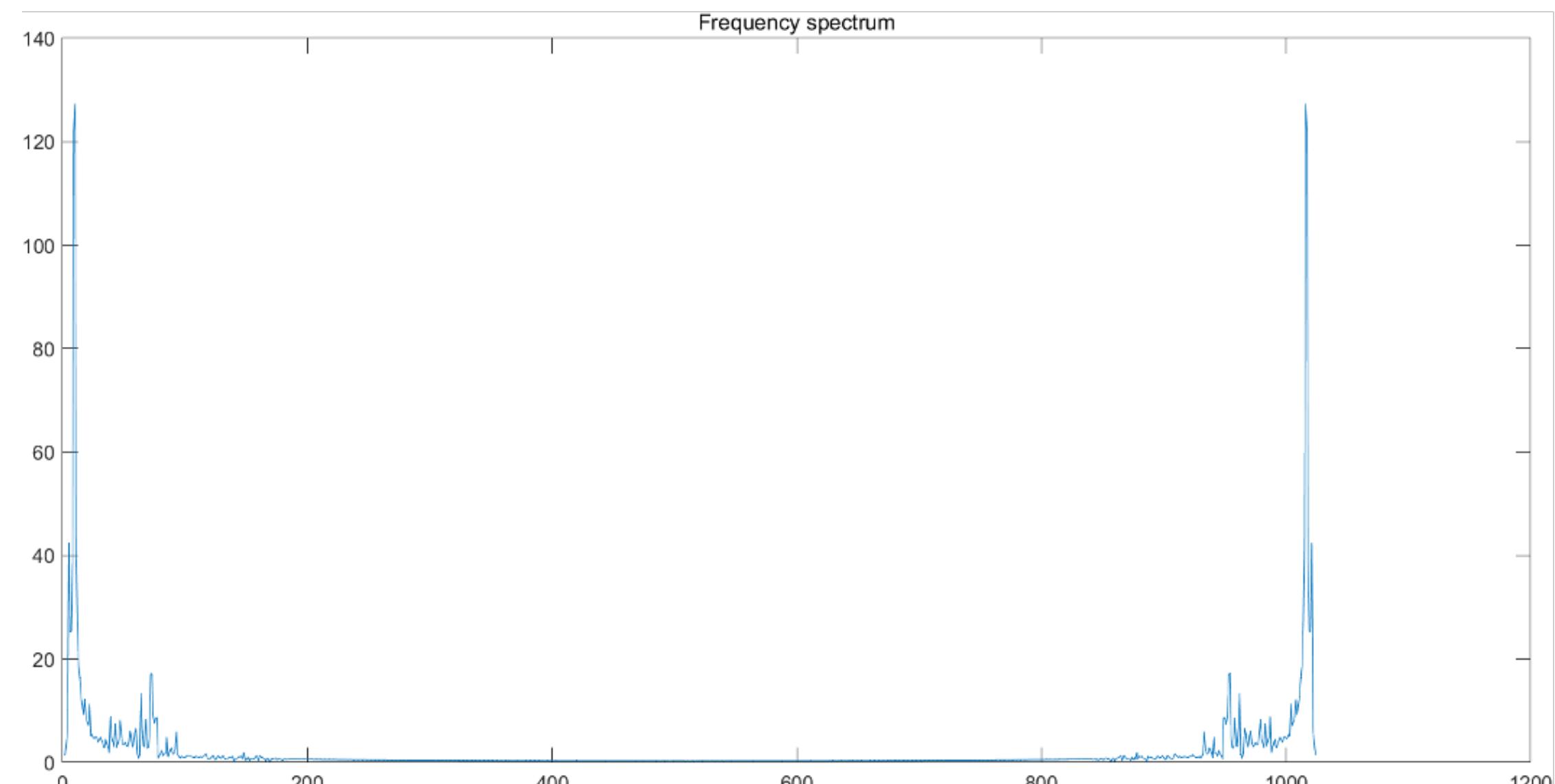
语音信号预处理

- 预加重: 使用预加重函数, 对高频分量加重处理
- 分帧加窗



频域特征提取

- 快速傅立叶变换, 通过蝶形变化实现
快速傅立叶变换得到频谱信号
- 谱线能量, 将FFT后频谱平方以用于
求解每一帧谱线的能量

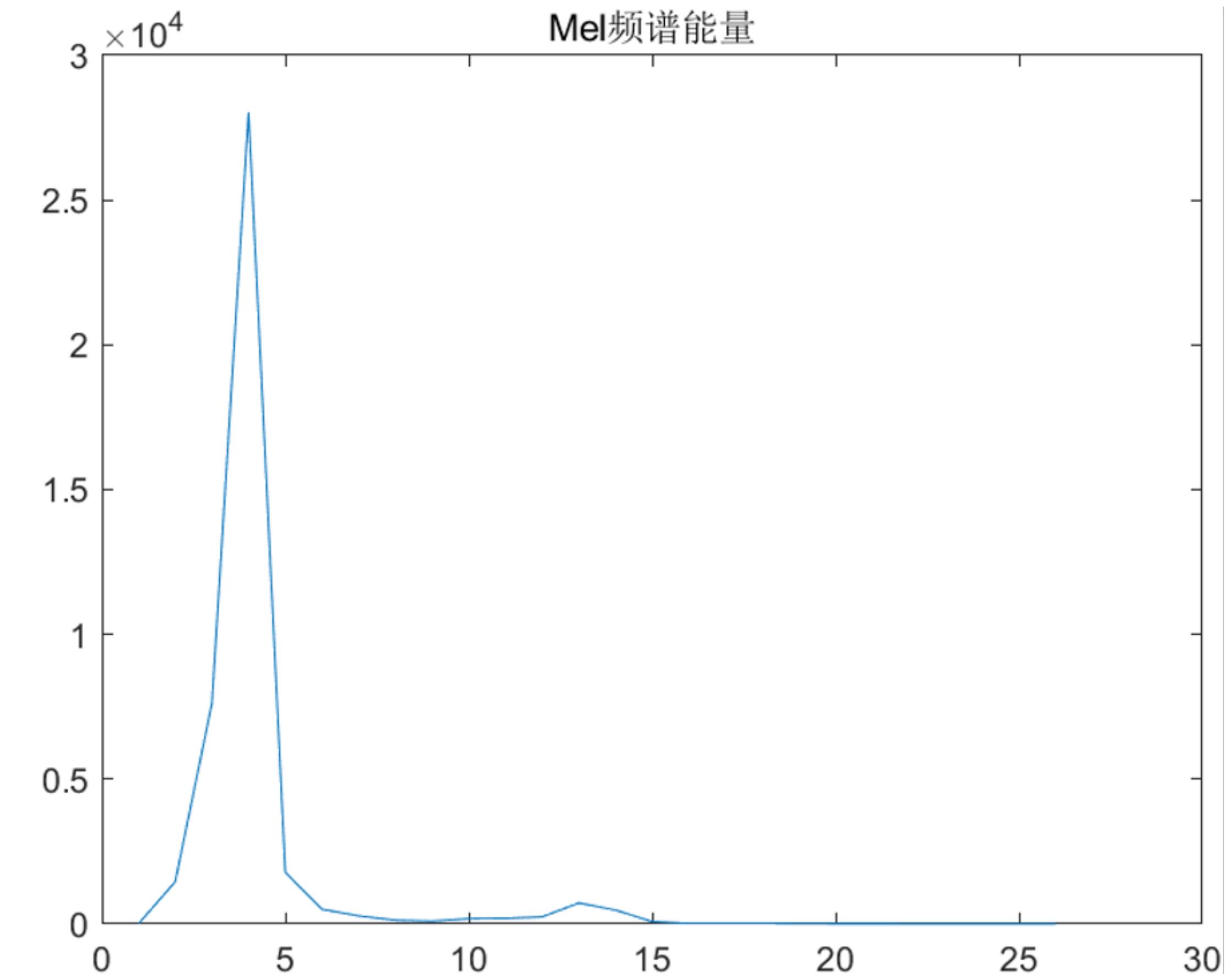


频域特征提取

- Mel滤波器组设计
- 选取Mel范围和滤波器个数设计

$$\bullet \quad H(m, k) = \begin{cases} \frac{k - f(m-1)}{f(m) - f(m-1)} & f(m-1) \leq k \leq f(m) \\ \frac{f(m+1) - k}{f(m+1) - f(m)} & f(m) \leq k \leq f(m+1) \end{cases}$$

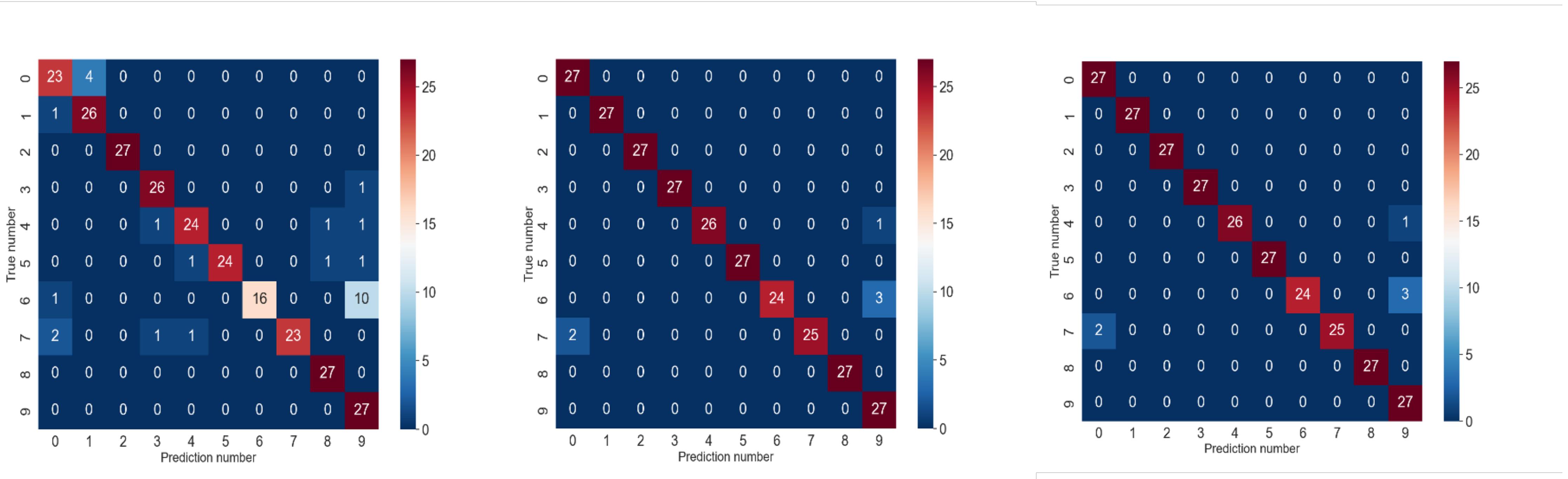
- 将谱线能量与Mel滤波器运算得到Mel能量
- 并由对数化的Mel能量得到MFCC



交融实验分析

基于MATLAB分析数据

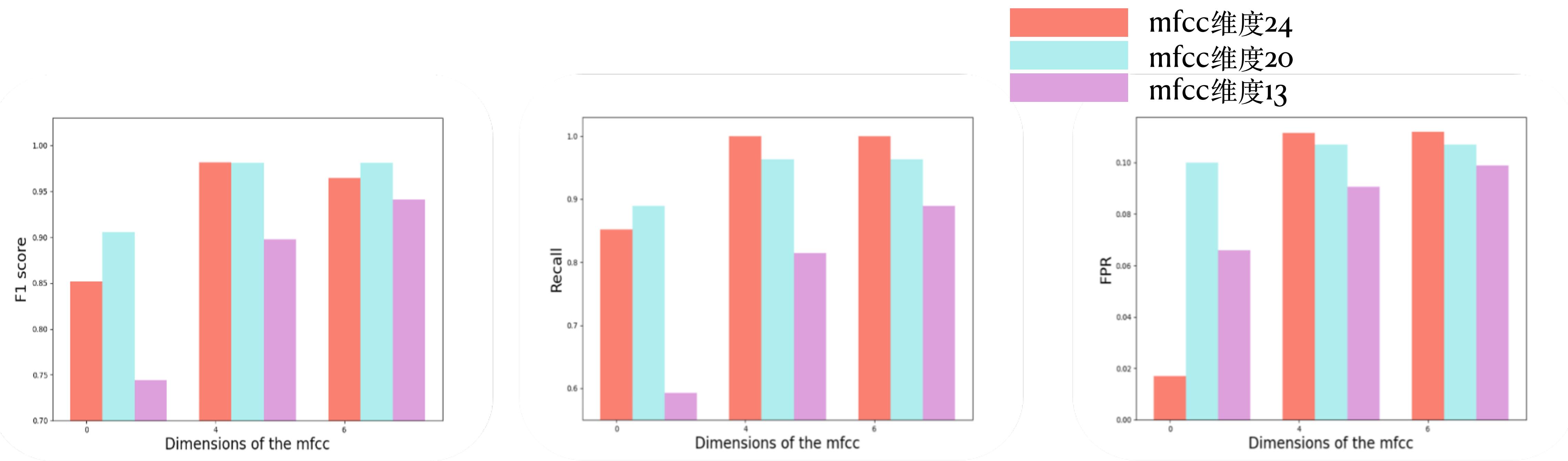
均选择矩形窗, 24Mel滤波器个数下混淆矩阵



交融实验分析

基于MATLAB分析数据

均选择boxcar窗, 24Mel滤波器个数下三分类评价指标

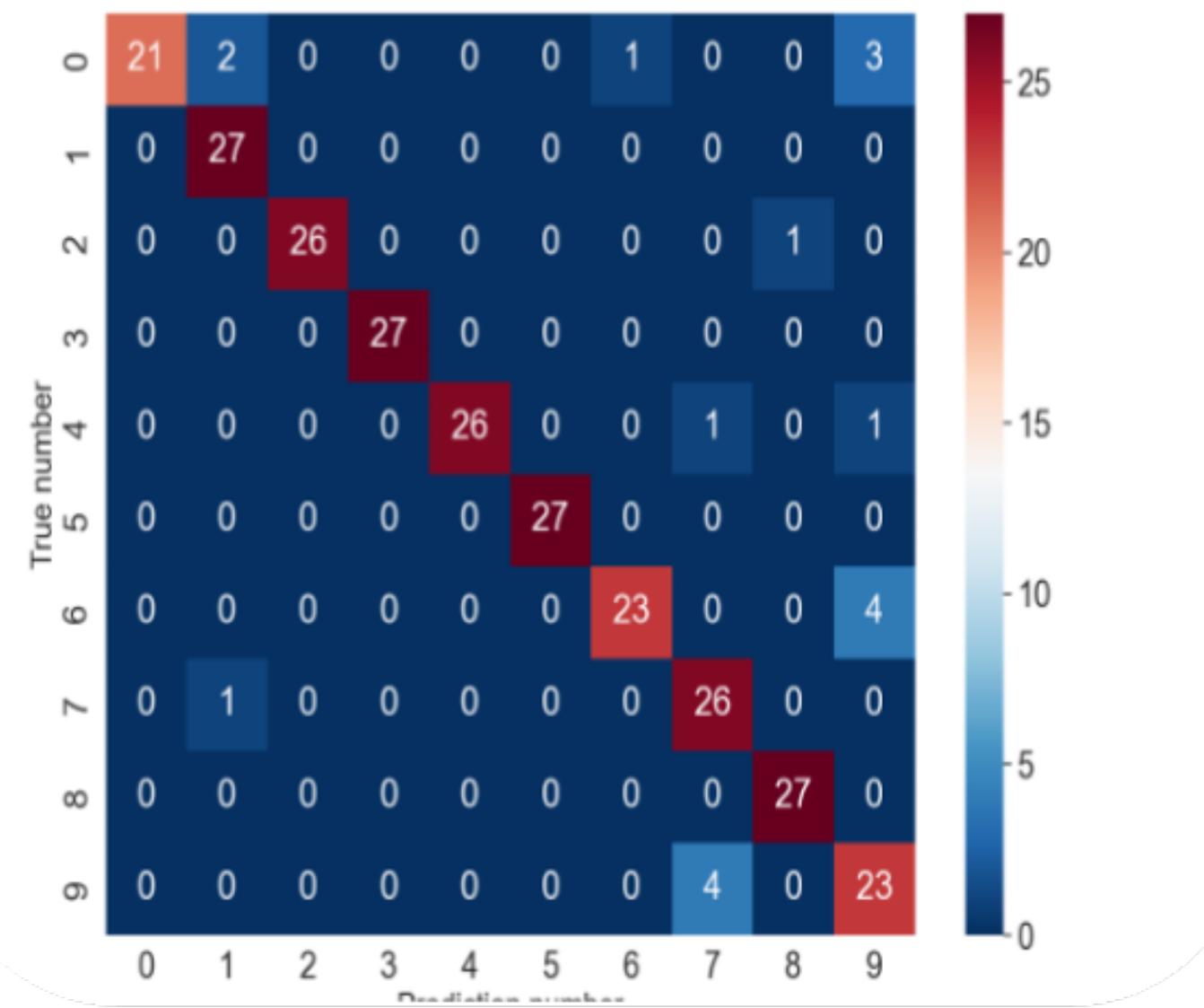


随着mfcc特征向量维度的增加，分类的结果准确性得到改善随后趋于稳定。结果表明，适当增加mfcc的特征向量维度，有助于提高识别的准确率。

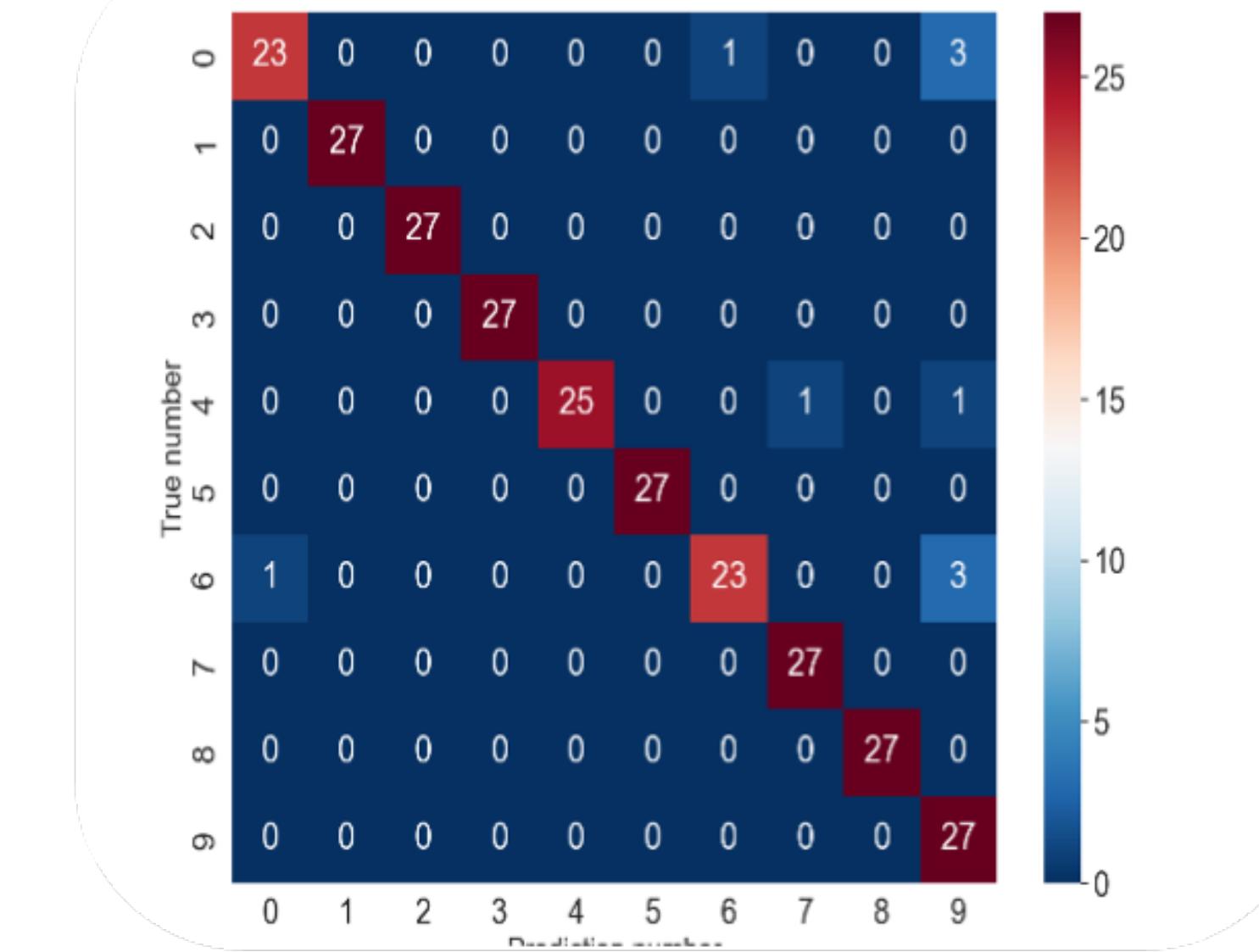
交融实验分析

基于MATLAB分析数据

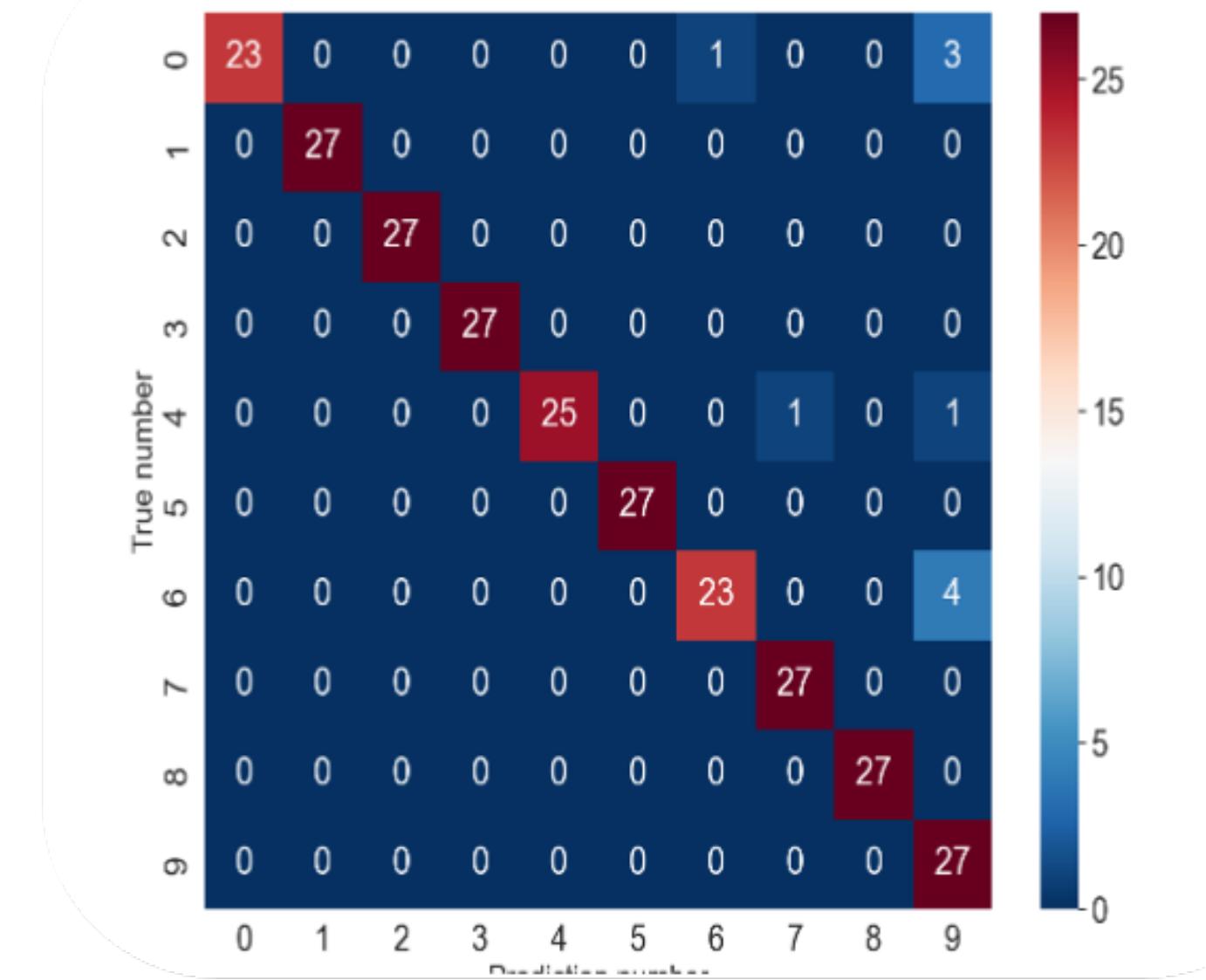
均选择mfcc维度20, 20Mel滤波器个数下不同窗函数混淆矩阵



boxcar窗



hamm窗



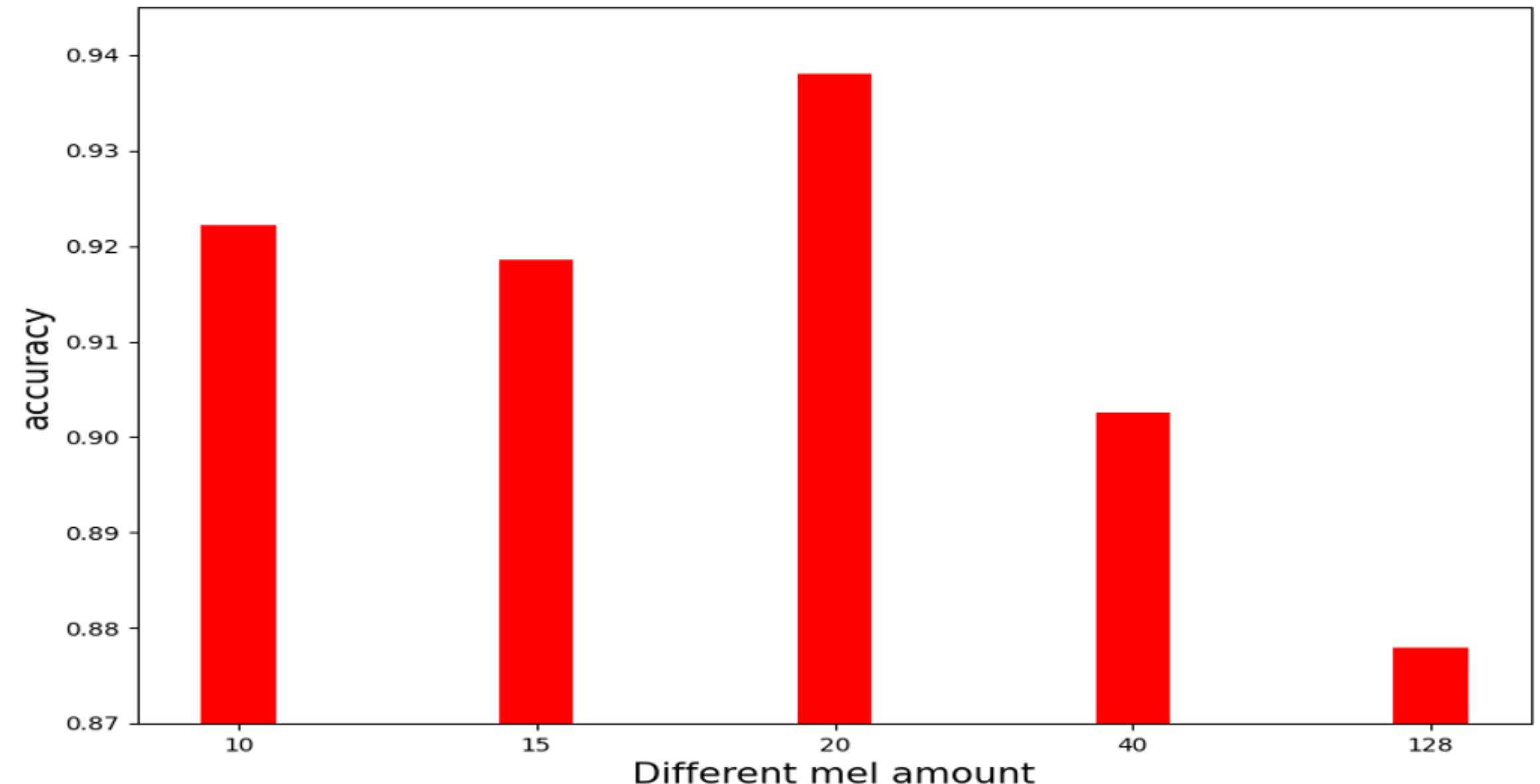
hanning窗

交融实验分析

基于MATLAB分析数据

均选择mfcc维度20, boxcar窗函数, 不同Mel滤波器个数

- Mel滤波器组数为128的分类结果在五者中最差
- 组数为20的结果最好
- 过分增加或减少Mel滤波器组数都不能提高准确率
- 组数为20即为本实验较为合适的参数



Thanks