

## 语音信号处理：第七章作业讲解



主讲人 彭震东



# 作业题目

本题将给定两种不同的麦克风阵列构型，以及在每种构型下，每个麦克风的录音数据。录音数据的长度约为 12s，在此期间声源的方位固定不变。请根据阵列接收到的语音数据，编写程序，计算声源的入射角 $\theta$ 。 $\theta$  定义为声源入射方向（可以看做连接平面直角坐标系原点和声源之间的线段）与 x 轴正向的夹角，取值范围限制在  $0^\circ \sim 360^\circ$  之间。

为了简化问题，以下两种场景均为远场场景，也就是说声源到达麦克风阵列中心的距离远远大于麦克风阵列本身的几何尺寸。所以，在远场假设下，声源定位任务不需要精确定位出声源的坐标，而是定位出声源的入射角 $\theta$ 即可。

编程语言不限，使用算法不限。请提交完整的代码，和两种场景下各自的定位结果 $\theta$ 。

- 6mic均匀线性阵列：135° 或者 225°
- 6mic均匀环形阵列：60°
- 以 SRP 为例：

将(7.4)式扩展成更一般的形式：

$$Y(\omega, \mathbf{k}_s) = \sum_{i=0}^{N-1} G_i(\omega) X_i(\omega) e^{j\omega \Delta_i} \quad (7.5)$$

计算SRP：

$$P(\mathbf{k}) = \int_{-\infty}^{\infty} |Y(\omega, \mathbf{k})|^2 d\omega \quad (7.6)$$

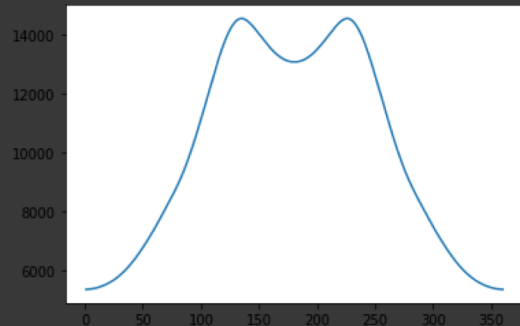
Maximizing SRP:

$$\hat{\mathbf{k}}_s = \underset{\mathbf{k}}{\operatorname{argmax}} P(\mathbf{k}) \quad (7.7)$$

# 声源定位

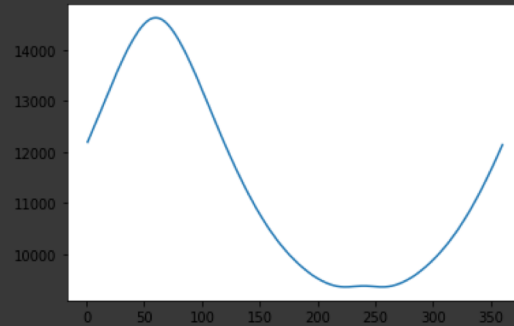
```
1 path = '作业/1_uniform_linear_array'
2 fs, X = read_waves(path)
3 N = X.shape[0]
4 X_fft = np.fft.rfft(X, axis=0)
5 omega = 2 * np.pi * np.tile(np.arange(N / 2 + 1), (M, 1)).T * fs / N
6
7 P = []
8 for i in tqdm(theta[:180]):
9     tau = np.arange(M) * d * np.cos(np.deg2rad(i)) / c
10    Y = np.sum(X_fft * np.exp(-1j * omega * tau), axis=1)
11    P.append(np.linalg.norm(Y))
12 k = np.argmax(P) + 1
13 print(k, -k % 360)
14 plt.plot(theta, P + P[::-1])
15 plt.show()
```

100% | 180/180 [00:10<00:00, 16.61it/s]  
135 225



```
1 path = '作业/2_uniform_circular_array'
2 fs, X = read_waves(path)
3 N = X.shape[0]
4 X_fft = np.fft.rfft(X, axis=0)
5 omega = 2 * np.pi * np.tile(np.arange(N / 2 + 1), (M, 1)).T * fs / N
6
7 P = []
8 for i in tqdm(theta):
9     tau = d * np.cos(np.deg2rad(i - np.arange(M) / M * 360)) / c
10    Y = np.sum(X_fft * np.exp(-1j * omega * tau), axis=1)
11    P.append(np.linalg.norm(Y))
12 k = np.argmax(P) + 1
13 print(k)
14 plt.plot(theta, P)
15 plt.show()
```

100% | 360/360 [00:23<00:00, 15.12it/s]  
60



- 以 SRP-PHAT 为例:

借鉴GCC方法中的PHAT加权思想, 定义SRP-PHAT方法:

$$P(\mathbf{k}) = \sum_{i=0}^{N-1} \sum_{l=0}^{N-1} \int_{-\infty}^{\infty} \Psi_{il}(\omega) X_i(\omega) X_l^*(\omega) e^{j\omega\tau_{il}} d\omega \quad (7.8)$$

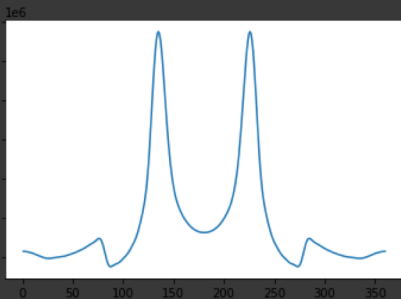
其中 $\Psi_{il}(\omega)$ 为PHAT加权函数:

$$\Psi_{il}(\omega) = G_i(\omega) G_l^*(\omega) = \frac{1}{|X_i(\omega) X_l^*(\omega)|} \quad (7.9)$$

# 声源定位

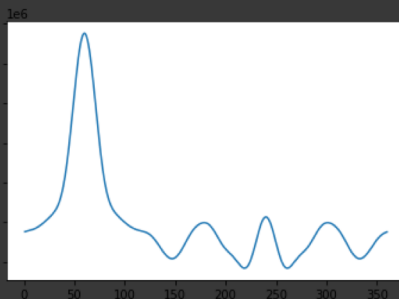
```
1 path = '作业/1_uniform_linear_array'
2 fs, X = read_waves(path)
3 N = X.shape[0]
4 X_fft = np.fft.rfft(X, axis=0)
5 omega = 2 * np.pi * np.arange(N / 2 + 1) * fs / N
6
7 P = [0] * 180
8 for i in tqdm(theta[:180]):
9     tau = np.arange(M) * d * np.cos(np.deg2rad(i)) / c
10    tau = np.tile(tau, (M, 1)) - tau[:, np.newaxis]
11    for m in range(M):
12        for n in range(M):
13            XX = X_fft[:, m] * np.conj(X_fft[:, n])
14            XX = XX / np.abs(XX) * np.exp(-1j * omega * tau[n, m])
15            P[i - 1] += np.sum(XX).real
16 k = np.argmax(P) + 1
17 print(k, -k % 360)
18 plt.plot(theta, P + P[::-1])
19 plt.show()
```

100% | 180/180 [01:01<00:00, 2.92it/s] 135 225



```
1 path = '作业/2_uniform_circular_array'
2 fs, X = read_waves(path)
3 N = X.shape[0]
4 X_fft = np.fft.rfft(X, axis=0)
5 omega = 2 * np.pi * np.arange(N / 2 + 1) * fs / N
6
7 P = [0] * 360
8 for i in tqdm(theta):
9     tau = d * np.cos(np.deg2rad(i - np.arange(M) / M * 360)) / c
10    tau = np.tile(tau, (M, 1)) - tau[:, np.newaxis]
11    for m in range(M):
12        for n in range(M):
13            XX = X_fft[:, m] * np.conj(X_fft[:, n])
14            XX = XX / np.abs(XX) * np.exp(-1j * omega * tau[n, m])
15            P[i - 1] += np.sum(XX).real
16 k = np.argmax(P) + 1
17 print(k)
18 plt.plot(theta, P)
19 plt.show()
```

100% | 360/360 [02:00<00:00, 2.99it/s] 60



# 作业问题

---

- 此次作业提交的人数过少
- SRP 的难点在于离散傅里叶变换输出的  $\omega$ 
  - 离散傅里叶变换
  - 采样定理

**感谢各位聆听**  
Thanks for Listening

