

利用广义相关算法估计远场声源角度

郝千越¹, 胡开哲², 刘雨场³

(¹无 85, 2018011153; ²无 84, 2018013326; ³无 81, 2018010651)

1 设计思路

1.1 总体设计

本实验模拟麦克风阵列声源定位, 输入数据为环形麦克风阵列中四个麦克风采集到的音频, 据此推断出声源方向, 其设计方案主要分为以下三个阶段:

- 阶段 1: 通过时频域综合方法对输入语音信号进行处理, 滤除杂音得到质量较高的音频;
- 阶段 2: 通过广义相关的方法, 对四个输入音频进行滑动相关, 得到音频之间的时间差;
- 阶段 3: 利用麦克风阵列的几何关系, 根据四路输入音频的时间差得到声源方位。

1.2 音频处理

音频处理主要包括低通滤波和谱减法降噪两部分。

根据前期的测试结果, 推断真实信号为语音, 故首先使用 1500Hz 低通滤波去除部分噪声。从频域语谱图观测, 噪声分布近似高斯噪声, 可使用谱减法进一步降噪。

谱减法基于频域变换进行降噪。假设开始一段时间(0.25s)的音频为基底噪声, 通过对音频加窗、分段, 并比较基底噪声和每段音频的谱相似度判断当前音频是否含有语音信息。若当前音频为噪声, 则将其平滑后加入基底噪声。为了更好地利用噪声的频域信息, 我们采用了非线性谱减法算法¹, 对得到的频谱添加核函数进行映射, 随后在映射域将原信号与噪声信号相减得到降噪信号。

1.3 广义相关算法

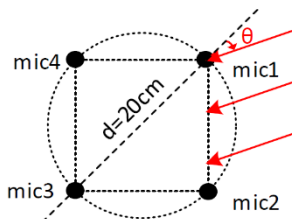
相关算法即将两个信号进行相对滑动并计算其积分, 在积分值最大位置即为两信号重叠最大的位置。利用这一方法可以判定两个信号的时间差, 从而进一步应用于声源定位。广义相关算法是在相关算法上进一步加入功率谱密度的归一化, 可以降低噪声对时间差检测精度的影响, 其公式如下。同时容易得到其与信号功率谱密度 $S(\omega) = \mathcal{F}(R_0(\tau))$ 的关系, $R_0(\tau)$ 为信号的互相关函数。

$$R(\tau) = \mathcal{F}^{-1} \left(\frac{X_1(\omega)X_2^*(\omega)}{|X_1(\omega)X_2^*(\omega)|} \right) = \mathcal{F}^{-1} \left(\frac{S(\omega)}{|S(\omega)|} \right)$$

实现该算法时可以使用快速傅里叶变换函数直接按照第一个等号右侧表达式计算, 也可利用 MATLAB 中的 `xcorr` 函数先得到 $R_0(\tau)$ 后再按照第二个等号右侧表达式计算, 本设计中使用后者。

1.4 时间差与方位确定

根据文档中描述的麦克风阵列结构, 建立各麦克风受到声音的时间差于声源角度 θ 的函数如下, 其中 d 为对角线长度, d_1 为边长, v 为空气中的声速。



$$t_{mic2} - t_{mic1} = f(\theta) = \begin{cases} \sin(45^\circ - \theta) d_1 / v, & 0^\circ < \theta < 180^\circ \\ \sin(\theta - 225^\circ) d_1 / v, & 180^\circ < \theta < 360^\circ \end{cases}$$

$$\begin{cases} t_{mic3} - t_{mic1} = \cos(\theta) d / v \\ t_{mic4} - t_{mic1} = f[(\theta + 45^\circ) \bmod 360^\circ] \end{cases} \quad 0^\circ < \theta < 360^\circ$$

由于数学上三角函数的周期性, 或反映在空间中的结构对称性, 只凭两个麦克风的音频时间差无法唯一确定声源角度, 因此麦克风阵列中需要设置多个麦克风。利用两对连线互不平行的麦克风 (如 $mic1 \& mic3$ 、 $mic2 \& mic4$), 即可唯一确定声源方向。

在初始设计中, 只计算 2、3、4 号麦克风与一号麦克风的时间差, 使用反三角函数直接推算出声源角度。通过对实验结果的观察, 这一方法存在以下两个严重问题:

¹ Poruba, Jiri. "Speech enhancement based on nonlinear spectral subtraction." Proceedings of the Fourth IEEE International Caracas Conference on Devices, Circuits and Systems (Cat. No. 02TH8611). IEEE, 2002.

- 由于滤波操作无法完全消除噪声，因此广义相关不能保证得到合理的结果（结果对噪声敏感度很高）。一些情况下，广义相关会给出不可能出现的时间差（即大于最大对角线长对应的时间差），导致定位结果严重偏差。
- 在无损情况下，四个麦克风之间的时间差只有三个自由度。但存在误差时，计算两两之间共六个时间差对于消除误差、避免错误有重要意义。初始设计中并未利用这一点。

因此，结合测试结果只可能为整数的特点，最终采用以下方法：首先计算出理论上的麦克风两两之间共六个时间差；其次由广义相关得到四个声源信号两两之间六个实际时间差，去除超出理论最大值的时间差（最差情况下仍保留有大于两个值，能够唯一定位）；最后用剩余时间差与表中理论值匹配，选取实际时间差向量与理论时间差向量归一化后内积最大的角度作为最终结果。

2 实验结果

连接上述模块，在训练集上进行测试，14 个测试样本的平均误差为 4.3 度，可以证明本设计的有效性。在测试集上运行的结果见 result.txt。

3 结果分析与讨论

3.1 滤波与降噪

实验证明，降噪过程对提高训练集准确度有较明显的效果，对比不进行降噪，误差降低了 10 度左右。

除谱减法降噪外，另一种实用的降噪算法是基于最小均方误差(lms)的自适应滤波。假设有带噪信号 $y = x + n$ 和纯噪声信号 n' ，自适应滤波使用梯度下降算法设计滤波器，通过噪声信号 n' 拟合带噪信号 y 。由于 n' 与 n 相关而与 x 无关，因此经过多轮迭代后可以通过 n' 拟合出 n ，此时拟合 y 的残差即为信号 x 。

实践中，自适应滤波的降噪效果并不理想，可能的原因为难以获得纯净的噪声信号，单纯提取开始一段时间得到的噪声较少不利于拟合等，因此我们选择了基于谱相似度和谱减法的噪声提取算法。

3.2 定位过程与冗余信息的利用

下面讨论另一种利用冗余信息的尝试：将 4 个麦克风之间两两时间差记为 $t_{mici} - t_{micj} = t_{ij}$, $i, j = 1, 2, 3, 4; i \neq j$ 。最终希望得到 2、3、4 号麦克风与 1 号麦克风之间经过优化的时间差用于定位推断，记为 t_{41}^* 、 t_{31}^* 、 t_{21}^* ，则有如下过定关系：

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} t_{21}^* \\ t_{31}^* \\ t_{41}^* \end{bmatrix} = \begin{bmatrix} t_{21} \\ t_{31} \\ t_{41} \\ t_{32} \\ t_{42} \\ t_{43} \end{bmatrix}$$

尝试用最小二乘法求解该过定方程得到 t_{41}^* 、 t_{31}^* 、 t_{21}^* ，然而发现训练集上的准确率反而下降，分析原因为：六个时间差中存在广义相关判断失误的值，在最小二乘法中直接将这值也用于最终结果，增大了误差。前述设计中一定程度上去掉了野值的影响，表现较好。

一些专业场合使用的麦克风阵列包含更多数量的麦克风，因而具有更多冗余信息，设计合理的阵列空间结构、有效地利用冗余信息消除误差是获得更好的声学表现的关键。

附录 文件清单

文件名	说明	文件名	说明
设计报告.pdf	本文件	corr.m	广义相关算法
test.m	顶层文件	corr_fun.m	广义相关函数
audio_process.m	声源处理与降噪	orient.m	时间差声源定位
voice_filter.m	声源低通滤波器	以上.m文件均位于 code 文件夹中	
spec_sub.m	基于谱减法的降噪依赖函数	result.txt	测试结果

本实验代码在 Windows10，MATLAB R2020a 上调试无误。将 code 文件夹与存放测试数据的 test 文件夹置于同级位置，在 matlab 中运行 test.m 文件，即可于 test 文件夹中获得测试结果 result.txt。