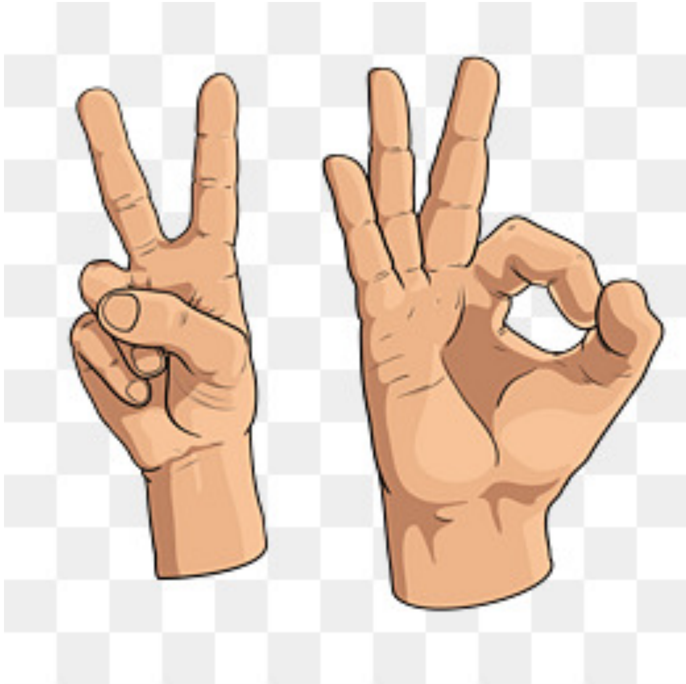
本节内容

传感器数据预处理技术



问题

传感器测量数据

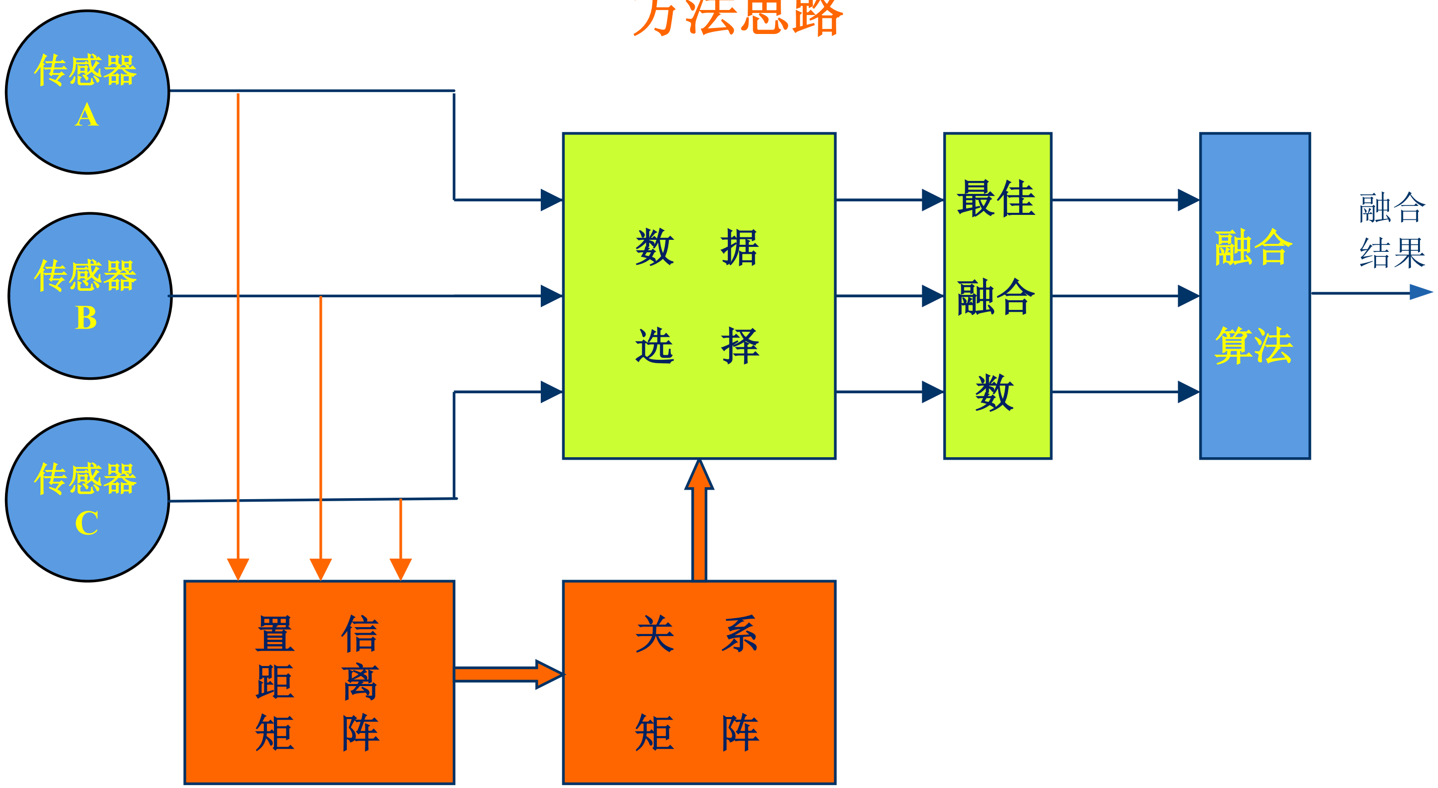
- 不够准确

数据融合

- 被测量数据的真实值是多少

传感器测量数据——预处理

方法思路



基本理论和方法—置信距离和置信距离矩阵

- **利用多个传感器测量某参数的过程中有两个随机变量，**
  - 被测参数 μ，
  - 每个传感器的输出X_i，$i=1$，2，…，m。
  - 一般认为它们服从正态分布，用x_i表示第i个测量值的一次测量输出，它是随机变量X_i的一次取样。
- **设：**

$$\mu \sim N\left(\mu_0, \sigma_0^2\right)$$
$$X_k \sim N\left(\mu, \sigma_k^2\right)$$

基本理论和方法—置信距离和置信距离矩阵

- 为对传感器输出数据进行选择，必须对其可靠性进行估计，为此定义各数据间的置信距离。
- 用X_i、X_j表示第i个和第j个传感器的输出，则其一次读数x_i和x_j之间的置信距离定义为：

$$d_{ij} = 2\int_{x_i}^{x_j} p_i\left(x \middle| x_i\right) dx$$

$$d_{ji} = 2\int_{x_j}^{x_i} p_j\left(x \middle| x_j\right) dx$$

基本理论和方法—置信距离和置信距离矩阵

- 若X_i、X_j服从正态分布，则上式中：

$$p_i(x|x_i)=\frac{1}{\sqrt{2\pi}\sigma_i}\exp\left\{-\frac{1}{2}\left(\frac{x-x_i}{\sigma_i}\right)^2\right\}$$

$$p_j(x|x_j)=\frac{1}{\sqrt{2\pi}\sigma_j}\exp\left\{-\frac{1}{2}\left(\frac{x-x_j}{\sigma_j}\right)^2\right\}$$

故可知：

- 当 $x_i = x_j$ 时，$d_{ij} = d_{ji} = 0$
- 当 $x_i >> x_j$或$x_j >> x_i$ 时，$d_{ij} = d_{ji} = 1$

基本理论和方法—置信距离和置信距离矩阵

- 置信距离矩阵：对m个传感器的一次测量数据，利用上述方法可以分别计算任意两个传感器数据之间的置信距离 $d_{ij},\quad i,j=1,2,\cdots,\quad m$ 得到一个 m X m 矩阵。

$$D_m=\begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1m} \\ d_{21} & d_{22} & \cdots & d_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ d_{m1} & d_{m2} & \cdots & d_{mm} \end{bmatrix}$$

基本理论和方法—关系矩阵和数据选择

- 根据具体问题选择合适的临界值β_{ij}由d_{ij}对数据的可靠性进行判定。

$$r_{ij}=\begin{cases}1 & d_{ij}\le\beta_{ij}\\ 0 & d_{ij}>\beta_{ij}\end{cases}$$

- 由此得到一个二值矩阵，称为关系矩阵。

$$R_m=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m}\\ r_{21} & r_{22} & \cdots & r_{2m}\\ \cdots & \cdots & \cdots & \cdots\\ r_{m1} & r_{m2} & \cdots & r_{mm}\end{bmatrix}$$

基本理论和方法—基于Bayes估计的数据融合算法

- 设被测参数 $\mu \sim N\left(\mu_0, \sigma_0^2\right)$ ，第k个传感器的测量数据 $X_k \sim N\left(\mu, \sigma_k^2\right)$ ，经过删选，选择l个数据作为最佳融合数。融合结果 $\hat{\mu}$ 为：

$$\hat{\mu} = \frac{\sum_{k=1}^{l} \frac{x_k}{\sigma_k^2} + \frac{\mu_0}{\sigma_0^2}}{\sum_{k=1}^{l} \frac{1}{\sigma_k^2} + \frac{1}{\sigma_0^2}}$$

基于Bayes估计的数据融合一般步骤

① 计算m个传感器数据的置信距离矩阵，为简化计算，当测试数据服从正态分布时可利用误差函数计算置信距离。

$$d_{ij} = erf\left(\frac{x_j - x_i}{\sqrt{2}\sigma_i}\right)$$

$$erf(\theta) = \frac{2}{\pi}\int_0^{\theta} e^{-u^2} du$$

基于Bayes估计的数据融合一般步骤

② 选择合适的距离临界值，由置信距离矩阵产生关系矩阵，表示第j个传感器对第i个传感器的支持。

$$r_{ij}=\begin{cases}1 & d_{ij}\le \beta_{ij}\\ 0 & d_{ij}>\beta_{ij}\end{cases}$$

③ 由关系矩阵对多传感器数据进行选择，产生最佳融合数。

如果一个传感器被一组传感器支持，则它的读数是有效的。
否则它的读数无效，在融合中不考虑。

基于Bayes估计的数据融合一般步骤

④ 将 μ_0、σ_0^2 和最佳融合数对应的 x_k、σ_k^2 代入Bayes融合估计公式求的参数估计值。

$$\hat{\mu} = \frac{\sum_{k=1}^{l} \frac{x_k}{\sigma_k^2} + \frac{\mu_0}{\sigma_0^2}}{\sum_{k=1}^{l} \frac{1}{\sigma_k^2} + \frac{1}{\sigma_0^2}}$$

举例计算

- 利用8个传感器对一个恒温槽的温度进行测量，已知恒温槽温度满足正态分布，其中 μ_0=850.50℃，σ_0^2 =4.5025

8个传感器的测量结果如下：

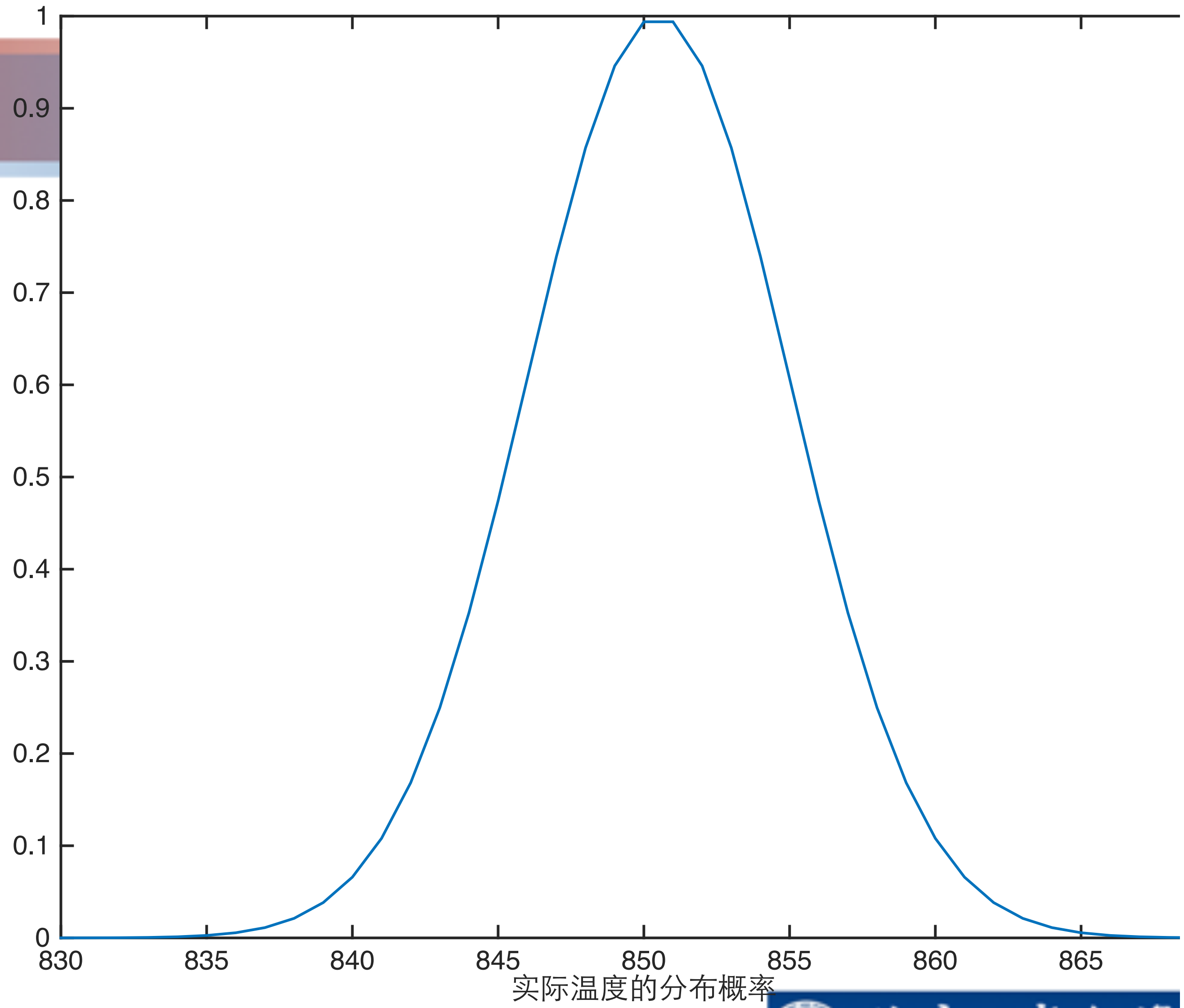
传感器编号	1	2	3	4	5	6	7	8
方差	25.73	23.81	24.95	25.75	35.65	21.33	23.94	22.96
测量值	848.1	850.5	851.9	849.9	854.6	849.3	848.0	848.3

恒温槽温度的分布概率

注意：

这是每次测量的实际真实值的概率分布，

它和测量的真实值不同。



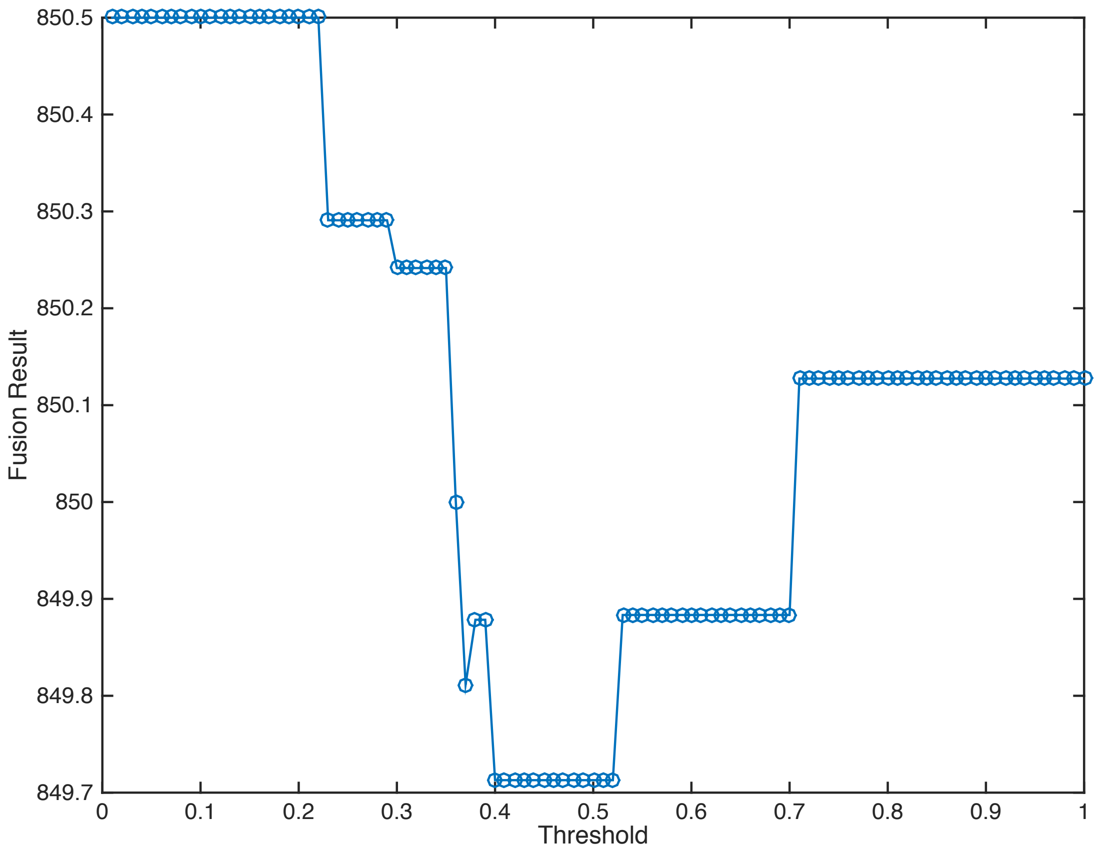
程序

- measurements=[848.1,850.5,851.9,849.9,854.6,849.3,848.0,848.3]';
- covv=[25.73,23.81,24.95,25.75,35.65,21.33,23.94,22.96]';
- RealValue=850.5;
- RealCov=4.5025;

```
for i=1:8
    for j=1:8
        d(i,j)=erf((measurements(j)-measurements(i))/(sqrt(2)*covv(i)))
    end
End
Threshold=0.4;
R=(abs(d)-Threshold*ones(8))<0
SupportNumber=5;
support=(sum(R,2)>SupportNumber)
FusionData=(support'*(measurements./covv)+RealValue/RealCov)/(support'*(1./covv)+1/RealCov)
```

融合结果

- **849.7129**
- 修改程序研究以下问题
  - 阈值和融合结果的关系
  - 支持传感器个数与融合结果的关系



考察阈值从0.01到1、要求最少的传感器个数1到8，融合结果如何？？