

A Specific Speech Control Home System

Yan Peng, Jingkuo Zhang, Runxue Chen, Shuang Ren

School of Computer, Sichuan University of Science & Engineering, Zigong Sichuan
Email: 1468007872@qq.com

Received: Sep. 28th, 2017; accepted: Oct. 10th, 2017; published: Oct. 17th, 2017

Abstract

The application of speech control in smart home brings great convenience to life. However, conventional independent speech control will make incorrect speech triggering. Based on the non-specific speech recognition, this paper designs a specific human voice recognition home control system by improving the corresponding algorithm. This system gets speech features by MFCC algorithm and matches speech feature by DTW algorithm. Host computer is based on Android platform. This paper uses Android Studio tools and NDK assembly to achieve the development of software and algorithm. Lower computer is based on Zigbee platform used IAR to achieve lower computer development of smart home. As a result, this system can solve specific speech control with a higher recognition rate and more human operation.

Keywords

Home System, Specific Speech, Speech Control, Android Platform, Zigbee Platform

一种特定人语音家居控制系统

彭 龔, 张景扩, 陈闰雪, 任 爽

四川理工学院计算机学院, 四川 自贡
Email: 1468007872@qq.com

收稿日期: 2017年9月28日; 录用日期: 2017年10月10日; 发布日期: 2017年10月17日

摘 要

语音控制在智能家居的应用为人民生活带来极大便利, 但常规的非特定人语音控制会产生语音的误触发[1]。本文在非特定语音识别的基础上, 通过改进相应算法设计了一种特定人语音识别家居控制系统。系统采用MFCC算法获取语音特征, 通过DTW算法实现语音特征匹配。上位机基于安卓平台, 利用Android Studio工具并结合NDK组件实现相应软件与算法开发。下位机基于Zigbee平台, 利用IAR完成家居系统下位机的开发。系统可完成特定人的语音控制, 具有较高的识别率, 操作更加人性化。

关键词

家居系统, 特定语音, 语音控制, 安卓平台, Zigbee平台

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

语音控制技术突破传统手动操作的局限, 让人与机器之间的沟通交流更智能化[2]。语音控制发展迅速, 出现了许多出色的语音控制系统。其中语音控制家居系统发展前景广阔, 针对家居控制的各项要求中, 控制的准确性及安全性极为重要。本文在传统的语音家居控制系统基础上增加特定人语音识别, 在非特定语音识别的基础上, 通过改进相应算法设计出一种特定人语音识别家居控制系统, 可减少意外因素所带来的语音误触。

2. 特定人语音家居控制系统设计原理

特定人语音家居控制系统主要由移动终端控制和家居下位机控制组成(如图 1)。移动终端基于安卓平台, 由用户发送语音指令至移动平台, 然后利用 MFCC(梅尔频率倒谱系数)算法获取语音特征, 并利用 DTW(动态时间规整)算法实现语音特征匹配, 最后经 Wi-Fi 将指令送至下位机。下位机基于 Zigbee 平台, 作为系统执行者, 首先接收来自上位机的 Wi-Fi 指令信号, 然后对信号识别, 实现对家电的控制。

3. 语音识别算法

3.1. 语音特征参数提取

语音信号经过预处理之后, 提取语音特征的目的是为了提取说话人的语音特征进而确定说话人身份[3]。常用的特征参数提取算法有: 基于人耳听觉特性的 MFCC(梅尔倒谱系数)、基于声道特性的 LPCC(线性预测倒谱系数)。本文采用 MFCC 算法。MFCC 算法在孤立词识别上, 相比基于声道模型的 LPCC 算法具有更好的鲁棒性, 更符合人的听觉特征, 当信噪比降低时仍然具有较好的识别性[4]。图 2 为 MFCC 特征参数提取原理, 获得 MFCC 特征参数方法如下:

1) 先对输入信号进行预处理(分帧、加窗、预加重), 然后作离散傅立叶变换得到语音信号的频谱值。通过以上方式达到提高初始语音稳定性的效果并且方便后续语音特征参数的提取与匹配, 并规整语音信号的频谱。设语音信号的离散傅里叶变换为:

$$X_a(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi nk/N}, 0 \leq k \leq N \quad (1)$$

N 表示傅立叶变换的点数, $x(n)$ 表示语音输入信号。

2) 利用频谱幅度的平方获得能量谱。

3) 利用三角形滤波器组分析上述步骤获得的能量谱。

定义一个有 M 个三角滤波器的滤波器组, 中心频率为 $f(m, m=1, 2, \dots, M)$, 本文取 $M=24$ 。 m 值的减小会致 $f(m)$ 之间的间隔减小, 反之亦然。三角滤波器的频率响应定义为:

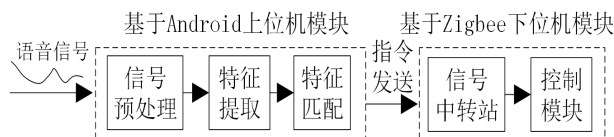


Figure 1. Principle of specific speech control home system

图 1. 特定人语音家居控制系统原理

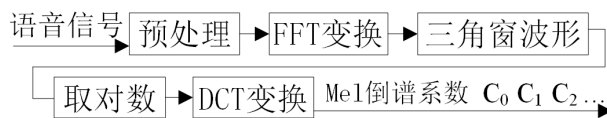


Figure 2. Principle of MFCC extracting characteristic parameters

图 2. MFCC 特征参数提取原理

$$H_m(k) = \begin{cases} 0, k < f(m-1) \\ \frac{2(k-f(m-1))}{(f(m+1)-f(m-1))(f(m)-f(m-1))}, f(m-1) \leq k \leq f(m) \\ \frac{2(f(m+1)-k)}{(f(m+1)-f(m-1))(f(m+1)-f(m))}, f(m) \leq k \leq f(m+1) \\ 0, k \geq f(m+1) \end{cases} \quad (2)$$

其中 $\sum_{m=0}^{M-1} H_m(k) = 1$

4) 根据(1)与(2)的结果计算 MFCC 对数能量:

$$s(m) = \ln \left(\sum_{k=0}^{N-1} |X_a(k)|^2 X_m(k) \right), 0 \leq m \leq M \quad (3)$$

5) 经离散余弦变换得到 MFCC 参数:

$$C(n) = \sum_{m=0}^{N-1} S(m) \cos \left(\frac{\pi n(m-0.5)}{M} \right), 0 \leq n \leq M \quad (4)$$

2.2. 特征参数匹配

为解决说话人与模板语音的问题, 本文采用特征参数匹配算法(如图 3)。经过 MFCC 算法提取得到相应的参数即匹配中所使用的特征参数, 再根据特征参数建立相应模板完成说话人语音匹配。在时间序列中, 若语音信号波形长度相等则语音较容易匹配[5]。如图 4 所示, 需要比较相似性的两段时间序列的长度可能并不相等, 在语音识别领域, 语音信号存在一定的随机性, 比如 a 点会对应到 b 点。传统的欧几里得测距无法有效求出两个时间序列间的距离, 利用 DTW 算法可找到 a 点实际应匹配的 c 点。

DTW 利用非线性规整技术可有效完成孤立词的匹配, 通过规划函数 $i_m = \Phi(i_n)$, 将参考模板的时间轴 m 与测试信号的时间轴 n 实现非线性地映射, 并使该函数满足:

$$D = \min_{\Phi(i_n)} \sum_{i_n=1}^N d(T(i_n), R(\Phi(i_n))) \quad (5)$$

D 代表最优时间规整下两矢量间的长度。DTW 算法利用局部最佳化的处理寻找最优路径。在最优路径中, 信号特征变量的累积失真最小, 有效减少时间不统一所带来的误差。

DTW 算法原理如图 5, 横轴为测试模板的各个帧号 $n = 1 \sim N$, 纵轴为参考模板的各帧 $m = 1 \sim M$ 。

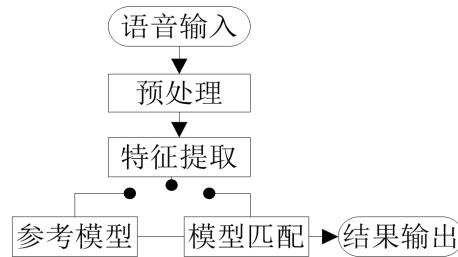


Figure 3. Flow of characteristic parameter matching
图 3. 特征参数匹配流程

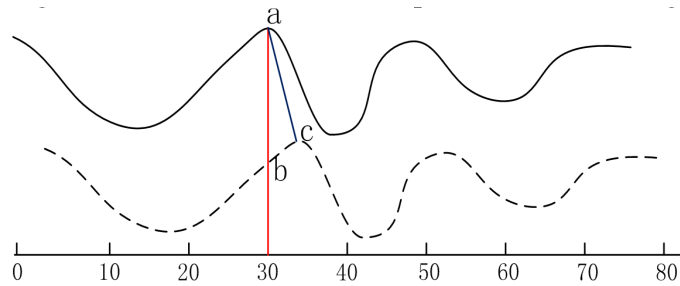


Figure 4. DTW algorithm principle
图 4. DTW 算法原理

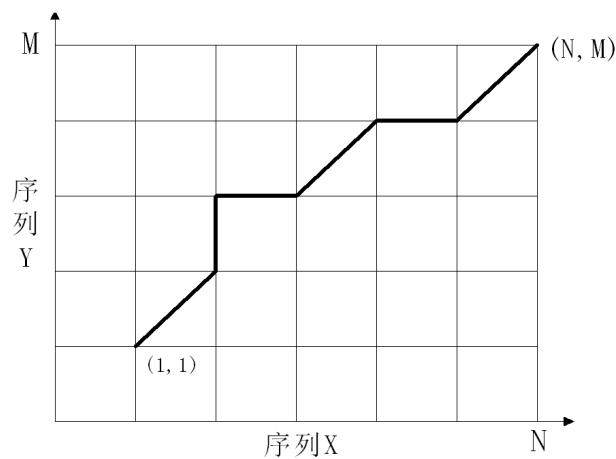


Figure 5. Principle of DTW algorithm
图 5. DTW 算法原理

DTW 算法首先分析两个模板各帧之间的长度，接着在帧匹配距离矩阵中寻求最优路径。寻求最优路径方式描述为：从图 5 中(1,1)点开始，到达点 (i_n, i_m) 有三种方式(如图 6, 分别为 (i_{n-1}, i_m) 、 (i_{n-1}, i_{m-1}) 和 (i_{n-1}, i_{m-2}))。从这三种方式找到距离 (i_n, i_m) 最近的方式，获得累积距离：

$$D(i_n, i_m) = d(T(i_n), R(i_m)) + \min\{D(i_{n-1}, i_m), D(i_{n-1}, i_{m-1}), D(i_{n-1}, i_{m-2})\} \quad (6)$$

这样从点(1,1)出发至(N,M)反复计算获得最优路径。待测语音和模板语音进行匹配，获得的最优路径距离为识别结果[6]。

常规的 DTW 匹配算法存在效率低下问题。论文利用放宽端点限制 DTW 算法(如图 7)：利用先验知识估计点 (i, j) ，路径规划分为从 (i, j) 至点(N,M)以及从点 (i, j) 至点(1,1)，这样点 (i, j) 分段向(N,M)和(1,1)计算。由此解决 DTW 算法端点对齐上存在的不足，并可有效缩小搜索空间。

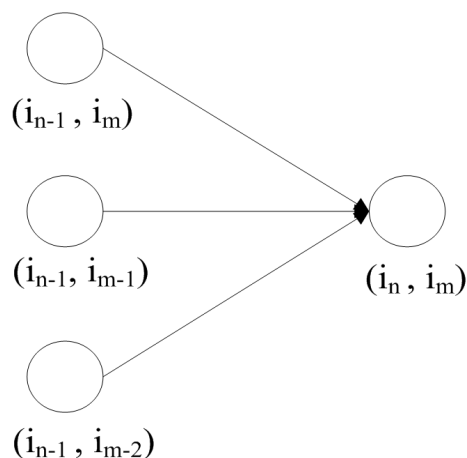


Figure 6. Local constraint path

图 6. 局部约束路径

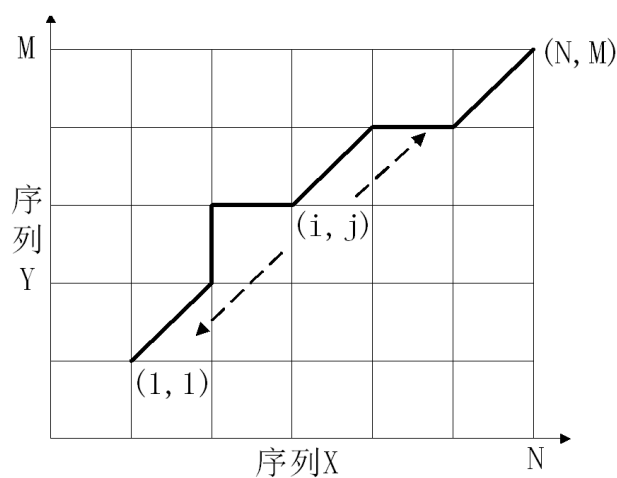


Figure 7. Principle of relaxing endpoint restrictions DTW algorithm

图 7. 放宽端点限制 DTW 算法原理

3. 基于 Android 平台上位机软件设计

本文利用 Android Studio 2.2.3 (JDK1.8, Gradle 2.2.3)完成上位机软件开发。在软件设计中，首先选择操作模式(手动模式和语音模式)，语音操作模式需录入 5 次语音命令以保证特征提取的精度，利用 MFCC 算法及 DTW 算法对语音的特征提取和语音匹配，最后将指令发送至下位机。上位机软件流程如图 8。

上位机软件主程序包括：界面设定、语音输入、NDK 语音处理以及 Wi-Fi 传输模块。具体完成步骤如下：

1) 界面设定

上位机界面主要包含模式切换界面以及语音控制界面。其中语音控制界面包含语音特征录入、语音特征删除及语音输入。

2) 语音输入

通过搭载 Android 移动设备的麦克风完成录音工作，语音录制须开启两个权限：

```
<uses-permission android:name="android.permission.RECORD_AUDIO">
```

```
<uses-permission android:name="android.permission.WRITE_EXTERNAL_STORAGE">
```

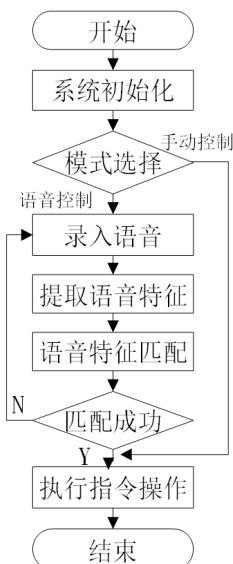


Figure 8. Flow of upper computer software design

图 8. 上位机软件设计流程

完成语音传输的步骤:

- ① 先调用 `setAudioSource()` 完成初始化状态
- ② 调用 `setOutputFormat()` 设定输出格式, 完成设定输出, 完成数据源组态。
- ③ 设定 `setAudioEncoder()` 编码器, 以及 `setOutFile()` 输出文件。
- ④ 调用 `prepare()` 进入准备完成状态
- ⑤ 调用 `start()` 进行录音

3) NDK 语音处理

NDK 利用 C/C++ 等原生代码执行部分程序处理语音信号和图形图像识别等[7]。利用 NDK 开发有利于防止代码反编译且方便外部代码移植。NDK 开发步骤:

- ① 打开 Android Studio 其中 SDK Manager, 勾选其中的 CMake、NDK、LLDB。
- ② 创建工程, 选中 Include C++ Support, 然后点击 Finish。若出现报错则检查上述步骤。
- ③ 进入主界面后先在 MainActivity 定义相应的函数, 然后进入 cpp 下完成 C/C++ 的编写。
- ④ 最后在 Java 框架下调用所定义的函数。

4) Wi-Fi 传输

Wi-Fi 传输类似于蓝牙一样在两个设备之间直接建立数据传输通道。Wi-Fi 的有效距离较长, 传输速率及穿墙性优于蓝牙传输。在 Android SDK 中提供了相应的 API, 通过这些 API 可以发现其他支持 Wi-Fi Direct 的设备, 也可以接收其他支持 Wi-Fi Direct 的设备。Wi-Fi 传输开发如下:

- ① 创建监听 Wi-Fi Direct 动作的广播接收器

```
public class WiFiDirectBroadcastReceiver extends BroadCastReceiver
{public void onReceive(Context context,Intent intent){    } }
```

- ② 注册广播接收器

注册广播接收器需在 AndroidManifest.xml 中限制 Android 版本至少为 4.0。且添加与 Wi-Fi 相关的权限。利用 BroadCastReceiver 相应组件完成注册。

- ③ 发现并连接设备

使用 Wi-Fi Direct 进行通信。首先要搜索周围的 Wi-Fi 设备, 调用 `WiFiP2pManager.discoverPeers` 方法完成这项工作。当搜索到 Wi-Fi 设备后, 可以选择其中设备进行连接。每一个设备都是 `WiFiP2pDevice` 对象, 然后利用 `WiFiP2pDevice` 对象存储设备相关的信息。最后调用 `WiFiP2pManager.connect` 方法连接设备。

④ 数据传输

当设备之间建立连接后, 可以调用 `WiFiP2pManager.requestConnectionInfo` 方法获取作为服务端设备信息。根据主机信息, 先创建 `ServerSocket` 对象, 然后使用 `ServerSocket.accept` 方法等待另一端连接。

4. 下位机设计

4.1. 下位机硬件设计

下位机负责与上位机进行数据通信和家电控制。下位机硬件机构如图 9。下位机一部分是充当信号中转站, 收到 Wi-Fi 数据, 然后按照 Zigbee 的协议进行封包, 在设置好通道参数的前提下, 再通过 Zigbee 芯片发送至相应节点; 另一部分用作终端节点, 主要负责数据的采集及控制命令的执行[8]。

4.2. 下位机软件设计

上位机和下位机间通信过程是由上位机发出请求通信, 下位机完成执行[9]。终端节点与协调器的通信是通过终端节点将信息发送至协调器, 再经串口送至上位机。下位机软件设计包括终端节点软件和协调器软件的设计。软件设计主要分为组网与联网、端点通信和节点间数据通信, 程序流程如图 10。

5. 实验分析

本文利用搭载 Android7.0 的安卓智机为测试工具, 录音采用手机麦克风, 提取语音 MFCC 参数储存

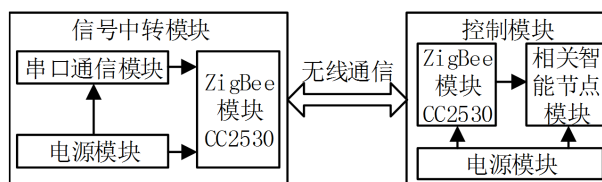


Figure 9. Hardware structure of lower computer

图 9. 下位机硬件系统结构

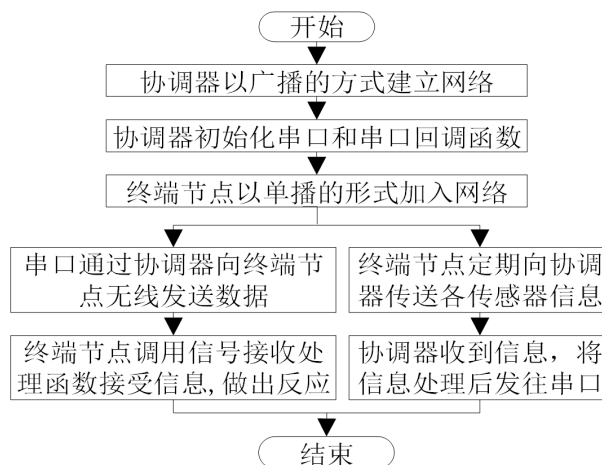


Figure 10. Flow of lower computer program

图 10. 下位机程序流程图

Table 1. Results of specific speech control home system (indoor quiet condition)**表 1.** 特定人语音家居识别系统结果(室内安静状态)

测试语音	室内安静下语音识别成功率/%		
	人物一	人物二	人物三
打开台灯	89%	19%	18%
关闭窗帘	87%	21%	17%

Table 2. Results of specific speech control home system (indoor noisy condition)**表 2.** 特定人语音家居识别系统结果(室内嘈杂状态)

测试语音	室内嘈杂下语音识别成功率/%		
	人物一	人物二	人物三
打开台灯	86%	23%	21%
关闭窗帘	84%	21%	22%

在手机内存上。系统下位机采用 Zigbee 开发工具箱(CITE-T-ZA)完成对智能家居的模拟。采用语音预加重系数 $a = 0.9$ ，汉明窗作为窗函数[10]，采样频率 9 khz、18 big，帧移 85，语音每帧 28 ms。实验对于孤立词各训练 5 次，并在室内相对安静与嘈杂环境进行 100 次语音控制测试。测试不同人物的传输成功率以及不同指令的传输成功率[11] (其中人物一为主说话人，其他为参照对象)，解锁成功率为说出四字短语到下位机执行操作的成功度，结果如下：

据表 1 与表 2 可观测出：特定人语音的识别正确率较参照对象有明显的区别；且室内识别成功率高于室外。根据本文的改进算法，针对特定人的语音识别能够准确识别特定人的语音，进而防止非特定人语音误触的发生，且在室内外均能准确的识别特定人的语音。

6. 结束语

本文基于 MFCC 和 DTW 设计出一种特定人语音识别家居控制系统。相较于传统的语音控制模式，系统可完成特定人的语音控制，具有较高的识别率、安全性及稳定性，语音操作更加人性化。

基金项目

自贡市科技局苗子工程(2016CXM08)；四川省科技厅科技支撑计划(2012GZX0086)。

参考文献 (References)

- [1] 吴桐. 一种特定人语音识别的方法[P]. 中国专利, 103236261. 2013-8-7
- [2] 赵智琦, 房建东. 一种改进的特定人语音识别系统及算法研究[J]. 电子设计工程, 2014, 22(16): 31-37.
- [3] Chahuara, P., Portet, F. and Vacher, M. (2017) Context-Aware Decision Making under Uncertainty for Voice-Based Control of Smart Home. *Expert Systems with Applications*, **75**, 63-79. <https://doi.org/10.1016/j.eswa.2017.01.014>
- [4] Mishra, S., Bhowmick, A. and Shrotriya, M.C. (2016) Hindi Vowel Classification Using QCN-MFCC Features. *Perspectives in Science*, **8**, 28-31. <https://doi.org/10.1016/j.pisc.2016.01.010>
- [5] Le, H.L., Hai, N.T., Thuyen, N.V., et al. (2015) MFCC-DTW Algorithm for Speech Recognition in an Intelligent Wheelchair. *5th International Conference on Biomedical Engineering in Vietnam*, Ho Chi Minh, 11-14 January 2015, 417-421.
- [6] Xie, C., Cao, X. and He, L. (2012) Algorithm of Abnormal Audio Recognition Based on Improved MFCC. *Procedia*

Engineering, **29**, 731-737. <https://doi.org/10.1016/j.proeng.2012.01.032>

- [7] Hsu, C.J., Huang, K.S., Yang, C.B., *et al.* (2015) Flexible Dynamic Time Warping for Time Series Classification. *Procedia Computer Science*, **51**, 2838-2842. <https://doi.org/10.1016/j.procs.2015.05.444>
- [8] Papcun, P., Zolotova, I. and Tafsi, K. (2016) Control and Teleoperation of Robot Khepera via Android Mobile Device through Bluetooth and WiFi. *IFAC-Papers Online*, **49**, 188-193. <https://doi.org/10.1016/j.ifacol.2016.12.032>
- [9] Liu, Z.Y. (2014) Hardware Design of Smart Home System based on Zigbee Wireless Sensor Network. *Aasri Procedia*, **8**, 75-81. <https://doi.org/10.1016/j.aasri.2014.08.013>
- [10] 郝博麟. 基于 Zigbee 的智能家居系统应用平台的设计与实现[D]: [硕士学位论文]. 长春: 吉林大学, 2016.
- [11] Chahuara, P., Portet, F. and Vacher, M. (2017) Context-Aware Decision Making under Uncertainty for Voice-Based Control of Smart Home. *Expert Systems with Applications*, **75**, 63-79. <https://doi.org/10.1016/j.eswa.2017.01.014>

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2161-8801, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: csa@hanspub.org