**基于DSP和FPGA的机器人声控系统设计与实现**

1系统简介

机器人听觉系统主要是对人的声音进行语音识别并做出判断，然后输出相应的动作指令控制头部和手臂的动作，传统的机器人听觉系统一般是以PC机为平台对机器人进行控制，其特点是用一台计算机作为机器人的信息处理核心通过接口电路对机器人进行控制，虽然处理能力比较强大，语音库比较完备，系统更新以及功能拓展比较容易，但是比较笨重，不利于机器人的小型化和复杂条件下进行工作，此外功耗大、成本高。

本次设计采用了性价比较高的数字信号处理芯片TMS320VC5509作为语音识别处理器，具有较快的处理速度，使机器人在脱机状态下，独立完成复杂的语音信号处理和动作指令控制，FPGA系统的开发降低了时序控制电路和逻辑电路在PCB板所占的面积[1]，使机器人的"大脑"的语音处理部分微型化、低功耗。一个体积小、低功耗、高速度能完成特定范围语音识别和动作指令的机器人系统的研制具有很大的实际意义。

2系统硬件总体设计

系统的硬件功能是实现语音指令的采集和步进电机的驱动控制，为系统软件提供开发和调试平台。如图1所示。



系统硬件分为语音信号的采集和播放，基于DSP的语音识别，FPGA动作指令控制、步进电机及其驱动、DSP外接闪存芯片，JTAG口仿真调试和键盘控制几个部分。工作流程是麦克风将人的语音信号转化为模拟信号，在经过音频芯片TLV320AIC23量化转化成数字信号输入DSP。DSP完成识别后，输出动作指令。

FPGA根据DSP输入的动作指令产生正确的正反转信号和准确的脉冲给步进电机驱动芯片，驱动芯片提供步进电机的驱动信号，控制步进电机的转动。片外FLASH用于存储系统程序和语音库并完成系统的上电加载。JTAG口用于与PC机进行联机在线仿真，键盘则用于参数调整和功能的切换。

3语音识别系统设计

3．1语音信号的特点

语音信号的频率成分主要分布在300～3400Hz之间，根据采样定理选择信号的采样率为8 kHz。语音信号的一个特点在于他的"短时性"，有时在一个短时段呈现随机噪声的特性，而另一段表现周期信号的特性，或二者兼而有之。语音信号的特征是随时间变化的，只有一段时间内，信号才表现稳定一致的特征，一般来说短时段可取5～50ms，因此语音信号的处理要建立在其"短时性"上[2]，系统将语音信号帧长设为20ms，帧移设为10ms，则每帧数据为160×16b。

3．2语音信号的采集和播放

语音采集和播放芯片采用的是TI公司生产的TLV320AIC23B，TLV320AIC23B的模数转换(ADC)和数模转换(DAC)部件高度集成在芯片内部，芯片采用8k采样率，单声道模拟信号输入，双声道输出。TLV320AIC23具有可编程特性，DSP可通过控制接口来编辑该器件的控制寄存器，而且能够编译SPI，I2C两种规格的接口，TLV320AIC23B与DSP5509的电路连接如图2所示。



DSP采用I2C口对TLV320AIC23的寄存器进行设置。当MODE=O时，为I2C规格的接口，DSP采用主发送模式，通过I2C口对地址为0000000～0001111的11个寄存器进行初始化。I2C模式下，数据是分为3个8b写入的。而TLV320AIC23有7位地址和9位数据，也就是说，需要把数据项上面的最高位补充到第二个8B中的最后一位。

MCBSP串口通过6个引脚CLKX，CLKR，FSX，FSR，DR和CX与TLV320AIC23相连。数据经MCBSP串口与外设的通信通过DR和DX引脚传输，控制同步信号则由CLKX，CLKR，FSX，FSR四个引脚实现。将MCBSP串口设置为DSP Mode模式，然后使串口的接收器和发送器同步，并且由TLV320AIC23的帧同步信号LRCIN，LRCOUT启动串口传输，同时将发送接收的数据字长设定为32 b(左声道16 b，右声道16 b)单帧模式。

3．3语音识别程序模块的设计

为了实现机器人对非特定人语音指令的识别，系统采用非特定人的孤立词识别系统。非特定人的语音识别是指语音模型由不同年龄、不同性别、不同口音的人进行训练，在识别时不需要训练就可以识别说话人的语音[2]。系统分为预加重和加窗，短点检测，特征提取，与语音库的模式匹配和训练几个部分。

3．3．1语音信号的预加重和加窗

预加重处理主要是去除声门激励和口鼻辐射的影响，预加重数字滤波H(Z)=1一KZ-1，其中是为预加重系数，接近1，本系统中k取0．95。对语音序列X(n)进行预加重，得到预加重后的语音序列x(n)： x(n)=X(n)一kX(n一1)(1)

系统采用一个有限长度的汉明窗在语音序列上进行滑动，用以截取帧长为20 ms，帧移设为10 ms的语音信号，采用汉明窗可以有效减少信号特征的丢失。

3．3．2端点检测

端点检测在词与词之间有足够时间间隙的情况下检测出词的首末点，一般采用检测短时能量分布，方程为：



其中，x(n)为汉明窗截取语音序列，序列长度为160，所以N取160，为对于无音信号E(n)很小，而对于有音信号E(n)会迅速增大为某一数值，由此可以区分词的起始点和结束点。

3．3．3特征向量提取

特征向量是提取语音信号中的有效信息，用于进一步的分析处理。目前常用的特征参数包括线性预测倒谱系数LPCC、美尔倒谱系数MFCC等。语音信号特征向量采用Mel频率倒谱系数MFCC(Mel Frequency Cepstrum Coeficient的提取，MFCC参数是基于人的听觉特性的，他利用人听觉的临界带效应[3]，采用MEL倒谱分析技术对语音信号处理得到MEL倒谱系数矢量序列，用MEL倒谱系数表示输入语音的频谱。在语音频谱范围内设置若干个具有三角形或正弦形滤波特性的带通滤波器，然后将语音能量谱通过该滤波器组，求各个滤波器输出，对其取对数，并做离散余弦变换(DCT)，即可得到MFCC系数。MFCC系数的变换式可简化为：



其中，i为三角滤波器的个数，本系统选P为16，F(k)为各个滤波器的输出数据，M为数据长度。

3．3．4语音信号的模式匹配和训练

模型训练即将特征向量进行训练建立模板，模式匹配即将当前特征向量与语音库中的模板进行匹配得出结果。语音库的模式匹配和训练采用隐马尔可夫模型HMM(Hidden Markov Models)，他是一种统计随机过程统计特性的概率模型一个双重随机过程，因为隐马尔可夫模型能够很好地描述语音信号的非平稳性和可变性，因此得到广泛的使用[4]。

HMM的基本算法有3种：Viterbi算法，前向一后向算法，Baum-Welch算法。本次设计使用Viterbi算法进行状态判别，将采集语音的特征向量与语音库的模型进行模式匹配。Baum-Welch算法用来解决语音信号的训练，由于模型的观测特征是帧间独立的，从而可以使用Baum-Welch算法进行HMM模型的训练。

3．4语音识别程序的DSP开发

DSP的开发环境为CCS3．1及。DSP／BIOS，将语音识别和训练程序分别做成模块，定义为不同的函数，在程序中调用。定义语音识别器函数为int Recognizer(int Micin)，识别结果输出函数为int Result(void)，语音训练器函数为int Train(int Tmode，int Audiod)，动作指令输入函数为int Keyin(int Action[5])。

语音识别器的作用是将当前语音输入变换成语音特征向量，并对语音库的模板进行匹配并输出结果，语音应答输出函数将获取的语音识别结果对应的语音应答输出，语音训练是将多个不同年龄、不同性别、不同口音的人语音指令输入转化为训练库的模板。为防止样本错误，每个人的语音指令需要训练2次，对于2次输入用用欧氏距离去进行模式匹配，若2次输入相似度达到95％，则加入样本集。语音应答输入函数是为每个语音库中模板输入对立的语音输出，以达到语言应答目的。系统工作状态为执行语言识别子程序，训练时执行外部中断，执行训练函数，取得数据库模板，训练完毕返回。程序框图如图3所示。



4机器人的动作控制系统设计

4．1 FPGA逻辑设计

系统通过语音控制机器人头部动作，头部运动分为上下和左右运动2个自由度，需要2个步进电机控制，DSF完成语音识别以后，输出相应的动作指令，动作执行结束后，DSP发出归零指令，头部回到初试状态。FPGA的作用是提供DSP接口逻辑，设置存储DSP指令的RAM块，同时产生步进电机驱动脉冲控制步进电机转动方向和角度。

FPGA器件为动作指令控制单元，设计采用FLEXlOKE芯片，接收DSP数据后并行控制2路步进电机。FPGA内部结构逻辑如图4所示，FPGA内部设置2个元件为电机脉冲发生器，控制电机的工作脉冲以及正反转。AO～A7为DSP数据输入端口，WR为数据写端口，P1，P2为2个步进电机驱动芯片脉冲输入口，L1，L2为电机正反转控制口，ENABLE为使能信号。



RAM1和RAM2分别为2个步进电机的指令寄存器，电机脉冲发生器发出与RAM中相应数量的方波脉冲。DSP通过DO～D8数据端输出8位指令，其中。D8为RAM选择，为1时选择RAM1，为0时选择RAM0，DO～D7为输出电机角度，电极上下和左右旋转角度为120°，精度为1°，初始值都为60°，DO～D7的范围为00000000～11111000，初始值为00111100。FPGA作为步进脉冲发生器，通过时钟周期配置控制电机转速，与初始值对应坐标决定正反转。系统动作指令程序如图5所示。



其中R1为DSP指令寄存器，R2为当前坐标寄存器，通过DSP的输出坐标与FPGA的当前坐标进行差值运算来确定步进电机的旋转方向和旋转角度，优点是可以根据新的输入指令的变化，结束当前动作以运行新的指令，指令执行完毕后，系统清零，步进电机回到初始状态。

4．2 FPGA逻辑仿真

FPGA以MAX-PLUSⅡ开发平台，用语言为VHDL语言对上述逻辑功能进行设计，并通过JTAG接口进行了调试，FLEXl0KE芯片能够根据DSP输出指令输出正确的正反转信号和脉冲波形。

4．3 步进电机驱动设计

FPGA通过P1，L1，P2，L2输出控制控制步进电机驱动芯片。步进电机驱动采用的是东芝公司生产的单片正弦细分二相步进电机驱动专用芯片TA8435H，FPGA与TA8435H电路连接如图6。



由于FLEX1OKE和TMS320VC5509工作电压为3．3 V，而TA8435H为5 V和25 V，所以管脚连接使用光电耦合器件TLP521，使两边电压隔离。CLK1为时钟输入脚，CW／CCW为正反转控制脚，A，A，B，B为二相步进电机输入。

5 结 语

系统充分利用了DSP的高处理速度和可扩展的片外存储空间，具有高速、实时、识别率高的特点并支持大的语音库，FPGA的使用使系统电路获得简化，一片FLEXl0KE芯片可以完成2个步进电机的时序控制。虽然在处理速度和语音库的存储容量上与PC机系统具有一定的差距，但在机器人的微型化、低功耗和特定功能实现上，以DSP和FPGA为核心的嵌入式系统无疑具有广阔的前景。