

基于 i.MX21 的可视电话单芯片解决方案

1. 概述

1964 年，美国贝尔实验室提出了第一个可视电话解决方案。但是由于受各种技术条件的限制，可视电话一直没有取得实质性进展。八十年代末，随着通信、计算机、语音和视频编解码技术的不断发展，可视电话在世界各国得到了迅速发展。为了实现互联互通，以推动可视电话和视频会议系统的发展，国际电信联盟(ITU)在 80 年代就成立一个研究视频会议的小组，对基于不同的网络环境，制定了一系列视频会议标准。传统的会议电视系统大多遵循最早的 ITU-T H.320 标准，这是基于 N*64K 电路系统的标准规范。随着 TCP/IP 技术的广泛应用，及 Internet 迅速普及，人们逐渐认识到，将来的网络对于使用者来说将只会以一种形式存在，那就是 TCP/IP 网络。因此，ITU 从 96 年开始，将视频会议标准制定的工作重点转向 ITU-T H.323 标准，这是一个基于 TCP/IP 网络的视频会议系统标准。



很快，在 97 年 ITU-T H.323 建议第一版正式发布。准确的说，第一版应该称作基于 LAN 这种没有服务质量保证的网络的视频会议系统规范。98 年 ITU-T H.323 建议第二版本又很快推出，第二版更加完善、全面，称作基于分组包协议的网络上的多媒体通信系统规范。

值得注意的是基于 SIP 协议的可视电话正在逐步得到发展。SIP 是由 IETF 提出来的一个应用控制（信令）协议。正如名字所隐含的----用于发起会话。它可用来创建、修改以及终结多个参与者参加的多媒体会话进程。参与会话的成员可以通过组播方式、单播连网或者两者结合的形式进行通信。

国内厂家从 90 年代初开始进行可视电话的研发工作，这两年有更多的厂家加入视讯产品的研制工作。之所以有这么多厂商被吸引到这个领域是因为可视电话有着巨大的市场潜力。资料显示我国电话装机总量如今已位居世界第一。目前，我国电话用户已达 2.5 亿，而且每年正以接近 2000 万户的速度增长，未来 5 年将达到 4 亿户，其中城市电话普及率达 80%以上。以 1.2 亿城市电话用户计，即使有 10%先富起来的人购买可视电话，就是 1200 万部，市场规模即可达到 50 亿元左右。可视电话主要有两个市场：面向中小企业和家庭的个人可视系统和面向集团用户的高清综合电话会议系统。到目前为止，视讯的用户基本是面对企业，然而可视电话整个行业的发展光靠企业用户是远远不够的，只有可视电话进入家庭才能带来行业的爆炸式增长。

目前可视电话的应用还处于集团用户方面，家庭应用处于萌芽阶段，主要是两个方面的原因：销售渠道问题和终端价格问题。随着联通“宝视通”业务的开通和网通、铁通等运营商介入视频通讯市场，以及单芯片解决方案的出现，制约可视电话进入家庭的两方面因素将会解决，可视电话的爆炸式增长将指日可待。

2. 可视电话的结构与核心技术

可视电话的基本结构包括视频输入/输出单元、视频编解码器、语音输入/输出单元、语音编解码器、延时单元、数据处理单元(可选)、系统控制单元、多媒体数据复用/解复用单元和网络接口单元。

语音和视频压缩技术是可视电话的核心技术。可视电话作为一种消费产品,要想走入寻常百姓家,必须能够提供足够好的语音和视频质量,同时占用的信道带宽要尽量小。语音编码技术和视频编码技术的发展就是围绕着上述两点展开的:在保证压缩后语音和图象质量的同时,尽量提高压缩效率。

语音通信是可视电话最基本的功能。受网络条件的限制,可视电话通常工作在较低码率下。为了适应这种低码率语音应用,ITU-T 推出了 G.72X 系列语音压缩标准。其中 G.723.1、G.728、G.729 和 G.729A,在可视电话中得到了广泛应用。

G.723.1 能够产生两种速率的码流,高速率编码器使用多脉冲最大自然量化(MP-MLQ)算法,低速率编码器使用代数码激励线性预测(ACELP)算法。G.729A 是 G.729 的简化版本,G.729A 算法复杂度与 G.729 相比降低了 50%,语音质量略有降低,两种标准编码后的码流可互相解码。当可视电话与普通电话通信时,采用 G.711 标准。G.711 为 PCM 编码,只对语音信号进行采样和量化,产生 64kbit/s 的码流。G.711 编码后的语音质量高,缺点是占用的带宽也很高。在实际选择语音压缩标准时,要综合考虑带宽、时延、算法复杂度等各种因素。

视频压缩是多媒体应用中的核心技术,ITU-T 推出的低码率视频压缩标准对推动可视电话的发展和实用化起到了重要的促进作用。H.261 是 ITU-T 推出的第一个低码率视频压缩标准,码率为 $p \times 64\text{kbit/s}$,其中 $p=1 \sim 30$,图像格式为 CIF 和 QCIF。H.261 压缩编码算法的基本思想是利用预测编码减少时间冗余度,利用变换编码减少空间冗余度。算法主要由运动估计、运动补偿、DCT 变换、量化和霍夫曼编码构成。每帧图像分成图像层、宏块组(GOB)层、宏块(MB)层、块(Block)层共 4 个层次来处理,分为 I 帧和 P 帧。后来推出的 H.263、H.264 标准继承了 H.261 的基本思想,在 H.261 的基础上提出了一些改进。

与 H.261 相比,H.263 在以下几个方面做出了改进:更多的图像格式、半像素运动估计、不同的 GOB 结构、四个可选模式、减少的头信息开销、采用不同的 VLC 表等。在相同的图像质量下,因为 H.263 在运动估计及编码方面的改进,H.263 编码后的码率大约比 H.261 低 30%。为进一步提高 H.263 的编码效率和抗误码性能,ITU-T 在 H.263 的基础上,增加了一些选项,修改后的版本被称之为 H.263+、H.263++。目前,H.263 是可视电话中应用的最广泛的视频压缩标准。

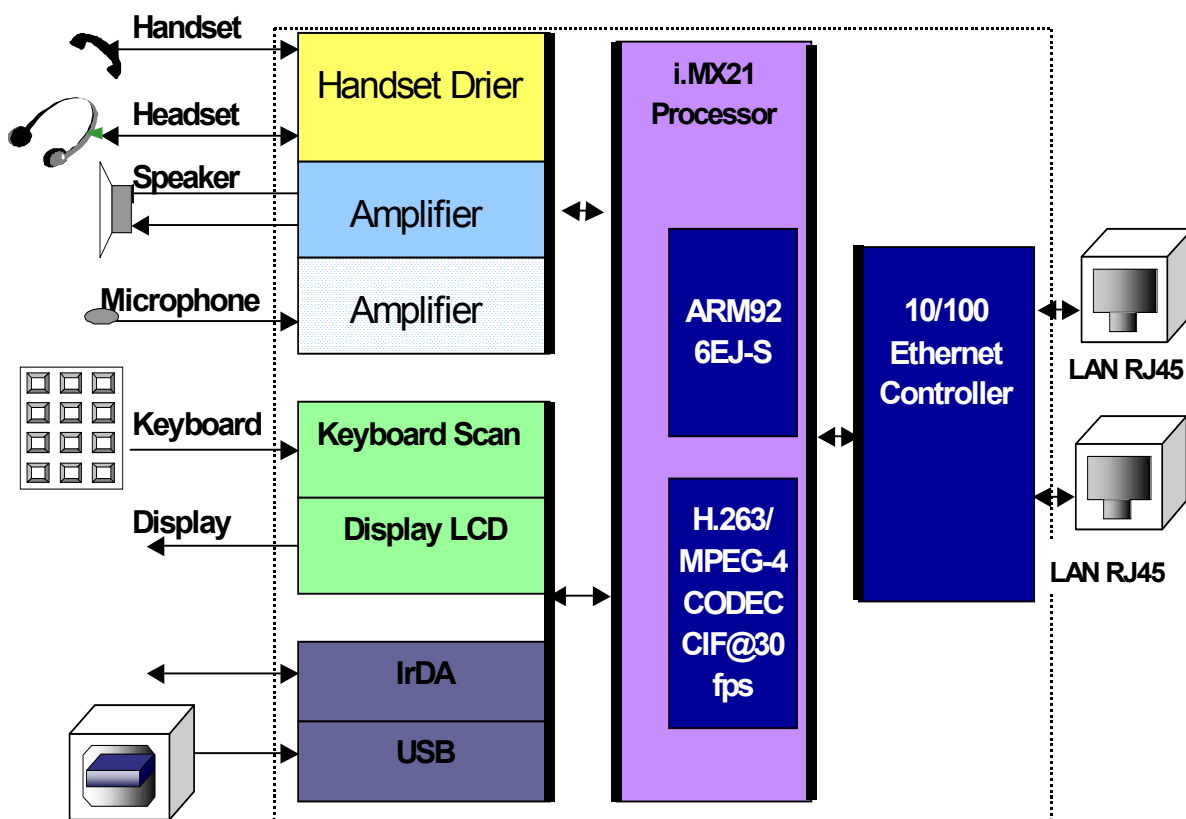
2003 年,ITU-T 通过了一个新的视频编码标准,即 H.264 标准。H.264 与 H.263 相比具有灵活的宏块和块的分割方式,运动估计精度进一步提高,可采用 1/4 或 1/8 像素精度的运动估计。H.261 和 H.263 采用的是 DCT 变换,而 H.264 采用的是类似于 DCT 的整数变换。在相同的重建图像质量下,H.264 编码后的码率比 H.263 低 50%。H.264 在提高编码效率的同时,计算复杂度也大大增加。据估计,编码的计算复杂度大约相当于 H.263 的三倍,解码复杂度大约相当于 H.263 的两倍。

ITU-T 推出的 H.32X 系列标准,具有相同的系统框架。不同之处在于面向的网络不同,因此具有不同的网络接口,不同的信令过程,以及为适应不同的网络而优化设计的包结构。复用协议规定了视频数据、语音数据等的打包标准,而控制协议的作用是在终端之间协商通信方式,如视频编码标准的协商,语音编码标准的协商,信道带宽的协商等。

3. 基于 i.MX21 的可视电话解决方案

长期以来，可视电话由于受芯片技术的制约价格居高不下，而单芯片解决方案的出现是解决该问题的关键。i.MX21 芯片售价仅\$17，而且外设接口非常丰富，单芯片即可完成音视频的编解码、LCD 显示、键盘等功能，相比 DSP 方案大大降低 BOM 成本。

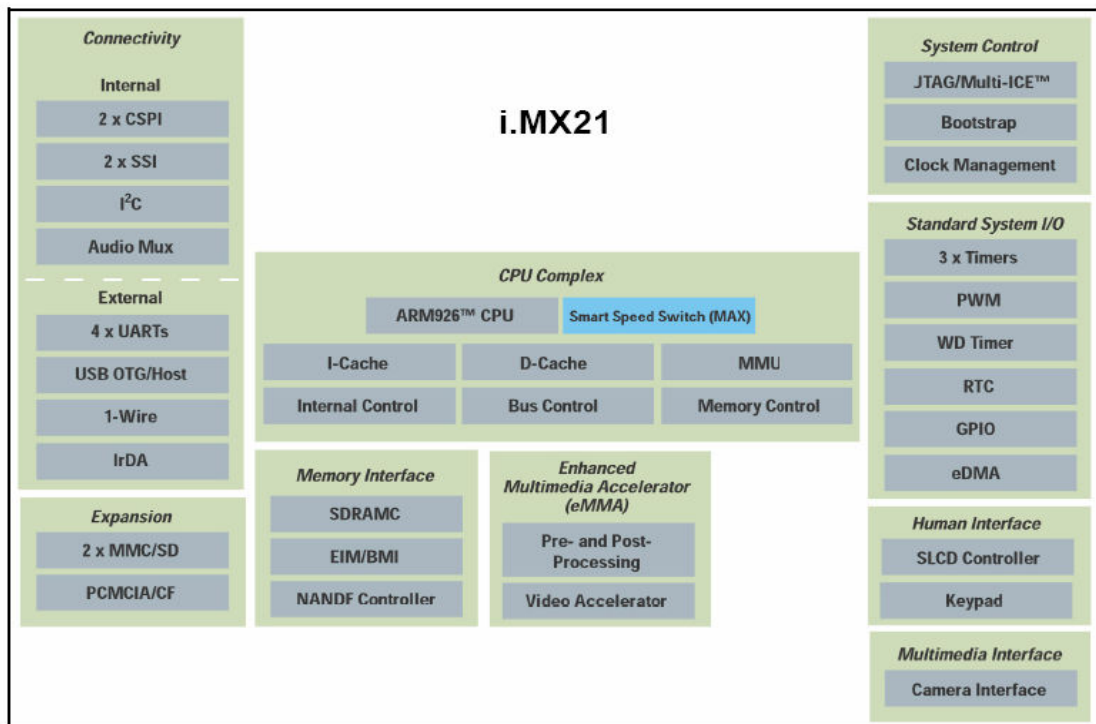
基于 i.MX21 的单芯片可视电话解决方案硬件结构框图如下图所示。



i.MX21 是 Freescale 公司推出的最新一代多媒体处理器，其主频最高可达 400MHz，主要特点如下：

- ARM926EJ-S 内核
- 支持 ARM Jazelle 技术用于 JAVA 加速
- 支持 USB-OTG 功能
- H.263 硬件编解码，CIF@30fps
- MPEG-4 硬件编解码，CIF@30fps

i.MX21 处理器结构框图如下图所示。

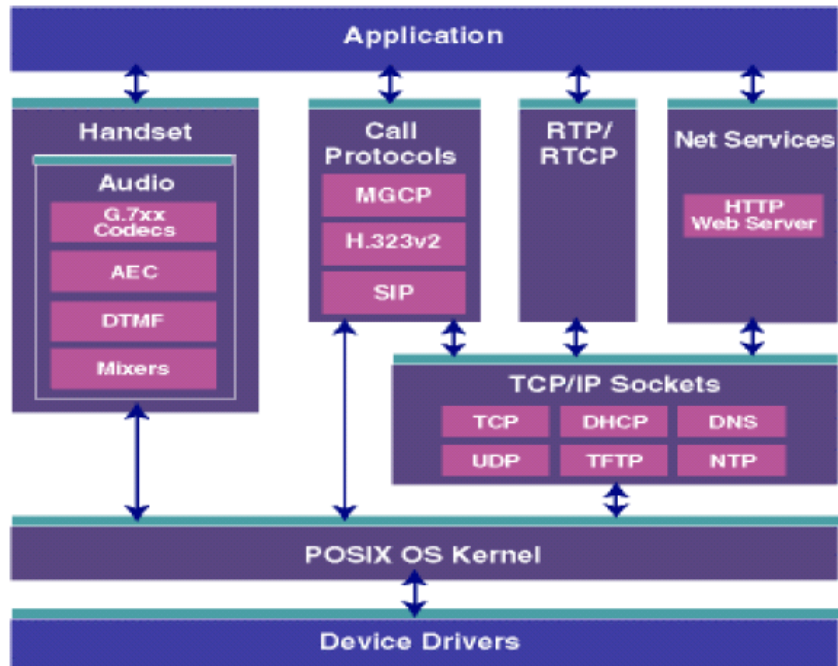


除以上特征 i.MX21 还包括以下外设接口：

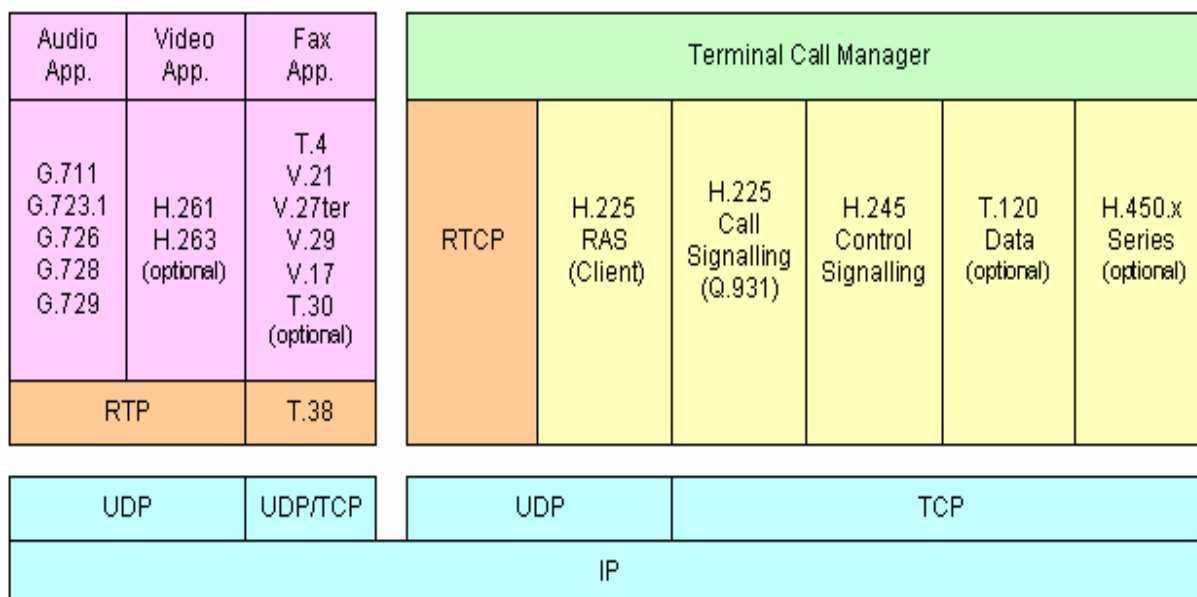
- 支持 BootStrap 调试模式
- 低功耗，支持 3VDC 和 1.8VDC 供电
- PCMCIA/CF 接口，支持 TrueIDE 模式
- NAND Flash 接口
- SDRAM 控制器
- EIM 外部 32 位总线接口
- MMC/SD 接口
- 4 个 UART 接口，其中一个支持 IrDA 接口
- MultiICE JTAG 调试接口
- 一个 USB OTG 2.0 接口，可以配置为主接口和从接口
- 2 个 USB2.0 主接口
- 2 路 IIS 接口，支持 AC97 音频接口
- RGB666 LCD 接口
- 智能 LCD 接口
- 1wire 接口
- 2 个 SPI 接口
- GPIO 接口
- 8x8 键盘接口
- 看门狗电路
- RTC 实时钟
- 3 个通用定时器
- PWM 控制器
- I2C 接口
- 随机数生成加速器

- 安全启动控制器
- 16 通道 DMA 控制器
- CMOS 图像传感器接口
- 内嵌 MPEG-4/H.263 编解码器，同时编解码速度为 CIF@30fps

从以上介绍可以看出 i.MX21 单芯片完全可以满足可视电话硬件要求。基于 i.MX21 的单芯片可视电话解决方案的软件框图如下图所示。

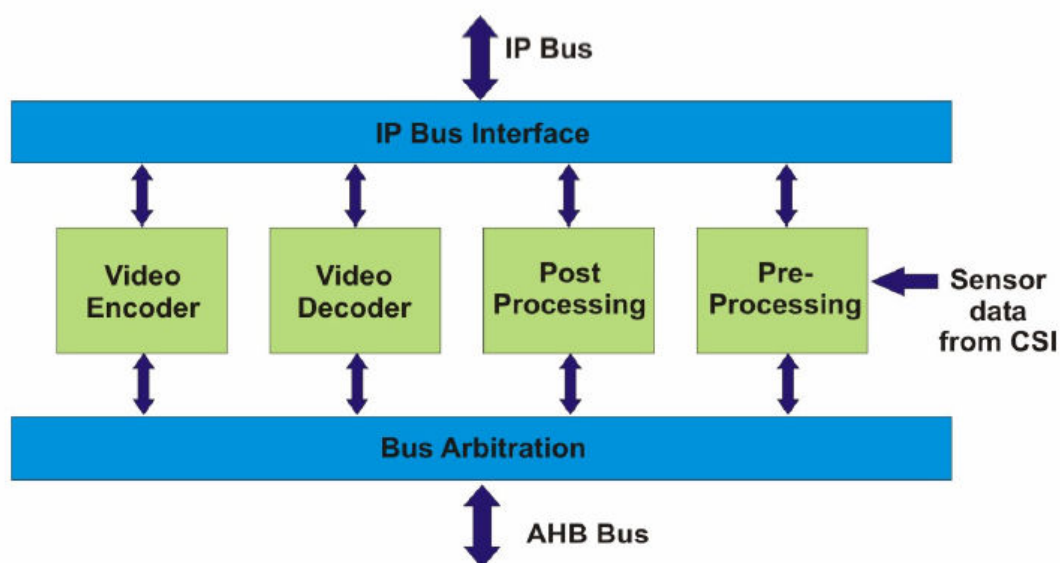


可视电话软件可以分为应用程序模块、协议控制模块、音频编解码模块、视频编解码模块等几部分。应用程序模块完成终端的用户人机界面功能及收发 E-mail、上网等增值功能，协议控制模块负责信令交互等功能，目前可视电话主要是基于 H.323 协议族，但基于 SIP 的框架逐渐得到应用，下图所示是基于 H.323 协议集的软件框架。



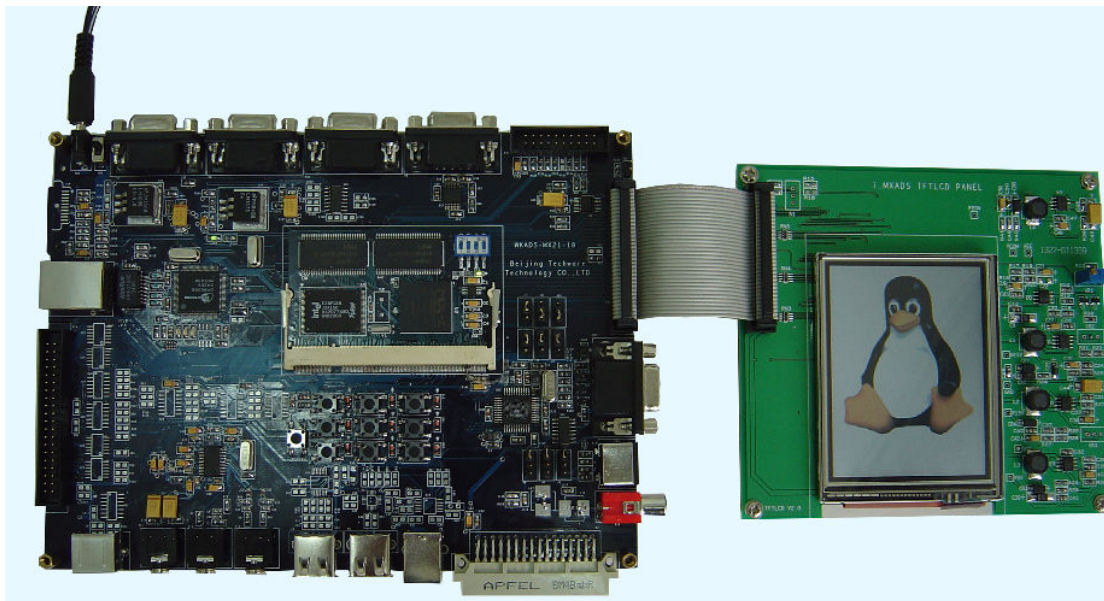
H.323 Terminal

视频编解码是可视电话的核心技术，需要占用大量的 CPU 时间。以一个无线可视电话终端为例，如果视频图像采用 QCIF @15fps 格式，终端需要同时进行 MPEG-4 编码、MPEG-4 解码、MPEG-4 后处理、图像采集、MPEG-4 预处理、音频编码、音频解码、回声消除等功能的处理，再加上操作系统本身的开销，在 ARM926 处理器上此类任务需要 400MHz 左右的带宽。如果利用 i.MX 21 的多媒体加速器模块，则可以降到 160MHz 左右。当然如果进行 CIF@30fps 的传输，处理器负担也不需要增加太多负荷，因为视频编解码是由硬件实现的，不占 CPU 开销。I.MX21 的多媒体加速器模块包括预处理、编码、解码、后处理等功能，同时还包括从 CMOS 图像传感器接口到 LCD 的专用 DMA，以实现本地图像的显示。i.MX21 的多媒体加速器模块的框图如下图所示。



4. 利用 WKADS-MX21-18 加速产品开发进程

针对目前的可视电话市场，亿维东方推出了一套可视电话开发平台 WKADS-MX21-18。该平台基于 Freescale i.MX21 处理器，包括实现可视电话所需要的全部硬件功能，并且提供了完善的嵌入式 Linux 操作系统和驱动程序，方便用户二次开发，用户可以在我们的开发板上直接进行应用程序的开发，而不需要把大量时间浪费在平台的搭建上。而且该平台可以提供硬件的参考设计，大大加速用户产品的上市时间，缩短研发周期。



WKADS-MX21-18 开发平台包括以下配置：

- MC9328MX21 龙珠处理器，基于 ARM926EJ-S 核，主频最高可达 400MHz，是龙珠系列最新一代基于 ARM 核的产品。
- 供电：9VDC, 500mA
- 两个时钟源：32kHz 和 16MHz
- 一片 16MBytes 16-bit Intel StrataFlash
- 两片 32Mbytes 16-bit SDRAM，共 64Mbytes
- 4 个 DB9 插座，支持四个 RS232 接口
- MultiICE JTAG 调试接口
- 10/100M 自适应以太网控制器，RJ45 插座
- 一个 USB OTG 接口，可以配置为主接口或从接口
- 一个 USB 2.0 主接口
- 一个耳机音频输出接口
- 一个麦克风音频输入接口
- 一个 Line 音频输入接口
- LCD 接口，支持 Sharp HR-TFT 260k 真彩液晶
- 串行模拟前端芯片，支持触摸屏输入
- 一个 CRT 显示器接口
- 一个 S-Video 视频输出接口

- 一个 CVBS 视频输出接口，可以输出至电视机
- 一个 PS2 键盘接口
- 一个 1wire 接口
- 一个 SD/MMC 卡接口
- 2.5"IDE 硬盘接口
- CMOS 图像传感器接口
- 系统功能设置拨码开关
- 电源、网络 and 通用 I/O 指示灯
- 3x3 键盘
- 一个红外接口
- 核心板+底板结构，方便用户二次开发
- 尺寸：核心板 47.6mm x 41mm，底板 150mm x 200mm

5. 结论

可视电话可以满足人们对音视频的需求，具有非常好的市场前景。然而由于各种因素的制约，目前可视电话还主要是企业用户，远远没有进入家庭，这中间主要是运营商基础服务和终端价格的问题。然而我们可以看到可视电话的应用环境正在得到改善。

随着中国联通第一期容量达 10 万户的可视电话网络测试完毕，中国第一个可视电话商用网就将正式投入使用。这意味着，可视电话这一原本被利用为集团用户间电话会议的技术手段，今后将逐步降低门槛走向更多的寻常百姓家。中国网通也已经推出“宽带视讯”产品，这是基于网通光纤宽带网资源而提供的一项应用服务，利用数字技术同步传输视频和语音信号，实现一边看图像一边通话的效果。

另外随着 i.MX21 等 SOC 的出现，单芯片解决方案逐渐得到推广，BOM 成本有望控制在千元左右，可视电话终端价格会逐步回落，最终进入家庭，这是大势所趋。