

文章编号:1004-9037(2012)06-0665-05

# 基于TMS320DM6446 EVM的多摄像机跟踪系统

何 迪 彭弘铭 李 凯 员鹏飞

(西安电子科技大学通信工程学院,西安,710071)

**摘要:**本系统利用三帧差法识别运动目标,质心法标定运动目标的质心,利用双摄像机交叉视野的几何约束,通过求解单应矩阵实现两个摄像机之间的目标匹配,实现运动目标的连续跟踪。和单摄像机跟踪相比,双摄像机跟踪的视野范围大,可以对目标持续跟踪,在一定程度上解决目标遮挡、场景混乱等问题。经过优化系统每秒可以处理30帧图像。跟踪的检测准确率达到92.1%,取得了较好的跟踪效果。

**关键词:**多摄像机;嵌入式;目标检测与跟踪

**中图分类号:**TN911.73

**文献标识码:**A

## Multi-camera Tracking System Based on TMS320DM6446 EVM

*He Di, Peng Hongming, Li Kai, Yuan Pengfei*

(School of Telecommunications Engineering, Xidian University, Xi'an, 710071, China)

**Abstract:** The system uses the three-image difference method to identify moving targets, and the centroid method to demarcate the target centroid. It achieves the goal of target matching between two cameras by solving the homography matrix based on the geometric constraints of the double cross-camera vision, thus continuously tracking and monitoring moving goals. Compared with the single camera tracking, the two-camera tracking with a larger view can track targets continuously by solving the problems of temporal occlusion and scene confusion to some extent. After optimization, the system can deal with 30 frames image per second. Experimental results show that the detection accuracy rate of object tracking achieves 92.1% and it obtains a good tracking performance.

**Key words:** multi-camera; embedded; object identification and tracking

## 引 言

传统的跟踪一般使用单摄像机,而单摄像机跟踪不能很好地实现视野范围广、目标遮挡、时空跟踪连续性等场景监控的目的。多摄像机跟踪可以有效解决上述问题,通过多个摄像机之间的相互协同合作,扩大监控系统的检测范围,增大观测角度,持续跟踪目标,有效提高了目标检测、跟踪的准确性和连贯性。摄像机与高性能的嵌入式芯片相结合就可以在监控的最前端对视频信息自动分析处理,然后将有价值的信息通过网络系统传输到监控中心或传输到其他摄像机前端进行信息交换,实现多个摄像机之间的协同跟踪<sup>[1-2]</sup>。使用多摄像机跟踪系

统时首先需要解决的是在不同的摄像机中观察到运动目标是否是同一目标的问题。因为同一个目标会在不同摄像机中产生不同的观测值,需要让这些对应同一目标的不同观测值相互联系匹配起来,所以多摄像机之间的目标匹配是多摄像机协同工作的关键<sup>[3]</sup>。

两个摄像机组成的双目系统是现在多摄像跟踪系统研究的热点,双目跟踪技术一般用在机器人仿生学中,利用双摄像头模拟人眼的功能,实现目标物体的三维重建<sup>[4]</sup>。或者是利用双目摄像机进行目标的识别跟踪,使得机器人可以躲避障碍物接近跟踪的目标。

本设计采用两个摄像头,分别连接一块TMS320DM6446EVM,两个TMS320DM6446-

EVM 之间通过网线连接在路由器上, EVM 开发板可以对连接自己的摄像头采集的视频数据直接处理, 然后将处理完的数据, 通过网络传输给另一块 EVM 开发板进行信息交互, 利用单应约束方法进行目标匹配, 以克服传统单摄像机跟踪的缺点。

## 1 算法介绍

### 1.1 帧差法

帧差法<sup>[5]</sup>是利用相邻帧图像之间的相关性, 通过两帧或多帧之间的差分, 获取短时间内图像中变化比较剧烈的像素点。最简单的差分算法是采用相邻两帧进行差分, 公式为

$$\text{Mask} = \begin{cases} 1 & |I_{(x,y)} - B_{(x,y)}| > T_d \\ 0 & \text{其他} \end{cases} \quad (1)$$

式中: Mask 为差值后的二值化图像;  $I_{(x,y)}$  表示像素点  $(x,y)$  在  $t$  时刻的灰度值;  $B_{(x,y)}$  表示相同像素点在  $t-1$  时刻, 也就是上一帧的灰度值。取差值后像素值变化大于阈值的判定为前景图像, 变化小于阈值的像素点判定为背景图像。  $T_d$  代表了阈值, 它的大小决定了运动检测的灵敏性。

帧差法优点是计算量小, 运算速度快, 可以满足跟踪系统实时性的要求。缺点是物体的运动速度不能太快, 如果速度过快, 那么物体在两帧之间的位移就会变的很大, 造成的结果是出现双目标, 会造成物体内部的“空洞”。

为了改进两帧差法, 研究了三帧差分检测算法, 它是将相邻的三帧图像做差分运算, 选取三帧连续的图像, 分别两个为一组进行差分

$$\text{Mask}_1 = \begin{cases} 1 & |I_{(x,y)} - B_{(x,y)}| > T_d \\ 0 & \text{其他} \end{cases} \quad (2)$$

$$\text{Mask}_2 = \begin{cases} 1 & |I_{(x,y)} - B_{(x,y)}| > T_d \\ 0 & \text{其他} \end{cases} \quad (3)$$

得出两幅二值化图像, 将两幅二值图像对应的像素点相“与”, 得到最终的二值图像。

$$\text{Mask} = \text{Mask}_1 \& \text{Mask}_2 \quad (4)$$

与二帧差分法相比, 三帧差分可以有效地解决运动目标的遮挡和背景重现问题, 并且在一定的程度上抑制了光照、阴影以及噪声的影响, 空洞现象有所改善, 本文采用这种方法进行运动目标检测。

### 1.2 多摄像机目标跟踪

在多摄像机协同问题中, 可以利用的约束条件可以分为

(1) 利用目标自身的特性作为约束条件, 提取目标的颜色、外形、边缘等特征。但由于光线、视角

不同, 这些特征是不可靠的, 同一目标在不同摄像机中可能颜色形状等观测值都不相同。

(2) 利用目标的位置信息作为约束条件, 提取目标进入或离开场景大致位置的分布信息, 以及目标穿过摄像机视野的大概时间等, 这类约束可靠性差, 随机性比较强, 一般需要通过大量观测数据进行统计学习来确定, 一般用于摄像机场景不重叠的情况。

(3) 几何约束<sup>[3]</sup>, 在摄像机场景重叠的情况下, 几何约束依靠摄影几何约束得到对极线, 如图 1 所示。采用对极线约束或者单应约束进行对应点的匹配。这类约束条件是最有效最可靠的, 可以用数学矩阵准确地描述摄像机间匹配情况。本文采用几何约束中的单应约束来完成摄像机之间的匹配。

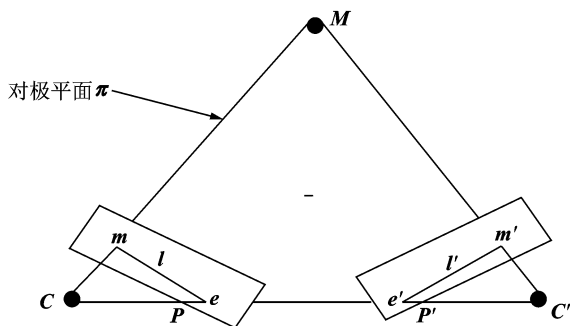


图 1 视觉系统

#### 1.2.1 单应约束

有两个摄像机共同组成一个视觉系统, 如图 1 所示,  $P$  和  $P'$  分别为两个摄像机显示的场面画面,  $C$  和  $C'$  是两个摄像机的中心,  $M$  为场景中的任意一点, 分别连接  $CM$  和  $C'M$ , 那么这两条直线与两个场景画面的产生交点  $m$  和  $m'$ , 这两个交点就是  $M$  点在两个摄像机画面中的投影点(像点)。再连  $C$  和  $C'$  点, 这条连接两个摄像机中心点直线称为基线, 基线与两个画面的交点为  $e$  和  $e'$ , 这两个点称为基点。两个摄像机的中点  $C, C'$  和  $M$  点组成了一个平面  $\pi$ , 称为对极面。连接  $m$  点和  $e$  点组成一条直线, 这条直线在另外一个画面中的投影就是  $m'$  和  $e'$  两点确定的直线  $l'$ , 这两条直线称之为对极线。

从图中可以看出,  $m$  点位于平面  $\pi$  和场景画面  $P$  的相交线  $l$  上, 无论  $X$  点怎么移动,  $m$  点始终是在  $l$  线上的。同理  $m'$  点始终也是位于  $l'$  直线上的。因此当  $m$  点确定, 寻找它的对应点  $m'$  时, 只需要在极线  $l'$  上寻找即可。这就是重要的几何约束, 因为可以把对应点的范围从整个显示画面缩小到极线上。而且所有的对极线都要穿过极点  $e$  和  $e'$ 。

从推导上可以看出, 对极约束是两个摄像机场

景视图之间的射影几何关系,与目标自身的结构无关,只与摄像机本身的参数有关。

可以定义点  $m$  与  $m'$  的映射关系表达式为

$$m'^T F m = 0 \quad (5)$$

式中:  $F$  为一个  $3 \times 3$ , 秩为 2 的齐次矩阵, 称为基本矩阵。设  $m = [u, v, 1]^T$ ,  $m' = [u', v', 1]^T$ , 并将  $F$  的元素组成一个矢量  $f$ , 则式(5)可以表示为

$$(u'u, u'v, u'u', v'u, v'v, v'u', u, v, 1)f = 0 \quad (6)$$

给定存在视野交叉的两幅图像, 那么两幅图像之间存在着对极几何约束, 但是从上面的推导中可以看出, 对极几何约束不是图像点之间的一一对应, 而是图像点和极线之间的对应关系。所以如果要构成空间点之间的对应关系, 需要得到空间点另外的部分信息。从射影几何可以得到, 如果空间点位于同一平面上且该平面不通过二个摄像机的任一光心, 那么两幅图像对应点之间存在一一对应关系, 那么这种一一对应关系可以用单应矩阵<sup>[6]</sup>表示。

令  $x = (x, y, z)^T$ ,  $x' = (x', y', z')^T$  为同一场景点在摄像机两个不同视野坐标系下的坐标,  $m = (x_1, y_1, 1)$ ,  $m' = (x'_1, y'_1, 1)$  为两幅图像的对应点, 由摄像机线性模型可以得到

$$m \cong Kx, m' \cong Kx' \quad (7)$$

其中, “ $\cong$ ”表示相差一个非零标量因子,  $K$  为摄像机内参数矩阵。

令  $\pi$  为一个平面, 其规范法向量为  $n$ , 则对所有空间点  $x \in \pi$ , 有  $n^T x = 1$ , 于是可以推导出

$$x' = Rx + t = (R + tn^T)x \quad (8)$$

从而,  $m' \cong Hm$ , 其中,  $H = K(R + tn^T)K^{-1}$ , 是  $3 \times 3$  阶的矩阵;  $R$  为世界坐标相对于摄像机坐标的  $3 \times 3$  阶旋转矩阵;  $t$  为世界坐标相对于摄像机坐标的平移大小, 是一个三维向量。上式还可以写为

$$m' = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} m \quad (9)$$

从单应矩阵  $H$  的表达式可以看出, 单应关系包含了摄像机内部参数和外部参数, 可以描述出不同摄像机坐标系之间的旋转、平移关系, 可以实现两个平面之间的二维映射。

### 1.2.2 单应矩阵的求解

将单应矩阵写成  $9 \times 1$  的向量形式  $h = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, 1)$ , 单应矩阵有 8 个未知数, 每一对对应点  $m$  和  $m'$  可以得到

$$\begin{cases} x_1 h_{11} + y_1 h_{12} + h_{13} - x'_1 x_1 h_{31} - x'_1 y_1 h_{32} = x'_1 \\ x_1 h_{21} + y_1 h_{22} + h_{23} - y'_1 x_1 h_{31} - y'_1 y_1 h_{32} = y'_1 \end{cases} \quad (10)$$

式(10)一共有 8 个未知数, 所以至少需要 4 对对应点来求解映射矩阵, 假设知道  $n$  组对应点  $(x_1, y_1)$ ,  $(x'_1, y'_1)$ ,  $\dots$ ,  $(x_n, y_n)$ ,  $(x'_n, y'_n)$ , 代入式(10)后如式(11)所示

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1 x_1 & -x'_1 y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1 x_1 & -y'_1 y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x'_2 x_2 & -x'_2 y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y'_2 x_2 & -y'_2 y_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_n & y_n & 1 & 0 & 0 & 0 & -x'_n x_n & -x'_n y_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -y'_n x_n & -y'_n y_n \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = \begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ \dots \\ \dots \\ \dots \\ x'_n \\ y'_n \end{bmatrix} \quad (11)$$

$$H = (A^T A)^{-1} A^T b \quad (12)$$

当  $n$  大于 4 时, 无法直接采取求逆矩阵的方法得到单应矩阵, 可以采用奇异值分解(Singular value decomposition, SVD)。若  $A$  为  $m \times n$  阵, 对  $A$  进行 SVD 分解得  $A = U * S * V^T$ , 则  $U$  为  $m \times n$  阵,  $V$  为  $n \times n$  阵。奇异值在  $S$  的对角线上, 非负且按降序排列。对  $A$  求逆只是简单的求倒数即可。

## 2 算法在 DSP 上的实现

### 2.1 DSP 平台的选择

本方案选用 TMS320DM6446 EVM 作为开发平台。TMS320DM6446 数字媒体处理器是基于 ARM926 处理器与 TMS320C64x+DSP 内核的高集成度 SoC。该芯片包含 ARM 子系统、DSP 子系统、视频处理子系统、系统控制模块; 还有外部存储接口、电源管理、外围控制模块和交换中心资源等模块<sup>[7]</sup>。

TMS320DM6446 EVM 包含: 256MB 的 DDR2 DRAM; 1 个复合视频输入接口; 4 个 DAC 输出接口; 音频输入接口 1 个, 双通道输出接口 2 个; S-video 输入输出各 1 个接口; USB2.0 接口 1 个; 10/100 MB 网口 1 个; 9 针串口 1 个; 媒体卡接口(SD 卡, xD 卡, SM 卡, MS 卡, MMC 卡); 16 MB 的 NOR Flash, 64MB 的 NAND Flash, 4MB 的 SRAM 等。

### 2.2 算法在 DSP 上的流程与优化方案

#### 2.2.1 实现流程

两块 TMS320DM6446 开发板通过套接字 Socket 线程<sup>[8]</sup>, 可以发送目标的质心坐标。事先计算出两个摄像头交叉视野的单应矩阵, 当摄像头 1 内有运动目标出现时, 摄像单元 1 可以发送目标质

心坐标给摄像单元2,摄像单元2计算出摄像头1中目标质心在摄像头2中的位置坐标,这样就可以完成运动目标在两个摄像机间的匹配。当目标走出摄像头1后,摄像单元1无法发送质心坐标,这时摄像单元2判定出目标离开摄像头1完全进入摄像头2的视野,依照上一帧计算出的匹配质心位置继续跟踪目标标定目标的质心。软件流程图如图2,3所示。

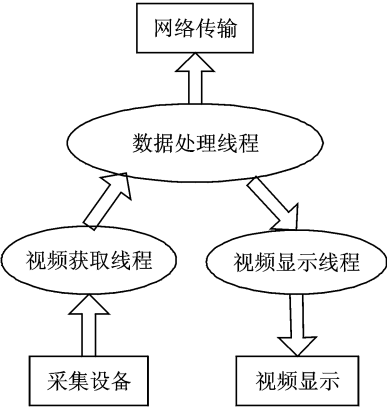


图2 ARM 软件流程图

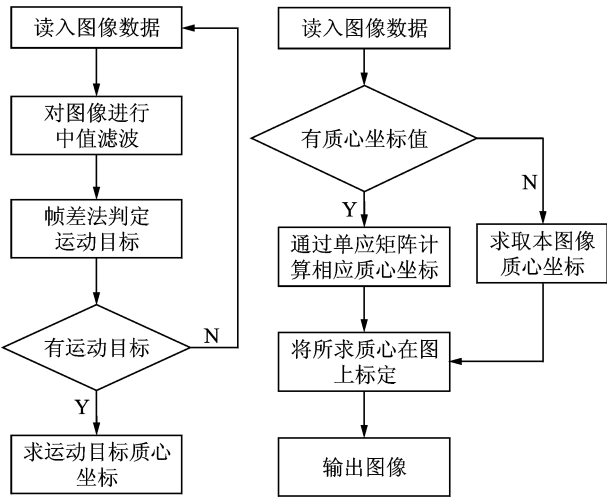


图3 DSP 软件流程图

2.2.2 优化方案

使用互斥锁完成共享内存间的同步问题,避免线程间的竞争以及帧上下不一致的问题。减少循环的次数,尽量在少的循环内完成多个的操作。使用静态变量解决堆栈溢出的问题。数组变量尽量改为指针变量,减少寻址时间。编译器选项优化:命令-mv6400,使代码利用DSP的附加硬件和指令执行任务;命令-pm 把所有的C语言文件合成一个文件,是程序级别的优化;命令-o3 最大限度优化循环,是文件级别的优化。

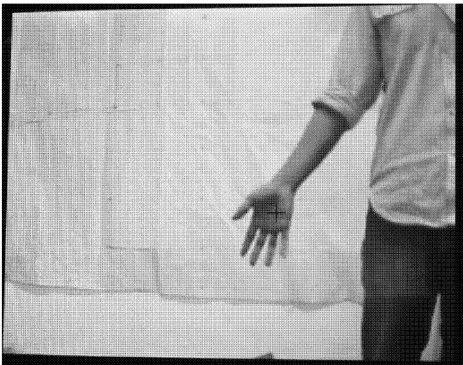
对于算法匹配值误差大,研究后发现原来算法中是有一个常数系数的,所以坐标值大的点,计算

出的匹配点误差也大。需要将计算出的匹配点第三维坐标值与1相除得出比例系数,然后将比例系数乘以匹配点前两维坐标值即可缩小误差。改进算法后,经过大量实验统计误差值在4个像素值以内。

3 性能和结果

选取一面有标定点的墙面作为场景模型,在墙壁对面放置两个摄像机 $C_1, C_2$ 。输入视频分辨率为 $640 \times 480$ 。两个画面各自选取5个对应点(图4和图5中黑色点)作为映射矩阵求解单应矩阵参数。当摄像机 $C_1$ 中出现目标之后,采用三帧差法对目标进行检测,质心法标注目标的质心坐标,通过套接字将 $C_1$ 质心坐标信息传输给 $C_2$ 摄像机,通过单应矩阵计算出 $C_1$ 目标质心所对应的 $C_2$ 中目标的位置,从而实现双摄像机的匹配。当目标走出摄像机 $C_1$ 视野范围之后, $C_2$ 可以持续地跟踪目标。

图4~6为大量实验中的3组典型实验结果,在摄像机 $C_1$ 检测出运动目标质心(黑色十字)后,相应的摄像机 $C_2$ 利用已求解的单应矩阵计算出 $C_1$ 目标在 $C_2$ 图中的相应位置(白色十字),从而完成同一目标在两个摄像机中的匹配工作。从实验结果可以看出,实验都具有较高的检测精度。根据实验数据计算跟踪的检测准确率为92.1%。系统每秒可以处理30帧图像,达到了较好的跟踪效果。

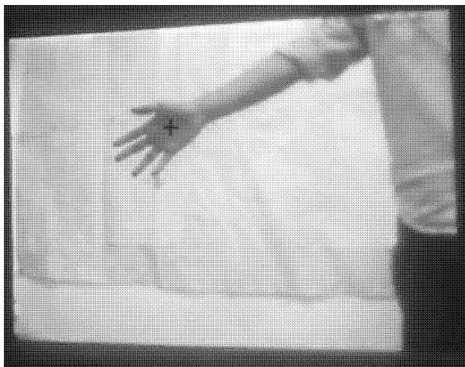


(a)  $C_1$ 的跟踪结果图



(b)  $C_2$ 的跟踪结果图

图4 实验结果1

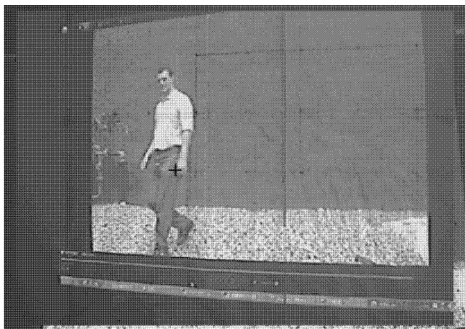


(a)  $C_1$ 的跟踪结果图

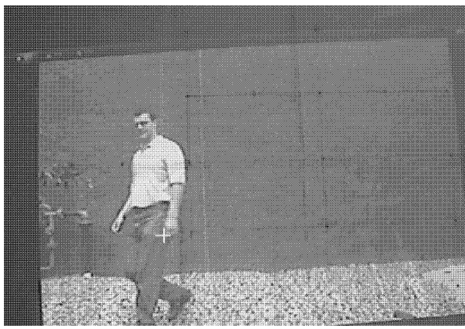


(b)  $C_2$ 的跟踪结果图

图5 实验结果2



(a)  $C_1$ 的跟踪结果图



(b)  $C_2$ 的跟踪结果图

图6 实验结果3

## 4 结束语

本文利用双摄像头、双嵌入TMS320DM6446 EVM 式开发板对场景行人进行持续监控,开发板之间通过网络通信传送数据。可以在双摄像头之间

连续地跟踪一个运动目标,达到预期目标。相比单摄像头跟踪,双摄像头跟踪的视野大,可以对目标持续跟踪,在一定程度上可以解决目标遮挡、场景混乱等问题。而且多摄像机系统还可以组成智能跟踪系统,将多个摄像机之间获取的目标特征进行融合,对目标行为和特征进行分析理解。所以多摄像机跟踪系统在民用和军用中都有着极大的应用前景和研究价值。

### 参考文献:

- [1] 胡伏原,张艳宁,李映,等.多摄像头协同感知系统的设计与实现[J].中国图像图形学报,2006,11(12):1849-1853.  
Hu Fuyuan, Zhang Yanning, Li Ying, et al. Design and implementation of multi-sensor cooperative perception system[J]. Journal of Image and Graphics, 2006, 11(12): 1849-1853.
- [2] 陈炳文,王文伟,秦前清.多摄像机目标跟踪系统综述[J].微型机与应用.2010,29(17):1-3.  
Chen Bingwen, Wang Wenwei, Qin Qianqing. Overview of target target tracking system based on multi-camera [J]. Microcomputer & Its Applications, 2010, 29(17): 1-3.
- [3] 梁华.多摄像机视频监控中运动目标检测与跟踪[D].国防科技大学电子科学与工程学院,2009:4.  
Liang Hua. Moving target detection and tracking in multi-camera video surveillance[D]. School of Electronic Science and Engineering, National University of Defense Technology, 2009:4.
- [4] 侯志强,韩崇昭.视觉跟踪技术综述[J].自动化学报,2006,32(4):603-617.  
Hou Zhiqiang, Han Chongzhao. A survey of visual tracking[J]. Acta Automatica Sinica, 2006, 32(4): 603-617.
- [5] Collins R T, Lipton A J, Kanade T, et al. A system for video surveillance and monitoring: VSAM final report[R]. CMU • RI- 11 000 • 12, Pittsburgh, PA: Carnegie Mellon University, 2000.
- [6] Velipasalar Wolf. Multiple object tracking and occlusion handling by information exchange between uncalibrated cameras[C]//IEEE International Conference on Processing. Italy: IEEE, 2005: 418-421.
- [7] Texas Instruments. TMS320C64x technical overview [EB/OL]. 2011-10-16. <http://www.ti.com>.
- [8] Texas Instruments. Codec engine application developer user guide [EB/OL]. 2009-08-06. <http://www.ti.com>.

作者简介:何迪(1980-),男,讲师,研究方向:图像信号处理,E-mail: penghm307@163.com;彭弘铭(1987-),男,硕士,研究方向:图像信号处理;李凯(1986-),男,硕士,研究方向:图像信号处理;员鹏飞(1987-),男,硕士,研究方向:图像信号处理。