基于达芬奇平台的智能视频汽车安全驾驶系统①

Intelligent Video System in Safe Driving Based on Davinci Platform

叶 林 陈岳林 (桂林电子科技大学 机电工程学院 广西 桂林 541004)

摘 要：提出用德州仪器生产达芬奇开发板做为硬件平台来使用，同时结合相应的算法。在某些特定的条件下如汽车、行人、车道等特点目标的跟踪识别。实现该技术，并实现移动性和实时性，满足辅助汽车安全驾驶的作用。所研究的智能视频系统所涉及研究的内容多，工作量大，很多东西不能在文中详尽描述。提出的观点和方法，可以供相关技术研究和后续开发作参考。

关键词：智能视频分析 达芬奇 DM6446 辅助驾驶 SIFT 算法

1 引言

智能视频技术的研究是当前计算机视觉领域的一个热点，该技术应用范围广泛，涉及军事、航天、交通、监控等领域。国内外许多专门研究机构都在此领域进行了大量研究，并取得一定成果。作为智能视频监控核心技术的智能视频分析技术更是备受关注。目前，美英等国家已经就这方面技术开展了大量相关项目的研究。如 1997 年美国国防高级研究项目署设立的视觉监控重大项目 VSAM(Visual Surveillance and Monitoring)[1]，该项目研究主要用于战场及普通民用场景进行监控的自动视频理解技术；实时视觉监控系统 W4[2]不仅能够定位人和分割出人的身体部分，而且通过建立外观模型来实现多人的跟踪，并可以检测人是否携带物体等简单行为；另外，英国 Reading 大学也针对车辆和行人的跟踪及其交互作用的识别展开了相关研究。智能视频分析技术中的动态视频的跟踪与识别是该项技术的一个难点，到目前为止还没有很好的解决方法；而在视频技术中对物体以及人的行为的预测估计是最具挑战性。

本文所涉及的应用领域是汽车安全驾驶系统，前人在该方面的研究多是集中在对道路边界的识别、路面状况的识别、汽车前面障碍物的判断和识别等上面。由于汽车是快速行驶中的物体，所以通过机器视觉所获得的环境信息多而且复杂，这就存在一个如何对所采集的环境信息进行分类和判断的问题。另外，由于汽车在行驶时造成图像的前景和背景都是运动的，所以所采集的图像数据十分复杂，造成了数据量大，计算的算法复杂且很难实现实时。本文将在下面具体阐述构建在达芬奇平台上的智能视频处理系统。

2 达芬奇开发平台的介绍

达芬奇技术是美国 TI 公司推出的集成了基于 DSP 的处理器系列、软件、工具以及支持开发广泛范围的优化数字视频终端设备的先驱，专门为高效、强大的数字视频量身定制。应用范围很广，如：数码相机、视频安全监控系统、高级医疗影像、便携式视频播放器、网络摄像机，视频服务器、图象处理等。

基于达芬奇结构的独特性和运算控制性能，本文选用达芬奇 DM6446 开发板做为硬件平台进行研究。其独特性在于一块开发板上集成 DSP 和 ARM 两个芯片，区别于其他只含有单个 DSP 芯片或者是单个 ARM 芯片。正是这种双核结构使得此开发平台能满足视频开发应用中对运算和控制能力上的要求，适合各种视频系统的开发。在使用过程中两个芯片有其不同的分工：DSP 芯片主频高，计算速度快，主要进行视频处理中的计算部分，而 ARM 芯片则是用来进行控制的。两个芯片之间的通信问题不需要考虑，两个芯片的协调工作，控制和运算分开使效率更高，从而降低了系

① 基金项目:国家自然科学基金(50375003)
收稿时间:2008-12-16

统的复杂程度，也为开发出创新高效的数字产品提供了条件，也使得芯片能够适应更多应用场合的需要。

而相比于其他开发板达芬奇 DM6446 运算能力的提升主要在于 DSP 子系统中 C64X+新型 DSP 内核的引入。该内核针对高分辨率视频编解码算法进行了很多方面的优化。DSP 部分则采用 594 MHz 主频的 C64X+DSP 核，运算速度快。

达芬奇 DM6446 开发板为我们提供了各种丰富的接口，完全可以满足本系统的开发。用达芬奇开发板可以实现系统的移动性，处理速度快，从而实现实时性。

如下图 1 是达芬奇 DM6446 开发板的功能模块图。

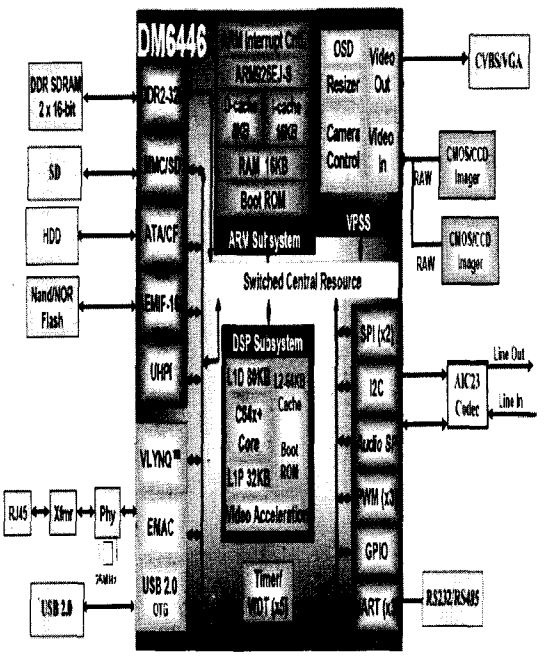

图 1　功能模块框图

DM6446 开发板是本系统的硬件的核心，在开发板的基础上添加上其它的硬件设备就够成为一个完整的视频系统。系统的实现大体过程就是：首先是摄像头数据的采集，其次是根据事先设计好的移植进开发板算法对数据进行处理，最后是通过开发板的运算实现功能和输出结果。

3　系统功能以及算法的分析

由于此智能视频系统是应用于辅助汽车安全驾驶，所以系统必须要对路面的信息进行判断分析。汽车在行驶过程中遇到的状况是比较复杂的，但是如果要对各类信息都进行判断区别也是很困难的。因此我们仅对特定的某些场景进行识别分辨，然后做出判断之后对驾驶员提醒，以实现辅助驾驶的目的。例如对车辆以及行人信息的判断、对红绿灯进行判断、对路上出现的突然状况做出反应。根据不同的天气状况对路面情况也有一定的识别能力，如大雨、大风、雪、雾天等各种恶劣天气。所以此智能视频系统必须有良好的实用性和鲁棒性，同时系统对实时性要求也是比较高的。而文献[3－6]都是类似方向应用的研究。

要增强系统的实时性和鲁棒性，无非是从两个方面着手：① 系统硬件平台的搭建要合理。本文在前面已经介绍，就是使用达芬奇 DM6446 嵌入式开发板做为硬件的平台，可以很好的实现要求。② 进行跟踪识别的算法。算法对系统功能的实现至关重要，一个好的算法对系统效率的提高是远远比硬件性能上提高要来的明显，所以对算法的研究改进非常重要。

本文中所设计的系统主要是在移动场景下对特定目标进行跟踪和识别，目前在此研究领域内相对跟踪而言识别是相对简单的，尤其是在动态场景下的跟踪是很困难，至今还是该领域研究的热点和难点。接下来重点研究移动场景下的跟踪算法，而在本系统中采用简化的 SIFT 算法，而实现系统的其它有关算法会在后续文章中说明。

目前跟踪识别算法比较多，较为成熟的是 SIFT 和 KLT 算法，很多算法对旋转、尺度缩放、亮度变化，对视角变化、仿射变换、噪声干扰等不能保持一定程度的稳定性，鲁棒性不好。

SIFT 算法即 Scale Invariant Feature Transform，即尺度不变特征变换，是由 David G. Lowe 最早在 1999 年提出的。

首先介绍尺度空间的概念，一幅图像的尺度空间被定义为函数 $L(x,y,\sigma)$，它是尺度变化的高斯函数 G 与图像 I 的卷积[7,8]：

$$L(x,y,\sigma)=G(x,y,\sigma)*I(x,y) \tag{1}$$

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \tag{2}$$

其中，$G(x,y,\sigma)$ 是尺度可变高斯函数，(x,y) 是空间坐标，σ 是尺度坐标，L 为尺度空间，σ 大小决定图像

的平滑程度，大尺度对应图像的概貌特征，小尺度对应图像的细节特征。

SIFT 算法主要包括 4 个步骤[9]：

①尺度空间极值检测。首先建立图像的 DOG 金字塔，在 DOG 尺度空间中的 26 个领域中检测极值，$D(x,y,\sigma)$ 是两个相邻尺度图像之差，即：

$$D(x,y,\sigma)=(G(x,y,K\sigma)-G(x,y,\sigma))*I(x,y)=L(x,y,K\sigma) \quad (3)$$

一个点如果在 DOG 尺度空间本层以及上下两层的 26 个领域中是最大或最小值时，就认为该点是图像在该尺度下的一个特征点，如图 2 所示。

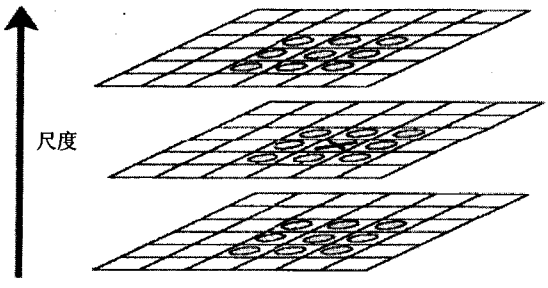

图 2 DOG 尺度空间特征点检测

②利用关键点邻域像素的梯度方向分布特性，为每个关键点指定方向参数，使算子具备旋转不变性。

$$m(x,y)=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2} \quad (4)$$

$$\theta(x,y)=\arctan\{[L(x,y+1)-L(x,y-1)]/[L(x+1,y)-L(x-1,y)]\} \quad (5)$$

公式(4)、(5)为处梯度的模值和方向公式。其中 L 所用的尺度为每个关键点各自所在的尺度。

实际计算时，在关键点为中心的邻域窗口内采样，并用直方图统计邻域像素的梯度方向。梯度直方图的范围是 0～360°，每 10° 代表一个方向，总共 36 个方向。直方图的峰值代表该特征点处邻域梯度的主方向，作为该关键点的方向。图 3 是采用 7 个方向时使用梯度直方图为特征点确定主方向的示例。

③生成 SIFT 特征向量，将坐标轴旋转到特征点方向，以保证旋转不变性。然后，以关键点为中心取 8×8 的窗口。图 5a)的中央黑点为当前关键点的位置，每个小格代表关键点邻域所在尺度空间的一个像素，利用公式(4)、(5) 求得每个像素 (i,j) 的梯度幅值 $m_{i,j}$ 与梯度方向 $\theta_{i,j}$，箭头方向代表该像素的梯度方向，箭

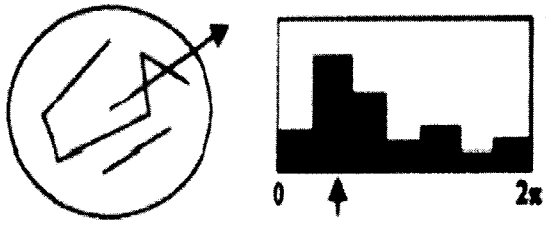

图 3 建立特征点方向

头长度代表梯度模值，然后用高斯窗口对其进行加权运算，每个像素对应一个向量，长度为 $G(\sigma',i,j)*m_{i,j}$，$G(\sigma',i,j)$ 为该像素点的高斯权值，方向为 $\theta_{i,j}$，图中圆圈代表高斯加权的范围，根据文献[4]高斯参数 σ' 取 3 倍特征点所在的尺度，即 3σ。每个向量在 2×2 的子窗口范围内投影到 8 个梯度方向上，绘制每个梯度方向的累加值，即可形成一个 8 维向量，如图 4(b)所示。Lowe 在实际应用中采用 4×4 个小窗口，这样每个特征点就用 128 维向量来表征。最后对特征向量规一化，去除光照影响。

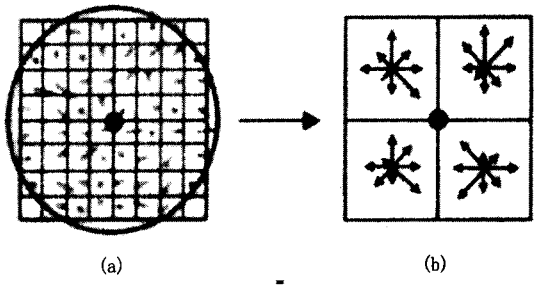

图 4 由特征点邻域梯度信息生成特征向量

④特征匹配。当两幅图像的 SIFT 特征向量生成后，就采用欧式距离作为两幅图像中关键点的相似性判定度量，当此距离小于某个阈值时就认为这两个点已匹配上。

以上是 SIFT 算法的基本原理，此算法已经有代码的实现。而在本文所研究的系统中应用此算法进行跟踪，通过软硬件的结合可以很好的满足要求，同时对算法进行改进能够更快的实现目标，有更好的实时性。改进的一系列算法会在后续文章中介绍。图 5 是用 SIFT 算法做实验的结果。

图 5 用 SIFT 算法的出前后两帧画面的匹配

4 总结

本文在调查研究了目前的汽车辅助驾驶系统的基本状况，了解了在此领域应用中的一些要求和实现难度。从而选择了 TI 公司的达芬奇 DM6446 开发板作为实现系统的硬件对象，同时对达芬奇技术及开发板进行了相应的介绍。本文还在应用领域的可行性和市场应用价值上进行分析阐述，此项目极具市场价值。本文创新点就是使用达芬奇开发板作为系统的硬件支持，DSP 和 ARM 双核结构可以使两个芯片更加协调的工作，容易实现系统要求的实时性，同时还考虑到了成本的问题，在研究中的开发板是研发套件，价格相对贵一些，但是此系统研发成功以后，成本会降低很多。再一个就是使用 SIFT 算法为主要的跟踪算法，移植到开发板上之后执行的效率比较高。困难点不足之处就是对系统实际的功能的实现需要对算法进行大量的研究和改进，对运动物体的运动趋势的预测比较难，这在后续的文章中会相应涉及到。

参考文献

1 Collins RT, A.J.L, Kanade T, et al. A System for Video Surveillance and Monitoring. IEEE Trans. on Computers, 2000,32(1):69 – 72.

2 Haritaoglu S, Harwood D, Davis LS. W4: Real-Time Surveillance of People and Their Activities. IEEE Trans. on Pattern Analysis and Machine Inteelligence, 2000,22(8):58 – 61.

3 方帅,迟健男,徐心和.视频监控中的运动目标跟踪算法.控制与决策, 2005,20:1388 – 1391.

4 胡志强.全景影像之移动物体侦测与追踪系统.台湾:中央大学图书馆, 2007:1 – 62.

5 岳昊,邵春福,赵熠.基于 BP 神经网络的行人和自行车交通识别方法.北京交通大学学报, 2008,32:46 – 49.

6 王臣豪,刘富强,田敏.基于 Blob 的车辆识别及其跟踪算法研究.同济大学学报, 2007,(6):37 – 42.

7 Lowe DG. Object recognition from local scaleinvariant features. International Conference on Computer Vision, Corfu, Greece. 1999,(9):1150 – 1157.

8 Lindeberg T. Scale space theory: A basic tool for analyzing structures at different scales. Journal of Applied Statistics, 1994,21:224 – 270.

9 Lowe DG. Distinctive image features from scaleinvariant keypoints. International Journal of Computer Vision, 2004,60(2):91 – 110.