

基于内容的图像检索在智能监控系统中的应用

石跃祥^{1,2}, B. Benhabib², 蔡自兴³

(1. 湘潭大学信息工程学院, 湘潭 411105; 2. 多伦多大学机械与工业工程系, 加拿大安大略湖 M5S 3G8;

3. 中南大学信息科学与工程学院, 长沙 410082)

摘要: 针对基于内容的图像检索与识别在智能监控中应用的问题, 建立了基于区域划分的颜色链码, 利用图像颜色的区域特征值来构成一个区域与整体的特征描述, 并将它应用到捕获图像的识别。捕获的图像, 通过梯度特征对来定位与检测, 对获取的脸部图像进行链码特征匹配与识别, 并将计算结果作为监控系统的判断与决策依据。运用这一方法的监控系统, 提高了监控系统的辨别对象能力, 拓展了智能监控研究领域。

关键词: 智能监控; 图像检索; 基于内容的图像检索

Application of Content-based Image Retrieval in Intelligent Monitor System

SHI Yuexiang^{1,2}, B. Benhabib², CAI Zixing³

(1. School of Information Engineering, Xiangtan University, Xiangtan 411105; 2. Department of Mechanical and Industrial Engineering, Toronto

University, Ontario, Canada M5S 3G8; 3. School of Information Science and Engineering, Central South University, Changsha 410082)

【Abstract】 This paper presents an application with content-based image retrieval in intelligent monitor and operator. Usually, color is used to describe the image, but not exquisite in detail of image. The chain coding is built by eigenvalue of region description that is partite from image. The grid feature couple is used to locate and detect the human face for the captured image. Extraction the feature and calculating the chain coding, the similarity function can decide the match degree for the processing object. This system enhances the intelligence in how to identify the object when the camera captures its image.

【Key words】 Intelligent monitor; Image retrieval; Content-based image retrieval(CBIR)

随着多媒体技术的发展和图像信息处理技术的完善, 图像信息获取的来源越来越广, 其数据量增长速度极快, 随之而来的是如何把图像检索与识别的技术应用于实践工作中, 文献[1]通过特定颜色目标来实现自动检测跟踪运动目标的方法, 而在医学上的应用有文献[2]通过建立彩色肺癌图像语义的描述模型来实现肺癌的诊断, 公安系统通过成相技术, 在图像数据库中来搜索疑犯特征及信息, 为案侦侦破提供支持决策。通过图像检索与识别技术来解决实际生产和应用工作中的问题, 将是图像处理技术有价值的实际应用。

本系统把图像检索与识别技术应用于监控系统中, 与一般的监控系统最大的区别在于什么对象需要报警操作, 什么对象不需要报警操作。它是在分析了图像特征的基础上, 检索“犯罪特征”/“特征”数据库, 通过分析、判断后再作出对捕获图像的对象采取何种行为处理的系统。并可以通过人机界面或者外接数据接口, 可进行特征数据库的更新和维护, 同时, 还可以自动记录监控对象的图像数据信息, 可以实现监视数据的记录、检索和维护等操作。

1 基于内容的图像链码描述方法

为了能够更好地以颜色来表达图像的特征, 虽然直方图能表达整体特征, 但无法描述局部特征, 因此在描述图像各局部特征时, 图像划分能形成区域表示, 再对划分后的图像的各区域特征进行描述, 基于区域模式的描述, 可以形成局部与整体的格局, 这样该描述可以充分地把图像特征表达出来。在这种需求下, 构建颜色特征的链码描述, 链码是一个

矢量, 它是将图像划分后的颜色特征描述按规定的方向进行组装, 此时的链码具有如下特点: (1)它是一个矢量, 具有表示特征值的功能, 同时也具有表示特征选择的方向性; (2)它分割了整体, 形成了局部与整体合二为一的对象描述方法。图像的区域化、描述和特征描述集组装示意结构如图1所示。

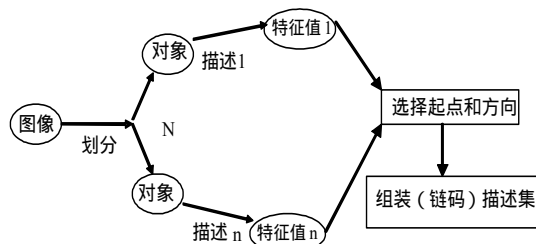


图1 图像的划分和描述

1.1 基于图像内容颜色链码的定义

基于图像颜色特征的描述最常见的方法是直方图, 而直方图反映的只是颜色概率密度函数, 它可能出现不同的图像具有相同的概率密度函数, 一种基于图像颜色特征的改进图像检索与识别方法——链码检索法能比较好地克服直方图直

基金项目: 国家自然科学基金资助重点项目(60234030); 国防科工委基金资助项目

作者简介: 石跃祥(1964-), 男, 博士、教授, 主研方向: 模式识别与智能系统; B. Benhabib, Professor; 蔡自兴, 教授、博导

收稿日期: 2006-06-11 **E-mail:** shiyx@xtu.edu.cn

接用于检索的某些不足。

链码 任意一个 $N \times N$ (N 能被3整除)像素的图像的链码是把图像分割成 3×3 个小块,求各个小块的颜色灰度平均值 $L_0L_1L_2L_3L_4L_5L_6L_7L_8$,以中心小块的灰度平均值 L_0 作为链码的起始位,在它周围的8个小块中选择一个灰度平均值大于 L_0 且只大于 L_0 的值 L_x ,将其排列在第2位,以此为起点,按逆时针方向将余下的灰度平均值区域连接起来所形成的一串值描述 $L_0L_x\dots$ 码叫链码。

例如将下图像分割为 3×3 小块后, $A_0 \sim A_8$ 分别表示了各小块灰度平均值,且 $A_i > A_j$,当 $i < j$ 时,则该图像的灰度链码为 $A_3A_4A_7A_8A_1A_6A_2A_0A_5$ 。如表1所示。其他情形可参照文献[3]处理。

表1 图像划分后的区域灰度平均值

A_6	A_1	A_8
A_2	A_3	A_7
A_0	A_5	A_4

1.2 基于内容的链码图像检索

通过利用阈值和采用相似性函数、相似性函数的加权处理来实现图像检索识别。当图像内容检索的对象与目标的描述链码(经归一化处理)的对应结点码值之差的绝对值之和 S ,当 S 大于 M 时,系统则对检索对象作舍弃,否则作相似处理, M 被称作相似匹配的阈值。

对于查询的图像对象 Q 和目标图像 G ,它们匹配的程度通过相似性公式来计算,这里分析一下对象与目标图像的两个链码,由于它们都是表达了图像的特征描述量,其相似性公式表示如下:

$$SIM(Q, G) = \frac{1}{N} \sum_{j=0}^{N-1} \left(1 - \frac{|q_j - g_j|}{\max\{q_j, g_j\}} \right) \quad (1)$$

其中, N 为图像划分的块数, q_j, g_j 分别表示对象与目标链码中的对应结点的值。首先计算 q_j 和 g_j 之间的绝对差,然后用二者的最大值归一化,为了反映相似性,用1减去归一化的差,接着除以 N ,得到链码相似的平均值,则在检索对象与目标时,已取得了决定性的依据。其加权处理的相似性公式如下:

$$SIM(Q, G) = \frac{1}{N} \sum_{j=0}^{N-1} \left(1 - \omega_j \frac{|q_j - g_j|}{\max\{q_j, g_j\}} \right) \quad (2)$$

其中, $\omega_j = S/N$,比较式(2)与式(1),不难发现,式(2)在计算的速度上比式(1)更快,分析 ω_j 与 $|q_j - g_j|$ 二者的积:

(1)当图像的对象 Q 与目标 G 相似程度较大时, S 相对来讲会比较小,当然 S/N 的值也会小, $|q_j - g_j|$ 也小,二者的积将会更小,这时的 $SIM(Q, G)$ 相对来讲就会高一些;

(2)而当图像的对象 Q 与目标 G 相似程度较小时, S 相对来讲会比较大,当然 S/N 的值也会大, $|q_j - g_j|$ 也大,二者的积将会大一些,这时的 $SIM(Q, G)$ 相对来讲就会低一些。其检索情形与对照如表2。

表2 各种位图检索情况与对照(%)

	1位	4位	8位	24位	颜色对法 ^[5]
旋转 90°	74.54	92.44	85.49	91.29	97
旋转 180°	89.73	92.44	100	100	100
放大 1倍	13.51	86.68	98.33	99.39	90
缩小 1/2	80.71	95.69	86.76	99.06	
旋转 5°				90.40	89

2 智能监控系统体系结构

监控系统采用单机监控点与中心服务器联接的方式,监控点负责监控地点的图像采集,捕获的图像可在单机上进行处理,同时通过网络与中心服务器进行数据的交换,实现监控的分布处理与数据的互动共享。其体系结构如图2所示。

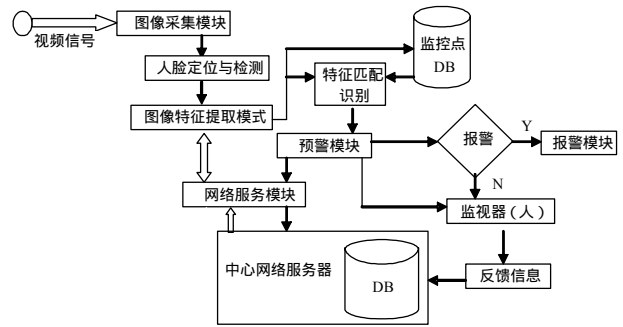


图2 智能监控系统结构

3 智能监控的实现

3.1 图像的采集

在智能监控中,图像获取是不可缺少的内容,它负责为监控系统提供监控对象的数字化信息。图像的获取也就是图像的数字化过程,即将图像采集到计算机中的过程,主要涉及成像图像的采集和数字化图像的存储技术。随着计算机与微电子特别是固体成像设备(电耦合设备(Charge Coupled Devices, CCD))的迅速发展,使得图像获取设备的成本显著降低,因而越来越普及。

以CCD技术为核心,目前图像采集设备有黑白摄像机、彩色摄像机、扫描仪、数码相机等,性能与价格主要取决于CCD的规格。除了这些常见的类型外,目前有许多厂商提供各种其他的专用设备,如显微摄像设备、红外摄像机、高速摄像机、胶片扫描器等。此外,遥感卫星、激光雷达等设备提供其他类型的数字图像。

目前,图像的数字化设备可分为2类:(1)使用图像采集卡或通过图像卡将模拟制式的视频信号(RS170/CCIR 黑白电视信号, PAL/NTSC 彩色电视信号, S-Video 视频信号等)采集到计算机;(2)摄像机本身带有数字化部件可以直接将数字图像通过计算机端口(如并口、USB接口)或标准设备(如磁盘驱动器)传送计算机。本实验系统使用的是USB接口的摄像头(型号:10Moons PCEye Camera)。

对于小摄像头的驱动有几种方法,如利用VFW(Video for Windows)的方法,即利用Visual C++所提供的Video for Windows库函数来实现从模拟视频源采集数字视频信号,并将其存储到文件中或者直接对视频缓存进行处理。本系统是通过使用DirectShow来驱动摄像头。DirectShow^[4]是微软公司在ActiveMovie和Video for Windows的基础上推出的新一代基于COM的多媒体开发包,它有着更好的灵活性,而且凡是符合微软DirectVideo规范的所有硬件,都可以做到自动的识别。

DirectShow的核心是一种被称为过滤器(Filter)的插件式模块系统,它被装配在过滤器图表(Filter Graph)中,用于管理在过滤器内部和过滤器之间处理的音视频数据流。过滤器图表管理器(Filter Graph Manager)负责把这些过滤器连接起来进行管理。DirectShow的体系结构如图3所示。

在过滤器图表中主要使用3种过滤器,主要功能是:

(1)源过滤器(Source Filter):主要把源端的数据读入到过滤器图表中;

- (2)变换过滤器(Transform Filter):按一定的编解码方法处理从源过滤器中接收到的数据,并把他们传送到提交过滤器中;
- (3)提交过滤器(Render Filter):提交数据到硬件设备或保存成文件。所有过滤器通过自身的针脚(Pin)相互连接起来。

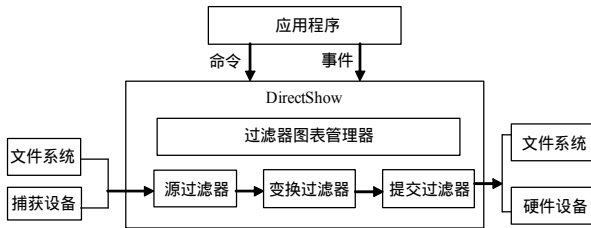


图3 DirectShow 的体系结构

其过程可分为下面几步:

1)利用 DirectShow 实现视频捕获的过程(VC++编程语言)

生成需要的构件:生成过滤器图表和图像捕获构造器,并使它们联合起来。

选择一个捕获设备。在下面的程序代码中,首先列举了可用的所有视频捕获设备,并选择列举序列中的第一个设备为使用设备。

将捕获的视频保存为文件的准备工作(如果只需要预览功能,可以去掉这一部分)。

视频图像的预览和存储。如果既要预览视频图像,又要将视频图像保存为视频文件,就必须引入 Smart Tee 过滤器,这个过滤器将捕获的数据复制为 2 条数据流:用于预览和用于存储。下面的方法完成了这个工作。它不仅可以在必要时,即预览和存储同时存在时添加 Smart Tee 过滤器;而且还可以自动地加载 WDM 设备所需要的任何能够支持的过滤器。

控制过滤器图表。对过滤器图表进行控制可以对捕获工作执行运行、结束和暂停的操作。

结束捕获操作,释放资源。

2)利用 DirectShow 实现图像捕获的过程(VC++编程语言)

在监控系统中,真正起作用的并非一段视频,而是一帧帧的图像,实现一帧图像的捕获是首先必须解决的问题。Sample Grabber 过滤器是一种变换过滤器,它可以从媒体流中攫取一帧图像。尽管 Sample Grabber 过滤器使用起来很复杂,但它几乎能支持所有的媒体类型,而且使用起来非常灵活,它是最佳选择。

和 都与视频捕获过程相同,即生成需要的构件和选择一个捕获设备。

加载 Sample Grabber。首先生成一个 Sample Grabber 过滤器,然后加入图表管理器。

设置媒体类型。在构建过滤器图表的其余部分之前,必须为 Sample Grabber 设置一个媒体类型。大多时候,只需设置主类型和子类型,有时也可能需要设置格式类型。下面的代码首先指定主类型为 MEDIATYPE_Video,然后设置了基于当前显示位深的子类型。

类似于视频捕获过程的第一步。

运行过滤器图表。Sample Grabber 工作包含两种模式:在将采样向下传送之前产生每个采样的拷贝,然后放到其缓冲;以回调方式进行处理数据,回调由应用程序定义。回调方式常常会影响工作效率,甚至死锁,所以采用缓冲模式。

得到缓冲区内存储的图像的格式。下面的代码将图像的格式保存在指针 pVih 中。

通过得到的捕获图像的数字化信息和图像的格式,即可直接处理这些信息,也可保存为图像文件,在操作中以 bmp 格式为例。

3)本系统的一些重要的系统指标

图像像素:352×288。该摄像头支持最大分辨率为 640×480,但帧率会很底。

帧率:30fps

3.2 人脸梯度特征定位与检测

若人脸图像用 x 表示, $n \times m$ 的矩阵 B 表示图像 x 的灰度值,矩阵 B 经边缘检测处理后,得到图像 x 的梯度矩阵 T ,其边缘梯度计算方法为

$$G[i, j] = |f[i, j] - f[i + 1, j + 1]| + |f[i + 1, j] - f[i, j + 1]| \quad (3)$$

$$G[i, j] = |G_x| + |G_y| \quad (4)$$

$$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (5)$$

在分析人脸图像中发现,眼球的灰度级较低,而眼球周围白色区域的灰度级较高,两个区域在其交界处灰度级产生强烈突变,这样形成梯度变化,而人的眼睛有一对,形成梯度对。在图像 x 的梯度矩阵 T 和灰度矩阵 B 中,梯度特征 f_{T_i} 存在的数目是不唯一的,为了更准确地定位眼睛的位置,此时,需要在灰度图像 B 的区域 Ω 中来计算所有的 f_{T_i} , ($i=1, 2, \dots, p$ p 为特征数)。在 f_{T_i} 中,若 $f_{T_i} \approx f_{T_j}$ 时 ($i \neq j$), 则 f_{T_i} 和 f_{T_j} 被确定为梯度对特征。

在梯度对特征确定的搜索过程中,采用的是 f_{T_i} 排序算法,即把 f_{T_i} 排序后,在排序队列中,相邻两两比较,当不存在比 $f_{T_i} \approx f_{T_j}$ 更相近的情形,此时的 i 与 j 被确定为梯度对特征所选定的位置^[6]。其实验检测见图 4。

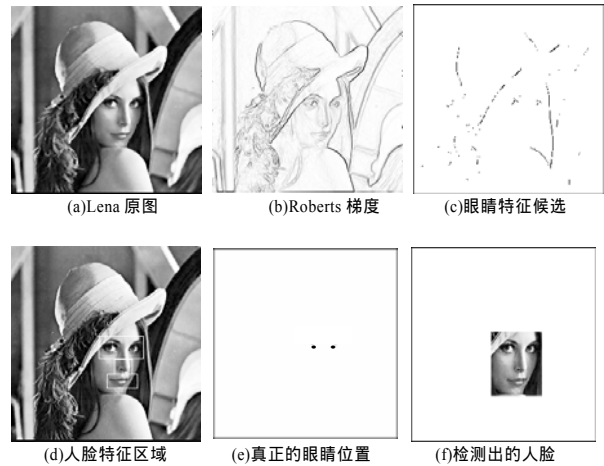


图4 基于梯度对特征的人脸定位与检测

表3 不同情形下的人脸检测成功率

对象	状况	
	简单背景	复杂背景
一人脸	125/127	86/89
二人脸	106/113	69/76
三人脸	96/108	65/75

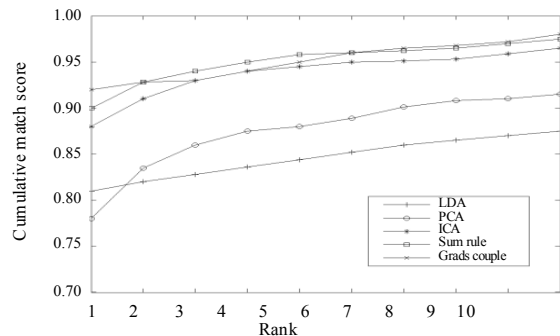


图5 累计匹配识别结果与特征曲线

(下转第 18 页)