

# 一种采用 AVS 的视频监控系统设计与实现

陈桂财<sup>1)</sup> 宋元征<sup>2)</sup> 王新<sup>2)</sup> 路红<sup>2)</sup>

<sup>1)</sup>(复旦大学软件学院, 上海 201103) <sup>2)</sup>(复旦大学计算机科学与工程系, 上海 200433)

**摘要** AVS 是我国具备自主知识产权的第 2 代信源编码标准, 是一套包含系统、视频、音频、媒体版权管理在内的完整标准体系, 在未来中国数字音视频产业中有广阔的应用前景。设计并实现了一套基于 AVS 视频编解码技术的视频监控系统, 可以对 AVS 视频流自动进行处理。该系统结合 AVS 视频编解码、背景/非背景分类、人脸检测与识别等技术, 可输入文本、图片和视频段对出现在监控视频中的人物进行身份的自动识别, 并引入两级索引结构以达到快速检索的目的, 在数字监控、门禁控制以及身份识别等方面有一定的应用价值。

**关键词** 基于 AVS 视频监控系统 自动化 背景/非背景分类 人脸检测与识别 索引结构

中图分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2009)08-1553-07

## Design and Implementation of an AVS-based Video Surveillance System

CHEN Gui-cai<sup>1)</sup>, SONG Yuan-zheng<sup>2)</sup>, WANG Xin<sup>2)</sup>, LU Hong<sup>2)</sup>

<sup>1)</sup>(Software School, Fudan University, Shanghai 201103)

<sup>2)</sup>(Department of Computer Science and Engineering, Fudan University, Shanghai 200433)

**Abstract** AVS is a fully intellectual property of second generation source coding and decoding standard owned by China. And AVS is a set of standard system which includes system, video, audio, and media copyright management. It will be potentially and widely used in digital video and audio industry of China. This paper designs and implements an AVS-based video surveillance system, which can automatically process AVS video streams. In the AVS-based video surveillance system, the AVS coding and decoding standards are combined with other processing techniques, such as background/non-background classification, face detection and recognition, and two-level indexing structure to automatically identify the persons who appeared in the surveillance videos. The videos can be retrieved quickly by inputting texts, pictures, and video segments. The AVS-based video surveillance system can be used in many areas such as digital surveillance, access control, and etc.

**Keywords** AVS-based, video surveillance system, automation, background/non-background classification, face detection and recognition, indexing structure

## 1 引言

目前安全问题已经受到广泛关注, 涌现了越来越多的视频监控系统, 如门禁系统、考勤系统以及身

份识别系统等等。视频监控系统可以让管理人员在控制室中观察到前端防范区域内所有人员活动情况并做记录, 为保安系统提供实时的图像、声音信息。但是, 传统的视频监控系统需要大量的人力资源开销, 对监控视频内容的检测、识别和理解完全依靠人

**基金项目:** 国家高技术研究发展计划(863)项目(863-2006AA01Z203); 自然科学基金项目(2006AA01Z203); 上海市重点攻关项目(07dz15004-1)

**收稿日期:** 2008-03-01; **改回日期:** 2008-05-07

**第一作者简介:** 陈桂财(1982 ~ ), 男。复旦大学软件学院计算机软件与理论专业硕士研究生。研究方向为视频分析和应用。  
E-mail: fducgc@hotmail.com

**通讯作者:** 王新, E-mail: xinw@fudan.edu.cn

工,降低了视频监控系统的工作效率,而且,安全性与准确性也缺乏保证。因此,对自动化视频监控系统的研究非常必要。无论是实时视频监控系统,还是对大量监控视频进行离线操作,如果预先通过计算机系统的自动处理,在保证返回内容的有效性的前提下,反馈给操作人员的信息量将远小于传统监控系统,从而大大节省了人力资源,同时也提高了视频监控系统的可靠性。

文献[1]~[3]提出了针对不同应用环境和功能的视频监控系统。视频压缩是视频监控系统的核心技术,需要根据监控系统的性质、结构和要求来选取压缩方法,然而目前还没有专用的数字视频监控系统视频压缩标准<sup>[4]</sup>。本文提出针对人物身份识别和检索的视频监控系统,包括实时与离线操作。在离线操作中,该系统采用我国自主知识产权的 AVS 视频编解码技术,直接处理 AVS 码流,顺应了视频编解码技术的发展。

## 2 AVS 标准介绍和优势

在 ISO/IEC 14496 MPEG-4, ITU-T H. 264, SMPTE VC1, AVS 等第 2 代信源编码标准中, AVS 是我国具备自主知识产权的编码标准,是一套包含系统、视频、音频、媒体版权管理在内的完整标准体系,具有很大的应用前景。

### 2.1 AVS 与其他编码标准比较

AVS 标准和 H. 264 标准具有相同的编码架构<sup>[5]</sup>。与其他标准相比, AVS 更注重应用的需要,提供不同的解决方案来适应不同的应用,以达到复杂度和效率适当的折衷,并且在提供相当性能的情况下,一定程度上降低编码复杂度<sup>[6]</sup>。

根据文献[6]~[9],列出了为 AVS 与 MPEG-4 AVC/H. 264 视频标准中影响性能和效率的主要技术模块的对比,如表 1 所示。

由文献[9]可知, AVS 相比于 H. 264, 在编码效率方面略低,平均有 0.11 dB 的损失;在逐行编码方面性能基本一致;在隔行编码方面,平均有 0.5 dB 的性能差距;大致估算,解码复杂度降低 70%, 编码复杂度降低 30%。

### 2.2 AVS 应用优势

在应用方面, AVS 的优势首先表现在它的知识产权清晰,尽可能多地采用公共知识和新技术,必需专利的数量不到相应国际标准的一半,而且绝大多

表 1 AVS 与 MPEG-4 AVC/H. 264 技术对比

Tab. 1 Technical comparison between AVS and MPEG-4 AVC/H. 264

视频编码标准	AVS 视频	MPEG-4 AVC/ H. 264 视频
帧内预测	基于 8×8 块 4 种色度预测模块 5 种亮度预测模式	基于 4×4 块 4 种亮度预测模式 9 种亮度预测模式
多帧预测	最多 2 帧,缓冲区 管理机制相对简单	最多 16 帧,缓冲区管 理机制复杂
变块大小运动 补偿	块运动估计:16× 16, 16×8, 8×16, 8×8	块运动估计:16×16, 16×8, 8×16, 8×8, 8× 4, 4×8, 4×4
1/4 像素运动 补偿	1/4 像素位置采用 4 拍滤波、线性插值, 1/2 像素位置采用 4 拍滤波	1/4 像素位置线性插 值,1/2 像素位置采用 6 拍滤波
B 帧宏块直接 编码	时域空域相结合	独立的空域或时域 预测模式
B 帧宏块双向 预测	对称预测,只编码 一个前向运动矢量, 由前向运动矢量导出 后向矢量	前后两个运动矢量 都要编码
熵编码	基于上下文自适应 的 2 维变长编码,编 码块系数过程中要进 行多码表切换	基于上下文的自适 应变长码;与周围块相 关性高,实现较复杂基 于上下文的自适应二 进制算术编码;计算 复杂
变换与量化	8×8 整数变换,编 码端进行变换归一 化,量化与变换归一 化相结合	4×4 整数变换,编解 码端都需要归一化,量 化与变换归一化相结 合

数专利都来源于中国会员;其次, AVS 采用了更合理的专利池的许可价格,且不对运营商收费。应用前景更广泛,可以针对不同应用选择不同的技术,定义不同的标准部分。目前国家和地方政府都在大力支持对 AVS 的应用推广。中国音视频编码标准组织也准备着手制定中国视频监控行业视频编码标准 (AVS-S)。

将 AVS 标准与背景/非背景分类,人脸检测与识别等技术相结合,对视频当中出现的人物进行身份识别以达到监控的目的。与其他压缩标准相比, AVS 能在对等的压缩比和复杂度的情况下,提供更佳图像性能,而且 AVS 视频流在背景/非背景分类方面的优势,使得计算机的智能化识别更迅捷准确。

### 3 系统框架及工作流程

#### 3.1 系统核心部分设计框架

首先采集按照 AVS 编码标准压缩的 AVS 码流,使用 AVS 码流解码过程中的压缩域信息来进行背景/非背景的分类。为了提高效率,当分类结果表明当前的帧不是背景时,才进行人脸检测。当检测到人脸时将人脸数据进行变换后与训练数据进行比较。在识别结果被反馈给用户之前,先计算置信度  $t$ ,  $t$  表明当前识别结果的可信程度。当置信度  $t$  小于阈值  $t_{min}$  时,认为该人脸不属于当前库中的数据,并将这个结果反馈给用户,经用户确认后将此新的人脸添加进库中。否则表明识别结果有较高的可信度,然后记录识别结果并对视频进行标注。图 1 是系统核心部分设计框图。

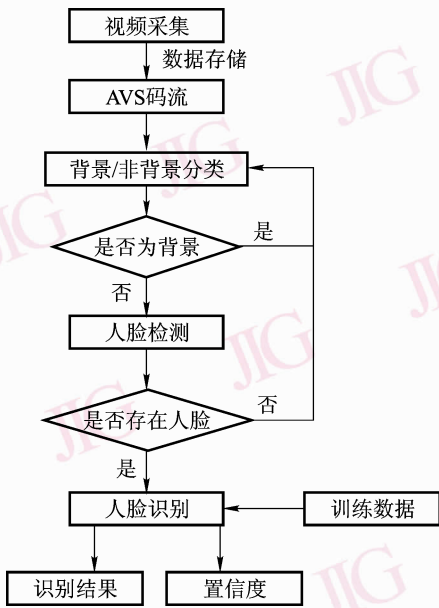


图 1 核心部分设计框图

Fig. 1 System core design

#### 3.2 系统工作流程

具体实现的系统主要由训练模块、标注模块和检索模块 3 个部分组成。

训练模块包括监控环境背景的训练和人脸库的训练。输入为 人脸样本库和背景样本库,提取得到各人脸特征和背景特征。

标注模块是对输入的监控视频进行自动标注。输入为训练模块得到的背景特征、人脸特征和待标注的监控视频,输出待标注监控视频的检索索引。

检索模块是对指定监控视频进行检索,包括图片查询、文本查询和视频查询。输入指定监控视频的索引及用户提交的图片、文本或小段视频,得到用户所提交内容在监控视频中相应的图像内容。图 2 所示为系统的主要组成模块、工作流程以及各个模块之间的逻辑关系。

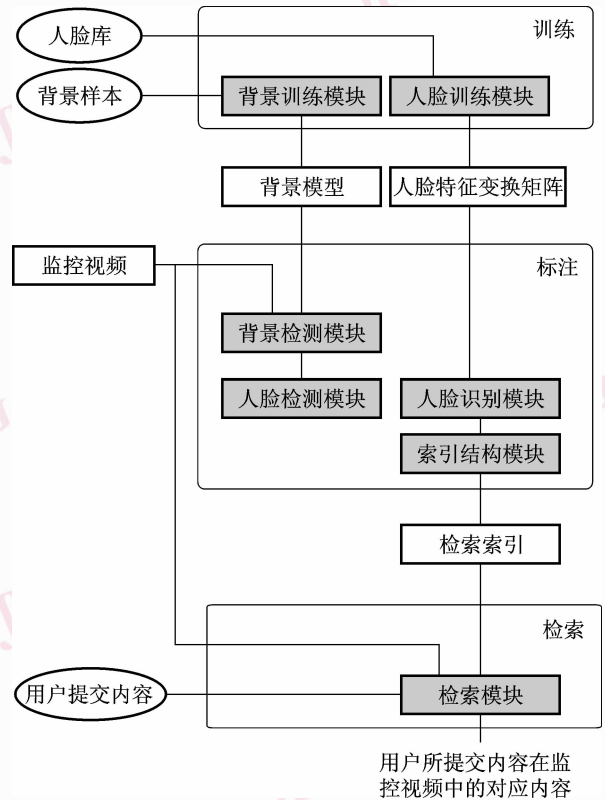


图 2 系统主要模块及工作流程

Fig. 2 System modules and workflow

### 4 实现方法

#### 4.1 背景/非背景分类

为了提高系统的处理效率,在系统当中不对 AVS 码流进行完全的解码,而是使用 AVS 的压缩域信息来进行背景/非背景分类。

背景/非背景分类是监控系统的第 1 步,功能是判断当前图像是否为背景,如果是背景就不进行后续的工作,因此可以提高整个系统的工作效率。系统适合应用于门禁控制等拍摄角度和背景相对固定,光线变化不太剧烈 的环境。因此可以认为拍摄角度是不变的,这是符合该视频监控系统应用特点的。

在 AVS 的压缩域当中,宏块的运动向量可以反映视频当中物体的运动。在背景片段中,图像是相对静止的,当有人出现时会使视频中引入更多的运动信息。文献[10]中提出使用 H. 264 的运动预测技术来进行背景/非背景的分类。将类似的算法用于 AVS 码流。设  $m_i$  为当前图像中一个宏块的运动向量,  $m_i = (\Delta x_i, \Delta y_i)$ ,  $0 \leq i \leq N-1$ 。  $N$  为当前图像中宏块的总数。用下式来计算当前图像中的运动强度。

$$MV = \frac{\sum_{i=0}^{N-1} size_i \times \sqrt{\Delta x_i^2 + \Delta y_i^2}}{\sum_{i=0}^{N-1} size_i} \quad (1)$$

式中,  $size_i$  表示第  $i$  个宏块的面积,  $\Delta x_i, \Delta y_i$  表示第  $i$  个宏块在两个方向上的运动向量。

单纯使用运动强度并不能完全表征当前图像中物体的运动状态,因此引入另一个参数( $MS$ )表示图像中运动的范围。

$$MS = \sum_{i=0}^{N-1} b_{-s_i} \quad (2)$$

$$b_{-s_i} = \begin{cases} size_i & \bar{m}_i \neq (0,0) \\ 0 & \text{其他} \end{cases}$$

在背景图像序列中,图像中不存在剧烈的运动,运动强度和运动范围都被限制在较小的数值。设置阈值  $mv_{\min}$  和  $ms_{\min}$ 。当满足下列条件时,判定当前图像属于背景。

$$MV < mv_{\min} \quad \text{且} \quad MS < ms_{\min}$$

进行背景/非背景分类的意义不仅仅在于提高了系统的效率,另一方面也搜集各监控点的统计信息,从而推断监控点的环境信息。例如通过统计非背景帧在监控序列当中的分布,就可以得知该监控点在什么时间段处于人流密集的状态,从而进一步对该监控点做出适当的部署。例如在人流相对密集的时间段提高录制的帧率,而在人流稀少的时间段降低录制的帧率等等。

#### 4.2 人脸检测和识别

经过背景检测,对判断不为背景的图像进行人脸检测。人脸检测采用 AdaBoost 算法<sup>[11]</sup>。但是为了提高系统的处理效率,不进行全局检测,而是进行局部检测。

从人脸检测中检测出的人脸图像进行尺寸统一缩放后,按照由左至右,由上至下扫描成样本向量,然后对样本向量进行降维。系统中采用经典的

PCA(主元分析)与 LDA(线性鉴别分析)结合的 Fisher-Face 算法进行人脸投影特征的抽取<sup>[12]</sup>。在使用 PCA 降维后的空间上使用 LDA,得到所检测人脸的特征向量。特征抽取后采用最小距离分类器与库中的人脸进行比对与识别。

设人脸  $f$  经过 Fisher-Face 特征抽取后的样本向量为  $f'$ ,  $f' = (u_0, u_1, \dots, u_k)$ , 然后计算其与训练样本的距离:

$$d(f', f'_i) = \sqrt{\sum_{j=0}^k (u_j - v_j)^2} \quad (3)$$

其中,  $f'_i = (v_0, v_1, \dots, v_k)$  表示库中的第  $i$  个训练样本,  $k$  为样本维数。  $d(f', f'_i)$  表示当前待识别样本与库中第  $i$  个训练样本的距离。

计算完  $f'$  与库中所有样本后,找出距离最小的前 5 个样本,  $f'_{i1}, f'_{i2}, \dots, f'_{i5}$ 。其中多数样本属于类  $c$ 。若 5 个样本各属一类,则以与  $f'$  距离最小的样本  $f'_{i1}$  所属类作为  $c$ 。用下式计算识别的置信度:

$$t = \frac{\sum_{j=1}^5 d(f', f'_{ij} | f'_{ij} \in c)}{\sum_{j=1}^5 d(f', f'_{ij})} \quad (4)$$

当置信度  $t$  小于阈值  $t_{\min}$  时,将  $f$  反馈给用户,经用户确认后将此新的人脸添加进库中,否则表示识别结果可靠并记录结果。

#### 4.3 索引结构

自动化身份识别是本系统的一大特色,在实际应用中将实时拍摄的视频进行背景/非背景分类、人脸检测、提取特征,并与库中已有的人脸特征进行对比,快速得到识别结果,记录出现人员的身份或者发出陌生人警报提示;或者在海量数据中自动查找特定人脸图像,在可接受的时间内得到查询结果。这就对查询速度有比较高的要求。

不同于一般数据的搜索,可以采用精确查找,人脸识别是对向量的查询,即使是同一个人在极短的间隔拍摄的两张人脸图像经过处理得到的特征向量也不会完全一样,所以人脸比对的时候采用的是模糊查找。计算两个向量之间的距离,距离越小,对应的两张人脸就越相似。虽然经过 PCA 的降维,但是所得到的人脸特征向量的维数仍然比较高,为获得向量之间的距离而需要的计算量还是很大。尤其是在所查询人脸特征数据库很大的情况下,所要花费的时间就会长得让人无法接受,也不符合实际应用的要求。

所以在系统中简单应用了两级索引结构。通过

预处理,采用最基本的 k-means 聚类法将特征向量库简单分为一定数目的聚类,并获得每个聚类的中心点,然后建立两级索引,一级索引为每个聚类的中心点,存储于一级文件;二级索引为每个中心对应聚类中的特征向量,每个聚类存储于单独的文件。查询时,先找到与所要查询向量最相似的中心点,再在这个中心点相对应的聚类里做顺序遍历。

假设库中所有特征向量有  $n$  个,计算目标特征向量和库中特征向量的时间花费是  $T_c$ 。如果不用索引结构,那么就要直接对库中所有特征向量进行遍历,所花费的时间是  $nT_c$ ;如果使用了索引结构,假设分了  $m$  个群(因为聚类可以离线完成,这里没有考虑聚类所花费的时间),相比较于 hierarchical clustering, k-means 聚类法对噪音点不是很敏感,倾向于得到聚类中数目大致接近的聚类结果,这里为了简化分析,在时间的理论分析时,可以认为每个群内特征向量数基本相等,即为  $n/m$ ,那么寻找最相似的中心点花费的时间为  $mT_c$ ,群内特征向量进行遍历的时间花费为  $nT_c/m$ ,总共时间花费为  $mT_c + nT_c/m$ ,在实际测试和应用中,一般情况  $n$  的数目都远远大于  $m$ ,那么使用索引结构查询一次花费的时间近似为  $nT_c/m$ 。所以可以得出结论,根据群的数目不同,查找速度可以获得相应的线性提升,这一点将在后面的实验结果中得到验证。

## 5 实验结果及用户界面

在实验中,背景/非背景分类结果以查全率( $R$ )与准确率( $P$ )来衡量:

$$R = \frac{N_c}{N_c + N_m}, P = \frac{N_c}{N_c + N_f} \quad (5)$$

其中,  $N_c$  表示正确检测到的背景帧数,  $N_m$  表示漏检的背景帧数,  $N_f$  表示误检的背景帧数。在实验当中阈值  $mv_{\min}$  和  $ms_{\min}$  的设置满足:

$$\{mv_{\min}, ms_{\min}\} = \operatorname{argmax}_{mv_{\min}, ms_{\min}} \left( \frac{R \times P}{R + P} \right) \quad (6)$$

分别使用 AVS 与 H. 264 编码标准对同一段 462 帧的定点监控视频进行编码,其中包含 272 幅纯背景帧。然后使用 B 帧进行背景/非背景分类,其结果如表 2 所示。通过对比发现,通过 AVS 编码得到的码流中,由于 B 帧的宏块采用直接模式来得到运动向量,在背景段落当中产生了大量的零向量,因此检测到了绝大多数的背景帧,从而获得了较高

表 2 背景/非背景分类试验结果

Tab. 2 Result of background detection experiment

	$R$	$P$
AVS	97.7	93.5
H. 264	90.0	95.7

单位:%

的查全率,其查准率与 H. 264 编码的结果相近。这对于监控系统来说是十分重要的。AVS 视频流在背景/非背景分类方面的优势,使得计算机的智能化识别更迅速准确。

从实际测试中发现,人脸识别的结果与用于识别的最小人脸尺寸有关,图 3 为一段大小为  $640 \times 480$  的监控视频的人脸识别实验结果。

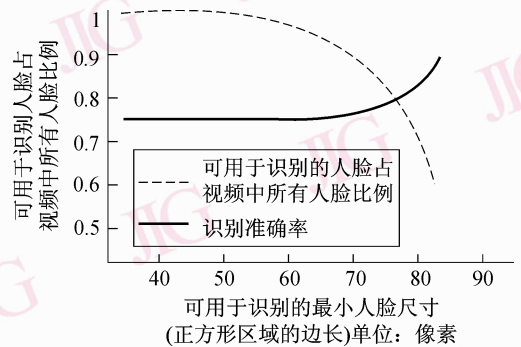


图 3 人脸识别准确与人脸大小之间的关系

Fig. 3 Relation between the precision of face recognition and the size of face

当用于识别的人脸的尺寸比较大时,人脸图像更细致,可以得到更好的识别结果。但是当采用越大尺寸的人脸样本时,可用于识别的人脸数目所占视频中出现的的所有人脸数目的比例也就相应地减小,图 3 中两条曲线体现了这一点,实线表示的是识别准确率,虚线表示的是所占比例。

根据设计要求比较完整地实现了系统功能,并提供一个操作性较强的用户界面。图 4 为系统的用户界面。

用户界面的最右边一列按钮用以打开指定视频,控制播放视频,进行背景训练和人脸训练,对指定视频进行标注,打开指定的特征向量数据库,对当前查询结果绘图,控制小段视频查询。中间上边是视频播放和图片显示区域;下面是查询结果绘图区和小段视频即时处理框。最左边上方是图片查询、姓名查询、视频查询输入框;中间是当前查询结果的

统计数据输出栏;下方是当前查询结果详细列表;列表右边的曲线表示的则是结果与查询目标的相似度。从图 4 中可以看出查询结果与查询目标的相似度曲线一开始趋势比较平缓(实线开始部分),从实线虚线相交处 A 开始其斜率有一个明显变化的过程,此后又趋于平缓(虚线后面部分),可见 A 处后面的数据与目标向量的差距很大,置信度比较低。

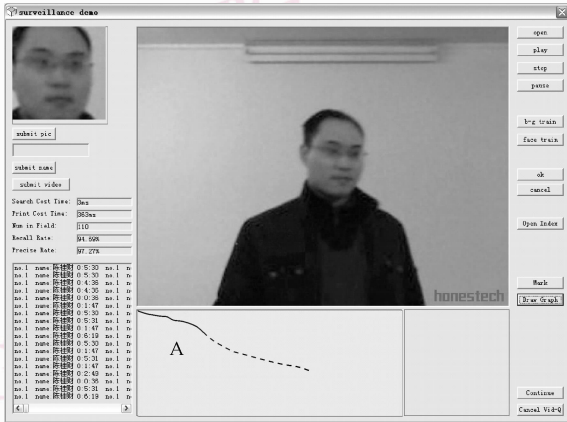


图 4 用户界面

Fig. 4 Interface

在进行人脸查询整体测试的时候,先根据相似度得到一定数目查询结果,然后在这些结果中取一定的阈值( $L$ )来对查全率( $R$ )与准确率( $P$ )进行对比,如果特征向量与目标向量之间的距离小于等于 $L$ ,则为有效统计数据(图中实线部分),如果特征向量与目标向量之间的距离大于 $L$ 则为无效统计数据(图中虚线部分)。根据下式计算得

$$L = \lceil (d_m - d_0)f + d_0 \rceil \quad (7)$$

其中, $d_0$ 表示目标向量和遍历范围中之最相似的特征向量之间的距离, $d_m$ 表示结果目标向量和遍历范围中之最不相似的特征向量之间的距离, $f$ 是可调节的梯度因子。统计所有样本人脸查询结果,系统性能达到设计要求的查全率和准确率均达 85% 以上。

在一个有 3 055 500 个样本向量库中进行检索速度实验。实验方法如下:取固定的属于 5 个不同个体的 5 个查询向量,分别在不调用索引,调用分群数目为 5 的索引,调用分群数目为 10 的索引这 3 种情况下,多次查询取平均时间(不计算聚类的开销)来比较查询速度(在前文的分析中已经提到,分群处理离线未完成,不影响系统在线运行的查询速度)。实验结果如表 3 所示。

表 3 检索速度实验结果

Tab. 3 Result of searching speed experiment

样本所属个体	查询时间(ms)		
	不分群的	群数目为 5	群数目为 10
A	415	133	68
B	403	127	71
C	414	141	70
D	411	145	70
E	405	135	68

从表中可以看出,当不进行聚类时,由于遍历范围是一样大的,所以 5 个样本查询时间变化幅度不大,当聚类数目过多,遍历范围都很小,由于检索速度过快和系统时间函数精度的问题,看起来变化幅度更小了,从聚类数目为 5 的时间结果推测 5 个样本查询时间, A, B, E 在库中数目较为接近, C, D 数目较为接近,符合实际情况;使用索引结构,系统的检索速度会明显增加,排除系统开销因素,速度增加大致呈线性关系。同样可以看出,在大规模查找比对中,检索系统速度是令人满意的。

## 6 结 论

针对视频监控系统的需求特征和要求,利用 AVS 的特点,并结合背景/非背景分类、人脸检测和识别、索引结构等技术,提出、设计并实现了一个完整的视频监控系统。该系统可以应用于门禁控制和自动考勤等拍摄角度和背景相对固定,光线变化不太剧烈的环境,可以自动识别身份,也可以进行大规模的人脸图像比对查询。在实际应用中,该系统体现出较好的自动化处理功能,在大规模身份识别中,该系统有较快的响应时间和准确率。且实验结果显示,与基于 H. 264 的监控系统相比,基于 AVS 的视频监控系统可以获得更高的查全率,而且 AVS 视频流在背景/非背景分类方面具有更大的优势,更加适合视频监控系统的要求。

进一步研究工作包括采用效率更高、更节省空间、更适合监控需求的码流存储方案,采用更合适的索引结构提高查询速度,以及进一步完善系统功能。

## 参考文献 (References)

- 1 Leo M, D' Orazio T, Spagnolo P. Human activity recognition for automatic visual surveillance of wide areas[A]. In: Proceedings of

- the ACM 2nd international workshop on Video surveillance & sensor networks[C], Now York, NY, USA: 2004: 124-130.
- 2 Wang Y F. A video analysis framework for soft biometry security surveillance [A]. In: Proceedings of the 3rd ACM International Workshop on Video Surveillance & Sensor Networks[C], Singapore, 2005: 71-78.
  - 3 Jain M D. A video surveillance system under varying environmental conditions[A]. In: Proceedings of the 24th IASTED International Conference on Signal Processing, Pattern Recognition, and Applications [C], Innsbruck, Austria, 2006: 32-37.
  - 4 Li Lei, Fan Rui-xia. Application of video compression in video monitoring system[J]. Journal of Shandong Institute of Light Industry (Natural Science Edition), 2006, **20**(1): 72-76. [李蕾, 范瑞霞. 视频压缩标准在视频监控系统中的应用[J]. 山东轻工业学院学报(自然科学版), 2006, **20**(1): 72-76.]
  - 5 Ma Si-wei. Review and application prospect of technique of AVS video coding[J]. Information Technology Letter, 2005, **3**(10): 20-25. [马思伟. AVS 视频编码技术回顾与应用展望[J]. 信息技术快报, 2005, **3**(10): 20-25.]
  - 6 Gao Wen, Reader Cliff, Wu Feng, *et al.* AVS-The Chinese Next-Generation Video Coding Standard[R/OL]. <http://www.avs.org.cn/reference.asp>, 2006-07-22/2008-04-23.
  - 7 GB/T 20090.2-2006, Information Technology—Advanced Coding of Audio and Video (Part 2): Video[S]. [GB/T 20090.2-2006, 信息技术——先进音视频编码第二部分: 视频[S].]
  - 8 Fan Liang, Ma Si-wei, Wu Feng. Overview of AVS video standard [A]. In: Proceedings of the 2004 IEEE International Conference on Multimedia and Expo(ICME)[C], Taipei, China, 2004: 423-426.
  - 9 Huang Tie-jun. Compare between AVS and MPEG[R/OL], <http://www.avs.org.cn/reference.asp>, 2006-07-10/2008-04-23. [黄铁军. 视频编码国家标准 AVS 与国际标准 MPEG 的比较[R/OL], <http://www.avs.org.cn/reference.asp>, 2006-07-10/2008-04-23.]
  - 10 Hui Hui, Liu Han, Wu Ya-li, *et al.* Video surveillance method based on H. 264 standard [J]. Computer Applications, 2005, **25**(11): 131-133. [惠穗, 刘涵, 吴亚丽等. 一种基于视频编码标准 H. 264 的智能视频监控技术[J]. 计算机应用, 2005, **25**(11): 131-133.]
  - 11 Freund Y, Schapire R E. A decision-theoretic generalization of online learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, **55**(1): 119-139.
  - 12 Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs fisherfaces: recognition using class specific linear projection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, **19**(7): 711-720.