

自然保护区海量视频数据快速分析与检索^①

谢琳^{1,2}, 张磊², 李健²

¹(中国科学院大学, 北京 100049)

²(中国科学院 计算机网络信息中心, 北京 100190)

通讯作者: 谢琳, E-mail: xielin@cnic.cn

摘要: 近年来, 随着视频监控系统在自然保护区的大量部署, 如何有效利用日益增加的海量视频监控数据成为亟待解决的难题. 通过基于图像相似度的关键帧提取算法对海量视频数据进行清洗和压缩, 同时利用基于深度学习的目标检测算法提取关键帧中的有效视频信息, 并提供多种基于内容的视频检索方式, 自动对用户提交的检索内容进行分析与处理, 从而快速检索出感兴趣的视频. 通过对青海湖野生动物视频监控数据进行分析与检索, 验证了该系统的有效性.

关键词: 视频分析; 视频检索; 关键帧提取; 深度学习; 目标检测

引用格式: 谢琳, 张磊, 李健. 自然保护区海量视频数据快速分析与检索. 计算机系统应用, 2018, 27(4): 63-69. <http://www.c-s-a.org.cn/1003-3254/6291.html>

Rapid Analysis and Retrieval of Massive Video Data in Nature Reserves

XIE Lin^{1,2}, ZHANG Lei², LI Jian²

¹(University of Chinese Academy of Sciences, Beijing 100049, China)

²(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: In recent years, a large number of video surveillance systems are deployed in the nature reserves, so it has become an urgent problem how to effectively use the increasing mass of video surveillance data. In this study, an efficient algorithm for key frame extraction based on image similarity is used to clean and compress the massive video data. At the same time, an object detection algorithm based on deep learning is used to extract valid video information. In addition, the system provides a variety of content-based video retrieval methods. It automatically analyzes and processes the search contents submitted by the user so as to quickly retrieve the video of interest. This study analyzes and retrieves the video surveillance data of wild animals in Qinghai Lake, which verifies the correctness of the proposed system.

Key words: video analysis; video retrieval; key frame extraction; deep learning; object detection

近几年来, 随着多媒体技术的蓬勃发展, 摄像设备的日益普及, 视频监控系统广泛应用在安防、城市交通、物业管理等多个领域, 视频监控数据呈现爆炸式的增长, 现已成为大数据时代的主要研究对象之一. 为了更好地实现对自然保护区的监控以及对野生动物的追踪与观察, 大部分自然保护区也安置了视频监控设备, 对自然保护区进行 24 小时不间断的、全方位的、

多角度的视频监控, 视频数量非常大. 例如, 自 2005 年以来, 中国科学院计算机网络信息中心与青海湖保护区管理局合作, 共同建设了青海湖野外网络视频监控系统, 每年青海湖核心区的视频数据高达几十 TB^[1]. 因此如何自动快速地分析和处理自然保护区海量视频数据, 实现对有效视频数据的充分利用, 成为一个亟待解决的重要问题.

① 收稿时间: 2017-07-14; 修改时间: 2017-07-28; 采用时间: 2017-08-14; csa 在线出版时间: 2018-03-31

迄今为止,国内外许多学者对视频处理技术进行了大量的研究,产生多种视频数据提取方法. Wofit^[2]提出基于光流分析提取关键帧,这种方法很好的考虑了运动特征,但是计算量非常大. Hanjalic 等人^[3]对聚类有较早的研究,通过多重分割聚类去除视频的冗余部分,并将聚类中心作为关键帧,该方法具有一定的通用性. Shahraray 等人^[4]提出基于镜头边界法,将切分得到的镜头中的首帧(或尾帧)作为关键帧. 文献^[5]中 Liu Gentao 等人提出了基于镜头的关键帧提取架构,结合了镜头边界检测算法和在镜头内的关键帧提取算法. 这些方法提取的关键帧能够有效的表达原始视频的内容,适合多镜头多场景的视频分析. 但是,自然保护区的视频监控系统面向特有的监控对象与自然环境,视频拍摄场景相对固定,没有明显的镜头切换,并且所产生的视频数据中往往含有大量近似于静止的视频片段,具有很大的可压缩空间. 现有的视频处理技术并没有针对这一特点做出相应的改进,因此并不完全适用于自然保护区视频数据的分析. 针对这一特殊需求,参考目前已有的视频数据处理方法,设计并实现了自然保护区海量视频数据分析与检索系统. 该系统中包括视频要素信息提取、视频关键帧提取、视频关键帧目标检测、建立视频索引等一系列处理流程,充分实现了对自然保护区视频数据的自动分析与利用,取得了良好的实践效果.

1 系统总体架构

自然保护区海量视频数据分析与检索系统涵盖了视频数据存储、分析、利用的完整处理流程,实现了数据存储、数据分析、数据利用等核心功能. 其中,数据分析包括视频数据的清洗、压缩以及有效视频数据的提取,数据利用则通过建立视频索引实现基于视频内容的检索操作,并将视频分析与检索结果进行可视化展示. 按照上述设计思路,自然保护区海量视频数据快速分析与检索系统的总体架构如图1所示. 该系统主要分为以下几个模块:

(1) 数据存储: 通过分布式数据平台将原始视频数据录入到数据库中. 然后利用分布式数据平台进行视频处理任务和视频检索任务的分配与并行处理,并将其产生的数据存储到数据库中.

(2) 数据分析: 视频处理引擎主要实现对视频数据的分析和处理. 首先获取视频拍摄地经纬度以及拍摄日期等视频要素信息; 然后利用基于视频帧图像相似

度的关键帧提取算法,提取视频中的关键帧数据,同时通过分析视频流之间的变化,将原始监控视频划分成静态视频段和动态视频段(即下文所指的有效视频段),并剔除其中不感兴趣的静态视频片段,实现对视频的清洗与压缩; 最后利用 YOLO 目标检测算法^[6,7]提取视频关键帧中包含的野生动物物种、数量等关键信息.

(3) 数据利用: 视频检索引擎为分析得到的有效视频数据建立多种索引,提供视频标签检索、视频要素信息检索以及视频图像检索等多种基于内容的视频检索操作. 可视化展示与检索模块主要为用户提供可视化服务,即通过 Web GIS 可视化展示自然保护区动植物的分布以及用户可以通过可视化窗口进行视频检索.

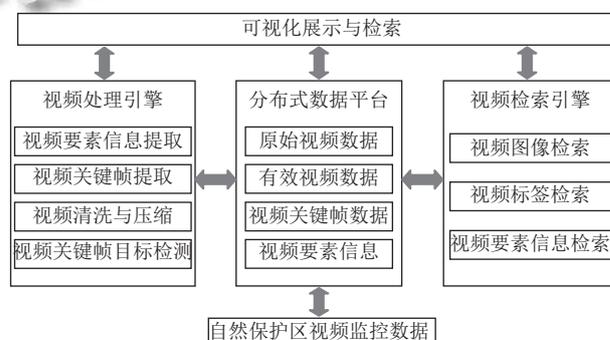


图1 自然保护区海量视频数据快速分析与检索系统总体框架

在本系统中,视频处理引擎和视频检索引擎是视频数据处理与检索的主要模块,实现了整个系统的核心功能,下面将重点介绍这两个功能模块.

2 视频处理引擎

目前自然保护区大多数采用高清红外摄像机、激光红外摄像机等监控设备,计算机可以采集视频拍摄地的经纬度以及拍摄日期等视频要素信息,并将它们存储在数据库中. 但是对于自然保护区的监测人员和研究者,更加关注监控视频中的动态视频片段和野生动物物种相关的要素信息,因此从海量视频片段中提取能够代表视频变化情况的关键帧以及关键帧包含的视频要素信息就显得尤为重要. 视频数据是由一系列连续的、渐变的帧组成,关键帧就是用于描述一段视频的关键图像帧,它通常能反映视频的主要内容,具有代表性. 本文针对自然保护区监控视频的特点,首先利用基于图像相似度的视频关键帧提取算法提取视频关键帧,然后进行视频片段的划分,剔除视频数据中的静

态视频片段,实现视频数据的清洗和压缩,有效地减少视频数据量.另外,本文采用基于深度学习的目标检测算法对视频关键帧进行进一步分析与处理,提取视频关键帧中包含的野生动物物种类别和数量等要素信息.

2.1 基于图像相似度的视频关键帧提取算法

通过对多段监控视频进行采样分析发现,在未进行镜头调整的情况下,一段连续视频的视频帧图像信息特征值总是与其前后几帧相近,处于一种连续渐变的状态,但是当这种渐变积累到一定程度时,视频帧的图像特征值会发生显著的变化^[8].因此,利用相似图像的特征值相近这一特性,本文提出基于图像相似度的视频关键帧提取算法.首先将视频划分成视频帧图像,然后提取图像的特征向量,通过图像的特征向量计算当前帧与后续帧之间的特征距离,如果特征距离大于某一阈值,则当前帧被选为关键帧,这样减少了关键帧的冗余度,提高了关键帧的代表性.

基于内容的关键帧提取技术主要利用的是图像的底层视觉特征,包括颜色、形状、纹理、边缘等.颜色特征^[9]是图像最基本的视觉特征,主要有RGB颜色空间、HSV颜色空间和HSI颜色空间.与RGB颜色空间相比,HSV颜色空间能较好地反映人眼对颜色的感知,同时又能够方便地与RGB颜色空间进行转换.在颜色特征提取的过程中,首先将RGB颜色空间的各个像素转换到HSV颜色空间内,然后将HSV空间进行非均匀量化,形成颜色特征向量.但是颜色直方图缺乏像素空间信息,本文在提取颜色特征的同时,采用LBP纹理描述子^[10]来描述图像的纹理特征,并对视频帧图像进行纹理特征的提取.采用综合特征^[11]计算特征距离,提高了关键帧提取的准确度,有效的避免了漏检和误检的问题.

视频关键帧提取算法的具体流程如下:

(1) 对于监控视频片段 $\{\pi_1, \pi_2, \dots, \pi_n\}$,将 π_i 划分为视频帧图像,同时将每帧视频图像划分成 $M \times N$ 的图像块^[12].

(2) 提取视频帧图像块的特征向量,记为 $I_p(m, n) = (x_{p1}, x_{p2}, \dots, x_{ps})$,其中 p 代表视频帧的位置, s 代表特征向量的维度.

(3) 将视频片段 π_i 的起始帧作为参考帧和关键帧,计算参考帧与后续帧对应图像块之间的特征距离.

$$D_{ij}(m, n) = \sqrt{\sum_{k=1}^s (x_{ik} - x_{jk})^2} \quad (1)$$

其中 i 为参考帧, j 为后续帧, $D_{ij}(m, n)$ 是参考帧和后续帧对应的第 m 行 n 列图像块的特征距离.

(4) 设 T_1 是局部阈值,如果 $D_{ij}(m, n) > T_1$,说明此图像块发生了显著变化;如果 $D_{ij}(m, n) < T_1$,则说明此图像块的内容基本没有发生变化.然后统计视频帧图像中发生显著变化的图像块的数目,记为 C_j .

(5) 将发生显著变化的图像块的数目 C_j 进行归一化.

$$C_j^* = \frac{C_j}{M \times N} \quad (2)$$

(6) 设 T_2 是全局阈值,比较 C_j^* 和 T_2 ,如果 $C_j^* > T_2$,说明后续帧相对于参考帧发生了巨大变化,将第 j 帧作为关键帧和参考帧,重复上述操作;如果 $C_j^* < T_2$,说明后续帧与参考帧相似,继续计算下一帧,直到视频片段 π_i 的所有帧计算完毕,此时即可得到视频片段 π_i 的所有关键帧.

图2中,(a)示例了参考帧 i 与后续帧 j 对应图像块特征距离的计算过程,(b)示例了根据设定的阈值判断图像块是否发生显著变化所生成的二维矩阵,其中0表示图像块未发生显著变化,1表示图像块发生显著变化.

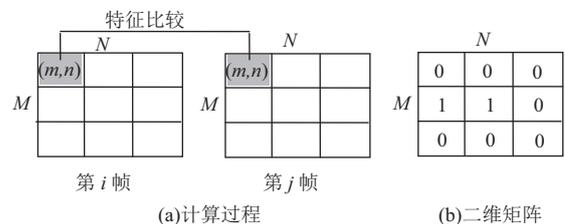


图2 参考帧与后续帧对应图像块特征

2.2 视频片段划分

一般情况下,实施监控的摄像机是静止不动的,即监控视频具有静止的背景,不存在明显的视频结构和镜头的切换,因此对于自然保护区的监控视频,视频流之间的差异往往是由视频中变化的目标造成的.根据视频流之间的差异性,视频数据可以分为静态视频片段和动态视频片段.对于静态视频片段,视频流之间存在的差异比较小,而对于动态视频片段,不同的视频帧包含着不同的视觉内容,内容相近的视频帧可以用关键帧表达.自然保护区的视频监控系统面向特定的监控对象,所产生的视频数据中往往会产生大量的静态视频片段,这些静态视频片段对于监控对象的观察与追踪没有太大价值,因此剔除视频数据中的静态视频

片段,实现视频数据的清洗和压缩显得尤为重要。

利用图像相似度的视频关键帧提取算法,可以得到原始监控视频的关键帧集合。由于关键帧的特性,视频中不会出现两段连续的静态视频片段,因此本文根据这一特点进行视频片段划分。首先计算当前关键帧与相邻关键帧之间间隔的视频帧数目,如果数目小于某一阈值,则说明视频流处于变化状态,当前关键帧与相邻关键帧之间间隔的视频帧属于动态视频片段,反之,属于静态视频片段。具体的算法流程如下:

(1) 对于视频关键帧集合 $f = \{f_1, f_2, \dots, f_n\}, f_k (k=1, 2, \dots, n)$ 表示视频帧的位置,根据公式(3)计算相邻关键帧之间的视频帧数目。

$$N(f_i, f_{i+1}) = f_{i+1} - f_i \quad (3)$$

其中 f_i 为当前关键帧位置, f_{i+1} 为相邻关键帧位置, $N(f_i, f_{i+1})$ 为当前关键帧与相邻关键帧之间间隔的视频帧数目。

(2) 设 T_3 是阈值,如果 $N(f_i, f_{i+1}) < T_3$,说明视频流处于变化状态,则当前关键帧 f_i 以及 f_i 与相邻关键帧 f_{i+1} 之间的视频帧属于动态视频段 π_{Dk} ,然后将 f_{i+1} 作为当前关键帧, f_{i+2} 作为相邻关键帧,重复上述操作;如果 $N(f_i, f_{i+1}) > T_3$,说明视频流变化较小,则 f_i 以及 f_i 与 f_{i+1} 之间的视频帧属于静态视频段 π_{Sg} ,然后继续比较 $N(f_{i+1}, f_{i+2})$,此时 f_{i+1} 作为新的动态视频段的第一帧。以此类推,直到所有关键帧计算完毕,此时即可得到所有动态视频片段和静态视频片段。

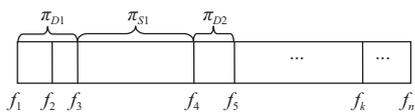


图3 视频片段划分示意图

图3 示例了一段视频的划分过程。 $\{f_1, f_2, \dots, f_n\}$ 是这段视频的关键帧集合,首先比较 f_1 和 f_2 ,由于 $N(f_1, f_2) < T_3$,则 f_1 以及 f_1 与 f_2 之间的视频帧属于动态视频段 π_{D1} ,接着比较 f_2 和 f_3 ,同样由于 $N(f_2, f_3) < T_3$, f_2 以及 f_2 与 f_3 之间的视频帧也属于动态视频段 π_{D1} ,然后比较 f_3 和 f_4 ,由于 $N(f_3, f_4) > T_3$,则 f_3 以及 f_3 与 f_4 之间的视频帧属于静态视频段 π_{S1} ,继续比较下一帧 f_5 ,此时 f_4 作为新的动态视频段 π_{D2} 的第一帧。以此类推,直到所有关键帧计算完毕,即可得到示例视频段的所有动态视频片段和静态视频片段。

2.3 视频关键帧目标检测

视频关键帧能有效的反映视频的关键内容,包含许多视频要素信息,例如野生动物的物种类别、数量等信息,这些信息的提取与保存将为视频检索引擎提供强大的数据支持。随着深度学习在图像与视频分析领域的广泛应用,2015年 Joseph Redmon 等人^[6]提出了一种新的目标检测方法 YOLO,可以快速检测出目标的类别和位置,为我们提取视频关键帧的要素信息提供了新的思路。本文采用 YOLO 方法对关键帧图像进行目标检测,提取出关键帧图像中包含的野生动物物种类别和数量信息,将这些信息作为描述视频的要素信息存入数据库,为视频检索引擎的实现奠定基础。

本文首先利用自然保护区视频监控数据构建野生动物目标检测的训练数据集,然后使用 YOLO 对训练数据集进行训练,形成特定自然保护区的目标检测特征模型。接着利用 YOLO 训练出的模型对提取的视频关键帧图像进行目标检测,将检测出的目标类别和目标数量存入数据库。经过关键帧图像目标检测处理,关键帧图像数据就可以转化为数据库中方便检索的文本数据。视频关键帧目标检测的实现过程如图4所示。

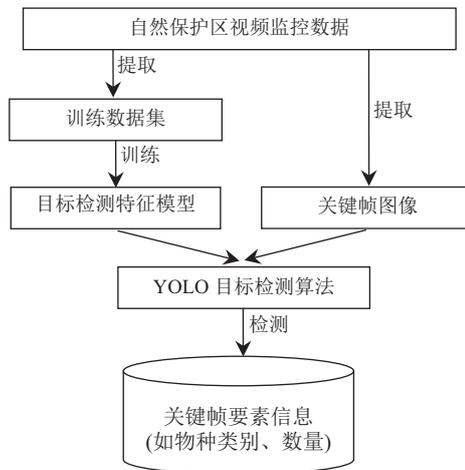


图4 视频关键帧目标检测流程图

3 视频检索引擎

视频检索引擎将利用视频的各种属性与关键信息快速寻找用户感兴趣的视频数据,达到检索的目的,极大地提高了视频查找效率。

3.1 视频标签检索

监控视频存入数据库时,为了方便管理,会为每段视频标注能够尽可能描述视频内容的文字,这样用户通过视频内容关键字进行检索时,系统将快速返回用

户感兴趣的视频片段,以及每段视频包含的关键帧、拍摄时间、经纬度等视频要素信息。

3.2 视频要素信息检索

对于监控视频来说,视频的信息量非常大,不同身份的用户对于视频信息的关注点不同,例如有些研究者主要研究不同野生动物的分布状况,而有些研究者则更加关注气候因素的影响,因此实现视频要素信息的检索是非常必要的。本系统为视频处理引擎提取的每个视频要素信息建立索引,加快查询速度,减少系统的I/O操作。用户可以根据自身需要,通过GPS经纬度、拍摄日期以及物种类别等视频要素信息方便快捷的检索出感兴趣的视频段。

3.3 视频图像检索

随着视频信息量的日渐增加,仅仅依靠视频标签以及视频要素信息进行视频内容的检索已经无法满足用户的需要,尤其对于视频来说,视频内容描述的准确性严重影响检索结果的准确性,时常出现检索结果无法满足用户需求的情况。因此,以图搜图成为视频图像检索中应用非常广泛的方法。本文利用感知哈希算法^[13],对视频的每一个关键帧生成一个“指纹”字符串,将其作为关键帧数据的图像特征存入数据库,当用户提交检索图片时也会产生相应的一个“指纹”字符串,将其与数据库中关键帧的指纹特征比较,结果越接近,说明图片越相近,从而快速检索出用户感兴趣的视频段,实现过程如图5所示。这种图像检索方法实现非常简单,能够快速检索出目标视频段,但是感知哈希算法对图像特征的表达不充分,影响检索结果的准确性。因此,在后续工作中,将结合SIFT和CNN的方法提取图像特征向量,增加图像检索的速度和准确率。

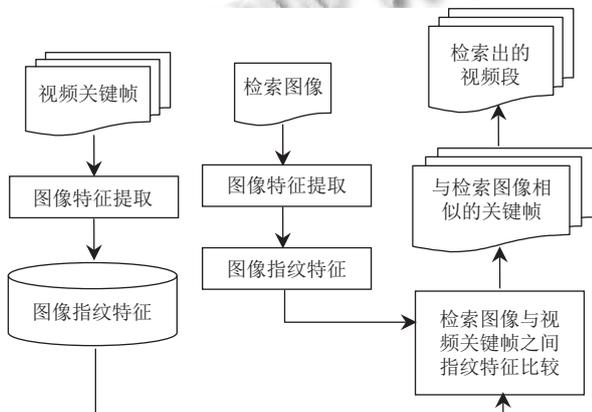


图5 视频图像检索流程图

4 实验结果与分析

为了验证本文中视频数据分析与检索方法的有效性,以青海湖野生动物监控视频为测试样本进行实验。由于视频处理过程中计算量非常大,并且计算任务重复性高,因此实验过程中利用4台计算机组成的HTCondor集群^[14]对视频数据进行处理,相对于单机而言,视频处理速度有了显著提高。

4.1 视频关键帧提取与视频片段划分实验

本实验首先利用基于图像相似度的关键帧提取算法提取青海湖黑颈鹤监控视频的关键帧,然后利用关键帧的特性进行视频片段的划分。



图6 视频关键帧提取与视频片段划分实验结果

由于监控视频的背景基本静止,视频中黑颈鹤的运动使得视频帧的图像特征值发生变化。通过视频帧图像特征值的比较,可以标注出视频帧变化的区域。当视频帧变化区域面积达到设定的阈值时,则认为当前帧为视频的关键帧,此时可以记录下关键帧的拍摄时间以及在视频片段中的位置,并将关键帧有效信息存入数据库中。图6的结果显示本文方法能够有效地从野生动物监视视频中提取关键帧,通过这些关键帧的内容,可以迅速地了解原视频中两只黑颈鹤的动作变化,如图6(a)-(d)所示。通过关键帧在视频中的位置,可以快速划分视频段,根据视频段的帧数判断出静态视频段和动态视频段。

为了进一步检验文中所述算法的效果,选择了三

段不同类型的视频进行实验,在实验中分别采用目测和文中所述的基于图像相似度的关键帧提取算法进行关键帧提取.实验数据和结果如表1.

表1 不同类型视频的关键帧提取结果

视频类型	总帧数	目测结果	实验结果
黑颈鹤活动视频	4759	36	47
青海湖风景视频	883	1	1
车辆行驶视频	2676	41	54

根据实验结果,可以发现对于黑颈鹤活动视频和青海湖风景视频,视频中含有大量的静态视频片段,采用基于图像相似度的关键帧提取算法可以有效地提取关键帧,并且视频压缩率比较高.而对于车辆行驶视频,视频中物体快速运动,造成视频帧间特征值相似性小,提取的视频帧数目较多,视频压缩率低.因此,针对自然保护区监控视频的特点,本文提出的关键帧提取算法可以快速准确的提取出视频中的关键帧,实现对视频的清洗和压缩,有效的减少了视频数据量.

4.2 视频关键帧目标检测实验

本实验对青海湖野生动物物种图像进行建模,然后利用YOLO目标检测方法对青海湖野生动物监控视频的关键帧图像进行检测.图7中是青海湖野生动物关键帧图像的检测结果,根据标注出的标签可以得到图像中包含的物种类别,同时根据目标框的位置可以统计出各个类别物种的数量.由图7(b)可以看出,本文中使用的目标检测方法对小目标也有较好的检测结果.本实验很好的验证了利用视频关键帧目标检测方法提取视频要素信息的有效性.

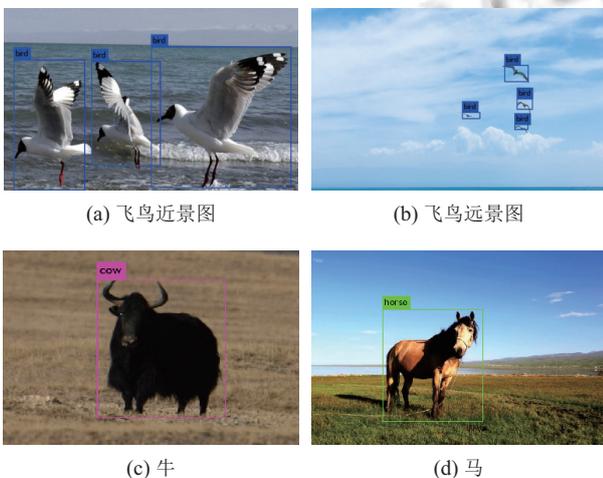


图7 视频关键帧目标检测实验结果

4.3 视频压缩与检索实验

视频关键帧的提取是实现视频压缩与检索的关键环节.通过视频关键帧的提取,可以快速删除视频中的静态视频片段,同时利用关键帧图像可以快速浏览视频内容,筛选出用户感兴趣的视频.

图8展示了青海湖鸟类监控视频压缩与检索的结果.由图8中原始视频帧数和提取关键帧数的统计结果可以看出,通过视频关键帧的提取,有效的压缩了自然保护区视频监控数据的数据量,节省了存储空间.同时,通过浏览视频关键帧信息,可以快速筛选出用户感兴趣的视频.另外,用户通过输入检索信息,系统自动检索出数据库中与检索信息相匹配的关键帧,进而检索出相应的视频段,达到检索的目的,极大的提高了视频查找的效率.视频处理与检索的过程中,由于视频文件都是大文件,计算机单节点的数据传输压力非常大,本系统利用Hadoop分布式系统^[15],有效地缓解了视频处理过程中单节点的数据传输压力,解决了计算节点传输数据时占用大量网络带宽的瓶颈问题,大大提高了海量视频数据分析与检索过程中文件传输的效率.

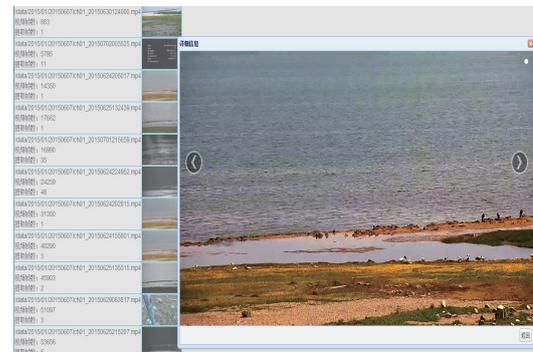


图8 视频压缩与检索实验结果

5 结束语

本文提出的视频关键帧提取算法有效地实现了海量监控视频的清洗和压缩,同时利用YOLO目标检测算法提取视频关键帧要素信息,为视频内容检索提供方便快捷的途径,并且将提取的有效视频段、视频关键帧和视频要素信息进行关联与整合,使用户可以通过图像、视频标签、视频要素信息检索出感兴趣的视频段、关键帧图像以及视频描述信息等多种数据.这样不仅为监控人员减少了视频处理与分析的工作量,也为快速检索视频内容提供了平台.同时,为了提高视

频处理与检索的速度,本文采用分布式数据平台,缓解了单节点传输数据文件的压力.近年来,随着深度学习的快速发展,在后续工作中,将改进视频图像检索的图像匹配方法,利用 Facebook 发布的相似性检索算法 Faiss^[16]进行图像检索,提高视频图像检索的速度和准确率.同时扩展视频处理引擎和视频检索引擎的功能,实现视频中的物体识别和目标检测,使之成为一个功能更加强大的视频数据分析与检索系统.

参考文献

- 1 谢慕哲,罗泽,阎保平.视频鸟类行为研究中基于尺度不变特征变换的形态分类算法.科研信息化技术与应用,2014,5(3):87-94.[doi:10.11871/j.issn.1674-9480.2014.03.011]
- 2 Wolf W. Key frame selection by motion analysis. Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing. Atlanta, GA, USA. 1996, 2: 1228-1231.
- 3 Hanjalic A, Zhang HJ. An integrated scheme for automated video abstraction based on unsupervised cluster-validity analysis. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(8): 1280-1289. [doi:10.1109/76.809162]
- 4 Shahraray B, Gibbon DC. Automatic generation of pictorial transcripts of video programs. Proceedings of the SPIE Volume 2417, Multimedia Computing and Networking. San Jose, CA, USA. 1995, 2417: 512-518.
- 5 Liu GT, Wen XM, Zheng W, *et al.* Shot boundary detection and keyframe extraction based on scale invariant feature transform. Proceedings of the 8th IEEE/ACIS International Conference on Computer and Information Science. Shanghai, China. 2009. 1126-1130.
- 6 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. 2016. 779-788.
- 7 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. arXiv: 1612.08242v1, Dec. 2016.
- 8 曹长青.基于内容的视频检索中关键帧提取算法研究[硕士学位论文].太原:太原理工大学,2013.
- 9 Rao AB, Srihari RK, Zhang ZF. Spatial color histograms for content-based image retrieval. Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence. Washington, DC, USA. 1999. 183-186.
- 10 Wang XY, Han TX, Yan SC. An HOG-LBP human detector with partial occlusion handling. Proceedings of the 12th International Conference on Computer Vision. Kyoto, Japan. 2009. 32-39.
- 11 张萌.视频检索中关键帧的提取和特征匹配的研究[硕士学位论文].北京:北京邮电大学,2012.
- 12 Cao CQ, Chen ZH, Xie G, *et al.* Key frame extraction based on frame blocks differential accumulation. Proceedings of the 24th Chinese Control and Decision Conference. Taiyuan, China. 2012. 3621-3625.
- 13 张慧,张海滨,李琼,等.基于人类视觉系统的图像感知哈希算法.电子学报,2008,36(12A):30-34.
- 14 卞涛,罗泽,马永征.基于Hadoop的分布式视频处理.科研信息化技术与应用,2016,7(4):61-69.
- 15 周文琼,王乐球,叶玫.云环境下Hadoop平台的作业调度算法.计算机系统应用,2014,23(5):177-181.
- 16 Johnson J, Douze M, Jégou H. Billion-scale similarity search with GPUs. arXiv: 1702.08734v1, Feb. 2017.