



视频大数据 智慧计算方法

◎ 朱定局 著

 科学出版社

视频大数据智慧计算方法

朱定局 著

国家科技支撑计划课题(No. 2015BAH42F01)

国家社会科学基金重大项目(No. 14ZDB101)

国家自然科学基金(No. 61105133)

科学出版社

北 京

内 容 简 介

本书以空间为纬度来研究视频大数据中图像的识别、重建,以时间为经度来研究视频大数据中帧间的切分、转码,以时空为经纬度来研究视频大数据内容的监控、审查、分类。本书的原创性在于:提出并研究给出了视频大数据智慧计算方法,具体又提出并研究给出了视频大数据空间智慧计算方法、视频大数据时间智慧计算方法、视频大数据时空编解码智慧计算方法、视频大数据时空监控智慧计算方法、视频大数据时空审查智慧计算方法、视频大数据时空分类智慧计算方法。

本书可作为大学和院所相关专业的教学用书和研究用书,同时还可以供政府部门和企事业单位参考。

图书在版编目(CIP)数据

视频大数据智慧计算方法/朱定局著. —北京:科学出版社,2015.3

ISBN 978-7-03-043461-6

I. ①视… II. ①朱… III. ①视听资料-数据处理-计算方法
IV. ①G255.73-32

中国版本图书馆CIP数据核字(2015)第087328号

责任编辑:杨州燕 傅保军 / 责任校对:张小霞

责任印制:赵 博 / 封面设计:北京图阅盛世文化传媒有限公司

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

新科印刷厂印刷

科学出版社发行 各地新华书店经销

*

2015年3月第 一 版 开本:787×1092 1/16

2015年3月第一次印刷 印张:12

字数:285 000

定价:69.00元

(如有印装质量问题,我社负责调换)

前 言

视频大数据是一种运动的图像集。视频大数据中的每一帧图像都代表了一个空间集合(如第2章),而帧间的变化就代表了时间,很多帧间变化的集合就构成了一个时间集合(如第3章)。从视频中抓出任意一帧图像,那么这个图像不能称为视频;而同一个空间的多个时相点的图像集,即使不将其合成为一个视频文件,这个有序的多时相点的图像集(如5.3节),也可以被看成视频,因为它具备和视频同等的特性。同时,视频大数据是图像、声音、文字的有机结合,所以是一种多媒体,本书在3.2节及4.2节中直接将视频扩展称为多媒体,便是取意于此。

第2章,先抛开时间,只从空间的角度来看视频大数据,即图像大数据。第2章提出使得视频大数据空间计算更加智慧的方法,2.1节利用识别知识库,2.2节利用图像样本库,2.3节利用三维模型库,加上自动搜索和匹配的算法,使得视频大数据空间计算(更直白地说就是图像的识别重建)更加自动化、普适化、精准化,再加上并行计算,又可使之快速化。使得在视频大数据空间计算(更直白地说就是图像大数据的识别重建)方面,计算机更加接近人脑对图像的处理重建能力,从而更接近人的智慧。

第3章,则抛开空间,只从时间的角度来看视频大数据,即各帧间的关系。第3章提出使得视频大数据时间计算更加智慧的方法,利用更有效的时间划分方法,使得一般视频和实时视频都能够更加快速又不影响质量地处理。通常来说,速度和质量总是矛盾的,难以兼得的,要想使之兼得,就要在方法中融入智慧。3.1节利用切片与转码所需的时间差,使得并行转码的速度能够最大化,同时通过先切分后合并的方式,最终得到的视频在质量上和用现有转码技术得到的视频质量上差异很小。3.2节利用重叠边界,使得视频分片转码的效果进一步提高,当然牺牲了一点速度,但由于重叠边界占整个视频的比例非常小,所以给速度带来的损失,也非常小,但给质量却带来了极大的提高,使得并行切片转码后得到的视频质量和用现有转码技术得到的视频在质量上不再有差异。3.3节给出了一种确定第一次切片和以后各次切片的时间长短的算法,确保了视频直播和实时视频的用户端播放的节目视频不会中断,从而既保证了速度,又保证了质量。第3章的方法充分体现速度与质量兼顾的智慧。

视频大数据的编解码是视频大数据加工、存储、传输中非常关键的环节。第3章,已经从时间划分的角度涉及视频大数据编解码。但第3章没有考虑到视频大数据编解码涉及终端环境、网络环境、服务器环境,第4章,针对这些环境因素,给出更加智慧的方法,4.1节使得切片转码能够适应不同外在环境,4.2节能够根据环境自动确定切分的粒度,得到视频大数据转码的速度和质量能得以进一步提高。4.3节借助图像库,使得视频大数据编解码的速度能得以更大的跨越,且不影响视频大数据编解码的质量。第4章,在方法中充分融入自适应和多方协同的智慧。

视频大数据的应用主要包括基于时空对视频大数据内容的监控(第5章)、审查(第6章)、

分类(第7章)。对视频大数据的监控,是在事先知道特定目标的前提下,在视频大数据内容中寻找特定的目标,如小偷(5.1)、异常行为(5.2)、违章建筑(5.3)。对视频大数据的审查,是在事先知道特定审查目标范围的前提下,在视频大数据内容中寻找符合特定审查目标范围的内容,如黄色视频、暴力视频。对视频的分类,是在事先知道划分标准和类型的前提下,从视频内容中找出某种类型的视频,或按照类型对视频进行切分。

第5章,在基于时空对视频大数据内容的监控方法中,融入协同、集成的智慧。通过视频信息与现实世界中对象的联动,扩展视频技术的应用范围;通过多类设备的同步,增加视频监控的灵活性;通过监测与仿真的结合,提高视频监控的可视化效果。

第6章,在基于时空对视频大数据内容的审查方法中通过冗余、融合、样本知识使得方法更加智慧。通过冗余增加系统处理与用户操作间的并发性(6.1节);通过融合图像、声音、文字等多类特征结合视频类型范围,提高审查的准确率(6.2节);通过样本知识来对视频进行审查,提高审查方法的普适性(6.3节)。

第7章,在基于时空对视频大数据内容的分类方法中通过相关性使得方法更加智慧。通过视频与各类事物的相关性来进行分类(7.1节),无需更多的专家知识;通过广告内容与视频内容的相关性,使广告的投放更为精准(7.2节);通过视频内部的相关性,对视频进行切分(7.3节);通过新闻视频与播报的相关性,对新闻视频进行切分(7.4节)。

本书的主要创新如下:

(1) 提出并研究给出了视频大数据智慧计算方法,具体包括视频大数据空间智慧计算方法、视频大数据时间智慧计算方法、视频大数据时空编解码智慧计算方法、视频大数据时空监控智慧计算方法、视频大数据时空审查智慧计算方法、视频大数据时空分类智慧计算方法。

(2) 提出并研究给出了视频大数据空间智慧计算方法,具体包括识别知识库下物体识别的方法、图像处理中物体自动识别并三维重建的方法、三维模型库下二维图像识别和物体重建的方法。

(3) 提出并研究给出了视频大数据时间智慧计算方法,具体包括视频分片并行转码的方法、基于重叠边界的多媒体切片转码的方法、直播视频和实时视频的分片转码的方法。

(4) 提出并研究给出了视频大数据时空编解码智慧计算方法,具体包括视频自适应转码的方法、基于自动切分的多媒体并行的方法、基于图像库的视频和图像的编码解码的方法。

(5) 提出并研究给出了视频大数据时空监控智慧计算方法,具体包括智慧城市视频监控的方法、多类视频设备同步监控的方法、自动监测仿真及其并行化处理的方法。

(6) 提出并研究给出了视频大数据时空审查智慧计算方法,具体包括视频数据云审查的方法、深入融合视频审查的方法、基于感知知识库的视频内容审查的方法。

(7) 提出并研究给出了视频大数据时空分类智慧计算方法,具体包括视频分类的方法、视频内容中自适应投放广告的方法、基于重复特征和相似特征的视频自动拆分的方法。

由于作者水平有限,书中难免有不妥甚至错误之处,恳请读者批评指正。

朱定局

2015年1月2日于华南师范大学

目 录

前言

第 1 章 视频大数据智慧计算绪论	1
1.1 视频大数据的特点	1
1.2 视频大数据对智慧计算的需求	2
1.3 视频大数据智慧计算的分类	4
第 2 章 视频大数据空间智慧计算方法	6
2.1 识别知识库下物体识别的方法	6
2.1.1 现有物体识别方法的不足	6
2.1.2 基于识别知识库的新方法的思路	7
2.1.3 基于识别知识库的新方法的实施方式	8
2.2 图像处理中物体自动识别并三维重建的方法	14
2.2.1 现有物体识别重建方法的不足	14
2.2.2 基于图像样本的新方法的思路	17
2.2.3 基于图像样本的新方法的实施方式	18
2.3 三维模型库下二维图像识别和物体重建的方法	28
2.3.1 现有图像识别重建方法的不足	29
2.3.2 基于三维模型库的新方法的思路	33
2.3.3 基于三维模型库的新方法的实施方式	34
第 3 章 视频时间智慧计算方法	44
3.1 视频分片并行转码的方法	44
3.1.1 现有视频转码方法的不足	44
3.1.2 基于分片并行的新方法的思路	45
3.1.3 基于分片并行的新方法的实施方式	45
3.2 基于重叠边界的多媒体切片转码的方法	47
3.2.1 现有多媒体切片转码方法的不足	47
3.2.2 基于重叠边界的新方法的思路	48
3.2.3 基于重叠边界的新方法的实施方式	48
3.3 直播视频和实时视频的分片转码的方法	53
3.3.1 现有直播视频和实时视频转码方法的不足	53
3.3.2 基于分片的新方法的思路	53
3.3.3 基于分片的新方法的实施方式	54
第 4 章 视频时空编解码智慧计算方法	57
4.1 视频自适应转码的方法	57

4.1.1	现有视频转码方法的不足	57
4.1.2	自适应的新方法的思路	57
4.1.3	自适应的新方法的实施方式	59
4.2	基于自动切分的多媒体并行的方法	66
4.2.1	现有多媒体并行方法的不足	66
4.2.2	基于自动切分的新方法的思路	66
4.2.3	基于自动切分的新方法的实施方式	69
4.3	基于图像库的视频和图像的编码解码的方法	74
4.3.1	现有视频图像编解码方法的不足	75
4.3.2	基于图像库的新方法的思路	75
4.3.3	基于图像库的新方法的实施方式	76
第5章	视频时空监控智慧计算方法	85
5.1	智慧城市视频监控的方法	85
5.1.1	现有城市视频监控方法的不足	85
5.1.2	自动化的新方法的思路	85
5.1.3	自动化的新方法的实施方式	87
5.2	多类视频设备同步监控的方法	95
5.2.1	现有视频监控方法的不足	95
5.2.2	多类设备同步的新方法的思路	95
5.2.3	多类设备同步的新方法的实施方式	96
5.3	自动监测仿真及其并行化处理的方法	100
5.3.1	现有监测仿真方法的不足	101
5.3.2	有机结合的新方法的思路	109
5.3.3	有机结合的新方法的实施方式	111
第6章	视频时空审查智慧计算方法	131
6.1	视频数据云审查的方法	131
6.1.1	现有视频数据审查方法的不足	131
6.1.2	基于云的新方法的思路	131
6.1.3	基于云的新方法的实施方式	132
6.2	深入融合视频审查的方法	134
6.2.1	现有视频审查方法的不足	134
6.2.2	深入融合的新方法的思路	135
6.2.3	深入融合的新方法的实施方式	136
6.3	基于感知知识库的视频内容审查的方法	144
6.3.1	现有视频内容审查方法的不足	145
6.3.2	基于感知知识库的新方法的思路	145
6.3.3	基于感知知识库的新方法的实施方式	147

第7章 视频时空处理智慧计算方法	159
7.1 基于熵的视频分类的方法	159
7.1.1 现有视频分类方法的不足	159
7.1.2 基于熵的新方法的思路	159
7.1.3 基于熵的新方法的实施方式	161
7.2 视频内容中自适应投放广告的方法	167
7.2.1 现有视频广告投放方法的不足	167
7.2.2 自适应的新方法的思路	167
7.2.3 自适应的新方法的实施方式	168
7.3 基于重复特征和相似特征的视频自动拆分的方法	174
7.3.1 现有视频拆分方法的不足	174
7.3.2 基于重复相似特征的新方法的思路	174
7.3.3 基于重复相似特征的新方法的实施方式	175
7.4 基于播报特征及播报员知识库的新闻自动拆条的方法	177
7.4.1 现有新闻视频拆条方法的不足	178
7.4.2 基于播报特征及播报员知识库的新方法的思路	178
7.4.3 基于播报特征及播报员知识库的新方法的实施方式	178
参考文献	183
后记	184

第 1 章 视频大数据智慧计算绪论

1.1 视频大数据的特点

视频大数据有三大特点：数据量大、结构复杂、全数据，如图 1.1 所示。

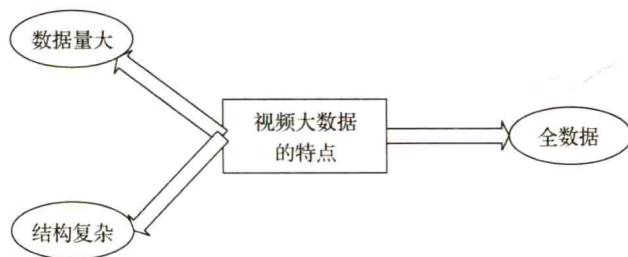


图 1.1 视频大数据的特点

视频大数据的第一个特点是数据量大，而且视频大数据的数据量增长速度非常快。每天有无数的人在向优酷等在线视频网站上传各种视频，每天都有无数的监控器、摄像头在录制各种各样的实时视频，每天电视台、电影院都在播放着不断更新的视频……。我们的世界是可视的、动态的，整个世界最原始的状态就是一个连续不断的、将会延续无数亿年的视频。

随着视频录制设备的普及，特别是智能手机也具备了录制视频的能力，现在已经进入视频大数据的时代。人们喜欢录制视频来记录生活中的点点滴滴，学校喜欢用视频来记录教师上课的实况，政府喜欢用视频来监控社会的动态。一个文本文件或图像文件或声音文件，其数据量都一般在 M 级，但一个视频文件数据量一般都在 G 级，而有的高清视频能达到 T 级。每个视频都含有非常丰富的信息，都要占用大量的存储空间，而视频的数量又非常庞大，这就使得视频成为各种格式的数据中数据量最为突出的大数据。

同时，随着互联网特别是移动互联网的发展，越来越多的视频被上传到网上进行分享、转码、加工、重组，从而使得从无数的原始视频中，又衍生出更多数据量更大的目标视频，这又进一步增加了视频大数据的数据量。

视频大数据的第二个特点是结构复杂。视频数据相对于文本、图像、声音这些单一的数据类型而言，是结构最为复杂的一种数据。视频大数据是一种集图像、声音、文字于一体的大数据，也是一种集时间维、空间维、本性维于一体的大数据。正是因为其组成成分的复杂性，以及不同维之间的关联性，使得其大数据的特性更为突出。正是由于视频结构的复杂性，所以以不同的方式来组织视频数据就形成了视频的不同编码格式和封装格式。视频的编码格式和封装格式之多远远超过了文本、图像、声音这些类型的文件数据，这也从一个侧面反映了视频结构的复杂性。不但视频内部的结构复杂，由于视频转码、视频加

工、视频重组、视频挖掘等技术的应用,使得不同视频数据之间也存在着千丝万缕的联系,形成了视频数据之间结构的复杂性。

视频大数据的第三个特点是全数据。视频所见即所得,是对现实世界的真实记录和复制,这一点不同于文本数据。因为文本数据是人类或人类程序对客观世界的记录,其中有人类对客观世界的抽象及取舍。全数据在数据的维度上没有任何损失,所以视频大数据的挖掘价值更大。视频大数据相对于文本等类型的大数据而言,其全数据性更为明显。当然,这也是相对的,因为视频的分辨率是有限的,而且大部分视频不是三维的、更不是全息的,因此视频也无法完全地复制现实,所以说视频大数据只是相对的全数据,但随着视频技术、存储技术的发展,视频的分辨率会越来越高,而且会朝着三维、全息的方向发展,所以视频大数据的全数据性会越来越高。

1.2 视频大数据对智慧计算的需求

大数据的价值在于其中所蕴含的知识,而大数据中的知识只有依靠智慧计算才能充分地发现和利用。因为大数据不同于传统的数据,大数据是未经采样和加工的全数据,因此其数据质量远远低于传统数据,而数据复杂性远远高于传统数据,这就对处理数据的计算方法提出了更高的要求,因此需要一种针对大数据的智慧计算方法。

特别是视频大数据对智慧计算有着更高的要求,因为视频大数据是最为难以处理的大数据类型之一。传统处理大数据的框架如 hadoop,比较擅长于处理文本大数据,但在处理视频大数据时就有些力不从心,因为视频大数据的数据内部、数据之间的关系非常复杂,同时其数据量也不是文本大数据可以同日而语的。所以急需研究专门针对视频大数据的更为智慧的计算方法,来针对视频大数据的结构复杂性、海量性、全数据性进行高效的处理,得到预期的结果和效果。

从应用的角度来看,视频大数据对智慧计算的需求包括七大需求:视频大数据识别的需求、视频大数据重建的需求、视频大数据重组的需求、视频大数据转码的需求、视频大数据监控的需求、视频大数据审查的需求和视频大数据处理的需求,如图 1.2 所示。

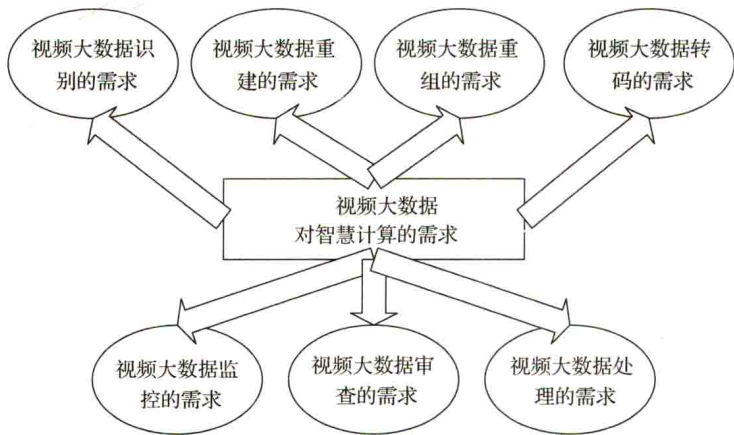


图 1.2 视频大数据对智慧计算的需求

视频大数据对智慧计算的需求,第一体现在对视频大数据识别的需求。因为视频如果不被识别,则毫无意义,但视频识别对计算机来说不是一件很容易的事情,虽然我们人类每天都在看电影、看电视,识别视频对于我们人类来说,非常容易。因此,在视频识别这一方面,计算机当前发展水平还远远不如人脑。但视频大数据的海量数据性,使得无法完全由人工来进行识别,所以急需一种更为智慧的计算来对视频进行自动高效的识别。

视频大数据对智慧计算的需求,第二体现在对视频大数据重建的需求。视频重建是基于视频识别所得到知识所进行的创造性活动。视频本身是源于现实,而视频重建的结果是对现实的重塑。因此,视频重建是对视频的有效利用。但视频重建比视频识别对计算机而言更为困难,因为视频重建是一种创造性的活动,而计算机最不擅长做的事情就是创造性的活动。所以当前视频重建一般都是人工来做的。例如,根据动画形象进行三维建模,都是由3D模型设计师进行手工设计。但是当面临视频大数据时,如果有大量的视频需要被重建,则无法完全由人工来进行重建,所以急需一种更为智慧的计算来对视频进行自动高效的重建。

视频大数据对智慧计算的需求,第三体现在对视频大数据重组的需求。因为视频是一个图像、声音、文字的时间序列,这个序列可以分片,可以合并,统称为重组。现实生活中对视频重组的需求是巨大而迫切的。例如,不同的人可能关心不同视频中的某个主题和片段,那么就需要把这些主题和片段切分出来并进行合并后交付给不同的用户。这个例子讲的是用户的需求。还有一种更大的需求,是对视频大数据处理的需求,因为为了加快大视频的处理速度,则需要采用分而治之的云计算模式进行处理,所以需要对视频进行分片后再合并。但视频的分片、合并并不像文本那么容易和简单,如果采用人工交互式地进行分片,那么就会影响对视频大数据的整体处理的自动化,因此急需一种更为智慧的计算来对视频进行自动高效的分片。

视频大数据对智慧计算的需求,第四体现在对视频大数据转码的需求。由于视频结构及其应用的复杂性,所以以不同的方式来组织视频数据就形成了视频的不同编码格式和封装格式。视频的编码格式和封装格式之远远地超过了文本、图像、声音这些类型的文件数据。特别是随着电视网、电脑网、手机网的互联互通,使得三网之间视频格式的互转变得非常频繁和泛化,转码可以使得视频大数据在三网之间无碍地流通,这就使得三网上的视频大数据对转码的需求与日俱增。当人们用不同手机、在不同带宽下观看或下载同一个视频源时,所得到的视频的编码格式和封装格式可能都会不同,在这些时候,都有转码的发生。但随着视频大数据的数据量的激增,传统上依靠改进转码硬件性能及转码串行算法的途径已经无法满足需求了,其结果必然会影响视频下载和在线播放的速度。因此急需一种更为智慧的计算来对视频进行自动高效的转码。

视频大数据对智慧计算的需求,第五体现在对视频大数据监控的需求。高速公路上、个人家庭中、公司中、超市中、ATM机旁等地方,监控摄像头无处不在,但这些“眼睛”真的在发挥作用吗?大部分的监控摄像头只能起到事后取证的作用,而无法像人眼那样看到异常立即报警。当前视频监控大数据没有做到实时的监控和处理,也没有做到事后的海量挖掘和分析。因此急需一种更为智慧的计算来对视频进行自动高效的监控。

视频大数据对智慧计算的需求,第六体现在对视频大数据审查的需求。现在网络上黄色视频、暴力视频严重危害了未成年人的身心健康,如果靠人工网络警察去一个个地看、一个个地查,总是有漏网之鱼的视频,因为视频的数量实在太大了,而且每一个视频人工也难以从头看到尾,即使从头看到尾也难以对非法内容毫无遗漏。因此迫切需要一种更为智慧的计算来对视频进行自动高效的审查。

视频大数据对智慧计算还有很多其他需求,统称为对视频大数据处理的需求,包括对视频分类的需求、对视频推荐的需求、对视频拆条的需求等。这些需求也随着视频大数据的数据量的攀升而越来越需要一种更为智慧的计算来对视频进行自动高效的处理。

1.3 视频大数据智慧计算的分类

视频大数据智慧计算方法,如果按照需求进行分类,则可以分为:视频大数据识别的方法、视频大数据重建的方法、视频大数据重组的方法、视频大数据转码的方法、视频大数据监控的方法、视频大数据审查的方法和视频大数据处理的方法,如图 1.3 所示。

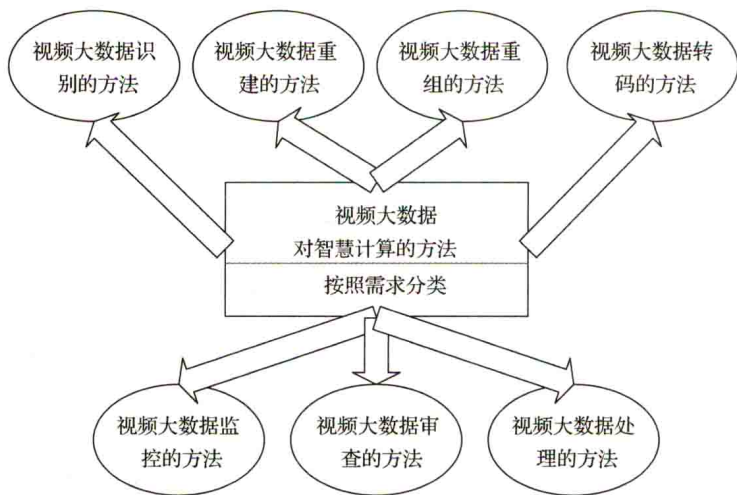


图 1.3 视频大数据智慧计算方法按照需求的分类

但如果从结合视频大数据的时空性三个维度来分类,则视频大数据智慧计算可以分为:视频大数据空间智慧计算方法、视频大数据时间智慧计算方法、视频大数据时空智慧计算方法,如图 1.4 所示。这里没有提到本性维。因为本性维是最基本的维,不管是哪种类型的计算,本性维都默认存在并且参与其中,因为如果脱离了视频的本性来讨论视频的时间性、空间性和时空性是没有意义,“皮之不存,毛将焉附?”

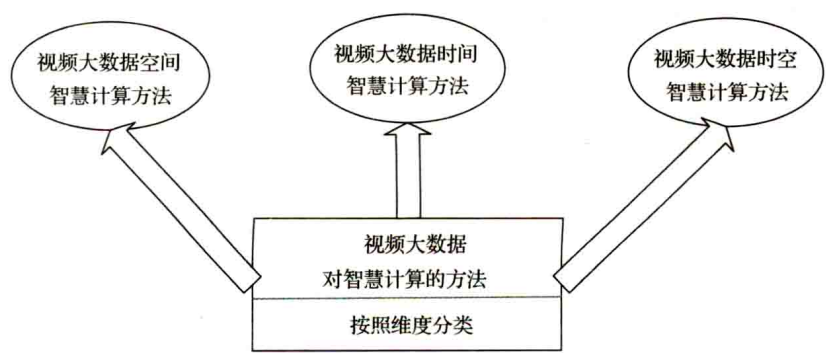


图 1.4 视频大数据智慧计算方法按照维度的分类

第2章 视频大数据空间智慧计算方法

如果不考虑时间维,则视频成为了一幅幅静止的图像,所以只考虑空间维的视频大数据实质上是图像大数据,所以视频大数据空间智慧计算方法实质上是图像大数据智慧计算方法。之所以要将视频大数据的空间维拿出来单独研究,是因为的确有不少视频实际应用(如视频大数据中图像的识别、重建)中只涉及视频的空间维,所以研究视频大数据空间智慧计算方法有其实际意义。将视频大数据的空间维拿出来单独研究的另一个原因是,只有将视频大数据的空间维(各幅图像)的智慧计算方法研究清楚了,才能与视频大数据的时间维(各帧)进行结合,从而为视频大数据的时空智慧计算方法奠定基础。

本章提出了使得视频空间计算更加智慧的方法,2.1节利用识别知识库,2.2节利用图像样本库,2.3节利用三维模型库,加上自动搜索和匹配的算法,使得视频空间计算(就是图像的识别重建)更加自动化、普适化、精准化,再加上并行计算,又可使之快速化,从而使得在视频空间计算方面,计算机更加接近人脑对图像的处理重建能力,因而更接近人的智慧。

2.1 识别知识库下物体识别的方法

本方法提供一种基于识别知识库的识别方法,用于识别多个类型目标,所述识别方法包括:从关联索引表中寻找与兴趣目标最接近的预期目标样本;根据关联索引表调用相应的预期目标特征提取规则,并根据该特征提取规则确定匹配窗口;根据关联索引表调用相应的特征提取规则,由所述匹配窗口对待识别对象进行特征提取;调用对应的特征比较规则并调用关联的特征比较程序,根据所述特征比较程序将预期目标样本的特征与匹配窗口中的对象的特征进行比较,并将特征比较的结果记为相似度;根据所述相似度确定识别结果。本方法提供的基于识别知识库的识别方法,预先在识别知识库中存储若干类型目标样本,实现识别多个类型目标。

2.1.1 现有物体识别方法的不足

现代信息处理技术的发展使得自动识别技术成为可能,比较常见的例子就是,将计算机信息处理技术应用到识别技术领域,使得无需人为操作从而实现自动化识别兴趣目标。

目前,自动识别方法有条码识别、IC卡识别、射频识别(RFID)、声音识别、视觉识别以及光学字符识别等。

识别的过程就是从对象中识别出目标,对所述对象中目标进行识别得到的结果被称为识别结果。

识别即是通过目标的特征将目标从对象中分离出来并判断出目标类型和目标属性的过程。识别的对象可以是图像、声音、视频、波谱、文本等中的一种类型或其混合。例如,

识别一个人,可以同时识别他的语音(声音识别)、面部相貌(图像识别)。再如识别草,可以同时识别它的可见光谱(图像识别)和红外光谱(波谱识别)。识别的输入为对象,输出为目标。

样本,是指用于识别的参照目标,其一般预先存储在所述识别系统中的数据库(也称识别知识库)。

特征,是指可以标志目标的属性。例如,图像特征是指可用以标志图像对象中目标的属性。

进行识别任务时,首先识别系统执行一个目标特征提取任务以获得一个特征结果;然后将特征结果与识别知识库匹配获得识别结果。

所述目标特征提取任务就是按照一定的目标特征提取规则(也称为目标特征提取程序)采集并进行计算。

以条码识别为例,条码是由一组规则排列的条、空以及相应的数字组成,条码阅读器识读这种由多个条、空组成的一组数据编码并译成一个二进制数或者十进制数,进而与数据库匹配获得需要的兴趣目标。

再以视觉识别为例,视觉识别系统获取一个视觉图像,通过一个特征抽取和分析的过程,自动识别出用于限定目标的标志、字符、编码或其他能够作为识别目标的基础呈现在视觉图像内的特征。

就声音识别技术而言,现有技术中一种典型的声音识别技术,就是用于计算机汉字输入的语音输入系统,识别系统采集语音对象获得目标特征并对照数据库识别出对象,但是这种语音输入系统单凭对语音对象的采集和匹配,其误码率很高。

不管采用何种识别方法,目前的识别系统都具有一个识别软件或者识别程序,其只能识别一种类型的目标。例如,只能识别条码、声音或者图像中的一种,而无法同时将输入的条码、声音、图像目标均识别出来。再如,只能识别一个人的语音、相貌或者指纹中的一种,而无法对人的语音、相貌以及指纹同时进行识别。

很明显,目前的识别系统只具有一个目标特征提取规则,只适合于单一类别的目标进行识别,如前述的条码识别,其识别系统只能识别对象中的条码,而对于对象中其他图像或者图案没有识别能力,很明显,单一类别的识别系统其识别精度较低。再如,一些安防设备如门禁系统,通过识别对象语音而决定是否打开门,其识别系统只识别对象的声音,这样的识别系统容易受到干扰就会出现识别错误,因此,其识别能力比较低。

因此,目前单一类别的识别从一定程度上可以识别出兴趣目标,但是识别精度不够高。

可见,现有技术还存在缺陷,有待于改进和发展。

2.1.2 基于识别知识库的新方法的思路

本方法的目的在于,提供一种基于识别知识库的识别方法,用于识别若干个类别的目标。

本方法的技术方案如下:

一种基于识别知识库的识别方法,其中,所述识别知识库包括一个预期目标样本库、一个预期目标样本特征库、一个特征提取规则库、一个特征提取程序库、一个特征比较规

则库和一个关联索引表,所述识别方法包括:

- A. 从关联索引表中寻找与兴趣目标最接近的预期目标样本;
- B. 根据关联索引表调用相应的预期目标特征提取规则,并根据该特征提取规则确定匹配窗口;
- C. 根据关联索引表调用相应的特征提取规则,由所述匹配窗口对待识别对象进行特征提取;
- D. 调用对应的特征比较规则并调用关联的特征比较程序,根据所述特征比较程序将预期目标样本的特征与匹配窗口中的对象的特征进行比较,并将特征比较的结果记为相似度;
- E. 根据所述相似度确定是否匹配。

所述的基于识别知识库的识别方法,其中,所述预期目标样本库、预期目标样本特征库、特征提取规则库、特征提取程序库和特征比较规则库采用分级结构,其分别包括若干类型目标的预期目标样本子库、预期目标样本特征子库、特征提取规则子库、特征提取程序子库和特征比较规则子库。

所述的基于识别知识库的识别方法,其中,所述步骤 E 中,所述相似度较小时,将所述待识别对象的特征更新至所述预期目标样本特征库中。

所述的基于识别知识库的识别方法,其中,所述识别知识库进一步包括一个目标知识库,其预先存储有对应目标样本的附加知识,所述步骤 E 中,相似度较大时,所述待识别对象与样本匹配,则输出目标样本特征及对应的目标样本的附加知识。

所述的基于识别知识库的识别方法,其中,步骤 A 中,如果未找到匹配预期目标样本,则执行步骤 A1,然后返回,

所述步骤 A1 为,将兴趣目标特征存至所述识别知识库,并更新识别知识库;

所述的基于识别知识库的识别方法,其中,所述步骤 A1 与步骤 B 并行进行。

所述的基于识别知识库的识别方法,其中,所述并行的实现方式为调用并行库、使用编译指导语句、并行编译器、使用多线程、使用任务的动态调度、使用分布式或网格基础设施。

本方法提供一种基于识别知识库的识别方法,与现有技术相比,所述识别知识库预先存储若干类型目标所对应的目标样本特征、目标特征提取规则、目标特征提取程序和目标特征比较程序等,可以实现对多个类型的目标进行识别,且所述识别方法可以对所述识别知识库进行更新,所述更新识别知识库和识别任务并行处理,不仅节省处理时间而且提高识别精度。

2.1.3 基于识别知识库的新方法的实施方式

图 2.1 为本方法方案的识别知识库的示意图。所述识别知识库包括一个预期目标样本库、一个预期目标样本特征库、一个预期目标知识库、一个特征提取规则库、一个特征提取程序库、一个特征比较规则库和一个关联索引表。

根据实际识别对象的需要,可以预先在所述预期目标样本库存储所需要的子库。所述子库不限,可为现有技术中普遍使用的波谱目标样本子库、声音目标样本子库、视频目

标样本子库、图像目标样本子库、化学目标样本子库、生物目标样本子库、物理目标样本子库等。

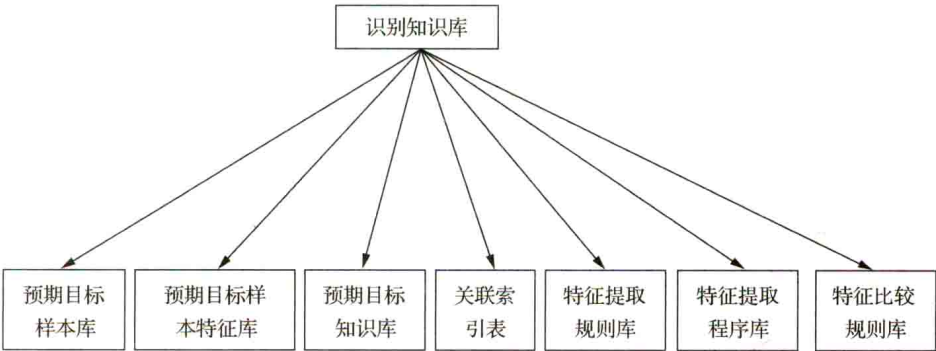


图 2.1 识别知识库的示意图

可以理解的是,所述预期目标样本库根据实际识别对象的需要来建立其需要的子库,如果实际识别对象为人,则所述预期目标样本库包括一个语音目标样本子库、一个面部图像目标样本子库等,当然可以理解的是,所述识别人子库不限于此,还可以进一步包括其他能够识别人身份的目标样本子库。如果实际识别知识库用于卫星识别用途,则所述预期目标样本库包括:一个预期水体样本子库、一个预期道路样本子库、一个预期建筑样本子库、一个预期树木样本子库、一个预期桥梁样本子库、一个地质影像目标样本子库等。所述目标样本子库由现有技术可以实现,所述预期样本库根据实际需要选择现有技术中已知的目标样本特征建立所述目标样本子库。

图 2.2 为图 2.1 的预期目标样本库的详细示意图。所述预期目标样本库包括若干不同类型的预期目标样本子库。

所述若干不同类型的预期目标样本子库具有多级结构,按照识别介质划分若干第一级子库,所述若干第一级子库为波谱目标样本子库、声音目标样本子库、视频目标样本子库、图像目标样本子库、化学目标样本子库、生物目标样本子库、物理目标样本子库等。

每个第一级子库按照采集方式分成若干第二级子库,图 2.2 给出第一级子库的图像目标样本子库按照采集方式分成若干第二级子库,分别为数码相片目标子库、医学影像目标样本子库、卫星影像目标样本子库、地质影像目标样本子库等。

每个第二级子库按照其目标类型分成若干第三级子库,图 2.2 给出第二级子库的卫星影像目标样本子库按照其目标类型分成若干个第三级子库,分别为预期水体样本及其子库、预期道路样本及其子库、预期建筑样本及其子库、预期树木样本子库、预期桥梁样本子库等。以此类推第四级子库、第五级子库等,其原理不再赘述。

所述预期目标样本特征库,按照所述预期目标样本库的多级结构,同样具有多级结构,其每一级的每一个目标样本特征子库对应于所述预期样本库的相应一级的相对应的一个目标预期样本子库。

同理,所述特征提取规则库、特征提取程序库和特征比较规则库,按照所述预期目标样本库的多级结构的划分方式,分别具有多级结构,其每一级的每一个子库(特征提取规