监控视频高效编码与智能分析

黄铁军　田永鸿　黄庆明　胡瑞敏　著

科学出版社

监控视频高效编码与智能分析

黄铁军　田永鸿　黄庆明　胡瑞敏　著

科学出版社
北京

内 容 简 介

千千万万的监控摄像头构成了时刻观测物理世界和人类社会的“视听感知网”，成为全球信息基础设施的重要组成部分。监控视频已成为全球大数据中体量最大的部分，对其高效编码能够节省巨额的存储成本和传输成本。监控视频数据中蕴含了丰富的信息，对其智能分析具有巨大的现实意义。本书在综述国内外相关研究进展的基础上，结合最新研究进展和标准制定，详细介绍了基于背景建模的高效视频编码方法和视觉对象的检测、跟踪、分析和识别技术，是相关研究和技术开发的重要参考资料。

本书可供图像视频处理分析和识别、视频监控等计算机应用和人工智能领域的科研人员、教师、研究生和工程技术人员阅读参考。

图书在版编目(CIP)数据

监控视频高效编码与智能分析 / 黄铁军等著. —北京：科学出版社，2016.4

ISBN 978-7-03-045662-5

Ⅰ. ①监… Ⅱ. ①黄… Ⅲ. ①监视控制-视频编码-研究 Ⅳ. ①TN762

中国版本图书馆 CIP 数据核字(2015)第 218517 号

责任编辑：任 静 / 责任校对：郭瑞芝

责任印制：张 倩 / 封面设计：迷底书装

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

新科印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 4 月第 一 版 开本：720×1 000 1/16

2016 年 4 月第一次印刷 印张：27 1/2

字数：539 000

定价：118.00 元

(如有印装质量问题，我社负责调换)

序

对数字图像处理的研究是从医学图像、文字图像、雷达图像等开始的，至今已经有六十多年的历史。近三十年来，由于数字电视和多媒体的发展，作为图像序列的视频逐步成为处理的主体，其中数字视频编码与数字视频分析已成为数字图像处理领域最热门的研究课题。但是，视频编码是基于信号处理理论的，处理对象是像素和像素块，使用的数学工具主要是正交变换、滤波器设计与熵编码；而视频分析是基于模式识别理论，处理对象是特征，使用的数学工具主要是特征提取与分类器设计以及机器学习等。因此，视频编码与视频分析一直像两条平行的铁轨，尽管共同承载视频处理的列车，但互不交叉，互不干预。

进入21世纪后，由于城市智能管理和防范恐怖袭击等需要，视频监控系统广泛部署，每天都在产生海量的视频记录数据，为破案和事后分析提供了很好的数据来源。据IDC 2012年的报告，监控视频已占全球大数据的一半以上。然而，监控视频直接采用传统视频编码技术存在两个问题：其一，现有视频编码技术与标准主要是为了满足数字电视和电影产业的需求而产生的，即满足场景不固定假设（电影电视场景都经常更换，电影要求一个场景持续时间平均不超过半分钟左右，以免观众视觉疲劳，注意力下降），由于当前的视频编码技术标准对于背景频繁变化的支持很强，但对于背景长时间不变的支持很弱，因此直接使用传统编码技术进行监控视频编码并不是最合理的；其二，现有视频监控系统对视频的处理过程是先编后解再识别，即摄像头采集的视频数据经过编码压缩后传输到数据中心/分中心进行存储，需要识别分析时再把这些压缩的视频数据解压后再分析与识别，这种做法由于在编码过程中过度压缩可能会丢掉一些有用特征，从而降低分析精度和识别率。传统视频编码在判断视频编码到解码的过程是否可被接受，采用主观评测方法对视频图像质量进行评分，以测试编码对视频的损伤，即用人的眼睛确认编码前后是否能看得出变化，看不出变化为优秀，看出的变化越大则质量越差。换句话说，判断编码质量好坏是利用人眼在时域上进行的。但是很多视频分析识别算法是在频域上进行的，其标准与数字电视和电影领域对视频质量的评价要求可能完全不同。因此，采用主观评测判断视频质量对于视频监控领域来说是不完全的，甚至南辕北辙，需要从根本上改变。

如何解决监控视频存在的上述两个问题？一是要针对场景基本不变的视频设计一种编码技术，二是要尽量在进行大压缩比（超过一百倍）视频编码之前或者编码的同时，有效提取后续视频分析所需要的视觉特征，以保证视频编码过程不会造成必要特征信息的丢失。经过几年工作，我们对第一个问题找到了解决方案。2008年，我牵头组织申请的国家基础研究（973计划）项目“基于视觉特性的视频编码理论与方法研究”，专

门设置了一个课题——面向智能监控的视频编码方法。项目研究成果之一是提出了针对监控视频的场景视频编码方法，不仅大幅度提高了监控视频的编码效率，还实现了背景和前景目标的分离。这个结果已经在 IEEE SA 1857 标准和我国自主制定的视频编码标准 AVS2 中发挥重要作用。标准工作组提供的相关测试结果表明，在视频损伤相同的条件下，编码高清晰度监控视频时，AVS2 的编码效率是最新国际标准 MPEG HEVC/ITU-T H.265 的两倍(AVS2 码流的码率为 H.265 的二分之一)，是谷歌最新标准 VP9 的三倍(AVS2 的码率为 VP9 的三分之一)，是目前市场上监控摄像机产品普遍采用的 MPEG AVC/H.264 标准的四倍(AVS2 码率为 H.264 的四分之一)。这个方法也为解决第二个问题打下了基础，因为可以从码流中直接提取出背景模型和前景区域，降低了特征提取的计算复杂度和实现成本，使得摄像头编码过程中进行特征提取成为可能。

2015 年 5 月，国家发改委、中央综治办、科技部、工信部、公安部、财政部、人社部、住建部、交通部联合发布《关于加强公共安全视频监控建设联网应用工作的若干意见》，就加强公共安全视频监控建设联网应用工作的指导思想、基本原则和主要目标及落实措施提出了明确意见，规划到 2020 年，基本实现“全域覆盖、全网共享、全时可用、全程可控”的公共安全视频监控建设联网应用，在加强治安防控、优化交通出行、服务城市管理、创新社会治理等方面取得显著成效。这样一个规模庞大的系统，不仅需要考虑视频云数据中心的架构和智能处理平台，还要考虑实现成本(包括存储成本、传输成本、处理成本和摄像头实现成本)，以及响应速度和使用便利性等。采用一套合理的标准规范体系是系统建设和应用成败之关键。例如，就存储而言，采用 AVS2，对比采用 H.264，可将监控系统存储和传输成本削减 75%，可将图像对象识别率提升 20~30%等。

北京大学、中国科学院计算技术研究所和武汉大学三个研究组在上述 973 项目之课题“面向智能监控的视频编码方法”支持下，以及国家自然科学基金重点课题“基于多摄像头协同的运动对象检测跟踪和异常行为分析”等支持下，已在国际期刊和重要国际会议上发表论文 100 多篇，申请了发明专利 60 多项。本书是这些成果的结晶，是项目组五年研究的总结，对于从事视频编码和分析研究的科研教学人员以及工程技术人员应该有很好的参考价值。

2014 年，973 计划对我们的“基于视觉特性的视频编码理论与方法研究”项目进行了滚动支持，持续五年。黄铁军教授领导的课题组继续承担相关课题，围绕场景视频编码，在利用云数据和视觉字典的编码、基于深度学习的编码和模拟生物初级视觉系统等方面开展研究。他们的工作已有新的进展，希望课题组面对视频大数据的重大需求和技术挑战，勇于创新，推进监控视频高效编码和智能分析研究前沿，并及时总结，推出本书的新版本。

高文　于北京大学

2016 年 4 月 18 日

目　录

第1章 绪 论

视频监控不仅是继数字电视、视频会议之后的一个新的大型视频应用，而且是视频技术和网络技术经过多年高速发展之后汇聚而成的一个具有变革性的大型信息系统：千千万万个摄像头通过宽带网络联系在一起，形成了一张覆盖全球的“视听感知网”，从此人类社会的运行状态都被海量的摄像头采集下来，物理世界和信息世界正在高度融合，构成人类生存的新大陆——赛博空间(Cyberspace)。

视听感知网对视频技术提出了全新的要求：需要在理论、方法、技术、标准、应用等层面长期开展大量研究开发工作。本章首先介绍“视听感知网”出现的技术背景，然后介绍智能视频分析的应用场景，并通过三个国际比赛的情况阐述目前的技术进展。我们是国内外第一个提出，也是第一个制定完成专门面向监控的视频编码标准的研究组，本章介绍了过去八年的相关研究进展。最后，对本书和各章安排情况作了简要介绍。

1.1 “视听感知网”不期而至

在信息技术发展的历史上，2009年可以说是一个分水岭。之前，信息技术给人们的印象首先是一个新的技术领域，它的作用曾被称为“(传统产业的)倍增器”，1986年863计划启动时，信息技术就是重要领域之一。2000年笔者参加国民经济和社会信息化报告起草时，有领导质疑“信息技术除了提高传统产业效率，还有什么用？”(因为说不清楚有什么用，就很难提出让国家大力支持这个方向)，有专家拿出“人类发明互联网，就像原始人发现了火”的比喻来证明其重要性，但无论是质疑还是比喻，应该说都还是在绕圈子。

2000年之后，随着互联网和移动通信技术的普及，信息基础设施作为人类社会重要基础设施之一的观念开始得到广泛认同。就像交通便利了物资的传送和使用，电力便利了能源的传送和使用，网络便利了信息的传送和使用。这种认识一方面可以说给了信息技术一个和既有技术和行业相提并论的“正宗地位”，另一方面还没摆脱人类思维的惯性，即总是从对原有物理世界增量的角度看问题，总是从对现状影响的角度评价新技术，而很难站在新技术本身的角度进行客观评价。

2009年信息技术跨过了一个分水岭，它不再仅仅是“倍增器”，也不再仅仅是“基础设施”，而是开始对已有物理世界进行反向渗透，信息世界和物理世界高度融合成为人类社会的未来发展方向，人类正式进入“赛博时代”。

今年以来，信息技术正在经历一次深刻变革：曾经独立存在的信息网络系统正在与物理世界系统通过各种感知通道越来越密切地联系在一起，信息物理系统(cyber-physical system，CPS)、物联网、智慧地球、网络司令部(cyber command)等概念从不同侧面描绘着这个正在浮现的场景——一张时时刻刻感知物理世界一举一动的大网正在形成。如果说采集多种物理信号的传感器网络构成了这张感知大网的神经末梢，那么正在迅速增长的摄像头无疑是这张大网的千千万万只眼睛。对于人类来说，超过70%的信息是从视觉系统获得的，同样对于这张感知大网，摄像头也将成为最重要的信息来源。

视听感知网正逐步变成活生生的现实，作为视听感知物理基础的摄像头正在广泛部署和迅速增长。由于公共管理和公共安全的需要，公共场合的摄像头正在以每年20%的速度快速增长，而居家养老和网络视频通信的流行，使得摄像头正在大量进入家庭。据统计，从2000年起，全球摄像终端市场年增长率保持在50%以上，到目前为止，我国大型城市的摄像头往往都在数十万个以上，全国已安装的视频监控摄像头约有2000万个，美国总统奥巴马就职典礼现场就部署了5000多个摄像头，其中还不包括典礼现场大量的摄像机和摄像手机。目前的主要问题是如何将这些分属不同部门、企业、小区甚至家庭、个人的“千万只眼睛”构成一个“复眼”，为城市管理、隐患发现、应急管理、国家安全等打造一张高效的视听感知网。如果能够成功实施，将为重大事件应急管理提供基础设施条件，成为发现安全隐患、提高事故处理速度、敏感场所安全防范的重要基础设施。

从互联网到物联网，从网络战到网络司令部(更准确地应该称为赛博司令部)，从人工智能到智慧地球，从视频监控到视听感知网……表面上看是概念之争，实质上是分处“2009年分水岭”两侧的不同观念。

20世纪是信息技术快速发展的时期，也是人类视听觉得到极大延伸的时期。人类超过90%的信息是通过视听觉获得的，视听媒体在人类的生存、进化、知识积累和文化发展中扮演着重要角色，也一直是信息学科的重要研究对象。1895年，意大利的马可尼和俄国的波波夫几乎同时发明了无线电，同年，卢米埃兄弟发明了电影。1926年，英国的贝尔德发明了电视，全球重大事件实况进入千家万户，之后，彩色电视、数字电视、高清电视纷至沓来，视听技术广泛渗透到人类社会政治、经济、军事、生活等方方面面。信息技术是20世纪与视听技术交织发展的另一个主旋律，在电子技术的推动下，通信、广播、计算机、互联网相继登场，人类进入信息时代，信息成为人类社会发展的中心要素。由于视听内容的数据量庞大，长期以来，只能采用“一对多”的广播方式传播，互联网由于带宽限制，初期传输的内容主要是文本和数据，视

听技术和信息技术这两个主旋律只能“交响”，未能“融合”。

21世纪以来，信息技术和视听技术开始交织融合，以视听媒体为主要传输和处理对象的未来媒体网络正浮出地平线。进入21世纪后，曾经独立发展的视听产业和信息产业开始快速融合。电视广播受到互联网的深刻影响，视频开始成为互联网流量的主体，图像、音频、视频和虚拟场景正在成为网络传输的主要内容。过去五年是我国互联网从Web到“Web+视频”时代的过渡期，接入网流量增加了6倍，城域网流量增加了22倍，目前正在从标清升级到高清甚至超高清和立体视频，即便节目数量不变也意味着带宽需要提高四倍乃至十多倍。思科公司曾预测，2013年90%的互联网流量来自视频，未来5年视听媒体数据将增长66倍。未来网络处理和传输的主要对象是视听信息，这是视听技术和信息技术这两大技术主旋律交织融合的结果。近年来，立体和超高清电视不断增强视听体验的逼真度，移动电视和无处不在的摄像头迅速提高了视听体验的便利性，不久的将来，沉浸式全景视听环境、物理世界和信息空间交融的增强现实系统将把人类社会带入赛博空间的新时空，政治、文化、安全、军事等领域将再次接受新技术的洗礼。

在这个网络技术和视听技术相互交融的时代，与到处渗透、无处不在的网络相比，摄像头也早已从当初的专业设备变成无处不在的传感器而散播开来。以视频监控系统为例，我国比其他国家发展的步伐更快，规模更加庞大。像其他国家一样，我国的视频监控行业在过去二十多年间共经历了三个阶段，分别是模拟闭路视频监控系统(CCTV)、“模拟-数字”监控系统(DVR、数字硬盘录像机)和全IP的网络视频监控系统。2004—2012年，数字监控在总体视频监控市场规模中所占比例从35.7%增长到了56.7%。与此同时，网络视频监控市场正在稳步增长，所占比例从2004年的7.4%增长到了2012年的28.2%。受平安城市建设、交通信息化建设、金融监控、安全生产、智能家居等各种项目建设与发展的带动，中国视频监控产品的需求量不断增加。2011年中国视频监控行业总体市场规模达到230.4亿元人民币，同比增长19.71%，2012—2015年保持21.52%的平均增长速度(水清木华研究中心，2012)。

与视频监控系统和摄像头数量快速增长直接相关的另一重要指数是摄像头密度的增加。据估计，仅2010年一年，我国新增监控摄像头的数量就超过1000万台，这主要是出于公共安全和城市管理的需要，由相关行业和部门部署，如设施安全(军队、边防、机关、港口、物资、武器库等)、财产安全(银行、商店、展馆、库房等)、公共安全(机场、体育场馆、社区、公园、监狱等)、交通安全(城市交通、高速公路、铁路/地铁、隧道等)、生产安全(矿井生产、施工现场、食品安全等)。与公共场合监控摄像头数量尚在千万量级相比，另一类摄像头的数量早就达到了亿台数量级，这就是个人计算机、平板电脑特别是手机配备的摄像头。我国个人计算机和平板电脑保有量约5亿台，手机保有量超过10亿台，随着网络视频通信的流行和居家养老的需要，摄像头成为标准配置，因此个人空间中摄像头的密度也在大幅度提高。在

上述两个趋势的联合作用下，公共空间和私有空间中不被摄像头覆盖的空间已经日渐减少。据估计，2008 年北京奥运会时全市监控摄像头数量达到了 50 万个，2012 年伦敦举办奥运会，全市摄像头数量突破百万，整个英国监控摄像头与人口的比例已经达到 1∶15。可以说，视听感知网并不需要刻意部署，已经不期而至。

1.2　监控视频的智能分析

比扑面而来的“视听感知网”更令人感兴趣的是，这么多摄像头“盯着你，我们可以知道什么？”这个问题是 IEEE 计算机学会新媒体门户 Computing Now（今日计算）2012 年 4 月的主题[①]（Dorée，2012）：

> 从每天离开家的那一刻起，你可能就已经被摄像机盯上了：无论是乘坐电梯，走在街上，在熟食店买杯咖啡，到银行取钱，还是穿过办公大楼。你在工作时，摄像机可能正在拍摄您家中发生的一切，记录保姆和孩子的互动，捕捉猫偷喝孩子碗里的饮料。你的形象可能出现在纽约时代广场相机广告牌上的人群中，路人会在电子商店的街边屏幕里看到你的购物过程，游戏机在客厅里分析你的手势，走进机场安检时你的脸又成为分析对象。
>
> 城市街道上最早安装摄像机是在 20 世纪 60 年代。20 世纪 70 年代闭路电视系统和视频录像机（VCR）这两种技术结合后，安全系统就开始在建筑物和公共场所中安营扎寨，今天已经习以为常。政府和执法部门最早把昂贵的监控系统部署到重点楼宇、公共场所、铁路、高速公路和体育场馆。不久之后，人们就在当地普通商店里看到了这类系统。随着各种技术的进步，包括数字多路复用器、数字视频录制、相机芯片以及互联网和无线技术，涌现出了性能和价格多种多样的解决方案。到了 20 世纪 90 年代，家用安全系统开始支持让用户通过网页界面远程控制摄像机的变焦和方向，而系统自己还可以自动检测场景中或指定区域的变化，并通过电话或电子邮件发送给用户。更昂贵的系统使用人脸识别技术来检测是否有未知用户登录计算机。监视视频屏幕的人工模式已经演进到自动流程：对多个视频流进行自动扫描，检测感兴趣的事件或对象，并采取相应行动。例如，一些智能家居系统在人摔倒或停止进食时可以自动报警，安全系统如果检测到门卫离岗就会提醒责任当局等。
>
> 当前的新研究走得更远：预测并阻止各类异常事件的发生。随着摄像

① 本书作者之一黄铁军是 IEEE Computing Now 咨询委员会委员、中文版负责人。这段文字摘自 Computing Now 2012 年第 4 期的客座编辑导言，作者为 Dorée，译者为黄铁军。

机无时无刻不在盯着我们：从我们的智能手机、平板电脑和笔记本电脑，到家里的游戏系统，再到环境中各种各样的摄像机，我们成为分析的对象，而我们可以从中受益。例如，智能家居系统发现我们可能倒下时会尝试协助，这类系统不仅可以区分人与物体，也可以识别行为模式，它不仅仅是识别面孔，还能识别微妙的面部表情。这些技术有望用于在公共场所检测犯罪意图，或者发现扑克玩家的心理活动，亦或一个求职者的谎言或伪装意图。多个学科的联合努力正在让我们更好地理解和使用捕捉到的视频影像。

目前的视频监控系统大多还处在信息孤岛状态，独立地完成对不同摄像机视频信号的处理、传输、控制与显示，多摄像头联动的网络视频监控技术与系统近年来已经得到了国内外政府和研究机构的广泛重视。2009 年 9 月，美国国防部高级研究计划局(DAPRA)发布了“全天候监视开发与分析系统(PerSEAS)”项目招标公告 DARPA09，规划中的 PerSEAS 是一套对采集的广域运动图像(WAMI)视频进行行为和事件自动检测和情报分析的系统，能够通过对数小时甚至数日的广域视频数据进行实时分析，检测具有威胁的活动(如聚众袭击、表现异常的恐怖活动)或者潜在的危险物并及时报警。在我国，曾经独立运营的视频监控系统正在走向联合，例如，北京市海淀区已经实现上万摄像头的集中联网，而有些省市在新建的省级视频监控系统从一开始就按照“一张网”进行规划和建设，全市、全省乃至全国范围的视听感知网正在逐步建设。

对于视听感知网来说，运动对象的检测与跟踪是异常发现、自动报警、目标检索、事件关注和快速响应等智能化功能的核心。在分布式监控摄像机网络所获取的监控视频数据中，运动目标间常见的关联有两种：一种是序列化关联(correlation in sequential surveillance videos)，即监控网络中有多个摄像机，摄像机监控范围不重叠，在任意时刻运动目标只出现在一个摄像机的监控范围之内；另一种情况是多视角关联(correlation in multi-view surveillance videos)，即任意时刻目标都出现在两个以上的摄像机监控范围之内，可以提供目标的多视监控视频。对于第一种情况，需要建立摄像机之间的关联，一旦运动目标移动至其监控范围的边界，需要通知下一个摄像机对该目标进行跟踪(典型的应用场景为高速公路)；而对后一种情况，可以在多个视角的监控视频中对运动目标进行关联分析，为监控人员提供不同视角的运动目标情况(如机场、银行等场合的监控及社区监控等)。

监控视频的智能分析涉及模式识别、计算机视觉、视频信号处理、机器学习等不同学科领域。多个国际比赛都围绕这个主题展开，其中较知名的包括监控视频事件检测比赛(TRECVid)、跟踪与监控性能评测比赛(PETS)、基于高级视频与信号技术的监控会议 AVSS 中的多摄像头人体跟踪挑战赛(Fiscus et al., 2009)。

为对不同的目标跟踪算法性能进行有效评价，IEEE PETS Workshop 提出了一系列可行的评价标准。PETS 近几年的主题是在拥挤的公共区域条件下，进行基于多

传感器的对象跟踪和事件检测，其中包括以拥挤人群个体计数和密度估计为内容的底层分析、以个体跟踪或群体跟踪为内容的中层分析，以及以特定人群流和事件检测为内容的高层分析。PETS'09 主要针对密集人群中的对象检测与跟踪进行测试。数据集为在英国雷丁大学校园环境中近 40 人活动的视频，包括取自 4～8 个摄像头的数据(将视频帧存储成图像，分训练集和测试集共约 22GB)。总体来说，参加 PETS'09 评估的系统实际检测与跟踪准确率还不高(平均检测跟踪精度只有 50%左右，最好的结果接近 60%)，离实际应用仍有较大差距。

为了推进多视角对象跟踪技术的研究，基于高级视频与信号技术的监控会议 AVSS 与 HOSDB、CPNI 和 NIST 在 2009 年联合举办了多摄像头的人体跟踪比赛(MCPT)。比赛分为三类任务：多摄像头单人跟踪(MCSPT)、单摄像头单人跟踪(SCSPT)以及双摄像头单人跟踪(CPSPT)。MCSPT 任务要求在某个视角的监控视频的最初 5 帧对某一人进行指定后，在多摄像头视野中对该人进行跟踪，包括摄像头无交叉视野时对该人的重新发现；SCSPT 任务专门对系统的跟踪稳定性进行测试；CPSPT 任务着重测试当人由一个视野进入另一个视野时，系统重新发现对象的能力。同 TRECVid 一样，比赛数据是来自伦敦 Gatwick 机场，共 44 小时。系统性能通过 F1、准确度、召回率和 NDCR 值来进行评测。由于场景极其复杂，同一个对象在不同的视角中的外表相差较大，而且不同视角中对象的尺度差别同样很大，而一些区域对象的分辨率很低，也给跟踪带来很大的困难，因此这一任务具有巨大的挑战性。

从目前视频监控领域三项主要的国际比赛结果来看，多摄像头视频监控分析已经引起了国际研究机构的广泛注意，但相关算法和技术还很不成熟。另外，由于比赛组织者的条件限制，比赛数据的摄像头数量还不足 10 个，远远没有达到视听感知网所描绘的成千上万乃至百万数量级摄像头的规模。

为了促进智能视频分析理论、方法与核心技术研究，我们课题组目前正在与相关单位合作，将建立千摄像头、万摄像头、十万摄像头、百万摄像头等级别的视听感知网验证平台，因此而产生的海量视频数据将成为相关研究和技术创新的宝贵源泉。

监控视频的智能分析是一个长期性的研究领域，那么到底有多少监控数据需要进行分析呢？根据 EMC 公司委托 IDC 做的研究报告，2012 年全球数据总量为 2.84ZB，到 2020 年，这个数字将上升到 40ZB，这是一个什么规模呢？如果把人类历史上说过的所有的话语都数字化，大约为 5EB，也就是说，2012 年全球数据量已经是这个数字的 500 多倍。IDC 把这样的一个大数据称为“数字宇宙(digital universe)”。

“数字宇宙”中的大部分数据并没有分析价值。据 IDC 估计，2012 年的数据中有 23%是有分析价值的，但是被标注的只有 3%，被分析的只有 0.5%。到 2020 年，大数据膨胀了 14 倍，其中有分析价值的数据所占比例增长到 33%，也就是达到了 13ZB，其中监控视频占到了 44%，处于绝对领先地位(其次是 25%的交易数据、20%