



专栏：信息安全

机器视觉技术、标准及安全产业应用

张园, 王慧芬, 王翰铭

(中国电信股份有限公司研究院, 上海 200122)

摘要: 聚焦机器视觉技术及标准化, 分析了机器视觉、计算机视觉与人类视觉的异同, 列举了与机器视觉相关的技术, 介绍了网络信息安全、智能监控和智能交通等机器视觉产业典型应用, 以及国际及国内主流标准化组织针对机器视觉技术、系统和应用的标准化工作进展、现状和布局。重点介绍了 VCM 和 DCM 等涉及机器视觉核心技术——机器视觉编码标准的工作, 以及机器视觉系统和应用标准化组织。在此基础上, 分析了机器视觉技术研究和标准化的未来发展方向。

关键词: 机器视觉; 标准化; 编码; 深度学习; 信息安全

中图分类号: TP277

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2020296

Machine vision technology, standardization, and application in safety industry

ZHANG Yuan, WANG Huifen, WANG Hanming

Research Institute of China Telecom Co., Ltd., Shanghai 200122, China

Abstract: The technology and standardization of machine vision were focused on, and the attributes and relations of machine vision, computer vision, and human vision were analyzed. The machine vision related technologies and typical applications such as network information, security intelligent surveillance and intelligent transportation were introduced, as well as the latest development of international and domestic mainstream SDOs of machine vision technology, system and service. The development of the core machine vision technology, video coding for machine vision standards in VCM and DCM, and machine vision systems and application SDOs were elaborated. On this basis, the future development trend of machine vision technology and standardization was analyzed.

Key words: machine vision, standardization, coding, deep learning, information security

1 引言

随着 5G 和 AI 技术的发展, 服务于机器视觉的数据量不断增加, 预计未来 4 年将超过用于娱乐的人类视觉网络流量。针对机器视觉数据的优

化和处理已成为重要商业方向, 推动机器视觉和计算机视觉产业快速发展, 在升级传统产业的同时催生新应用领域。广大发达国家和发展中国家在积极布局相关技术和应用。研究报告显示, 2020 年全球机器视觉市场规模为 107 亿美元, 到 2025 年,

预计市场将增长至 127 亿美元, 预测期(2020—2025 年)内的年复合增长率为 13.6%。机器视觉将引爆行业数字化, 成为 ICT 行业新的蓝海领域。

面向人类的视频图像与面向机器的视频图像在数据采集、预处理、压缩编码、评估标准和应用场景上存在着较大差异。面向机器的视频技术研究和标准化正成为产业界和学术界热点。ITU-T、ISO/IEC JTC1、ISO、全国信息技术标准化技术委员会、DCM (Data Coding for Machine-intelligence)、AVS (Audio Video Coding Standard Workgroup of China) 等国际及国内标准组织相继开展与机器视觉相关的技术和应用标准研究。产业发展与标准化工作互相促进, 新产业生态化发展亟待统一和高效的标准引领。

本文结合实际产业升级和机器视觉相关技术及标准化工作, 列举了机器视觉的核心技术和典型应用, 介绍了相关标准组织的标准化工作成果和最新进展, 探讨了机器视觉产业未来可能的发展方向 and 标准化研究方向, 给出了机器视觉产业和标准化建议。

2 机器视觉技术基础

机器视觉囊括技术范围广泛, 包括但不限于: 图像视频采样技术、图像视频压缩编解码技术、图像视频传输技术、数据加密技术、云计算与边缘计算技术、深度学习技术。从数据采集端的高清日用/夜用数字摄像头, 到数据分析端的云平台, 再到执行决策端的机械臂、汽车等, 本文聚焦深度学习和图像视频编/解码核心技术。

2.1 深度学习

机器视觉的根本是图像识别算法。摄像头可类比人眼, 图像识别算法是负责分析处理和决策的大脑。早期的图像识别算法主要依靠人工设计的提取器进行特征提取, 提取器的设计需要图像处理的专业知识和对提取对象深入的研究与了解。早期图像算法的识别泛化能力与鲁棒性较差。

现在神经网络在图像识别任务上的高准确率与泛用性能使基于深度神经网络的图像识别算法替代基于人工提取器的手工算法。在几十年的时间内, 人工智能从单个神经元的感知机, 发展到如今各类神经网络。性能从无法区分异或的线性分类器, 进化为在复杂图像识别任务上超过人眼的算法黑盒。机器视觉已经可以胜任智慧城市、无人驾驶等复杂任务, 其根本原因是深度神经网络的进步。全连接神经网络是拟合函数的绝佳模型; 卷积神经网络在很大程度上激发了深度学习在图像任务上的潜能; 胶囊神经网络的出现弥补了卷积神经网络在图像位置信息上的缺失, 进一步提升了网络性能; 对抗生成网络依靠生成器与判别器在图像生成方向以假乱真; 图神经网络为深度学习提供了全新的思路。

神经网络的雏形——感知机早在 20 世纪 60 年代被提出, 21 世纪初逐渐显示出优异的性能。制约其发展的瓶颈不仅是模型的落后, 更多的是数据量以及算力的匮乏。巨量、优质且全面的训练集可以在很大程度上弥补模型的不足。数据的归纳采集与打标签需要大量时间与精力。另一方面, 随着网络层数的加深与单层节点的增多, 深度学习的计算量也随之增长。例如, 有名的 ResNet 有多达 50 101 152 200 层, 其所需计算量可想而知。如今许多机器视觉应用如无人驾驶、智慧工厂等都需要伪实时, 甚至实时反馈。在算力不足的情况下神经网络分析一张图像需要十几分钟, 成为机器视觉技术落地的一大技术瓶颈。高性能的边缘计算、云平台、压缩算法以及实时传输使得实时反馈成为现实。在数据质量与算力不断提升的当下, 神经网络的潜力有待进一步的开发。模型复杂所带来的计算量增加与过拟合问题可以被支撑神经网络的高性能云边平台与海量高质量数据所缓和, 甚至解决。

2.2 图像视频编/解码

传统图像视频编/解码技术是为了节省空间与网络传输资源, 已经过三十余年的发展, 在不



断更新迭代中趋于成熟，相关标准也日益完善。最为有名的两大标准组织是 ITU 与 ISO/IEC JTC1，而影响力最广、应用度最多、性能也最高的压缩标准是两个组织共同参与推出的 H.264/AVC、H.265/HEVC 与 H.266/VVC 标准。上述标准通过不断提升的压缩码率与还原后图像视频的清晰度成为大部分企业公认并采用的行业标准。

压缩编码的本质是去除冗余。人眼对色度的辨识度不及对亮度的辨识度，所以在对图像视频进行采样时会舍弃一部分色度信息。人眼对高频信息的敏感度不及低频信息，所以在对单帧图像进行量化前会采用离散余弦变换分离高低频信息，并对高频信息进行大幅度量化。单帧图像内的像素点之间是有关联的，所以帧内预测会只对像素之间的差值进行编码。帧与帧之间的图像在时域上也是有关联的，所以帧间预测在进行运动矢量估计后也只对对应像素的差值进行编码。

然而针对人类视觉的冗余并不等于针对机器视觉的冗余。人眼看到的色彩与形状，在计算机里被表示为像素值与概率分布。

机器视觉是用机器代替人眼进行测量和判断。机器视觉系统通过机器视觉产品将被摄取目标转换成图像信号，传送给专用的图像处理系统，得到被摄目标的形态信息，根据像素分布和亮度、颜色等信息，转变成数字化信号，图像系统对这些信号进行各种运算来抽取目标的特征，进而根据判别的结果来控制现场的设备动作。

机器视觉和计算机视觉没有明显界限，紧密地联系在一起，具有相同的理论基础，只是在具体应用中根据目标不同而有所区分。机器视觉偏重于计算机视觉技术工程化，自动获取和分析特定图像，控制相应行为。计算机视觉针对图像内容开展研究。计算机视觉与机器视觉都通过图像描述世界，对基本层的图像获取、图像处理，中层的图像分割、图像分析和高层的图像理解是相

同的。计算机视觉是机器视觉的理论和算法基础，使机器视觉应用从工业领域延展到非工业领域，以实现对事物的定位、检测、测量、识别和判断。目前产业界有时将机器视觉等同于计算机视觉。

与人类视觉相比，机器视觉在灰度分辨力、空间分辨力、彩色识别力、感知速度、感光范围、环境适应性、处理能力、客观性方面占据优势。人眼无法理解的特征图可以被神经网络理解并进行识别与预测。对大部分针对人类视觉的应用（如艺术欣赏、观看电影）而言，图像及视频的背景十分重要，会影响人类对图像视频的理解。然而机器在做识别任务时，背景并不重要，只有与识别任务相关的一小块关键特征比较重要。神经网络在判断一个人是否在刷牙时，关注热点仅仅在人嘴部的一小块区域，那么实际应用中除嘴部以外的区域都可以被重度量化甚至去除^[1]。

机器视觉与人类视觉存在很多共同点。神经网络最初被研制出来就是为了模仿人脑的神经元工作模式。胶囊神经网络被研制出来的灵感是人眼可以轻松识别不同角度与不同位置的物体，而卷积神经网络不行。机器视觉对图像视频高频信息的不敏感性也类似于人类视觉。这一点可以被神经网络能轻松识别压缩过的视频与图像所证实。

由于人类视觉冗余与机器视觉冗余在一些方面存在不同，行业如今需要一套针对机器视觉的压缩技术与标准。这套新的标准需要在充分利用机器视觉冗余的同时，引用针对人类视觉冗余的压缩算法来实现低码率与高识别率。这就是机器视觉编码（video coding for machine, VCM）标准出现的意义。本文后续会详细对 VCM 标准的现状进行介绍。

3 机器视觉产业应用

机器视觉的应用场景包括智慧城市、智能监控、智能交通、车联网、自动驾驶、网络信息安

全、智慧工厂、机器人、防疫、医疗电子和工业检测等，市场空间巨大。机器视觉+云边融合+5G将极大地促进行业数字化转型，构建有机的生态系统，打造开放共赢的产业链。

机器视觉产业向着智慧化与数字化方向发展，快速的行业增长带来新的机遇和挑战。科技巨头企业和独角兽公司承担中台和整合者角色，各大学术机构承担技术输送者角色，合力进行生态打造和企业赋能。本文将对当前成熟度最高，并成为最主要的落地应用场景的智能监控和智能交通，以及当前尚未成熟，但极具潜力的机器视觉在网络信息安全方面的应用以及暴露出的问题进行介绍。

3.1 网络信息安全

在传统网络领域，DoS (denial of service) 攻击是当下最严重的网络入侵行为之一，它能给计算机网络以及基于互联网的各种服务带来极其严重的威胁。攻击者通常针对网络协议或路由器弱点进行入侵，或对服务器输送巨量无效信息来占用网络带宽，使得受害者的网络瘫痪。更为致命的是，如今发起攻击所需的工具包在互联网上唾手可得且极易使用。这就导致一些不怀好意的攻击者可以在几乎没有互联网相关专业知识的条件下发起 DoS 攻击。

近年来发展出了许多针对 DoS 攻击的自卫手段，其一般分为检测、预防、缓和和响应 4 个模式。现有大部分检测算法都是基于机器学习和概率分析，而机器视觉在检测中的应用不多。其原

因是检测算法的输入一般为网络连接记录，也就是 TCP 数据包序列，而并非图像。所以基于数据而非图像的算法更加契合攻击检测。

来自澳大利亚的学术团体^[2]开拓性地在 DoS 攻击检测领域引入了机器视觉算法：把一维网络连接记录通过多元相关性分析 (multivariate correlation analysis) 转换为二维后，作为图像输入神经网络来判断系统是否受到 DoS 攻击。

现如今基于一维数据的算法仍占攻击检测的主导，基于二维信息的机器视觉算法在传统网络信息安全方向上的应用仍处于萌芽阶段。但将一维的网络传输数据有效地映射到二维空间中，并输入当今性能强大的神经网络（如卷积神经网络、胶囊神经网络）中，在借助机器视觉进行分类任务后达到超过传统检测方法效果时，机器视觉在网络安全上的应用将会丰富起来。

随着智能交通以及智慧城市的高速发展，机器视觉在信息安全领域也暴露出了全新的问题。其主要原因之一是神经网络很容易被攻击。一种常见的针对机器视觉分类任务的神经网络攻击方式为对抗攻击。这种攻击方式利用了机器视觉与人类视觉的不同，在针对性地轻微改变图像后，人眼察觉不到明显变化，但原本正常运行的神经网络分类模型在接收到被篡改的图像后将做出错误的分类预测。如图 1 所示^[3]，左侧是正常工作的神经网络可以以 100% 的概率辨别“停车”与“禁止通行”标识。右侧通过往广告牌图片上添加具有其他路标图片的概率分布的噪声，可使人眼轻

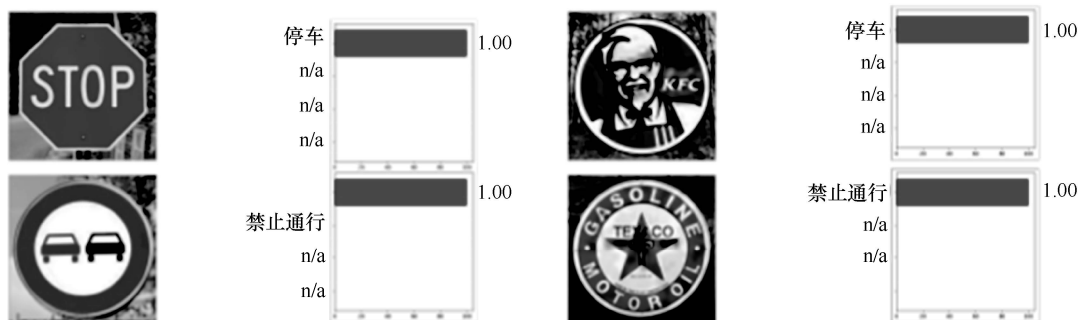


图 1 神经网络在原图与篡改图上的识别率



松辨认的广告牌标识被神经网络识别成其他标识（两张广告牌被以 100% 的概率识别成“停车”与“禁止通行”标识）。值得一提的是，恶意添加在图片上的噪声甚至比某些摄影设备的信道采样噪声更小。

在智慧城市与智能交通领域，摄像头采集的图像视频将是后续分析以及决策的重要输入。一张经过轻微改动的图片可能会造成极为严重的后果。例如，在自动驾驶领域，一张“禁止通行”标识如果被篡改成了“限速 80”标识，可能会酿成惨烈的车祸。

针对对抗攻击，现有多种解决方案。参考文献[4]研究发现当多个神经网络模型使用同样的数据集，以不同的图像压缩率进行训练时，得到的训练结果相差不多。当针对一个网络的识别图像被篡改时，该网络的判断会出现偏差。可是以不同图像压缩率训练的其他模型却不会受到如此大的影响。所以提出可以先并行地训练多个模型，在决策层多加一次判断，如果训练的多个模型对同一张图分类概率相差过大时，将采用另一个使用不同训练集的独立模型再做一次判断，以此来提高模型鲁棒性。

与使用多个模型来减少预测误差相比，参考文献[5]提出了可以在图像采集后，进入模型前进行预处理，缓和被恶意添加的噪声，使得图片可以成功地被神经网络识别分类。通过为图像添加高频信息来缓解被恶意添加的噪声，再通过小波去噪与全局池化进一步减缓噪声。上述方法采用了对整张图像的预处理。可以在通过分类热度映射（class activation mapping）获取图像中对分类结果影响最大的一部分之后，再对该区域进行预处理。

由于机器视觉技术的高速崛起，大部分研究力量都投放在提高神经网络模型准确率上，而在保护神经网络不受外界攻击方面的研究相对较少。针对机器视觉神经网络而生成对抗样本的方

法在不断增加，而完全抵御所有对抗样本的方法如今并不存在。神经网络的可解释性差，以及人类视觉与机器视觉的本质偏差导致深度学习在信息安全方面的漏洞难以被彻底发现及解决。所以依靠机器视觉的智慧城市和智能交通在真正落地之前还有很长的一段路要走。

3.2 智能监控

国内机器视觉产业应用中智慧监控产业占比接近一半。视频监控产业经过半个多世纪的演变，业已发展成熟，市场规模庞大、产业链高度细分、厂商分工结构清晰。广泛应用于平安城市和智慧城市建设、交通、环保、银行、保险、教育、医疗、物流等行业以及家庭和个人应用。欧美作为视频监控先进技术和应用最早地区，发展较为成熟稳定。中国视频监控产业的发展始于平安城市建设，随着天网工程、雪亮工程建设而逐步壮大，新兴云边融合、人工智能技术赋能安防，监控产业价值持续升级，向着智能化、工业互联网、安全隐私保护、云边融合方向发展。监控产业经历了从模拟到数字、从标清到高清、从人工监控到智能监控、从集中存储到云边融合的发展阶段，已经不仅仅是简单的视频传输，涉及数据采集、通信、处理、反馈的系统解决方案，朝着 IT、ICT 化方向演进。基于智能摄像机的监控产业催生了更大的机器视觉应用市场，是产业智慧化、规模化发展的必然趋势。

智能监控视频是我国的技术领先产业、重要战略方向和人工智能技术最成熟的落地应用，是典型的机器视觉和人机混合视觉应用场景。前端摄像机抓拍到的海量监控视频分发至智能客户端，执行目标识别、事件识别、密度检测等智能任务，部分场景需要人工回看和重建视频。

3.3 智能交通

机器视觉的另一个典型应用是智能交通，包括 V2X 和自动驾驶。机器视觉与 5G、增强学习、边缘计算、深度学习等一起服务于智慧交通应用。

其中以自动驾驶为主。自动驾驶是最为典型的机器视觉应用,视觉数据对于自动驾驶至关重要,除了摄像机,车辆传感器包括摄像机、红外线、微波雷达、激光雷达、红外线等,通过摄像头收集视觉数据,专注于道路上的所有信息,除此之外,点云数据也经常用于道路细节描绘。随着硬件技术发展和成本降低,摄像头的计算能力增加,可以执行智能分析任务。随着视频分析技术不断发展,驾驶辅助和自动驾驶的性能有所提升。结合 AI 预处理算法和边缘计算分析,通过视频感知进行对象检测、背景检测、道路检测、车辆检测、信号灯检测、行人检测等。智能任务包括多目标检测、目标分割、道路跟踪、行为识别、事件预测、光流等。

此外,典型机器视觉产业应用还包括健康医疗、防疫、智能眼镜、无人商店、智慧工厂、机器人、智慧农业、智慧渔业、智慧畜牧业、AR/VR 游戏等。机器视觉作为一项引领未来的战略技术,世界各发达国家纷纷在新一轮国际竞争中争取掌握主导权,出台规划和政策,对机器视觉核心技术、顶尖人才和标准规范等进行部署,加快促进机器视觉技术和产业发展。主要科技企业不断加大资金和人力投入,抢占发展制高点。

目前,我国机器视觉的发展已具备一定的技术和产业基础,在芯片、数据、平台、应用等领域集聚了一批高科技企业,在部分方向取得阶段性成果并向市场化发展。在视频编解码、视频监控、车联网、自动驾驶、工业互联网产业进行了技术及标准布局。

4 机器视觉标准化

4.1 标准化现状概述

机器视觉标准化包括机器视觉系统技术和应用的标准化^[6],涉及机器视觉系统的需求、架构、流程,数据的采集、处理、分发、表征、格式、编码、压缩、转换、存储、分析、挖掘、传输,

一致性测试以及安全和隐私。

当前对于图像视频数据的处理分析与应用研究在国际与国内都有极高关注度,具有一定的技术和标准化基础。但仍存在应用场景不具有通用性、缺乏统一标准和共性算法等问题,导致成本居高不下,影响产业的规模发展^[7]。

视频数据压缩编码技术已高度成熟,不断迭代。国际上相关主流标准组织包括 ISO/IEC JTC1 WG11 (MPEG) 标准工作组、ITU-T SG16,国内有 AVS 标准工作组,陆续推出了 MPEG-4、H.264 (AVC)、H.265 (HEVC)、AVS1、AVS2、AVS3 等多代标准,主要应用于广播电视、互联网视频等领域。传统方法针对人类视觉,注重信号保真度,无法匹配机器视觉对于智能任务的需求。

MPEG 标准工作组于 2019 年 7 月成立了机器视觉编码专题研究组,开展机器视觉及人机混合视觉编码的应用场景、标准化需求研究,收集数据集、技术可行性证据和技术方案,预计将于 2021 年正式立项 VCM 标准。

中国信息技术标准化委员会多媒体分委会于 2020 年 1 月成立了面向机器智能的数据编码标准工作组并召开了第一次工作组会议,制订了机器视觉编码及应用标准计划与目标。

点云数据表征和压缩技术处于起步阶段,MPEG 标准工作组和 AVS 标准工作组在制订相关标准,主要应用于三维地图、无人驾驶等领域。

机器视觉相关应用标准工作在 ITU-T SG16、IEC、CCSA 等多个国际和国内标准组织开展。ITU-T SG16 长期研究、制订智能监控、车联网等系统架构和应用标准,ISO TC22 SC33 和 ITU-T FG-AI4AD 正在开展自动驾驶标准化研究。

4.2 机器视觉编码标准

4.2.1 VCM

传统的视频压缩编码方法面向人类视觉,多用于娱乐用途,注重信号保真度,无法匹配面向智能分析任务的机器视觉应用需求。针对这一制



约机器视觉产业发展的主要问题，MPEG 标准工作组于 2019 年 7 月成立了机器视觉编码专题研究组，研究机器视觉智能应用所需的压缩编码技术。

MPEG VCM 标准组旨在定义一种压缩的视觉信息码流，服务于用于多种机器任务，同时保障高压压缩效率和机器智能任务性能。此压缩码流服务于机器视觉和人机混合视觉应用。

当前 VCM 架构如图 2 所示。输入视频信号经过 VCM 编码，输出压缩码流，经传输和解码，服务于机器视觉任务，部分应用也包括用于人类视觉的视频重建。VCM 内部架构分为特征提取和特征编码。此处“特征”为一种新的数据类型，可以包括视频、关键点、神经网络中间值、语义信息等数据。

当前面向三大应用场景用户案例视频监控、智慧交通、智慧城市做了具体需求分析和任务分解，主要包括对象检测、对象分割、图像重建、图像增强、对象跟踪、时间识别、时间预测、密度预测、事件搜索。将针对不同任务开展共性特征试验和基于特征的压缩编码试验。其他应用案例包括智能眼镜、无人商店、智慧工厂、机器人、智慧农业、智慧渔业、智慧畜牧业、AR/VR 游戏等。

目前 MPEG 已正式发布征集数据集和训练集，为发布征集证据做准备。同时进行证据收集和用户案例、需求完善工作。

4.2.2 DCM

为满足面向机器智能和人机混合应用的数据信息编码技术及标准化需求，开展标准的制订工作，全国信息技术标准化技术委员会多媒体分委会于 2020 年 1 月成立了面向机器智能的数据编码工作组，负责提出国内标准制订建议。同步组织中国专家参与国际标准化工作。工作组将开展面向机器智能和人机混合应用的数据信息编码研究，目前确定了两种面向人机混合应用的技术路线，成立两个研究组，针对两种技术路线分别研究技术可行性和标准化必要性：一是使用单独的编码流面向机器智能应用；二是使用共同的流面向人机混合智能应用。并开展系统框架和产业应用及其意义研究。

4.3 机器视觉应用标准

4.3.1 智能监控

ITU-T（国际电信联盟远程通信标准化组织）是国际三大标准组织之一，专门制订远程通信相关国际标准。ITU-T 的视频监控标准化工作从 2006 年开始，在第 16 研究组第 21 问题组（SG16 Q21）进行，2019 年 3 月成立 Q12 新课题组研究视频监控，独立举行工作会议推进相关标准发展。截至目前，发布和在研标准项目 20 个，涉及视频监控协议、移动视频监控、智能视频监控、点对点视频监控、视频监控系统互联、视频监控云存储、视频监控云计算、视频监控边缘计算、视频监控大数据应用、视频监控人脸识别、车牌识别等，详见表 1。

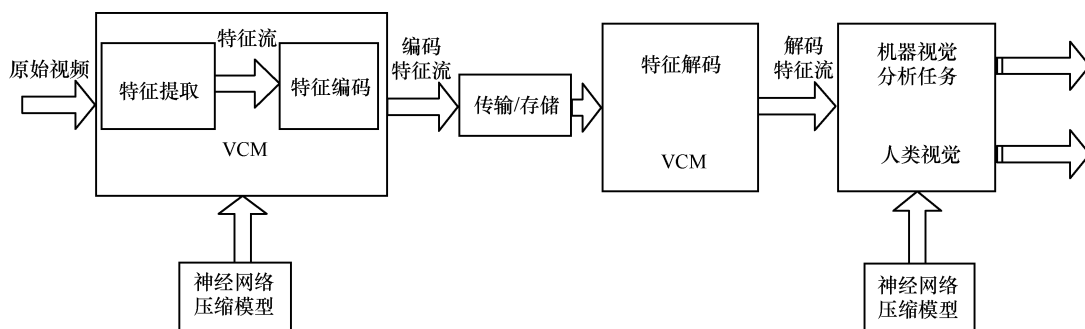


图 2 当前 VCM 架构

表 1 ITU-T 视频监控标准

标准类型	ITU-T 标准编号
视频监控基础架构	ITU-T F.743V2 (2019 年)、H.626V2 (2019 年)、H.627 (Rev.)
点对点视频监控	ITU-T H.626.4 (2018 年)
智能视频监控	ITU-T F.743.1 (2015 年)、H.626.5 (2019 年)
视频监控存储与计算	ITU-T F.743.2 (2016 年)、H.626.2 (2017 年)、F.743.8 (2019 年)、H.VSCC
视频监控系統互联	ITU-T F.743.3 (2016 年)、H.626.3 (2018 年)、F.ECVSReqs
视频监控大数据应用	ITU-T F.743.7 (2019 年)、H.VSBD
视频监控智能应用	ITU-T F.FRAVSReqs、F.VSVR

目前 MPEG 已经向 ITU-T Q12/16 发送联络函, 征询对于视频监控场景及需求意见, 后续可能以上下游或者联合工作的形式开展合作。

4.3.2 智能交通

ITU-T SG16 与智慧交通相关的标准化内容包括 UNECE WP29 联合国世界车辆法规协调论坛第 29 工作组、Q27 汽车网关、FG-VM 汽车多媒体焦点组以及 FG-AI4AD 人工智能自动驾驶焦点组。

联合国世界车辆法规协调论坛第 29 工作组 UNECE WP29 是国际驾驶类法规权威, 对国内标准制订具有影响。下设 GRVA 自动驾驶及车联网工作组、FRAV 自动化车辆测试工作和 TFCS/OTA 信息安全及软件升级子组, 分别负责自动驾驶范畴定义及分级标准、认证规范的制订以及汽车信息安全标准的制订。

FG-AI4AD 支持 AI 自动驾驶业务和应用的标准活动, 对接联合国 UNECE WP1、WP29 及各国自动驾驶车辆监管机构及政策制订机构, 制订相关测试标准。实现联合国提升安全、可持续发展的目标。依托 FG-AI4AD 研究全球自动驾驶安全共性需求, 通过合作共享和广泛参与, 研究制订全球统一的 AI 自动驾驶功能和性能需求和指标, 以及相关关键技术研究 and 应用。制订 AI 自动驾驶系统的标准化研究计划及路线图, 通过使能技术和关键任务研究, 分析自动

驾驶政策法规、标准化和评估测试的差距, 吸引全球自动驾驶及 AI 企业广泛参与, 研究可接受的驾驶行为、安全系统的先决条件, 研究制订测试工具、测试机制和测试方法。

国际标准化组织道路车辆委员会自动驾驶测试场景工作组 (ISO/TC22/SC33/WG9) 开展自动驾驶测试工作, 目前在研标准包括 ISO 34501 自动驾驶测试场景术语和定义标准, ISO 34502 工程框架和基于场景的安全评估过程, ISO 34503 自动驾驶系统运行设计域分类, ISO 34504 场景属性和分类, ISO PWI 测试场景评估。标准组面向自动驾驶测试场景的实际应用需求开展研究、制订标准。

电气和电子工程师协会自动驾驶系统视觉质量工作组 (IEEE SA P2020) 制订测试和衡量汽车图像质量的方法及指标。下设 7 个工作组, 分别进行图像质量需求及测试、LED 灯闪烁、图像观看质量、计算机视觉图像质量、摄像机子系统接口、图像质量安全和用户感知图像质量标准化工作。

5 结束语

未来机器视觉的研究和标准化将聚焦信息论和人工智能, 研究面向机器视觉、智能任务的新的信号和失真度量方法, 研究机器视觉数据的表征、编码、存储、传输、分析、安全和隐私保护,



开展面向机器视觉的通用特征、特征提取方法与特征压缩编码研究；针对人机混合应用场景研究视频重建与分析的联合优化。建立机器视觉数据标准及编码、传输系统与协议、安全隐私、合规性测试等系列标准体系。技术、应用、生态、标准形成有机整体，共同决定了产业的未来。

参考文献：

- [1] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016.
- [2] TAN Z, JAMDAGNI A, HE X, et al. Detection of denial-of-service attacks based on computer vision techniques[J]. IEEE Transactions on Computers, 2015, 64(9).
- [3] SITAWARIN C, BHAGOJI A N, MOSENIA A, et al. Rogue signs: deceiving traffic sign recognition with malicious ads and logos[Z]. 2018.
- [4] ISKHAKOV A, JHARKO E. Approach to security provision of machine vision for unmanned vehicles of “Smart City”[C]//Proceedings of 2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM). [S.l.:s.n.], 2020: 1-5.
- [5] MUSTAFA A, KHAN S H, HAYAT M, et al. Image super-resolution as a defense against adversarial attacks[J]. IEEE Transactions on Image Processing, 2019(99): 1.
- [6] 胡豆豆, 张艳霞, 张园. 视频监控领域标准化工作动态概述[J]. 电信技术, 2016, 1(5): 47-50.
HU D D, ZHANG Y X, ZHANG Y. Dynamic overview of standardization work in the field of video surveillance[J]. Telecommunications Technology, 2016, 1(5): 47-50.

- [7] 张园, 曹宁, 胡豆豆. 视频监控标准化最新进展[J]. 电信科学, 2018, 34(10): 130-136.

ZHANG Y, CAO N, HU D D. Visual surveillance standardization roadmap[J]. Telecommunications Technology, 2018, 34(10): 130-136.

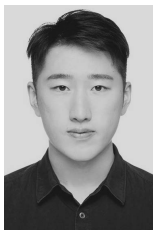
[作者简介]



张园（1986—），女，中国电信股份有限公司研究院工程师，MPEG VCM 联合召集人，ITU-T Q12/16 报告人，FG-AI4AD 副主任，全国信息技术标准化技术委员会多媒体与音/视频编码分技术委员会（SAC TC 28/SC 29）委员，主要研究方向为机器视觉、人工智能、视频监控、车联网。长期从事 MPEG、ITU-T、全国信息技术标准化技术委员会、CCSA 等标准组织机器视觉标准化工作，参与标准立项 8 项，标准发布 6 项，标准提案 30 余项，担任编辑人职务 6 项。



王慧芬（1986—），女，中国电信股份有限公司研究院工程师，主要研究方向为机器视觉、人工智能算法。



王翰铭（1995—），男，中国电信股份有限公司研究院工程师，主要研究方向为视频图像编解码、人工智能算法。