

# 半导体智能视觉系统芯片

## Semiconductor Smart Vision System Chips



吴南健 /WU Nanjian<sup>1, 2</sup>

(1. 中国科学院, 北京 100083;

2. 中国科学院大学, 北京 100049)

(1. Chinese Academy of Sciences, Beijing 100083, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

**摘要:** 介绍半导体智能视觉系统芯片的研究背景、基本概念、体系架构的演变、设计案例及今后的发展方向。认为半导体智能视觉系统芯片是一种支持图像获取、计算视觉和深度学习融合处理的典型边缘计算型视觉系统芯片, 具备实现或超越人类视觉系统的功能及其性能的潜力, 在高速运动目标的实时追踪、图像识别、智能交通、虚拟现实、机器人、生产线自动产品质量检测及各类智能化玩具等领域具有广泛的应用前景。

**关键词:** 智能; 视觉系统芯片; 图像传感; 视觉处理器; 计算视觉; 深度学习

**Abstract:** The research background, the concept of system chip, the evolution of chip architecture, the design case of vision system chip and the future development direction are introduced. Smart vision system chip is a typical edge computing vision system chip that can perform image acquisition, computational vision and deep learning processing. It can mimic or exceed human vision system in its function and performance, and can be widely applied into real-time tracking of moving objects, image recognition, intelligent transportation, virtual reality, robot, automatic product quality detection of production lines and various intelligent toys.

**Keywords:** intelligence; vision system chip; image sensor; vision processor; computational vision; deep learning

DOI: 10.12142/ZTETJ.202002006

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20200409.1412.002.html>

网络出版日期: 2020-04-09

收稿日期: 2020-02-18

视觉是人类感知外部世界的最重要手段, 视觉信息占到了人类获取外部环境信息总量的 80%, 视觉感知和视觉信息的智能化处理是人类智能活动的重要组成部分。如图 1 所示的人类视觉信息系统主要包括人眼视网膜和大脑视觉皮层<sup>[1]</sup>, 它们分别起到

了视觉传感器和视觉图像信息多层次并行处理器的功能。视觉系统能够感知外部世界的可见光视觉信号, 并且通过学习记忆来识别和认知外部的视觉信息。随着信息社会的发展和人工智能产业的再度崛起, 人工视觉信息的获取和快速高精度处理在人工智能产业化中的作用日趋凸显, 如何借鉴人的视觉系统结构和视觉信息处理模式来实现人工的智能视觉系统是当今人工智能领域面临的重要挑战。在硅片上实现能够进行视觉信息获取和智

能化处理的半导体智能视觉系统芯片成为了当今半导体信息领域中重要的热点研究课题之一。

半导体智能视觉系统芯片是一种集成了视觉图像传感器和智能视觉图像处理器的新型超大规模光电混合集成电路芯片, 它是一种典型的边缘计算型视觉系统芯片。半导体智能视觉系统芯片克服了传统视觉图像系统中数据串行传输和串行信息处理的速度限制瓶颈效应, 具有实现或超越人类视觉系统的功能及其性能的潜力, 在

**基金项目:** 国家重大科学研究计划(2011CB932902)、国家 02 科技重大专项(2011ZX02505-004)、国家自然科学基金重点项目(61234003、61434004)

高速运动目标的实时追踪、图像识别、智能交通、虚拟现实、生产线自动产品质量检测及各类智能化玩具等领域具有广泛的应用前景。智能视觉系统芯片是一种新型光电子、微电子融合的混合集成电路，它的工作原理和设计方法不同于传统的图像传感器芯片。智能视觉系统芯片的设计和应用研究富有挑战性。

C. MEAD<sup>[2]</sup>和K. AIZAWA<sup>[3]</sup>首先提出了一种基于半导体集成电路技术的神经形态视觉芯片，它以单芯片的形式实现图像传感和图像处理的功能，是一种非常有应用潜力的人工视觉系统芯片。在近30年的发展进程中，美国、

欧洲和日本等国家或地区纷纷投入了大量资金和研究人员支持大规模视觉芯片的科学技术研究。随着互补金属氧化物半导体（CMOS）图像传感器技术、智能化图像信息处理方法和光电融合集成电路技术的进步和研究人员不懈的努力，半导体智能视觉系统芯片的体系架构、信息处理模式和光电融合设计技术不断演化进步，取得了一系列重要的进展<sup>[4-12]</sup>。文章中，我们重点介绍和评述半导体智能视觉系统芯片体系架构的演变和最新研究成果。

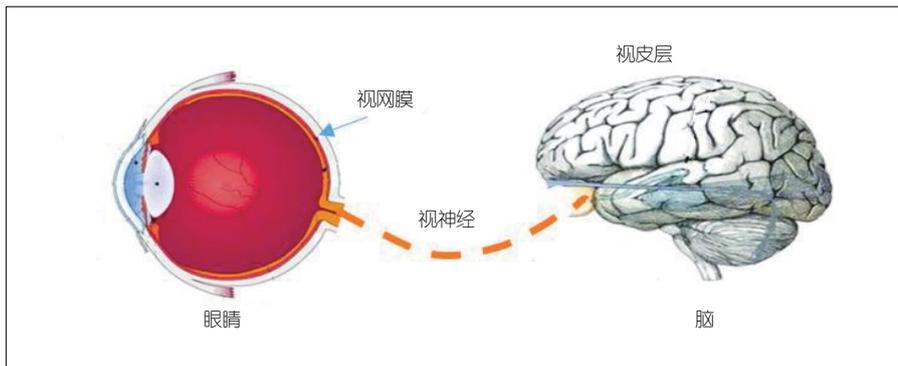
### 1 智能视觉系统芯片的概念

图2给出了智能视觉系统芯片的

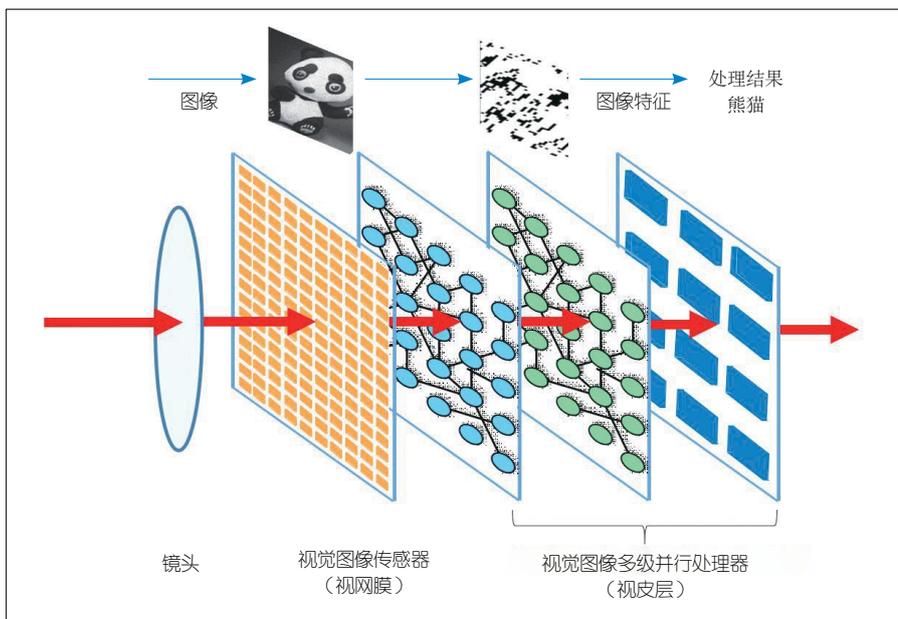
示意<sup>[1]</sup>，它由相当于视网膜的视觉图像传感器、相当于视皮层的多级并行处理网络和片上存储器组成。视觉图像传感器将捕获到的外界光学图像信号转换成电子图像信号，并进行图像信号的彩色处理、灰度调整、噪声消除和灵敏度控制等增强处理操作，然后电子图像信号传输给图像处理器进行多级图像信息并行处理。多级并行处理器接收来自视网膜的图像信号并进行图像信息处理，通过进行多层次视觉图像特征的提取和图像目标的检测，完成视觉信息的学习、记忆和识别功能。这种芯片的智能含义主要体现在两个方面：首先，芯片能够分析视觉图像传感器获得的图像质量，并判断图像是否能够满足后续信息处理的要求，然后优化传感器的拍摄条件进行图像获取；其次，芯片能够自动地对获取的图像进行智能化处理，实现图像信息的学习、记忆、识别分析和判断处理的功能。

目前的智能视觉系统芯片可分为两类：帧驱动型视觉系统芯片<sup>[4-9]</sup>和事件驱动型视觉芯片<sup>[10-12]</sup>。帧驱动型视觉系统芯片通过视觉图像传感器以帧的形式捕获图像，然后将图像数据传输到多级并行处理器中，多级并行处理器逐帧进行图像信息处理，最后输出处理结果。另一方面，事件驱动型视觉芯片感知场景中的时间或空间光信号变化进行动态视觉成像，输出脉冲视觉图像数据到多级并行处理网络中，然后脉冲型多级并行处理网络进行图像信息处理，最后输出处理结果。

这两类智能视觉系统芯片在设计方法、信号编码、操作模式和图像信息处理算法等方面有很大的不同。事件驱动型视觉芯片具有感知速度快、数据稀疏、类脑处理、能量利用率高等优点。其信号编码方式与人类视觉系统中的视觉信号编码方式相似，视



▲图1 人的视觉系统示意图



▲图2 智能视觉系统芯片的示意图

觉图像信息的表示、传输和处理都以脉冲形式来进行；因此，它具有时间分辨率高和视觉信息处理能效高的特点。但是，事件驱动型视觉信号时分处理困难，需要相对大规模的物理处理单元电路进行信息处理。

另一方面，帧驱动型视觉系统芯片在图像分辨率、静态目标检测、时间复用视觉图像信息处理、芯片面积等方面具有优势。在帧驱动型视觉系统芯片中像素电路紧凑，使得视觉图像传感器更加容易提高分辨率和填充因子。视觉图像传感器可以很好地感知静止和运动的目标，较好地支持视觉图像信息的时间复用处理模式和处理网络-存储器间的中间数据传输，方便地采用计算视觉和深度学习的方法进行视觉图像信息处理。由于篇幅的限制，本文中我们以帧驱动型视觉系统芯片为例，介绍智能视觉系统

芯片体系架构的演变和最新的研究成果。有关事件驱动型视觉芯片的研究内容细节可阅读文献 [1]、[10-12]。

### 2 智能视觉系统芯片体系架构

智能视觉系统芯片的概念提出出来后，各国研究人员开展了长期的视觉系统芯片研究。随着半导体视觉图像传感技术和图像信息处理技术的不断进步，智能视觉系统芯片的体系架构不断演进形成了如图 3 所示的 4 种不同架构，支撑着智能视觉系统芯片设计和应用技术的发展。智能视觉系统芯片主要包括视觉图像传感、图像增强处理、特征提取处理和图像检测识别处理的 4 个功能，能够完成智能化的图像感知和图像信息处理。早期的视觉系统芯片通常采用架构 I<sup>[5]</sup>，它由二维处理单元阵列组成，处理单元以单指令多数据方式进行图像获取和

信息处理操作。处理单元包括了光电二极管像素和图像信息处理电路，获取光电信号后处理电路就直接可以进行图像信息处理，可实现像素级的实时并行处理，具有工作速度高的特点。这种视觉系统芯片能完成图像信号增强处理、边缘检测、轮廓提取、块匹配、图像质心、光流计算、运动检测、测距和图像压缩等图像处理操作，可应用于目标跟踪、产品检测和人机交互等领域；然而，由于这种架构的芯片存在图像分辨率低、填充因子小和图像信息处理功能有限等缺点，实际应用推广困难。

为了克服这些困难，新的视觉系统芯片体系架构 II 被提出出来<sup>[6]</sup>。该芯片由图像传感器、二维处理单元阵列、行并行处理器阵列和微处理器单元构成，它们可以分别执行视觉图像传感、图像增强、特征信息提取和图像识别

功能	图像获取	图像增强	特征提取	图像识别	优缺点
处理方式	二维图像 $O(N \times N)$	原始图像 滤波图像 $O(N \times N)$	滤波图像 特征矢量 $O(N)$	特征矢量 识别结果 $O(K)$	
架构 I					优点：像素级并行 处理速度快 缺点：分辨率低 识别能力差
架构 II	像素	冯·诺依曼处理器 处理单元阵列	冯·诺依曼处理器 行处理器阵列	多核处理器	优点：多级并行处理 特征提取快 缺点：识别速度慢 算法依赖强
架构 III	像素	冯·诺依曼处理器 处理单元阵列	冯·诺依曼处理器 行处理器阵列	非冯·诺依曼处理器 神经网络	优点：多级并行处理 识别速度快 缺点：算法依赖强
架构 IV	像素 ADC SH	冯·诺依曼处理器 处理单元阵列	卷积神经网络	非冯·诺依曼处理器 神经网络	优点：多级并行处理 特征精度高快 识别精度高 缺点：计算量大

ADC: 模数转换      SH: 采样保持

▲图 3 智能视觉系统芯片架构的演进

的操作。架构Ⅱ克服了架构Ⅰ图像分辨率低的问题，将像素阵列和二维处理单元阵列分离，可以独立地设计像素阵列和二维处理单元阵列的规模和性能，并且图像传感器阵列和二维处理单元之间可以实现灵活的映射。但是，由于微处理器的能力无法与二维处理单元阵列和行并行处理器阵列的处理能力相匹配，限制了图像识别处理的速度。图像识别处理所需时间远大于前期图像处理所需时间；因此，根据 Amdahl 定律，图像识别处理将制约系统的性能。

为了提高图像识别的处理速度，我们进一步提出了新的视觉系统芯片体系架构Ⅲ<sup>[7]</sup>。它由图像传感器、二维处理单元阵列、行并行处理器阵列、自组织映射神经网络和微处理器单元构成，分别执行视觉图像传感、图像增强、特征信息提取、图像识别的操作和芯片控制。架构Ⅲ的特点是混合集成了冯·诺依曼型处理器和非冯·诺依曼型自组织映射神经网络处理器。自组织映射神经网络有效地提高了图像识别的处理速度，使图像识别处理的速度很好地匹配前期图像处理的速度，显著提高了视觉系统芯片的性能。其次，自组织映射神经网络电路和二维处理单元阵列电路可以在几个时钟周期内通过动态重构方式实现，从而大幅度节省了芯片的面积。

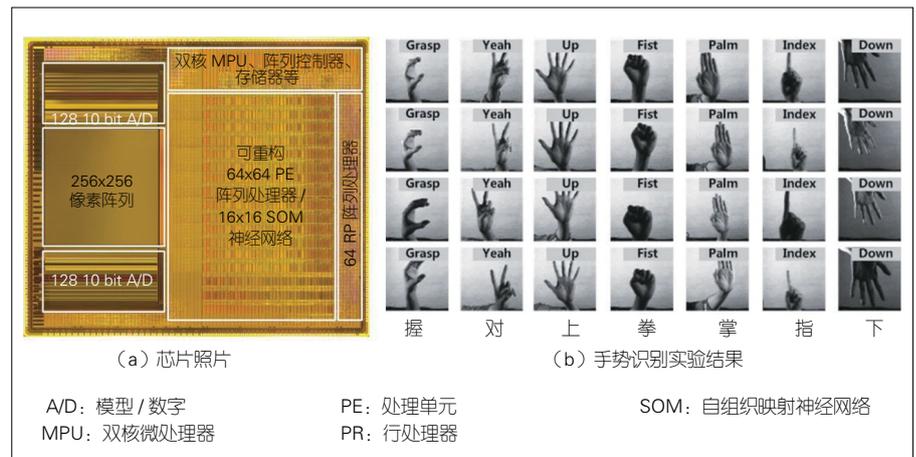
随着基于深度学习的图像信息处理技术的发展，架构Ⅳ的视觉系统芯片体系架构被提出来<sup>[9]</sup>。由于计算视觉的图像信息特征提取和图像识别的精度已被深度学习超越，该架构在特征提取处理阶段用卷积神经网络代替了冯·诺依曼型行并行处理器阵列。该架构充分考虑了计算机视觉和卷积神经网络的计算特点，可以支持卷积神经网络计算和计算机视觉算法计算，二维处理单元阵列、卷积神经网络电

路和检测分类神经网络电路可以通过动态重构技术实现时分复用处理。架构Ⅳ的视觉系统芯片的编程能力、图像信息处理能力和处理能率等各个方面都明显优于前面的视觉芯片，具备实用化应用的能力，在高速运动目标的实时追踪、图像识别、智能交通、虚拟现实、生产线自动产品质量检测及各类智能化玩具等领域具有广泛的应用前景。

### 3 智能视觉系统芯片的设计案例

图 4 给出了基于芯片架构Ⅲ的智能视觉系统芯片的照片和手势识别实验结果<sup>[7]</sup>。芯片集成了高速 CMOS 图像传感器、多级冯·诺依曼型并行处理器和非冯·诺依曼型自组织映射神经网络混合处理器，具备像素级、行级、矢量级和线程级多级并行的完整片上处理功能，可完成高速图像获

取、图像增强、图像特征提取和图像识别的各阶段图像处理任务。芯片的特点是：支持片上自组织映射网络在线训练；以约 2% 芯片总面积开销和 3 个时钟周期的代价实现了二维并行处理单元阵列和自组织映射神经网络的相互重构；是首款体系功能完整的冯·诺依曼/非冯·诺依曼混合处理视觉芯片。芯片采用 0.18 μm 图像传感器专用工艺流片实现，它包括一个 256×256 的像素阵列、128×2 路像素信号处理和 10 bit 模数转换器阵列、动态可重构的 64×64 处理单元阵列处理器/16×16 自组织映射神经网络处理器、64×1 的行处理器阵列、32 bit 精简指令集计算机 (RISC) 核的双核微处理器 (MPU) 和控制逻辑电路，芯片面积 9.8 mm×8.4 mm，系统时钟频率为 50 MHz，功耗为 630 MW。该视觉芯片达到了包括从图像采集到特征识别全



▲图 4 基于架构Ⅲ的智能视觉系统芯片及其测试结果

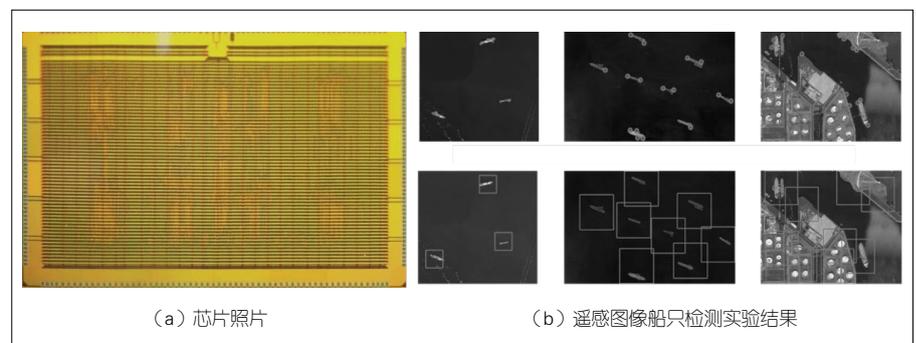


图 5 基于架Ⅳ的智能视觉系统芯片及测试结果

过程在内的 1 000 fps 系统级性能, 其中特征识别所耗时间小于 0.1 ms。

图 5 给出了基于芯片架构 IV 的智能视觉系统芯片的照片和遥感图像船只检测实验结果<sup>[9]</sup>。芯片采用 65 nm CMOS 工艺集成了多层次的并行处理单元阵列、局域存储器和片上互联网络电路。并行处理单元阵列包括 16 计算簇, 每一个计算簇主要包含 1 个 16 位标量处理单元、32 个互相连接的 16 位向量处理单元及本地存储器。标量处理单元可以灵活地获取和处理 32×32 图像块中任意数据, 每个向量处理单元可以进行像素级的图像处理, 并行处理单元阵列以单指令、多数据模式进行并行运算。RISC32 微处理器和系统总线以及其他相关模块是整个芯片的控制系统, 指挥数据通路和计算阵列协同工作。智能视觉系统芯片可以支持卷积神经网络计算和计算机视觉算法计算, 二维处理单元阵列、卷积神经网络电路和检测分类神经网络电路可以通过动态重构技术实现时分复用处理, 从而可以实现边缘计算型人脸检测和如图 5(b) 所示的遥感图像船只检测的实际应用。这种视觉芯片具有功能完整、处理速度快、功耗低和可灵活可编程的特点, 适合应用于对功耗、尺寸、实时性有较高要求的边缘计算等场合。

#### 4 结束语

本文中, 以我们研究小组长期的研究成果为例子, 介绍了半导体智能视觉系统芯片的研究背景、系统芯片的概念、芯片体系架构的演变、视觉系统芯片的设计案例。尽管视觉系统芯片技术已取得了长足进步, 但视觉系统芯片要真正进入到实用阶段仍然面临着诸多亟待解决的问题。首先,

要提高图像传感器的高分辨下成像速度、暗光条件下成像能力、成像动态范围和高精度三维成像能力, 来满足不断增长的应用需求; 其次, 目前视觉系统芯片的智能化处理能力相对于要求来说比较简单, 还不具备片上学习和图像传感器自适应控制的能力, 并且图像信息分析和处理的能力有限, 无法满足各种实际应用场景下完全自主工作的应用需求; 最后, 在视觉系统芯片集成的技术方面, 由于视觉图像传感器、智能图像处理器和存储器通常采用不同节点的 CMOS 工艺来实现, 并且智能图像处理器和存储器所用的工艺节点通常要比视觉图像传感器的更为先进, 要在单一的硅芯片上实现智能视觉系统芯片很难平衡视觉图像传感器的能力和图像处理器的性能之间的关系<sup>[12]</sup>。因此, 采用三维堆叠集成技术实现智能视觉系统芯片是今后发展的重要趋势。目前中国科学技术部的重点研究计划已经部署了相关的研究项目开展研究, 中国有望在视觉系统芯片领域取得重要的突破。

#### 致谢

本文中所介绍的芯片架构和芯片设计的案例是中国科学院半导体研究所超晶格国家重点实验室高速图像传感与图像信息处理课题组成员共同取得的成果, 在此对做出贡献的全体成员表示感谢!

#### 参考文献

- [1] WU N J. Neuromorphic Vision Chips[J]. Science China Information Sciences, 2018, 61(6): 060421. DOI:10.1007/s11432-017-9303-0
- [2] MEAD C. Neuromorphic Electronic Systems[J]. Proceedings of the IEEE, 1990, 78(10): 1629-1636. DOI:10.1109/5.58356
- [3] AIZAWA K. Computational Sensors-Vision VLSI[J]. IEICE Transaction Information System, 1999, E82D(3):580-588

- [4] KOMURO T, KAGAMI S, ISHIKAWA M. A Dynamically Reconfigurable SIMD Processor for a Vision Chip[J]. IEEE Journal of Solid-State Circuits, 2004, 39(1): 265-268. DOI:10.1109/jssc.2003.820876
- [5] MIAO W, LIN Q Y, ZHANG W C, et al. A Programmable SIMD Vision Chip for Real-Time Vision Applications[J]. IEEE Journal of Solid-State Circuits, 2008, 43(6): 1470-1479. DOI:10.1109/jssc.2008.923621
- [6] ZHANG W C, FU Q Y, WU N J. A Programmable Vision Chip Based on Multiple Levels of Parallel Processors[J]. IEEE Journal of Solid-State Circuits, 2011, 46(9): 2132-2147. DOI:10.1109/jssc.2011.2158024
- [7] SHI C, YANG J, HAN Y, et al. A 1000 Fps Vision Chip Based on a Dynamically Reconfigurable Hybrid Architecture Comprising a PE Array Processor and Self-Organizing Map Neural Network[J]. IEEE Journal of Solid-State Circuits, 2014, 49(9): 2067-2082. DOI:10.1109/jssc.2014.2332134
- [8] YANG J, YANG Y X, CHEN Z, et al. A Heterogeneous Parallel Processor for High-Speed Vision Chip[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(3): 746-758. DOI:10.1109/tcsvt.2016.2618753
- [9] 李鸿龙, 杨杰, 张忠星, 等. 用于实时目标检测的高速可编程视觉芯片[J]. 红外与激光工程. DOI: 10.3788/IRLA202049.20190553
- [10] CULURCIELLO E, ETIENNE-CUMMINGS R, BOAHEN K. A. A Biomorphic Digital Image Sensor[J]. IEEE Journal of Solid-State Circuits, 2003, 2(38):281-294. DOI:10.1109/JSSC.2002.807412
- [11] STROMATIAS E, SOTO M, SERRANO-GOTARREDINA T, et al. An Event-Driven Classifier for Spiking Neural Networks Fed with Synthetic or Dynamic Vision Sensor Data[J]. Frontiers in Neuroscience, 2017, (11):350. DOI:10.3389/fnins.2017.00350
- [12] YAMAZAKI T, KATAYAMA H, UEHARA S, et al. 4.9 A 1ms High-Speed Vision Chip with 3D-Stacked 140GOPS Column-Parallel PEs for Spatio-Temporal Image Processing[C]//2017 IEEE International Solid-State Circuits Conference(ISSCC). San Francisco, CA, USA: IEEE, 2017: 81-82

#### 作者简介



吴南健, 中国科学院半导体研究所研究员; 主要研究领域为高速图像传感与图像信息处理芯片设计; 已发表论文 300 余篇, 申请专利 100 余项。