

# 脉冲连续摄影原理与超高速高动态成像验证

黄铁军

(北京大学计算机学院多媒体信息处理全国重点实验室, 北京 100871)

**摘要:** 传统相机采用定时曝光方式获得静态图像或图像序列形式的视频, 不能有效表达极高速的光子流过程, 而且导致高速和高动态相互对立的“两难困境”。基于光电传感器像素独立的特点, 提出了连续摄影原理: 每个像素各自把接收的光子流转换成电子流, 连续测量光电流并转换成数字信息流, 实现像素级别的连续表达, 再按照像素空间排布组成序列阵列, 就是对像素平面入射光子流过程的连续表达。截取序列阵列任何一个时刻的状态就可得到该时刻的图像, 从而实现连续成像。

进而, 提出了把光电子流调制为脉冲序列的脉冲连续摄影原理: 像素从清空状态开始积累电荷, 达到额定阈值时产生一个脉冲作为积满标志并自动复位重新开始累积, 如此重复。一个脉冲积满所经历的时间称为它的脉宽, 与这个时段的光强成反比, 据此可以估计这一时段的光强。脉冲按照自然时序排列而成的脉冲序列就是对光电子流过程的数字化表达。各像素产生的脉冲流按照像素空间分布排列而成的脉冲流阵列称为视象, 蕴含了光过程丰富的时空信息, 从中可以生成任意时刻的图像, 实现超高速、高动态、无模糊连续成像, 解决了定时曝光成像的“两难困境”。

脉冲连续摄影原理的唯一参数是累积阈值  $Q$ , 对应积满一个脉冲所需光子/电子数, 唯一变量是脉冲累积时长  $\tau$ , 是最容易准确测量的物理量, 并可以利用自然时序简练表达。任意时刻的光强为  $Q/\tau$ , 由所处脉冲的宽度  $\tau$  决定,  $\tau$  的取值范围是  $(0, \infty)$ , 因此理论上可以表达任意强度的光, 动态范围无穷大。实际物理实现中, 强光表达的极限取决于电路最短读出时间, 暗光表达极限就是暗电流强度, 它累积为一个脉冲的时长  $\hat{\tau}$ ,  $\hat{\tau}/\tau$  即相机的动态范围, 采用常规光电器件和电路就能实现 160 dB 甚至 180 dB 超高动态成像。在要求成像时间灵敏度小于  $\hat{\tau}$  时, 可通过对累积电压亚阈值量化, 实现超高动态成像。

光子流和光电子流是符合泊松分布的离散随机过程, 往往是一个“分段线性”过程光强瞬时突变, 突变之间光强稳定。提出了实现脉冲流最简无损压缩的首脉冲编码方法: 只在光强变化时输出首脉冲发放时刻及其脉宽, 之后不输出脉冲就表示重复。采用首脉冲序列编码, 可在几乎不增加输出数据量条件下, 大幅提高时间计量精度, 显著提升系统的动态范围和时间灵敏度。

采用成熟 CMOS 光电器件和标准工艺, 研制了两款脉冲连续摄影芯片和脉冲相机, 空间分辨率分别为 10 万像素和 100 万像素, 采用 4 万赫兹同步脉冲输出, 即最短脉宽 25  $\mu\text{s}$ 。实拍试验验证了脉冲连续摄影原理的可实现性和超高速高动态无模糊成像性能。

脉冲连续摄影视象蕴含了光子流的时空信息, 替代图像和视频, 将从根本上重塑计算机视觉和视觉信息处理技术和产业。创立了脉冲视觉开源算法体系 SpikeCV, 用较低计算复杂度实现了速度比人眼快千倍的超高速目标检测跟踪识别系统。作为脉冲连续摄影的对称过程, 脉冲连续显示把超高速脉冲流调制为极高速光子流, 能够实现类似自然光的高速变化, 解决了传统显示系统因为低帧率带来的运动模糊和视觉疲劳眩晕等问题。连续摄影和连续显示相结合, 可实现近似玻璃的单向透明显示和超高速无介质光通信。

**关键词:** 光电传感器; 脉冲序列调制; 连续摄影; 高动态成像; 超高速成像; 相机; 视象

**中图分类号:** TB8; TP391

**文献标识码:** A

**文章编号:** 0372-2112(2022)12-2919-09

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20221075

## Spiking Continuous Photographing Principle and Demonstration on Ultrahigh Speed and High Dynamic Imaging

HUANG Tie-jun

(School of Computer Science, State Key Laboratory of Multimedia Information Processing, Peking University, Beijing 100871, China)

**Abstract:** Traditional camera accumulates photons during an exposure time window to generate a still image or a

video in the form of image sequences, which lose the temporal process of the photons flow, leads to an irreconcilable dilemma between high-dynamic and high-speed imaging. Based on the fact that the pixels of the photoelectric sensor are independent, a new continuous photographing principle is proposed: each pixel converts the received photon flow into an electronic flow independently, continuously measures the photoelectric current and converts it into a digital flow, and represents continuously at the pixel level. Then the sequence array ranked according to the pixel layout is a continuous representation of the photons shottng on the sensing plane. The image of any moment can be obtained by intercepting the state of the sequence array, so as to realize continuous imaging.

Furthermore, the spiking continuous photographing principle that modulates the photon flow into a spike sequence is proposed: for each pixel, accumulate charge from the reset state, generate a spike as a flag once the specified threshold is reached, reset and repeat. The duration that a spike takes to be fired is called its spiking width, which is inversely proportional to the light intensity during this period. Based on this, the light intensity during this period can be estimated. The spikes sequence with natural temporal order is a digital representation of the photon flow process. The array of the spike sequences according to the spatial layout of pixels is called "viform", which contains rich spatial and temporal information of the light process. An image at any moment can be calculated from viform, thereby ultrahigh-speed, high-dynamic and non-blurred continuous imaging is achieved, and the dilemma between high-dynamic and high-speed imaging of exposure imaging is solved completely.

The unique parameter of the spiking continuous photographing principle is the accumulation threshold  $Q$ , which corresponds to the number of photons/electrons required to fire a spike. The unique variable is the spike accumulation time  $\tau$ , which is the easiest to measure accurately and can be concisely represented by natural temporal order. The light intensity at any moment is  $Q/\tau$ , which is determined by the spike width  $\tau$ , which range is  $(0, \infty)$ . Thus, theoretically, arbitrary light intensity could be represented, and lead to infinite dynamic range imaging. In the physical implementation, the strongest light could be represented depending on the shortest readout time  $\hat{\tau}$  of the circuit, and the weakest light is the dark current, which can be accumulated as a spike within duration  $\hat{\tau}$ .  $\hat{\tau}/\tau$  makes up the imaging dynamic range. With conventional photoelectric devices and circuits, ultrahigh-dynamic imaging with 160 dB or even 180 dB can be achieved. When the imaging time sensitivity is required to be less than  $\hat{\tau}$ , ultrahigh-dynamic imaging can be achieved by sub-threshold quantization of the accumulated voltage.

Photon and photoelectron flow are discrete statistical process that follows a Poisson distribution and is often a "segmental linear" process light intensity changes suddenly while the intensity is stable between successive mutations. A method so-called first spike encoding is proposed: only the moment of the first spike firing and its spike width are output when the intensity changes, and no spike is output thereafter to indicate repetition. This is the optimal lossless compression method of the spike sequence, where the dynamic range and time sensitivity can be significantly improved by improving the time measurement accuracy with almost no increase in representation data volume.

Using mature CMOS photonic devices and standard processes, two spike continuous photographing chips and spike cameras with spatial resolutions of 0.1 million and 1 million pixels have been developed, respectively. The chips use 40 000 Hz synchronous spike output, with a minimum spike width of 25  $\mu$ s. Practical tests have verified the feasibility of the spike continuous photographing principle and its ability to achieve ultrahigh-speed, high-dynamic, and blur-free imaging.

Viform, the spike continuous photographing representation and the replacement of image and video, captures the temporal and spatial information of the photon flow, will fundamentally reshapes computer vision and visual information processing technology and industries. An open source algorithm framework, SpikeCV, achieves high-speed target detection, tracking, and recognition system that is faster than the human eye by thousands of times, with low computational complexity. As the symmetrical process of spike continuous photographing, spike continuous displaying modulates the ultrahigh-speed spike sequence into extremely high-speed photon flow, enabling high-speed display similar to natural light through glass, solving the motion blur and visual fatigue dizziness caused by low frame rate in traditional display systems. In addition, the combination of continuous photographing and continuous displaying can achieve a single-way transparent glass-like display and ultrahigh-speed light communication without medium.

**Key words:** photoelectric sensor; pulse sequence modulation; continuous photographing; high-dynamic imaging; ultrahigh-speed imaging; photons flow; viform

## 1 引言

光是宇宙中最快的物理过程,自古以来,精确记录和表达光过程就是人类最大的梦想之一.1827年法国人约瑟夫·尼埃普斯采用“日光蚀刻法”在敷有薄沥青的白蜡板上获得人类第一幅照片,开始了记录光的历

史,也铸造了“摄影”(photography)这个概念:使用某种专门设备进行影像记录的过程.1839年,法国人达盖尔发明银版摄影法,曝光时间降低到约30分钟,达盖尔成功后兴奋地大喊:“我抓住了光!我捕捉到了它的飞行!”前一句正确,后一句值得商榷:静态照片并不能记

录光的飞行过程。

1895年电影被发明出来,1925年电视被发明出来。前者采用胶片,后者直接用电信号,但都是用静态图像序列表达光过程:按照一定频率拍摄静态图像,采集的图像序列称为视频(video)。电影每秒24幅图像,电视每秒30幅图像。1990年代数码相机开始取代胶片相机,数字视频也被广泛应用,但仍然沿用了照相术静态成像(imaging)的基本原理。

然而,把高速变化的光过程“压扁”为静态图像,丢失了时域信息,造成高速成像和高动态成像相互对立的“两难困境”:缩短曝光时间可以减轻运动模糊,但光累积不充分,成像动态范围窄;增长曝光时间长可以提高动态范围,但会加重运动模糊。改进光学系统和光电器件性能,可以缓解这对矛盾,但不能根本解决。

视频拍摄时为了减轻运动模糊,每帧图像曝光时长 $\tau$ 往往小于相邻两帧实际时间间隔 $\tau'$ ,这意味着 $\tau'-\tau$ 时段内的光过程根本没有被采集到,完全丢失了,因此视频不是对光过程的完整记录。

静态成像在胶片时代是唯一选择。对于光电传感器来说,曝光成像不再是必然选择,这就是本文提出的连续摄影新原理:直接把超高速光子流转换成高速比特流,逼近光电器件物理极限,实现高动态、超高速、不模糊的连续成像。本文研制的芯片和相机验证了新原理的先进性。

## 2 相关工作

1905年,爱因斯坦提出光量子说,成功统一了波动说和微粒说:光子是粒子,每个光子是特定频率的波,光是独立的光子组成的粒子流。很多实验都清楚地表明,即便是完全稳定的单色光源,辐射出的光子流也不是均匀分布,而是服从泊松分布<sup>[1]</sup>,泊松分布的速率参数 $\lambda$ 对应的是单位时间内的光子数量。非单色光是由多种不同频率的光子流组合而成,根据泊松过程可加性,仍然是泊松过程,强度函数由各单色光子流强度函数相加得到。太阳之类的白热光源具有称为“相干时间”的量子尺度时间特性,在远大于量子尺度条件下(通常都满足这个条件),光子流分布仍然遵循泊松过程<sup>[2]</sup>。

1960年代固态光电传感器登上历史舞台。1990年代互补金属氧化物半导体(Complementary Metal Oxide Semiconductor, CMOS)实现了数码相机革命。CMOS图像传感器采用光电二极管把光子转换成携带电荷的光生载流子,正常光电流在皮安级到微安级,直接测量记录这种微弱光电流的代价很高,因此早在固态传感器研究初期,就提出了通过电容累积光生电流产生的电压变化 $\Delta V$ 来估计电流强度<sup>[3]</sup>,具体如下:

$$\int_0^{\tau} I(t) dt = \Delta V \times C \quad (1)$$

其中,左侧表示经历时间 $\tau$ 累积的电荷数, $I(t)$ 为光电流;右侧 $C$ 为感光电路的电容, $\Delta V$ 是经过累积时长 $\tau$ 产生的电压变化,两者相乘,就是电容器累积的电荷数。

数码相机继承了定时曝光成像原理,按固定时长把光电流累积为电容电压,也就是式(1)左侧的积分时长 $\tau$ 对所有像素都是一样的,这样只需要把各像素的累积电压 $\Delta V$ 进行量化,就能得到数字图像或图像序列。也因为如此,数码相机继承了传统相机固有的时间丢失缺陷和高动态和高速度不可兼得的“两难困境”。

然而,定时曝光成像并非光电传感器的必然选择,更非唯一选择。回到式(1),如果设 $\Delta V$ 为固定值,则测量光电流累积达到 $\Delta V$ 所需时长 $\tau$ ,同样可以估计光电流平均强度:

$$\bar{I} = \frac{\Delta V \times C}{\tau} \quad (2)$$

其中,电容 $C$ 是定制,电压变化 $\Delta V$ 是设定定制,两者乘积 $\Delta V \times C$ 也是定值,也就是累积光电荷总数为定值,因此这种模式称为定量模式。相比之下,传统相机采用定时曝光累积,称为定时模式。定量模式下,因为 $\Delta V \times C$ 是定值,因此电流估计值 $\bar{I}$ 完全由 $\tau$ 确定,二者成反比关系。

归根结底,累积电荷有两个可变量:累积时长 $\tau$ ,累积电荷量(和 $\Delta V$ 成正比)。定量模式和定时模式分别固定其中的一个,测量另一个。

光电传感器采用定量模式成像,称为脉冲调制(Pulse Modulation, PM)<sup>[4]</sup>,又具体分为脉冲宽度调制(Pulse Width Modulation, PWM)和脉冲频率调制(Pulse Frequency Modulation, PFM)。PWM图像传感器每个像素的值由它累积时长转换得到,PFM图像传感器则用曝光时段内发放的脉冲数量表示。

PWM思想最早可以追溯到Muller在1977年提出的 $I^2L$ 电路<sup>[5]</sup>,它把温度或光信号通过累积调制为与强度成反比的脉冲宽度:从不到1 ms直到10 s。Brajovic和Kanade设计了第一个PWM传感器原型<sup>[6]</sup>,阵列为 $21 \times 26$ ,用于低延迟视觉。PFM光电传感器概念最早可追溯到1982年Frohman提出的采用触发器电路翻转频率来表达输入光强<sup>[7]</sup>,1993年Tanaka等开发出第一款PFM图像传感器<sup>[8]</sup>。

在实际应用中,PWM设置的压差阈值较大,以有效区分部分不同强度的光。PFM设置的压差阈值较小,以有效区分强光(高频发放)和弱光(低频发放)。通过这样的设定,PWM和PFM就可以获得高动态范围。

尽管脉冲调制有高动态范围及其他优点,而且已经有多款脉冲调制图像传感器研制成功,但并未在数码相机领域得到规模应用,最重要的原因是脉冲调制只是作为静态成像的一种手段,并未跳出传统相机定时曝光静态成像原理。

### 3 脉冲连续摄影原理

传统胶片成像是光致化学反应过程,只能采用定时曝光方式. CMOS 光电传感器(相关原理和方法也适合其他光电传感器)采用光电二极管把光子转换成携带电荷的光生载流子,各像素是独立的,并不需要像胶片成像那样同步曝光,而是可以各自把接收的光子流转换成电子流. 通过连续测量光电流并转换成数字信息流,就实现了连续表达. 所有像素按照空间关系排布成阵列,各像素输出的数字流就是对照射到像素阵列平面上的光过程的一个连续表达,可从中获得任意时刻的图像,本文称之为“连续摄影”(continuous photographing).

#### 3.1 连续摄影

连续摄影是对光过程的一种新的表达,本身并不成像,但可以从中产生任意时刻的图像.

光强随时间变化时,对应的光子流是一个非时齐泊松过程,其速率函数 $\lambda(t)$ 就是光强变化过程. 光传感器的目的就是记录这个过程. 光电传感器光敏单元把光子流转换成光生载流子,光生载流子数和入射光子数的比例称为量子效率(最大值为1),这样,光子流转换为光生电荷流(简称光电流),对光子流速率函数 $\lambda(t)$ 的测量也就转换成了对光生电荷流 $I(t)$ 的测量.

实际光子流往往是一个分段变化过程. 光强变化来自光源的开关或能量改变或者反射物的移动或变化,往往是突变. 相比之下,突变之间的光强往往是相对稳定的,因此,可以通过监测光强变化对光电流进行分段表达. 具体来说,每个像素都配置一个光电流变化监测电路,当变化超过设定阈值时,启动电流强度测量电路并输出绝对强度值,如此循环,输出的绝对强度值序列就是对光电流的分段表达.

如前所述,直接测量记录大量微弱光电流的代价很高. 采用定量累积脉冲调制,可以用更为简洁的像素电路同时测量光强和光强变化,这就是脉冲序列调制连续摄影.

#### 3.2 脉冲序列调制

前面介绍过采用脉冲宽度调制和脉冲频率调制的成像方法. 与之不同,脉冲序列调制(Pulse Sequence Modulation, PSM)不以成像为目的,而是把光电流调制成一个脉冲序列,用这个序列作为光电流的数字化表达.

脉冲序列调制过程具体如下:光电传感器每个像素都作为独立的光探测器,独立采用脉冲调制把光电流转换成脉冲序列,像素从清空状态开始积累电荷,压差达到指定阈值时自动复位,重新开始积累,并产生一个脉冲作为积满标志,累积所经历的时间称为这个脉冲的脉宽,这样每个像素就把入射极高速光子流转换成超高速脉冲流,脉冲在时间轴上形成的序列就是对光电流过程的数字化表达. 各像素产生的脉冲序列按

照像素空间排布构成脉冲序列阵列,就是对像面入射光流的数字化表达.

图1展示了采用脉冲序列调制进行信号采集的过程. 图1(a)是像素产生的光电流,从中等强度升高再降低. 图1(b)是光电流的脉冲序列表达:脉冲1所在时刻是第一个脉冲累积完成的时刻,从起点到这个时刻的时长就是脉冲1的脉宽(箭头表示脉冲产生的时刻,高度没有意义);之后光电流强度保持稳定不变,因此之后的脉冲2至脉冲4的脉宽都和脉冲1相同;脉冲5的脉宽缩短了,因为累积过程中光电流在脉冲5累积期间突然提高;脉冲6~11的脉宽相同,因为它们累积期间光电流未变,又因为这段时间光电流强度是脉冲1~4期间的2倍,因此脉宽变窄为一半;脉冲12累积期间光电流变弱,因此脉宽变长;脉冲13和脉冲14的脉宽比脉冲1~4长一倍,因为累积期间光电流比脉冲1~4期间弱一倍.

脉冲序列调制作作为一种数字化方法,同时完成了采样量化,输出的脉冲序列和原始信号采用相同的时间坐标,每个脉冲只需要用一个比特表示. 需要强调的是,脉冲出现时刻是表达的关键:两个脉冲之间的时间间隔(即后一个脉冲的累积时长)和这段时间内信号的平均强度成反比. 图1(c)更好地表达了这个物理含义:每个脉冲的面积都是相同的,对应累积阈值,脉冲宽度(脉宽)即前述累积时长,脉冲高度就是这段时间内信号的平均强度(图1中把脉冲表示成了梯形,只是为了便于示意区分). 图1(d)描述的是为了降低输出码率的无损脉冲编码方法,5.2节专门介绍.

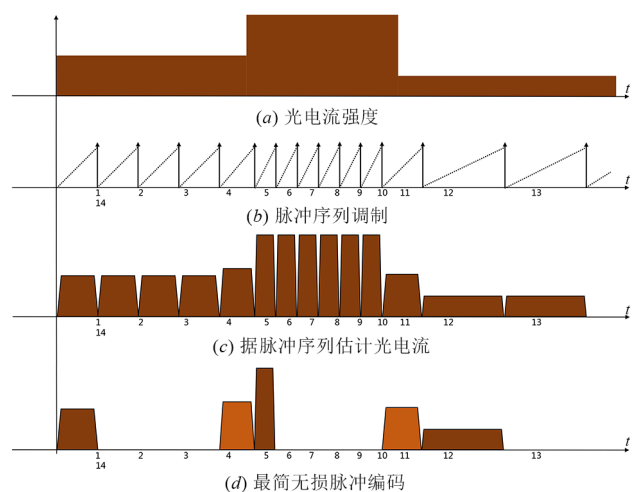


图1 脉冲序列调制原理

实际光电流波动更为复杂,从脉冲序列估计的平均值组成的分段过程,能够较好地逼近光电流变化过程. 需要说明的是,尽管脉冲5和脉冲12未能准确逼近原始光电流,它们仍然是有意义的:位于两个平稳段之间而且脉宽介于两者之间,说明该脉冲累积期间发生

了光强变化,这些脉冲将作为后续编码优化的关键脉冲,也可以用于运动目标检测等视觉任务。

采用脉冲调制对光电流进行连续表达的方法发明于2015年,2016年1月申请了发明专利,2019年至2022年相继获得中国专利<sup>[9]</sup>、美国专利<sup>[10]</sup>以及日本、韩国和欧洲专利。

### 3.3 脉冲连续成像

光电传感器是多个独立像素排布成的阵列面,相应地,各像素产生的脉冲流排布成的脉冲流阵列就是对入射到光敏阵列面的光子流的数字化表达,这种表达不同于传统的图像和视频,需要一个新名词,建议中文为“视象”,其中“视”表示“视觉、可见”,“象”来自《道德经》中的“大象无形”,指没有具体形态的形式,以区别于有具体形态的“图像”的“像”和“幻相”的“相”,对应英文建议使用“vform”。相应地,拍摄视象的光电设备称为“脉冲相机”(spike camera),也可称为“摄象机”或“象机”(vform camera)。

“视象”蕴含了光过程丰富的时空信息,从中可以生成任意时刻的图像,也就是实现连续成像。获得任意时刻图像的基本方案有两种(图2):脉冲累积成像和脉冲间隔成像。

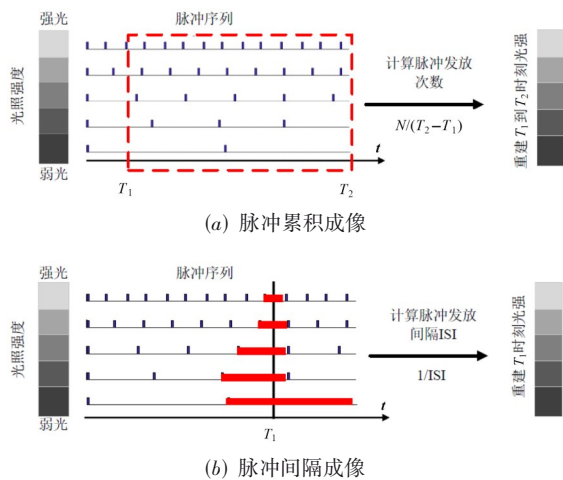


图2 获得任意时刻图像的两基本方法

#### 3.3.1 脉冲累积成像

脉冲累积成像继承了传统相机定时累积方法:要产生 $T_2$ 时刻的图像,则把其前一定时长时间窗口(图2(a))中的 $T_2 - T_1$ 的脉冲进行累积,脉冲越多,光强越大。

如果脉冲累积阈值 $Q$ 等于传统相机量化的级差,则两者成像能力相同。传统相机是把曝光累积得到的模拟值按照级差 $Q$ 转换成一个数值,脉冲序列调制每累积到 $Q$ 就加1,两者结果相同。

不同于传统相机曝光的物理过程,脉冲累积成像是在数字脉冲序列上的“虚拟曝光”, $T_2$ 及 $T_1$ 都可以自

由选择:通过滑动窗口实现连续成像,通过改变窗口大小可以调节图像动态范围。当然,虚拟曝光也继承了定时成像对高速运动模糊的缺点。

#### 3.3.2 脉冲间隔成像

图像是特定时刻所有像素的光强的空间排布。如果一个像素在该时刻所处脉冲的脉宽为 $\tau$ ,则其光强为 $Q/\tau$ 。也就是说光强和脉宽成反比,两个脉冲之间的间隔越短,光强越强,间隔越长,光越弱,如图2(b)所示,这就是脉冲间隔成像。显然,利用这种方法可以估计任意时刻光强,因此能够连续成像。

传统相机把累积光强线性量化为 $N$ 个级别,脉冲相机则是用 $M$ 个单位时间表示。假定 $M=N=256$ ,并设定脉冲累积阈值 $Q$ 为传统线性量化级差,则 $\tau=1$ 时对应光强256,脉宽 $\tau=256$ 时对应光强为1,两者表达的最强光和最弱光相同,因而动态范围相同。但是,脉宽 $\tau$ 的倒数到光强是非线性映射,表达能力更强:

(1) $\tau=1$ 是 $\tau=2$ 时光强的两倍,是 $\tau=3$ 的3倍,这种对强光的非线性有利于不产生过曝;

(2)通过缩小时间单位,很容易扩大动态范围,以表达更强的光;

(3)脉宽 $\tau$ 和脉宽 $\tau+1$ 两个相邻级别之间的光强之差是 $\frac{1}{\tau(\tau+1)}$ ,也就是随着脉宽增加,级差越来越小,因此,表达弱光的层次更细腻;

(4) $\tau$ 可以大于乃至远大于256,从而表达越来越弱的光。

人类视觉对亮度的感知也存在类似的幂律非线性映射,都是压缩强光表达层次,扩大弱光表达层次。这是因为人类视觉系统对光的表达也采用脉冲方式,脉冲累积时长也是人类视觉系统的重要编码机理之一。

## 4 表达性能的理论分析

成像追求的四个主要目标是:高清晰(高空间分辨率)、高速(高时间分辨率)、高动态(高光强分辨率)和高光谱(多光谱分辨率)。脉冲连续摄影在空间分辨率和光谱分辨率方面与传统相机相同:对空间分辨率的逼近取决于传感器面阵规模、像素尺寸等物理条件约束,对光谱的表达要依赖滤光片等技术的约束。

和传统相机曝光成像原理不同的是,脉冲连续摄影原理能够充分发挥光电器件物理潜力,无限逼近高速和高动态这两个目标。下面结合常规CMOS光电器件的能力进行定性和定量分析。

### 4.1 脉冲流逼近光子流

脉冲流阵列可以任意精度逼近光子流阵列。当脉冲累积阈值 $Q>1$ 时,每个脉冲对应 $Q$ 个光子,实现了物理空间到信息空间的 $Q \rightarrow 1$ 映射,两个脉冲之间的时间

间隔(或者说后一个脉冲的脉宽)是出现  $Q$  个光子的总时长. 当  $Q=1$  时, 脉冲流就是数字意义上的光子流, 一个光子用一个比特(脉冲)表示, 脉冲之间的时间间隔等于光子之间的时间间隔, 实现了物理空间到信息空间的一一映射. 因此, 脉冲流阵列是对光子流阵列的一个有效渐进逼近, 随着光电器件探测能力的提升, 可以不断逼近直到完整表达光子流物理过程.

#### 4.2 理论动态范围: 无穷大

脉冲连续摄影的唯一设定参数是电荷累积阈值  $Q$ , 记录的是脉冲累积时长  $\tau$  (即脉宽), 时间是最容易精确计量的变量, 通过两者就可以估计任意时刻的光强为  $Q/\tau$ .

用脉冲序列逼近光电流, 理论上可以表达任意强度的光:  $\tau$  的取值范围是  $(0, \infty)$ ,  $\tau$  越小表明光越强, 逼近 0 则意味着光强逼近无穷大;  $\tau$  越大表明光越弱,  $\tau$  逼近无穷大则意味着光强接近 0. 因此, 从理论上讲, 脉冲连续摄影可以实现任意高的高动态范围.

当然, 光子流和光电流是离散随机过程, 不能也不必要无穷逼近. 另外, 光电器件和电路存在暗电流, 电荷累积到阈值的时长  $\tau$  不可能无穷大, 脉冲信号读出时间也不可能无穷小.

#### 4.3 强光表达: 逼近电路信号读出能力

强光表达极限取决于最短脉宽和累积阈值  $Q$ .

如果累积阈值  $Q=100$ , 中午阳光直射条件下光亮度达到  $10^5$  lux,  $100 \mu\text{m}^2$  像素每秒射入的光子数超过  $4.09 \times 10^{10}$ , 设光电转换的量子效率为 0.7, 则积满 100 个电荷的时间为

$$\tau = \frac{100}{4.09 \times 10^{10} \times 0.7} \text{ s} = 0.35 \times 10^{-8} \text{ s} = 3.5 \text{ ns}.$$

这意味着要表达这种强光, 电路读出有效信号的最短时间应该达到纳秒(ns)级. 换句话说, 电路能实现的最短信号读出时间  $\check{\tau}$  决定了系统能够表达的最强光为  $Q/\check{\tau}$ . 如果要表达更强的光, 则需要增大累积阈值  $Q$ , 或者控制进光量.

#### 4.4 暗光表达: 逼近器件暗电流极限

光子流和光电流都是离散随机过程, 对光电流强度的有效测量必须持续一定的时间. 现代 CMOS 光电器件的暗电流强度可以控制在  $100 \text{ e}^-/\text{s}$  之下, 即便测量持续 10 ms, 也只能在概率意义上“碰到”一个光生载流子, 这是光电流固有的物理随机性, 这意味着要测量暗光产生的微弱光电流, 就必须花费更长的时间, 才能获得有效的测量结果, 而测量时长变长意味着测量的时间灵敏度降低, 这一性质可称为“光电测不准原理”: 微弱光电流的高精度测量和高时间灵敏度不能同时实现.

“光电测不准原理”的物理意义是显然的: 测量意味着获取准确的信息, 暗光条件下光子数量少, 获得准确的信息就需要更长时间. 极端情况下, 需要等待无限

长的时间, 才能得出“无光”的测量结论.

实际光电器件在无关条件下, 也会自发产生暗电流, 因此实际器件不会永远不输出脉冲. 若脉冲累积阈值为  $Q$ , 光电器件暗电流为  $I_d$  ( $\text{e}^-/\text{s}$ ), 则脉宽小于  $\hat{\tau} = Q/I_d$  的脉冲是信号, 超过这个值的为噪声.

对于一个采用确定的脉冲累积阈值  $Q$  进行连续摄影的光电传感器, 暗电流和电路读出速度决定了动态范围的上限和下限. 若最短读出时间为  $\check{\tau}$ , 最长信号脉宽为  $\hat{\tau}$ , 则其动态范围为

$$\text{DR} = \lg \left( \frac{\hat{\tau}}{\check{\tau}} \right) \times 20 = \lg \left( \frac{Q}{\check{\tau} I_d} \right) \times 20 \text{ dB}.$$

典型情况下,  $\check{\tau} = 10 \text{ ns}$ ,  $\hat{\tau} = 1 \text{ s}$ , 则动态范围为 8 个数量级, 160 dB, 利用目前成熟的 CMOS 工艺就能做到.

#### 4.5 累积阈值的调控作用和亚阈值量化

如前所述, 脉冲连续摄影唯一的调控参数是累积阈值  $Q$ , 在电路读出有效信号的最短时间  $\check{\tau}$  确定并作为时间精度单位的情况下, 能够表达的最强光就是  $Q/\check{\tau}$ . 如果可以通过改进光学系统或感光器件提升入射光强或量子效率, 则可以相应提高  $Q$  以表达更强的光, 扩大动态范围.

在光学系统和感光器件不变的情况下, 增大  $Q$  虽然也可以表达更强的光, 扩大动态范围, 但意味着脉宽等比例放大, 逼近光子流的时间精度降低, 成像的时间灵敏度降低. 这种情况下如果还需要保证时间灵敏度, 就需要对累积电压值进行亚阈值量化, 作为对采用脉宽估计光强的补充信息.

设电路能够实现的最短读出时间为  $\check{\tau}$ , 实际时间灵敏度要求累积时长最长为  $\check{\tau}$ , 则可在脉冲累积时长达到  $\check{\tau}$  但累积值仍未达到阈值  $Q$  时, 直接测量未满阱的累积压差并按  $N$  级进行量化, 这时传感器的动态范围为

$$\text{DR} = \lg \left( \frac{\check{\tau}}{\check{\tau}} \times N \right) \times 20 \text{ dB}.$$

例如,  $\check{\tau} = 10 \text{ ns}$ ,  $\check{\tau} = 10 \text{ ms}$ ,  $N = 100$ , 则动态范围为 8 个数量级, 160 dB.

自然场景光照从  $10^{-4}$  lux (黑夜) 到  $10^5$  lux (夏至阳光), 跨越 9 个数量级, 如果要全覆盖, 例如在外太空中同时观测太阳和微弱的星光, 则需要缩短  $\check{\tau}$ , 增加  $\hat{\tau}$ , 使得  $\frac{Q}{\check{\tau}} > I_s$ ,  $\frac{Q}{\hat{\tau}} < I_d$ , 其中  $I_s$  为最强光电流,  $I_d$  为暗电流, 这时动态范围为  $\lg \left( \frac{\hat{\tau}}{\check{\tau}} \right) \times 20 \text{ dB}$ . 取  $\check{\tau} = 10 \text{ ns}$ ,  $\hat{\tau} = 10 \text{ s}$ , 则动态范围为 180 dB. 如果要保证时间灵敏度达到  $\check{\tau} = 1 \text{ s}$ , 则需要把  $Q$  亚阈值量化为 10 个级别, 以保持 180 dB 的动态范围. 如果要保证时间灵敏度达到  $\check{\tau} = 0.1 \text{ s}$ , 则需要把  $Q$  亚阈值量化为 10 个级别, 以保持 180 dB 的动态范围.

## 5 脉冲流无损编码

本节先分析脉冲流固有的物理随机性,再根据这一性质,给出了一种容易电路实现的脉冲流无损编码方法.

### 5.1 脉冲流的物理随机性

入射光子流遵循泊松分布,光电转换时间灵敏度很高,延迟可视作定值,因此光生电子流保留了入射光子流的变化过程,可认为仍然近似遵循泊松分布.

脉冲视觉模型通过对光生电子流进行定量累积表达光生电流过程,每个脉冲实际上是“从0开始积满 $n$ 个电荷”.如果光生电荷流的速率为 $\lambda$ ,则相邻两个电荷出现的时间间隔服从参数为 $\lambda$ 的指数分布,累积 $n$ 个电荷的时长 $\tau$ (即脉冲脉宽)服从参数为 $(n, \lambda)$ 的伽马分布:

$$f(t) = \frac{\lambda e^{-\lambda t} (\lambda t)^{n-1}}{(n-1)!}.$$

伽马分布的期望为 $n/\lambda$ ,这是显然的:按照速率 $\lambda$ 出现了 $n$ 个电荷,所用时间为 $t=n/\lambda$ .

伽马分布的方差为 $n/\lambda^2$ ,标准差为 $\sqrt{n}/\lambda$ ,相对标准差(RSD)为 $\sqrt{1/n}$ .相对标准差和 $\lambda$ 无关,只和 $n$ 相关, $n$ 越大,相对标准差越小,这意味着随着 $n$ 的增大,量子尺度(光子/电子)的固有随机性越来越不明显.

当电荷累积阈值 $Q=1$ 时,RSD=100%,也就是光子/电子的随机性完全保留,这是显然的.当电荷累积阈值 $Q=100$ 时,RSD=10%, $Q=1\ 000$ 时,RSD=3.16%,光电流的固有的量子随机效应在脉冲层次仍然会有所表现,也就是即便光源完全稳定,脉冲序列也不是均匀分布,而是存在随机波动.累积阈值 $Q$ 提高到10 000以上,RSD降低到1%以下,量子层次的随机性才可以基本忽略.

考虑到表达的时间灵敏度, $Q$ 不能设置得过高,这意味着即使稳定光照,脉冲流也存在固有随机性,因此在对脉冲流进行编码压缩时,低于RSD的波动应视为固有特性而忽略,只有高于RSD的变化才有信号意义.

### 5.2 脉冲流最简无损编码

每个像素产生的脉冲流可以表示为脉冲出现时刻所形成的序列 $\{t_i\}$ ,或者脉宽序列 $\{\tau_i\}$ ,两者等价的.以前述最短脉宽1  $\mu\text{s}$ 为例,最强光条件下每秒产生100万个脉冲(1 Msps, sps为spikes per second).对于一个百万像素的脉冲视觉传感器,每秒产生的脉冲数最多为1 Tpsps,常规电路难以输出这么大的数据流.

如前所述,实际光子流往往是一个“分段线性”变化过程:光强瞬时突变,突变之间光强稳定.光强不变时脉冲周期性重复发放,但并不提供新的信息,因此只需要输出第一个脉冲(下称首脉冲)发放时刻及其脉

宽,之后不输出就表示重复.由于光强稳定时脉冲宽度仍然存在物理随机性,因此这里的“重复”是统计意义上的重复,需要根据累积阈值 $Q$ 估计偏差,脉宽波动在偏差范围内的都视为相同.

图1(d)例示了从脉冲序列中选择首脉冲的过程.脉冲1是第一个脉冲,自然作为首脉冲输出,其脉宽作为当前脉宽(图中箭头高度表示脉宽).脉冲2~4的脉宽与之相同,因此不需要输出.脉冲5产生时发现自己的脉宽与当前不同,因此作为第二个首脉冲输出,用其脉宽更新当前脉宽.脉冲6产生时也发现脉宽与当前脉宽不同,因此作为第三个首脉冲输出.脉冲7~11的脉宽与当前脉宽相同,不必输出.如此类推,脉冲序列就转换成了一个更为精简的首脉冲序列.就图1的例子来说,原有14个脉冲就精简为1,5,6,12,13五个首脉冲.如果只要保留绝对光强,不需要保留变化时刻,则脉冲5和脉冲12也可以删除,进一步精简为1,6,13三个首脉冲.

对于一个像素来说,首脉冲序列是个异步序列,可以表示为 $\{(t_i, \tau_i)\}$ ,其中 $t_i$ 是第 $i$ 个首脉冲的发生时刻, $\tau_i$ 是其脉宽.对于像素阵列来说,如果多个像素共享一个异步输出通道,则还需要增加像素位置 $(i, j)$ ,组成四元组 $(i, j, t, \tau)$ ,采用异步通信协议输出.

从首脉冲的定义和产生规则不难看出,首脉冲产生和光强变化有关,每次变化都用一个首脉冲表达,能够完整表达光过程.另外,每次光强变化都必须记录变化发生时刻和变化后的光强,首脉冲是必要的.因此,首脉冲序列是表达脉冲流充分有必要的最简无损编码.

根据脉冲连续成像原理,在累积阈值 $Q$ 确定的情况下,动态范围和时间灵敏度取决于时间计量精度.首脉冲编码方法只与光的变化有关,因此提高时间精度提高几乎不增加数据量(首脉冲数可能会少量增加,记录脉宽的位数也有少量增加),这对研制超高速和超高速动态传感器研制意义重大.

## 6 芯片研制与成像验证

2017年采用110 nm标准CMOS工艺和常规器件,第一款脉冲连续摄影芯片SpikeOne和相机被研制出来<sup>[11]</sup>,空间分辨率 $450 \times 200 = 10$ 万像素.2021年 $1\ 000 \times 1\ 000 = 100$ 万像素芯片SpikeTwo被研制出来.相机脉冲主频为4万赫兹,即最短脉宽为25  $\mu\text{s}$ ,每个像素每秒输出4万个比特.相机也可以工作在更低频率,例如2万赫兹.控制脉冲的电荷累积阈值可外部设置,支持16级动态调整.

采用百万像素脉冲相机拍摄挥手人物,频率采用2万赫兹模式(称为2万拍),即最短脉宽50  $\mu\text{s}$ .截取包含挥手动作的800拍.图3是采用脉冲累积成像方法

得到的六幅图像,分别是累积1,2,5,50,200,800拍的结果.前3幅图像“虚拟曝光”时间短,可以明显看出脉冲(白点,对应一组光子)的累积过程.第5幅“曝光”时长10 ms,已经出现运动模糊(挥手).第6幅“曝光”时长40 ms,运动模糊严重.

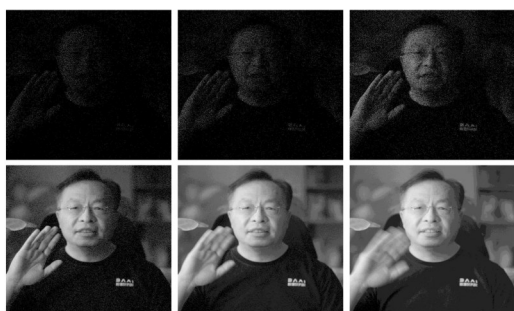


图3 脉冲累积成像示例

图4是脉冲间隔成像与脉冲累积成像的效果对比.图4(a)即图3的最后一幅图像.图4(b)是根据第800拍所在脉冲的宽度计算得到的瞬间图像,没有运动模糊.具体算法是先把脉宽归一化为 $\tau \in [0, 255]$ ,再用为 $255 - \tau$ 作为像素亮度,这样处理(而不是直接采用倒数)相当于进行了非线性伽马校正,模拟了人类视觉对亮度感知的非线性特性,显示效果更好.

图5展示了前800拍另外四个时刻的图像,并标识了10个典型像素对应的脉宽序列(图中脉宽乘以脉高等于累积阈值),可以形象地看出脉宽和像素亮度之间的反比关系.



(a) 脉冲累积成像(前800拍脉冲) (b) 脉冲间隔成像(根据第800拍累积,时长40 ms)  
所处境宽估计)

图4 脉冲累积成像和脉冲间隔成像效果对比

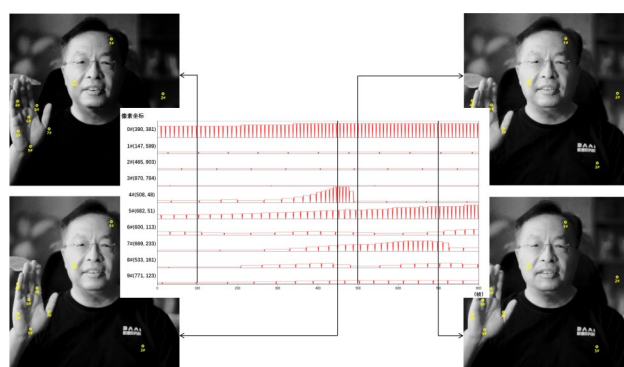


图5 从脉冲流中任选四个时刻进行无模糊瞬态成像

脉冲间隔成像从原理上保证了无运动模糊,这是脉冲连续摄影新原理的突出优势,特别适合超高速成像.图6展示了脉冲相机近距离拍摄6马赫超声速静风洞激波产生的过程<sup>[12]</sup>.

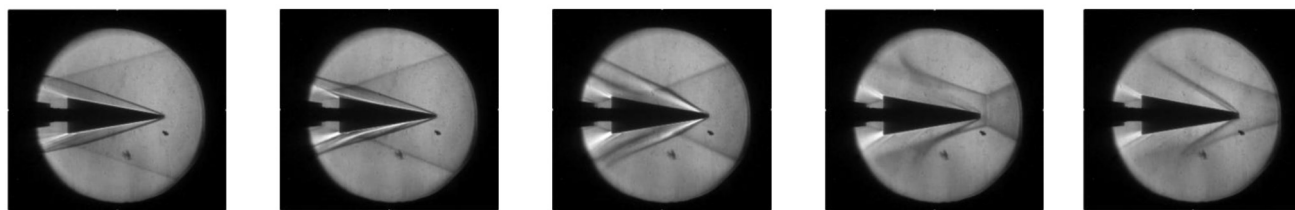


图6 脉冲相机(10万像素)拍摄6马赫超声速激波过程

## 7 结论与展望

照相术和电影的发明,开启了人类记录光的历史,也固化了人们的思维.脉冲连续摄影新原理打破了以静态图像为核心的传统成像模式,从原理上能够保证了超高速、高动态、无模糊连续成像,充分发挥CMOS光电器件的技术潜力,完成数码相机未曾完成的数字化革命,图像传感器跃升为高速视觉传感器,从超高速相机到拍照手机,从安防摄像头到车载传感器,整个产业都将重新洗牌.

视象将对视觉信息处理带来深远影响.视象蕴含了丰富的时空信息,可有效提高卫星遥感、断层扫描等

空间分辨率.视象高时间分辨率特性,对物理、化学、生命等学科连续观测具有重要意义.

视象(脉冲序列阵列)作为替代图像视频的视觉信息新表达,对计算机视觉发展影响巨大.视觉类似生物视觉脉冲,有利于实现光流、深度等特征的准确提取和高速目标检测跟踪等任务.笔者所在团队创立了脉冲视觉算法体系<sup>[13]</sup>并开源(<https://openi.pcl.ac.cn/Cordium/SpikeCV>),用较低复杂度实现了速度比人眼快千倍的超高速目标检测跟踪系统<sup>[14]</sup>,验证了“脉冲高速相机+脉冲神经网络=超高速机器视觉”的可行性和先进性.

显示是摄影的对称过程,脉冲连续显示把超高速脉冲流调制为极高速光子流. 现有电视、手机和虚拟现实应用的信号源都是视频,只有几十赫兹,最多 120 Hz,在一帧周期内,显示器件重复创新同样光强. 信号源变为视象后,显示器件潜力得到发挥,能够实现类似自然光的高速变化,解决低帧率刷新带来的运动模糊和视觉疲劳眩晕问题. 脉冲连续摄影采集的视象通过高速通道直连脉冲连续显示输出,实现单向透明光传输,用于虚拟现实眼镜,能够显著提高视觉感知的自然舒适度. 把脉冲连续显示作为可见光信号源,通过脉冲连续摄影进行高速采集,对超高速无介质光通信意义重大.

#### 参考文献

- [1] RODNEY L. The Quantum Theory of Light[M]. 3rd Edition. Oxford: Oxford University Press, 2000.
- [2] PHILIP N. From Photon to Neuron: Light, Imaging, Vision [M]. New Jersey: Princeton University Press, 2017.
- [3] WECKLER G P. Operation of p-n junction photodetectors in a photon flux integration mode[J]. IEEE Journal of Solid-State Circuits, 1967, SC-2(3): 65-73.
- [4] JUN O. Smart CMOS Image Sensors and its Applications [M]. 2nd Edition. Boca Raton: CRC Press, 2020.
- [5] Muller R. I<sup>2</sup>/L timing circuit for the 1ms-10s range[J]. IEEE Journal of Solid-State Circuits, 1977, 12(2): 139-143.
- [6] BRAJOVIC V, KANADE T. A sorting image sensor: An example of massively parallel intensity-to-time processing for low-latency computational sensors[C]//Proceedings of IEEE International Conference on Robotics and Automation. Minneapolis: IEEE, 1996: 1638-1643.
- [7] FROHMADER K P. A novel MOS compatible light intensity-to-frequency converter suited for monolithic integration[J]. IEEE Journal of Solid-State Circuits, 1982, 17(3): 588-591.
- [8] TANAKA K, ANDO F, TAKETOSHI K, et al. Novel digital photosensor cell in GaAs IC using conversion of light intensity to pulse frequency[J]. Japan Journal of Applied Physics, 1993, 32(11A): 5002-5007.
- [9] 黄铁军. 对时空信号进行编码的方法和装置: CN105681787B[P]. 2019-01-18.
- [10] HUANG Tiejun. Method and device for encoding space-time signals: US10523972B2[P]. 2019-12-31.
- [11] 黄铁军, 董思维. 超速电眼: 全时成像芯片重塑机器视觉[J]. 前沿科学, 2019, 13(2): 16-21.  
HUANG T J, DONG S W. Overspeed electric eye: Full-time imaging chip reshapes machine vision[J]. Frontier Science, 2019, 13(2): 16-21. (in Chinese)
- [12] 黄铁军. 脉冲视觉: 冲破图像的幻相[J]. 中国图象图形学会通讯, 2022, 1(1): 8-13.
- [13] 黄铁军, 余肇飞, 李源, 等. 脉冲视觉研究进展[J]. 中国图象图形学报, 2022, 27(6): 1823-1839.  
HUANG T J, YU Z F, LI Y, et al. Advances in spike vision[J]. Journal of Image and Graphics, 2022, 27(6): 1823-1839. (in Chinese)
- [14] HUANG T, ZHENG Y, YU Z, et al. 1000× faster camera and machine vision with ordinary devices[J]. Engineering, 2022, 15(11): 1-10.

#### 作者简介



黄铁军 男, 1970年出生, 河北大名人. 北京大学计算机学院教授, 北京智源人工智能研究院院长. 主要研究方向为视觉信息处理和类脑智能.  
E-mail: tjhuang@pku.edu.cn