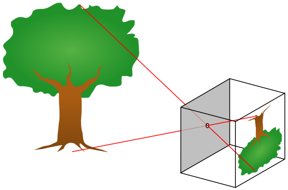
Event-based Computer Vision

Charles Clercq

Italian Institute of Technology
Institut des Systemes Intelligents et de robotique

November 30, 2011

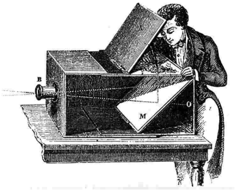
Pinhole camera



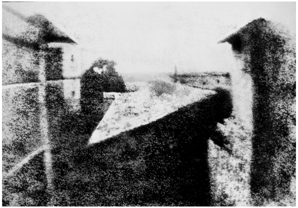
Principle

light rays from an object pass through a small hole to form an inverted image.

Camera obscura



(a) 18h Century, portable device for artists

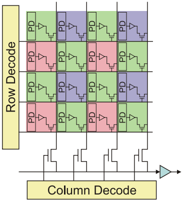


(b) First successful permanent photograph, by Nicéphore Niépce in 1826 at Saint-Loup-de-Varennes.

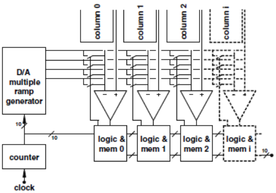
Principle

Optical device that projects an image of its surroundings on a screen.

Electronic device



(c) CMOS sensor

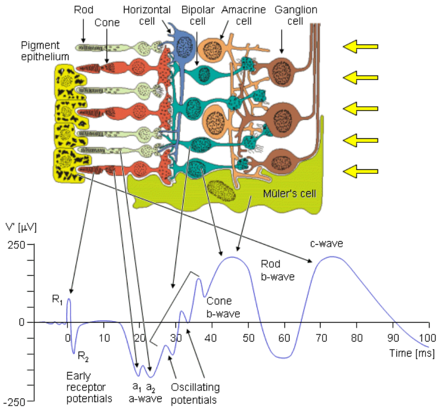


(d) Basic architecture of a column-parallel single-slope, multi-ramp analog-to-digital converter

Principle

The light is integrated through the photo-diode. The value of each element is read synchronously, triggered by the time.

Natural "device"



Visual system

Principle 1

The first steps in seeing begin in the retina, where a dense array of photoreceptor convert the incoming pattern of light into an electrochemical signal [...] *Nassi, Callaway 2009*

Principle 2

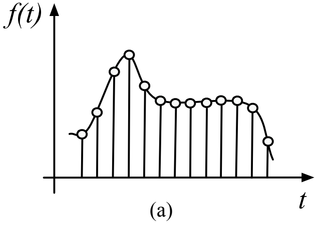
The strategy used by the mammalian visual system is to reduce the representation of the visual scene to a limited number of specialized, parallel output channels. [...] *Nassi, Callaway 2009*

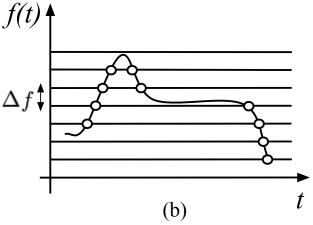
Encoding

Two possibilities

Two ways to encode the information:

- time-driven encoding,
- data-driven encoding.



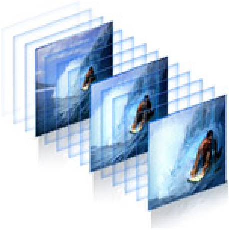


Frame-based Paradigm

Time-driven encoding

- fixed temporal interval, Δt,
  - $I_{seq} = \{I(t_0), I(t_0 + \Delta t), \ldots, I(t_0 + n\Delta t)\}$
  - $I_{seq} = \{f_{0,0}(t), f_{0,1}(t), \ldots, f_{M-1,N-1}(t)\}$

I_{seq} is composed of a set of discrete functions $f_{x,y}(t)$, obtained by sampling all of the pixels at the same time.



Frame-based issues

Disadvantage

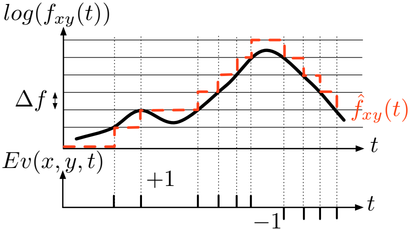
- Reduction of the dynamics,
- unnecessary redundant data,
- time and memory consuming.

The reason of those disadvantage come from the time sampling.
To resolve the problem, we have to **change the sampling process**

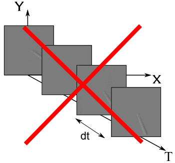
Event-based Paradigm

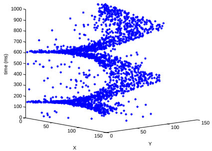
Sampling

- let t_k be set of times of the signal sampling,
  - $T = \{t_k \mid |\mathcal{F}(f_{x,y}(t_k)) - \mathcal{F}(f_{x,y}(t_{k-1}))| = \Delta f\}$.
- $\mathcal{F}$ is defined as a *log* function,
  1. provide a wide pixel range,
  2. make sensitive to the relative contrast.
- let $Ev(x, y, t)$ be the compact representation of $f_{x,y}$,
  - $Ev(x, y, t) = \delta(t, t_k) \cdot sign(f'_{x,y}(t))$,



Computation





Goal

- exploit the data-driven acquisition advantages
- avoid the time-based issues

Optical Flow

Definition

Apparent motion of displacements in the 3D space projected on the 2D plane of image sensors

Robotic field

The optical flow is an important tools used in the robotic field:

- object avoidance
- navigation ...

Methods

Plethora of different techniques

- differential methods,
- region-based matching,
- energy based,
- phased based.

Two distinct families

- Purely local (e.g Lucas and Kanade),
- Global (e.g Horn and Schunck),
  - take advantage of the global image structure.

Selection

Local method

Taking advantage of:

- high temporal resolution,
- sparse encoding.

Small displacement and assumption

- The light intensity conservation
  - $I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t)$
- Taylor series
  - $I(x+\delta x, y+\delta y, t+\delta t) = I(x, y, t)+\frac{\partial I}{\partial x}\delta x+\frac{\partial I}{\partial y}\delta y+\frac{\partial I}{\partial t}\delta t+H.O.T$
- The equation of the optical flow
  - $I_x V_x + I_y V_y = -I_t$

Least Square Method

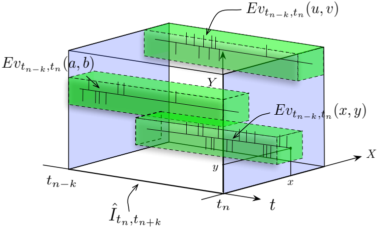
Estimation

- Local constancy
  - $$\begin{cases} I_x(q_1)V_x + I_y(q_1)V_y & = & -I_t(q_1) \\ & \dots & \\ I_x(q_n)V_x + I_y(q_n)V_y & = & -I_t(q_n) \end{cases},$$
- The Least Square Method
  - $$A = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \vdots & \vdots \\ I_x(q_n) & I_y(q_n) \end{bmatrix}, v = \begin{bmatrix} V_x \\ V_y \end{bmatrix}, b = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \vdots \\ -I_t(q_n) \end{bmatrix}.$$
  - $A^T A v = A^T b$
- To solve:
  - $v = (A^T A)^{-1} A^T b$

Event-based Optical Flow

Derivative

$$\begin{cases} \frac{\partial \hat{I}_{t_i,t_n}}{\partial x}(x,y) & \sim (Ev_{t_i,t_n}(x,y) - Ev_{t_i,t_n}(x-1,y))\Delta f \\ \frac{\partial \hat{I}_{t_i,t_n}}{\partial y}(x,y) & \sim (Ev_{t_i,t_n}(x,y) - Ev_{t_i,t_n}(x,y-1))\Delta f \\ \frac{\partial \hat{I}_{t_i,t_n}}{\partial t}(x,y) & \sim \frac{Ev_{t_i,t_n}(x,y) - Ev_{t_i,t_{n-k}}(x,y)}{t_n - t_{n-k}}\Delta f \end{cases}$$



Optical Flow Algorithm

Algorithm: Event-based optical flow

Require: $\mathbf{p} = [x, y]^t$

For each event occurring at time t at pixel $[x, y]^t$:

Define a $(n \times n)$ neighborhood N of $[x, y]^t$ and compute the partial derivatives:

- $grad(I)(\mathbf{p}) = \begin{bmatrix} \frac{\partial \hat{I}_{t_j, t_n}}{\partial x}(\mathbf{p}) \\ \frac{\partial \hat{I}_{t_j, t_n}}{\partial y}(\mathbf{p}) \end{bmatrix}$
- $\frac{\partial I}{\partial t}(\mathbf{p}) = \frac{Ev_{t_{n-k}, t_n}}{t_n - t_{n-k}} \Delta f$

Solve equation of the flow over $\mathbb{N}$ for $[v_x, v_y]^t$

Optical Flow Algorithm and noise

Algorithm: Event-based optical flow 2

Require: $\mathbf{p} = [x, y]^t$

For each event occurring at time t at pixel $[x, y]^t$:

if activity $Act_{t_{n-k},t_n}(x, y) > a$ **then**

Define a $(n \times n)$ neighborhood N of $[x, y]^t$ and compute the partial derivatives:

- $$grad(I)(\mathbf{p}) = \begin{bmatrix} \frac{\partial \hat{I}_{t_i,t_n}}{\partial x}(\mathbf{p}) \\ \frac{\partial \hat{I}_{t_i,t_n}}{\partial y}(\mathbf{p}) \end{bmatrix}$$
- $$\frac{\partial I}{\partial t}(\mathbf{p}) = \frac{Ev_{t_{n-k},t_n}}{t_n - t_{n-k}} \Delta f$$

Solve equation of the flow over $\mathbb{N}$ for $[v_x, v_y]^t$

else

set $v_x = 0$ and $v_y = 0$

end if

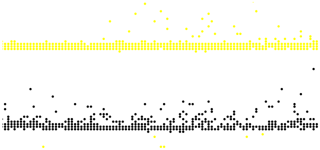
Noise issue

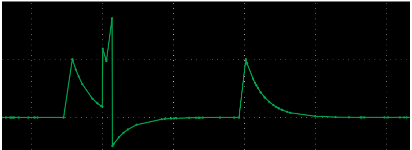
Simple filter

Non effective on the events by reverberation

- Accumulation and inhibition:
  - leaky integration over Δt,
  - comparison to a threshold,
  - hyper-polarization.

highly dependent on stimulus-driven activity
the parameters should be varied accordingly

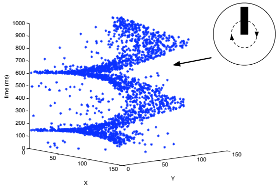


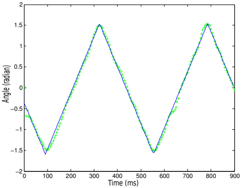


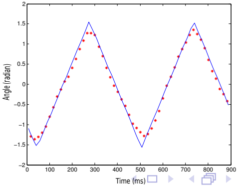
Orientation

Protocol

Use of a black bar painted on a white disk, rotating with a constant angular velocity



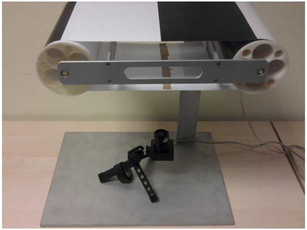


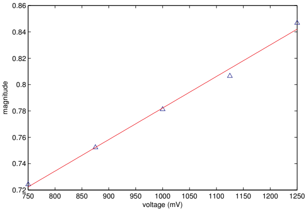


Amplitude

Protocol

Use of a moving pattern of bars whose translational speed can be accurately set by adjusting the supply voltage of a DC motor.

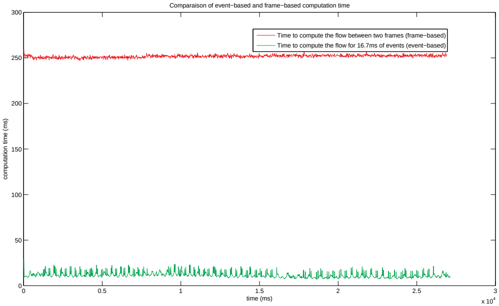




Computational cost

Protocol

Comparison of the computational times required by the frame-based and the event-based methods for computing optic flow



Computational cost

Comparison

- frame-based, $60 fps$ ($16.7 ms$)
  - number of processed pixels at each step: 16384,
  - mean computation time: $251.7 ms$

mean computation of a pixel: $15.4e^{-3} ms$

- event-based ($1e^{-3} ms$)
  - mean number of events (period of $16.7 ms$): 1340,
  - mean computation time: $9.65 ms$

mean computation of an event: $7.2e^{-3} ms$

Conclusion

The data-driven, *bioinspired* sampling

- removes redundancies,
- better captures the dynamic content of visual input.

event-based optical flow

The experiments showed that the new asynchronous event-based paradigm allows for

- high dynamic computation
- fast and low cost computation

Discussion

Further works

The method is sensible to noise, to deal with different method are explored:

- short term depression in input synapses,
- spike frequency adaptation of the leaky integration.