Denoising of Event-based Sensors with Spatial-temporal Correlation

[1]Jinjian Wu, [1]Chuanwei Ma, [2]Xiaojie Yu and [1]Guangming Shi

[1]School of Artificial Intelligence, Xidian University, Xi'an, China

[2]Beijing Institute of Space Mechanics & Electricity, Beijing, China

ICASSP 2020

Introduction

- Different from the traditional frame-based cameras, event-based cameras, such as the Dynamic Vision Sensor (DVS), are new kind of neuromorphic sensors.
- The output of DVS is in the form of address event representation (AER), which is completely different from traditional image and video. This means that the traditional image and video denoising method cannot be directly applied to event stream denoising.

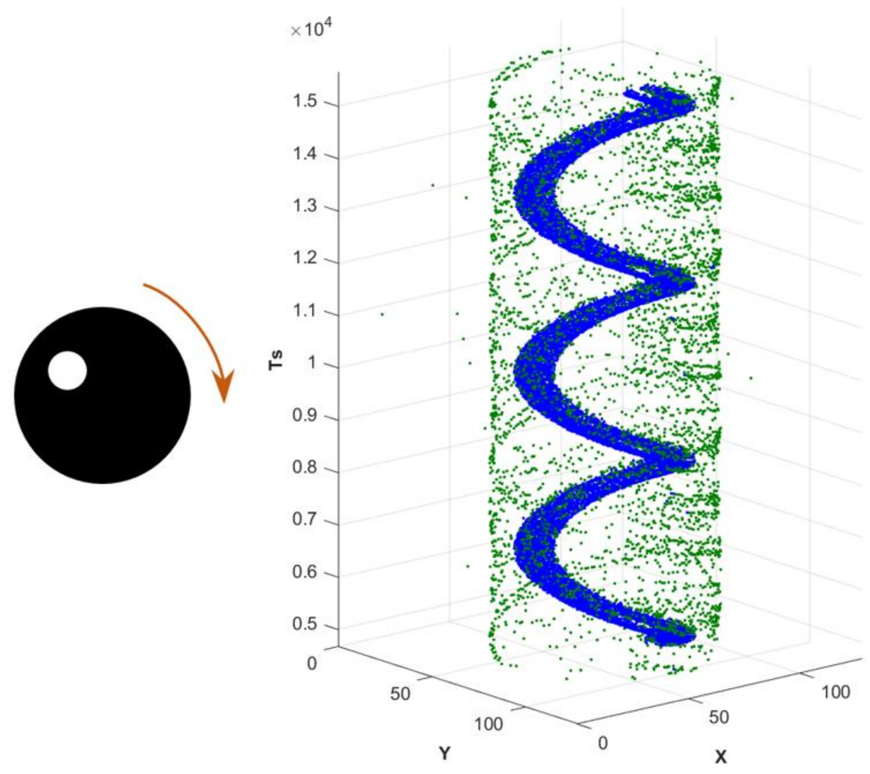


Fig.1 The event stream generated by DVS when shooting a rotating ball. Effective events are shown in blue and noise events are shown in green.

- In this paper, we propose a novel denoising method for constructing a probabilistic undirected graph model based on spatio-temporal correlation between events. Experiments show that our method can effectively remove noise events directly from the event stream and significantly improve event recognition rate

Motivation

- The regular movement of the object makes the change of the light intensity reflected by the object show a certain regularity, which means that the events in the effective event stream X generated by the movement of the object have a certain spatio-temporal correlation.

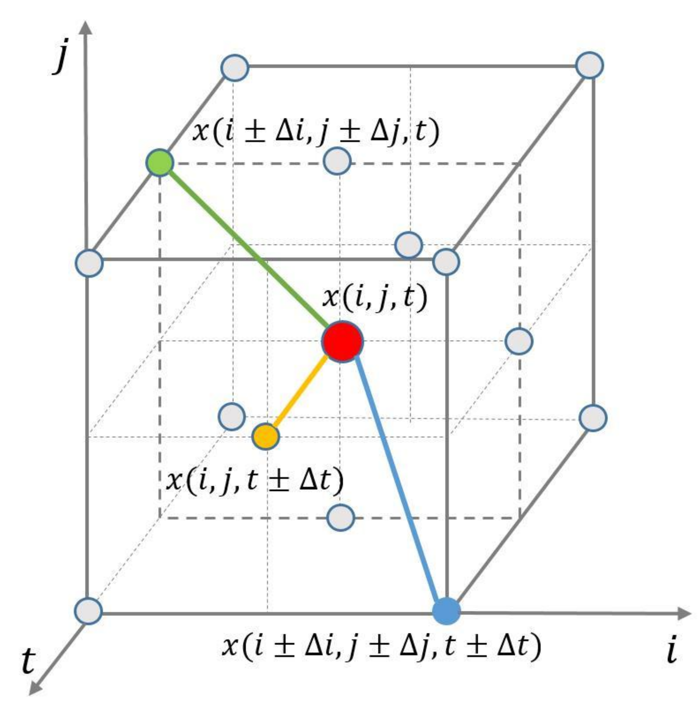


Fig. 2 Spatial and temporal correlations between effective events.

- The events in the noise event stream N generated by external environment interference and camera jitter have certain randomness. Therefore, the event stream Y output by DVS is considered to be obtained by adding N to X, which means that there is a strong correlation between Y and X.

$$Y = X + N$$

- Inspired by the traditional undirected graph model, we construct the joint probability distribution P(X,Y) to describe the above prior knowledge, with which the denoising problem is converted to a probability maximization problem.

Proposed Method

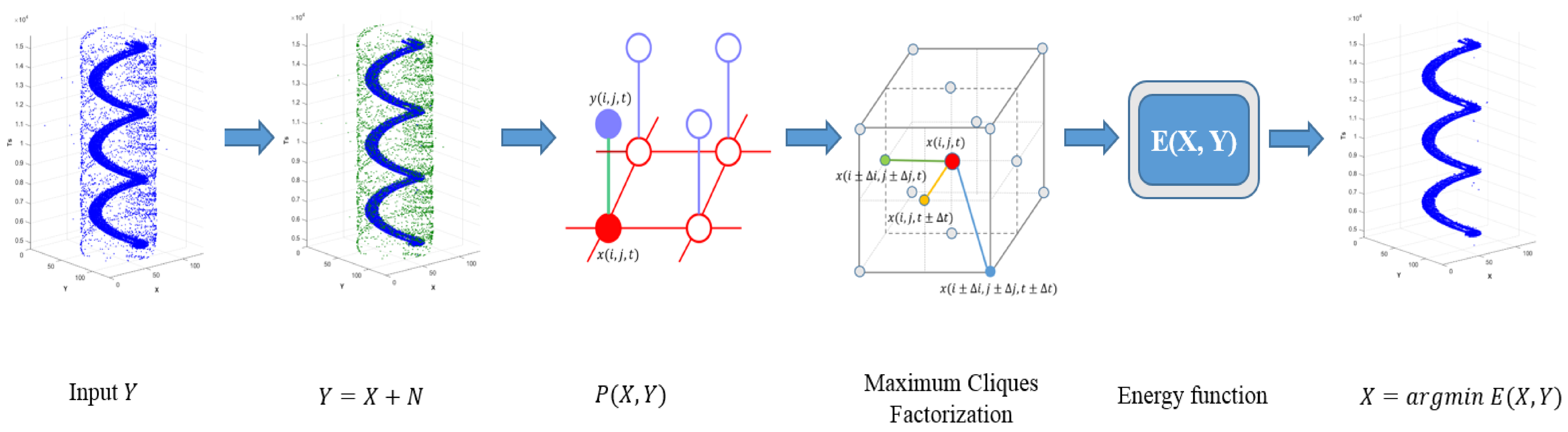


Fig.3 The construction process of probabilistic undirected graph model.

- Model construction
  - According to the factorization property of the probabilistic undirected graph, P(X,Y) can be decomposed into the product of potential functions on its maximum cliques.

$$P(X,Y) = \frac{1}{Z}\exp\{-E(X,Y)\}$$

  - The energy functions of the four largest clusters of the model are defined respectively .

$$E_{xy}(X_C Y_C) = \eta\left(2\left\|x_{i,j,t} - y_{i,j,t}\right\|_0 - 1\right),$$

$$E_s(X_C Y_C) = \frac{1}{Z_s\alpha}\left(2\left\|x_{i,j,t} - x_{i\pm\Delta i,j\pm\Delta j,t}\right\|_0 - 1\right),$$

$$E_t(X_C Y_C) = \frac{1}{Z_t\beta}\left(2\left\|x_{i,j,t} - x_{i,j,t\pm\Delta t}\right\|_0 - 1\right),$$

$$E_{st}(X_C Y_C) = \frac{1}{Z_{st}(\alpha+\beta)}\left(2\left\|x_{i,j,t} - x_{i\pm i,j\pm j,t\pm\Delta t}\right\|_0 - 1\right),$$

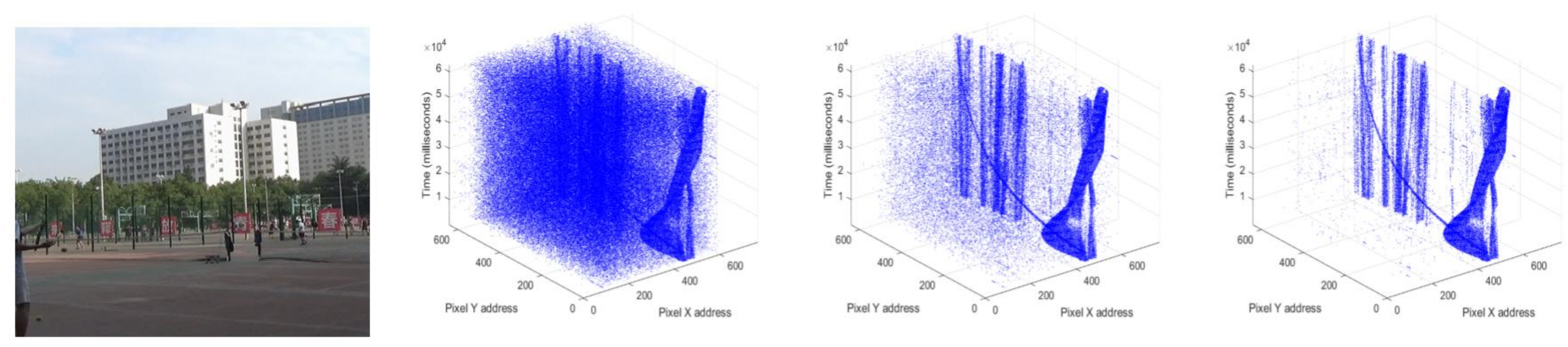
  - For better denoising effect, we hope to find an effective event stream X with a higher probability (ideally with the highest probability). This means finding an X that makes the model's energy lower (ideally with the lowest energy).

$$X = \arg\min E(X,Y)$$

- Model optimization
  - We first split the event stream into multiple 3D bloscks according to the number of events (dynamic time interval).
  - For the single event xi,j,t in each 3D block, the energy value of xi,j,t in two possible states (effective or noise) is calculated by fixing the state of the remaining events. The state with lower energy value is marked as current event state.

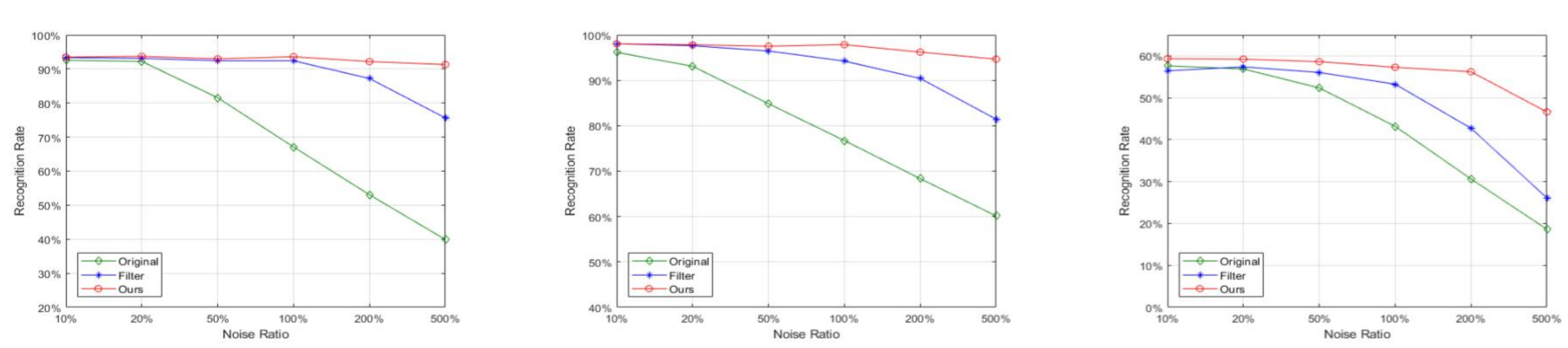
Experimental Results

- The denoising experiment in real scane.



(a) The experimental scene (b) 3D visualization (c) Filter (d) Ours

- The comparison of datasets recognition rate before and after denoising.



(a) Card (b) Posture (c) MNIST-DVS

- Comparison of recognition rate.

Dataset	Method	Noise Ratio					
		10%	20%	50%	100%	200%	500%
Card	Original	92.56%	92.23%	81.56%	67.08%	53.06%	39.89%
	Filter	93.35%	93.12%	92.43%	92.41%	87.21%	75.61%
	Ours	**93.48%**	**93.71%**	**92.99%**	**93.61%**	**92.18%**	**91.31%**
Posture	Original	96.17%	93.11%	84.86%	76.73%	68.34%	60.12%
	Filter	98.05%	97.63%	96.47%	94.24%	90.42%	81.44%
	Ours	**98.06%**	**97.83%**	**97.52%**	**97.87%**	**96.22%**	**94.68%**
MNIST-DVS	Original	57.63%	56.89%	52.37%	43.19%	30.63%	18.63%
	Filter	56.45%	57.39%	56.05%	53.25%	42.74%	26.09%
	Ours	**59.34%**	**59.23%**	**58.63%**	**57.29%**	**56.22%**	**46.62%**