جامعـــة خليفـــة
Khalifa University

Event-Based Camera Capture for High Amplitude and Abnormal Motion Events

Suhail Al Marzouqi
100045490

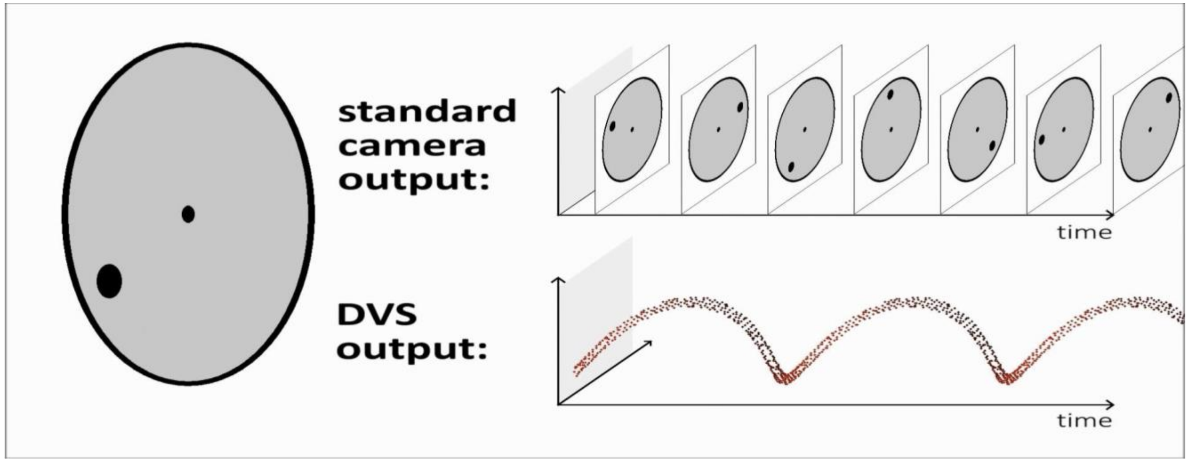
04 DECEMBER 2024

Outline

- Introduction
- Problem Statement and Objective
- Methodology
- Analysis and Discussion
- Conclusions and Future Work

Introduction

- **Motivation:**
  - **Traditional Frame-Based Cameras:**
    - Capture frames at fixed intervals (e.g., 30 fps).
    - Suffer from motion blur in fast-moving scenes.
    - Inefficient in handling high dynamic range (HDR) lighting conditions.
  - **Event-Based Cameras:**
    - Record brightness changes asynchronously for each pixel.
    - Output a stream of events in the form (x, y, t, p)
    - **High Temporal Resolution:** Events are recorded in microseconds.
    - **Low Latency:** No need to wait for the next frame.
    - **HDR Performance:** Handle extreme lighting variations effectively.



Problem Statement and Objective

- Motion segmentation in event-based data involves grouping events into Independently Moving Objects (IMOs).Challenges include:

  - **Sparsity of Data:**
    - Events are sparse compared to traditional pixel-based images.
  - **Asynchronous Nature:**
    - Events occur at different timestamps, requiring spatio-temporal alignment.
  - **Dynamic Environments:**
    - Objects can overlap, have non-rigid motion, or move at varying speeds.

Problem Statement and Objective

- The objective is to use motion segmentation with event-based cameras!

- **Key objectives:**
  - Leverage the asynchronous nature of events for real-time processing.
  - Construct a **spatio-temporal graph** to model relationships between events.
  - Use energy-based optimization to achieve precise segmentation.

Methodology Overview

- **Input:** Raw event data.
- **Steps**: Spatio-temporal graph construction, motion compensation, model fitting, and segmentation.
- **Output:** Segmented IMOs with clear labels.

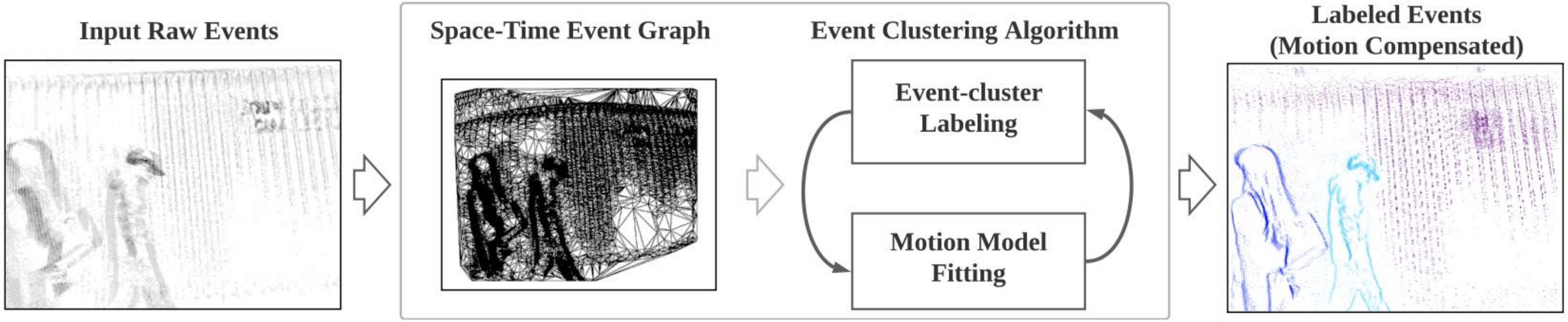


Figure adopted from [1]

Spatio-Temporal Graph Construction

- **Why Use a Graph?**
  - Event data from event-based cameras is sparse and asynchronous, making it challenging to process using traditional grid-based methods.
  - A spatio-temporal graph efficiently models the relationships between events in both space and time, enabling structured processing for motion segmentation.
- **1. Each event is represented as:**
  - $ei = (xi, yi, ti),$
  - (xi, yi) are spatial coordinates of the event on the image plane.
  - ti: Timestamp of the event.

Figure ac

Spatio-Temporal Graph Construction

- **2. Delaunay Triangulation**
  - Given events in a Volume **V** of size $W \times H \times \delta t$
  - Compute a binary image of event activity (i.e., the pixel is 1 if it contains at least one event, and it is 0 otherwise)
  - Delaunay triangulation on the non-zero pixels of this binary image which forms a 2D graph
- **3. Connection Principle**
  - To construct the 3D spatio-temporal graph, the concept of connection principle is used which extends the 2D graph into a 3D spatio-temporal graph. Each event (orange dot) connects to Its two temporally-closest events at the same pixel (red dots). Its neighboring events in the Delaunay triangulation (green dots). This results in a graph with 2 + 2N neighbors per event, where N is the number of neighboring edges from the Delaunay triangulatio

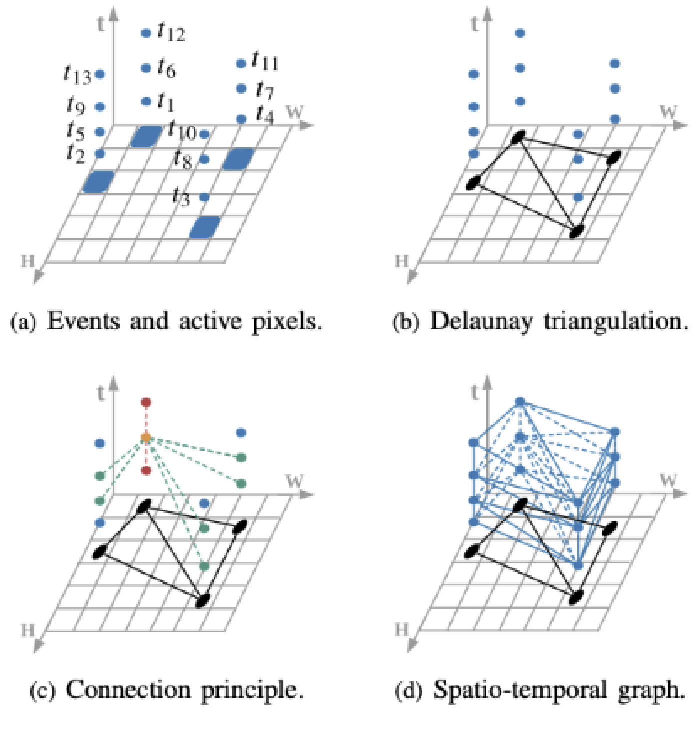


Fig. 1. Spatio-Temporal Graph Construction [7]

Figure adopted from [1]

Motion Compensation and Energy Formulation

- Motion compensation aligns events that belong to the same moving object by warping their spatial coordinates based on a motion model.
- Proper alignment ensures that events from the same object are spatially coherent, forming a sharp representation in the Image of Warped Events (IWE).

- The IWE aggregates the warped events over a given time window:

$$I(x;m) = \sum_{k=1}^{N_e} \delta(x - x'_k),$$

- x'_k: The warped spatial position of the $k-th$ event after applying the motion model mmm.
- $\delta(x - x'_k)$: Impulse function at the warped location.
- m: Parameters of the motion model (e.g., translation, rotation).

- When the motion model m aligns events correctly, the IWE becomes sharp, with high contrast in regions corresponding to moving objects. Misaligned events produce blurred IWEs, reducing contrast.

Energy-Based Segmentation Framework

- To achieve motion segmentation, the framework minimizes a global energy function

$$E(L, M) = E_D(L, M) + \lambda_P E_P(L) + \lambda_M E_M(L),$$

- Where**:**
  - L: Labels for event clusters.
  - M: Motion models for each cluster.
  - E_D: Data fidelity term.
  - E_P: Pairwise regularization term.
  - E_M: Label cost term.
  - λ_P, λ_M: Weights for regularization and label cost.

Energy-Based Segmentation Framework

- **Data fidelity** ensures that events fit their assigned motion model by maximizing the contrast of the IWE.

$$E_D(L, M) = \sum_{l \in L} \sum_{e_k \in C_l} \bar{I}(x'_k; m_l),$$

- where:
  - C_l: Cluster of events assigned to label l.
  - m_l: Motion model parameters for label l.
  - $\bar{I}\,(x'_k;\ m_l)$:Normalized and negated IWE contrast for events in cluster C_l.

Energy-Based Segmentation Framework

- **Pairwise Regularization Term** encourages spatial coherence by promoting smoothness between neighboring event labels.

$$E_P(L) = \sum_{(i,j)\in N} \delta_{L(i),L(j)},$$

- where:
  - N: Set of neighboring event pairs (defined by the spatio-temporal graph).
  - $\delta_{L(i),L(j)}$: Kronecker delta, equal to 1 if labels $L(i)$ and $L(j)$are the same, and 0 otherwise.

Energy-Based Segmentation Framework

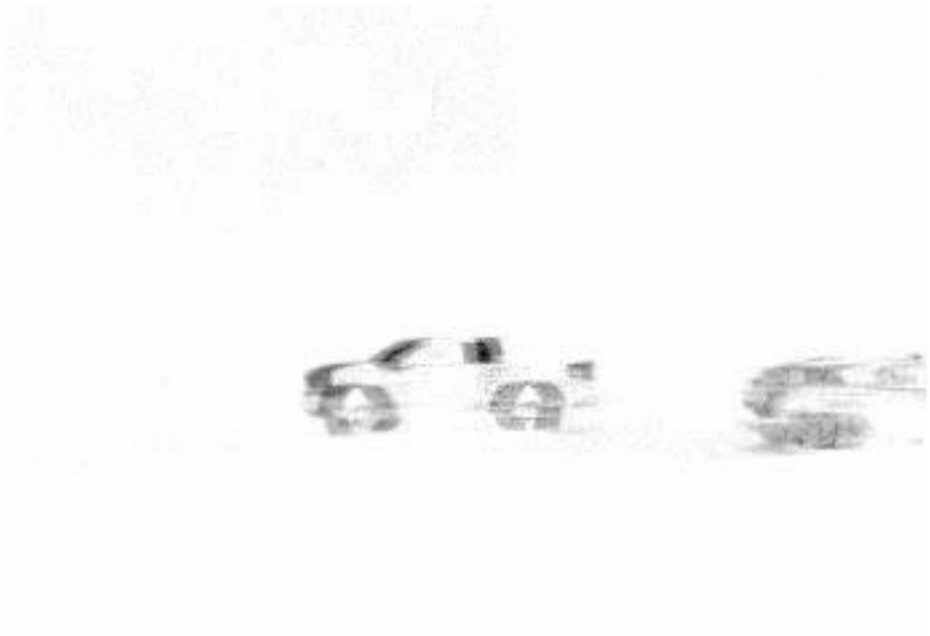
- **Label Cost Term** penalizes the creation of unnecessary clusters

$$E_M(L) = \sum_{l=1}^{N} \psi(l),$$

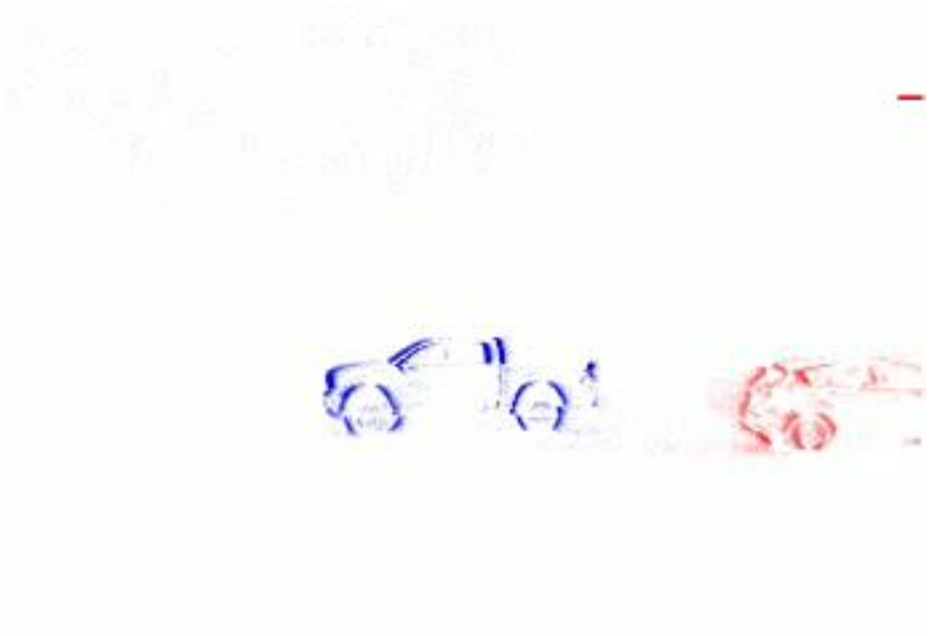
- where:
  - $\psi(l) = 1$ if cluster C_l is active (contains events), otherwise 0.

- The energy function integrates three essential components:
  - **Data Fidelity**: Ensures alignment of events to their motion model.
  - **Spatial Coherence**: Promotes smooth segmentation.
  - **Parsimony**: Avoids over-segmentation.
  - By alternating between optimizing L (event labels) and M (motion models), the framework achieves robust segmentation.

Results



Raw Events



Segmented Motion Events

Conclusion

- **Robust Framework for Motion Segmentation:**
  - Presented a novel energy-based framework that leverages spatio-temporal graphs and motion compensation for segmenting Independently Moving Objects (IMOs).
  - Demonstrated the ability to handle rigid, non-rigid, and high-speed motions effectively.
- **Integration of Key Concepts:**
  - Combined spatial and temporal information in a unified spatio-temporal graph.
  - Optimized segmentation using a multi-term energy function:
    - **Data Fidelity**: Ensuring events align with their motion model.
    - **Spatial Coherence**: Maintaining smoothness in event labeling.
    - **Label Cost**: Encouraging compact and efficient segmentation.
- **Extensive Validation:**
  - Evaluated on diverse datasets (e.g., EMSMC, DistSurf, EVIMO2, and custom sequences).
  - Addressed challenging scenarios, including high-speed events, HDR conditions, and complex overlapping motions.

Conclusion

- **Key Takeaway**
  - **Efficiency and Scalability:**
    - The framework is computationally efficient, supporting real-time applications with high-resolution event cameras.
    - Scalable to dynamic environments with multiple IMOs.

- **Limitations**
  - Sparse Event Regions:
    - Performance may degrade in areas with sparse events or minimal motion.
  - Subtle Non-Rigid Motion:
    - Challenges remain in distinguishing small deformations or overlapping motions with similar dynamics.

جامعة خليفة
Khalifa University

Thank You

ku.ac.ae