



Deteção de Objetos por Eventos

Marcelo Couto

28/07/2022

Enquadramento

- **Deteção de Objetos em Vídeo**

- estratégia comum envolve processamento individual de cada imagem
- Imagens normalmente demasiado complexas
- Processo computacionalmente pesado

- **Deteção de Objetos Baseada em Eventos**

- Utiliza Câmaras de Gravação de Eventos
- Nova abordagem
- Grande potencial

- Neste estágio pretende-se estudar a abordagem baseada na deteção de **eventos** e testar estratégias que permitam a sua aplicação para deteção de objetos



DAVIS240C

Objetivos

- **Fases:**

- Identificar e analisar **datasets**
- Estudar abordagens de **agregação de eventos**
- Estudar algoritmos de **deteção de objetos**

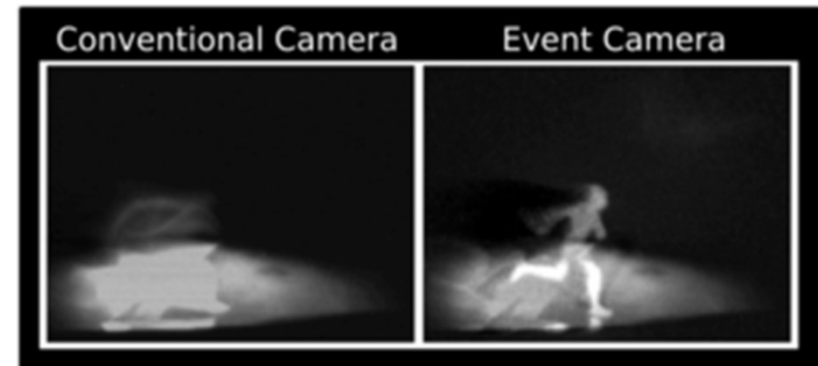
- **Objetivos:**

- Analisar o estado das tecnologias e dados disponíveis
- Investigar sobre o potencial das câmaras de eventos na deteção de objetos



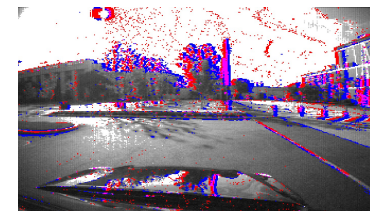
Trabalho desenvolvido – Câmara de Eventos

- **Câmara de gravação de vídeo RGB** – cria vídeo por agregação de várias *frames*, compostas pela informação da luminosidade e cor de cada pixel.
- **Câmara de eventos (sensores DVS)** – cada pixel independentemente regista eventos referentes a mudança de luminosidade
- **Vantagens**
 - Alta Gama Dinâmica (até 120dB)
 - Resolução Temporal Alta
 - Latência Baixa
 - Sem *Motion Blur*
 - Baixo Consumo
 - Banda Larga necessária reduzida
- **Desafios**
 - Resolução Espacial Baixa (240x180)
 - Falta de Algoritmos e Dados
 - Especialmente suscetíveis a ruído



Trabalho desenvolvido – Datasets

- De momento, existe um número **reduzido** de datasets disponíveis, principalmente com **dados de automobilismo**.
 - **Automobilismo**: GEN1 Automotive Detection Dataset, N-CARS Dataset, MVSEC
 - **Variado**: DVSMotion20, Event Camera Dataset
 - **Pose humana**: Dynamic Vision Sensor (DVS) 3D Human Pose Dataset
- Representação do evento: **(timestamp, x, y, polaridade)**
- Datasets Utilizados
 - **Event Camera Dataset**
 - Múltiplos cenários
 - DAVIS240C
 - Ficheiros de texto
 - **MVSEC**
 - Automobilismo e cenários urbanos
 - Sistema multi-sensores com DAVIS346B
 - Ficheiros h5py

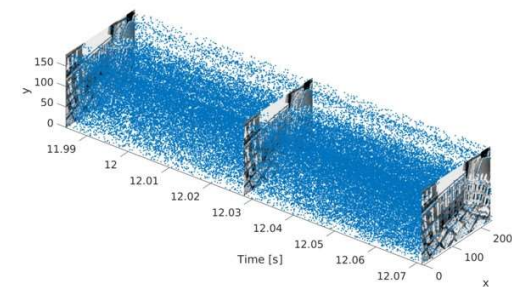
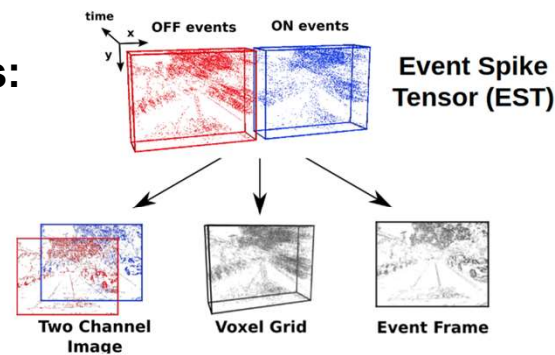


Trabalho desenvolvido – Abordagens de Agregação de Eventos

- A utilização direta de eventos na detecção de objetos geralmente não resulta.

- **Agregação de eventos:**

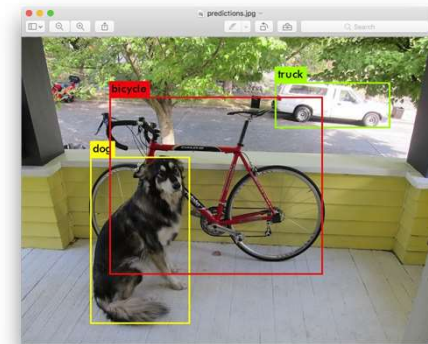
- Event Frames
- Voxel Grids
- Event Cubes
- Voxel Cubes
- GreyScale Frames



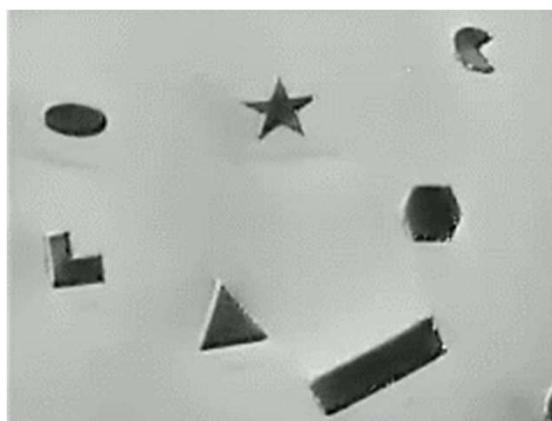
- **Event Frames** – foi desenvolvido um algoritmo que efetua a agregação de eventos em cada pixel e forma uma imagem de densidade.
- **GreyScale Frames** – foi utilizado um projeto externo. O algoritmo presente no mesmo utiliza uma **rede neuronal recorrente convolucional**.

Trabalho desenvolvido – Algoritmo de Detecção de Objetos

- **YOLO – You Only Look Once**
 - Algoritmo de deteção de objetos que se destaca pela sua precisão e rapidez
 - O **nome do algoritmo** tem origem no facto de que a rede neuronal é percorrida **apenas uma vez** para efetuar deteções
 - **YOLOv5** é um modelo pré-treinado e foi o modelo utilizado no trabalho
 - Usado em combinação com geração de imagens *GreyScale*
- Outras abordagens investigadas
 - Agregação em *Voxel Cubes* e deteção com *Spiking Neural Networks*
 - *Online Learning* com SVMs

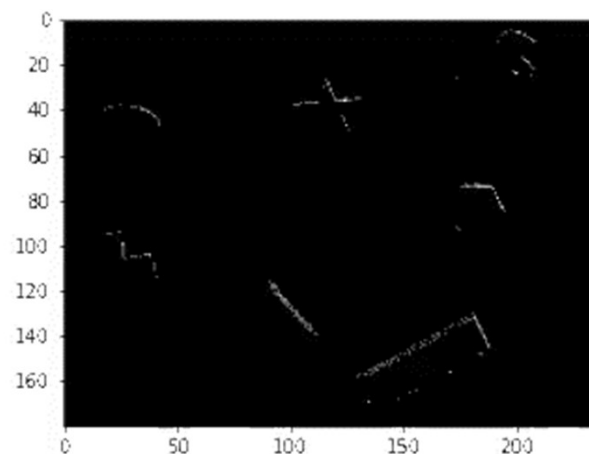


Resultados obtidos – Agregação de Eventos



GreyScale

- GreyScale
 - Algoritmo mais pesado
 - Imagens de melhor qualidade



EventFrames

- EventFrames
 - Algoritmo simples e leve
 - Imagens mais rudimentares

Resultados obtidos – Detecção de Objetos

Real-time Shape Tracking (C++)

Mueggler, E., Rebecq, H., Gallego, G., Delbruck, T., & Scaramuzza, D. (2017). The event-camera dataset and simulator: Event-based data for pose estimation, visual odometry, and SLAM. The International Journal of Robotics Research, 36(2), 142-149

- Imagens obtidas de projeto externo (e-TLD)
- Detecção efetuada com SVMs em Online Learning

Resultados obtidos – Detecção de Objetos



- Imagens geradas através de redes neuronais convolucionais
- Detecção utilizando YOLOv5

Resultados obtidos – Geral

- **Vantagens do método adotado**

- Reconstrução para GreyScale origina imagens de muito maior qualidade
- YOLOv5 é um modelo extremamente rápido e eficaz

- **Desvantagens**

- Reconstrução para GreyScale é um passo pesado comparado com os outros
 - Implica a introdução de um delay
- Abordagens com outras formas de agregação de eventos mais simples aproveitam melhor os benefícios das câmaras de eventos

Conclusões

- O trabalho desenvolvido foi principalmente de **cariz exploratório**
- Permitiu **avaliar** o corrente estado das **tecnologias e recursos**
 - Poucos datasets disponíveis
 - Projetos e algoritmos escassos e de difícil utilização
- Área **pouco desenvolvida** mas com **grande potencial**
 - Alta resolução temporal e gama dinâmica – grande potencial de deteção em cenários complexos
 - Baixas necessidades de banda larga e consumo de energia – algoritmos mais leves – permite a implementação dos mesmos mais perto dos sensores
- Resultados atingidos:
 - Reconstrução de vídeo através de eventos
 - Reconstrução de *frames* através de eventos
 - Deteção de Objetos utilizando um modelo YOLO e *frames* geradas através de agregação de eventos



ESTÁGIOS
de VERÃO 2022 @CTM

 **INESCTEC**