B.Tech project

INTELLIGENT TRANSPORT SYSTEMS

TIMELINE

01
Introduction

02
Problem Statement

03
Proposed Approach and Algorithm

04
Datasets

05
Implementation

06
Results

07
Applications

09
Conclusion

08
Future Directions

AUTONOMOUS DRIVING SYSTEMS

An autonomous car is a vehicle capable of sensing its environment and operating without human involvement. A human passenger is not required to take control of the vehicle at any time, nor is a human passenger required to be present in the vehicle at all. An autonomous car can go anywhere a traditional car goes and do everything that an experienced human driver does.

Levels of driving Automation

LEVEL 0

NO AUTOMATION
Manual control. The human performs all driving tasks (steering, acceleration, braking, etc.)

LEVEL 1

DRIVER ASSISTANCE
The vehicle features a single automated system (e.g. it monitors speed through cruise control).

LEVEL 2

PARTIAL AUTOMATION
ADAS. The vehicle can perform steering and acceleration. The human still monitors all tasks and can take control at any time.

LEVEL 3

CONDITIONAL AUTOMATION
Environmental detection capabilities. The vehicle can perform most driving tasks, but human override is still required

LEVEL 4

HIGH AUTOMATION
The vehicle performs all driving tasks under specific circumstances. Geofencing is required.
Human override is still an option.

LEVEL 5

FULL AUTOMATION
The vehicle performs all driving tasks under all conditions. Zero human attention or interaction is required.

ITS SYSTEMS

The Main Component of ITS Systems is real-time speed detection.

TWO OPTIONS

HARDWARE

LiDAR

Laser light to measure distances

- High Accuracy
- Minimum Human Dependency

- Price is too high
- Sensitive to weather conditions(rain and snow)

Acoustic wave pattern

Acoustic drive-by sounds recorded at a microphone.

- Works in severe weather

- Price is too high
- Low Range

SOFTWARE

No respectable research done in this field.

PROBLEM STATEMENT

Real-time Speed Detection

Need for a cost-effective, reliable, easy-to-scale software alternative

SOFTWARE ALTERNATIVE

- The Main Bone for speed estimation is to find the distance between the cars, obstacles, etc

How to predict the distance of the object using image processing?

METHODS TO FIND DEPTH OF OBJECTS IN IMAGE

Method 1

SfM Algorithm
(Structure from Motion)

Method 2

Monocular depth estimation using deep learning

Method 3

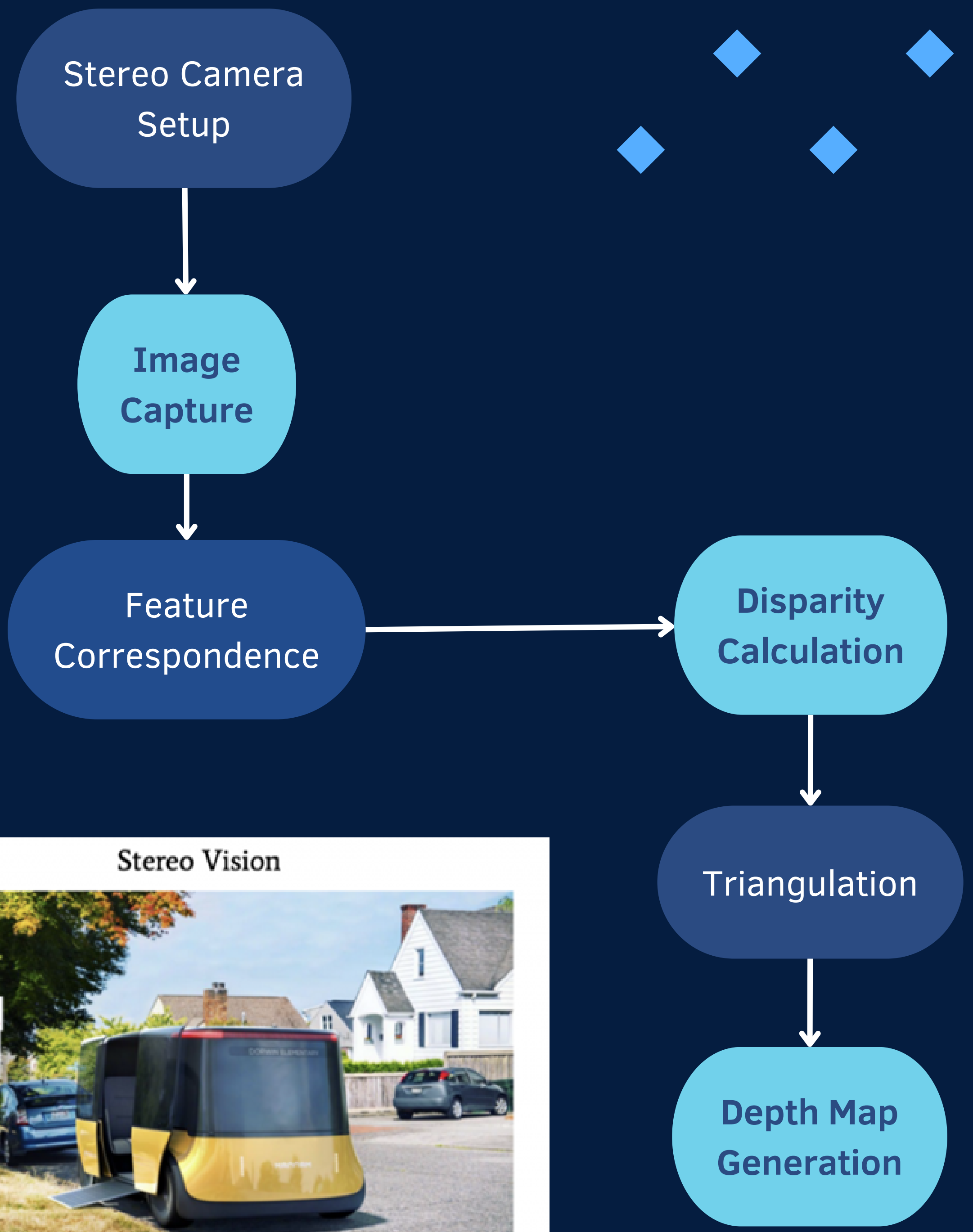
Stereo Camera Approach

METHOD 3

STEREO VISION

Ibrahim, O., ElGendy, H., & ElShafee, A. M. (2011). Speed Detection Camera System using Image ProcessingTechniques on Video Streams. International journal of computer and electrical engineering, 3(6), 771.

STEREO CAMERA APPROACH (STEREO VISION)



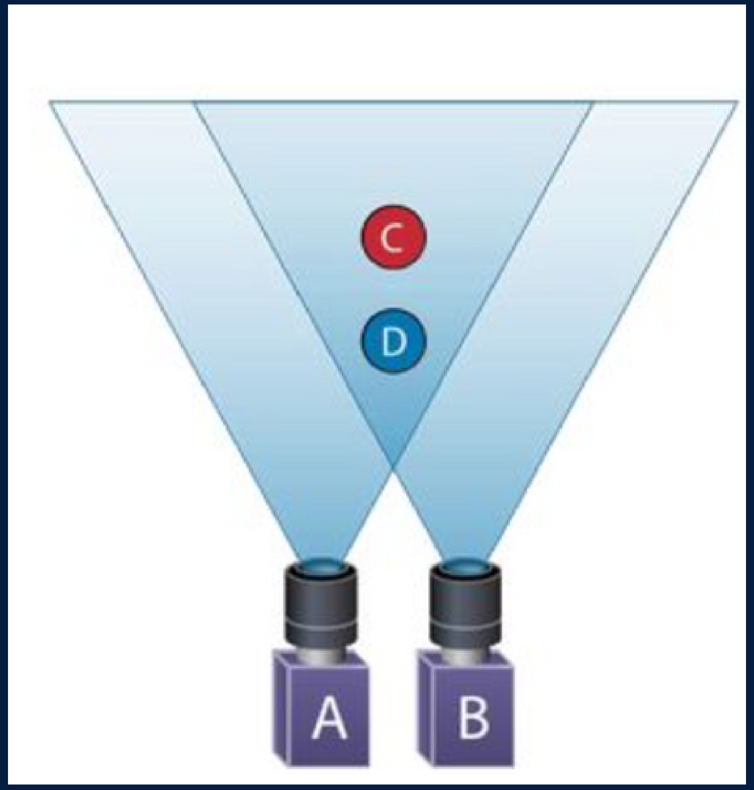
Technique used in computer vision and robotics to estimate the depth or distance to objects in a scene

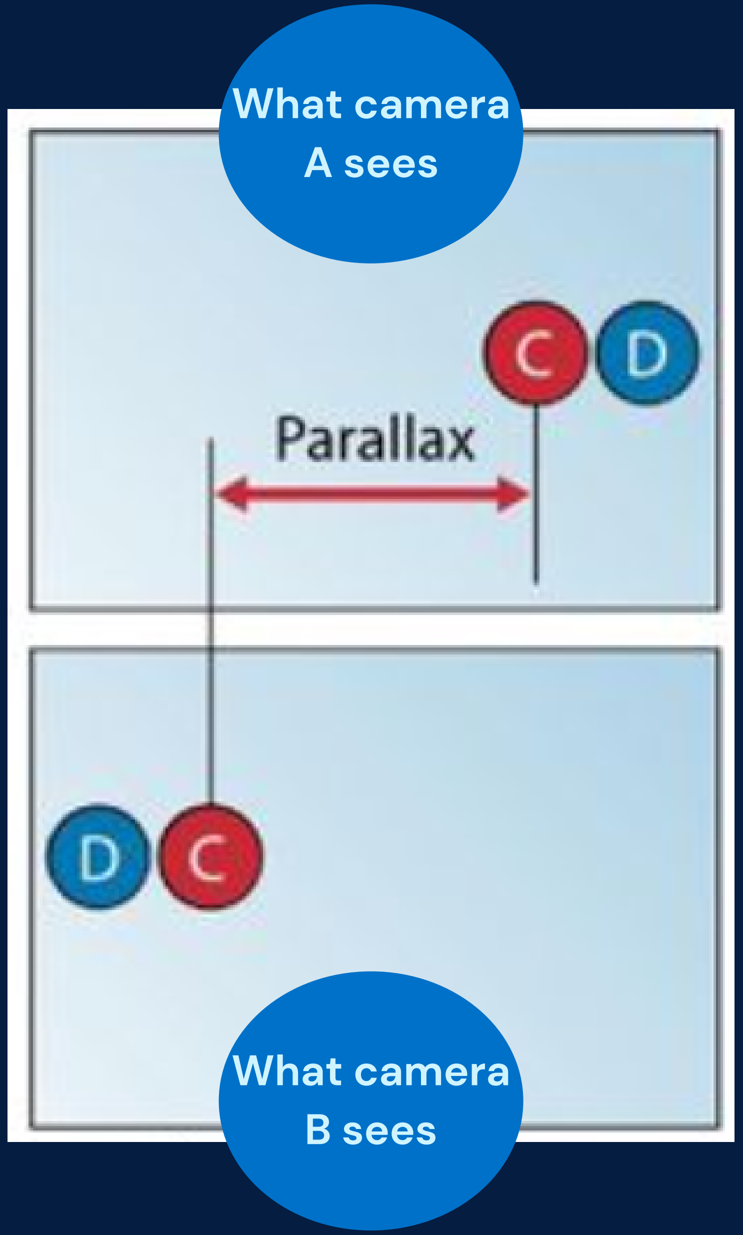
Relies on the principle of triangulation, similar to how human vision works with two eyes.

- **Accurate Depth Perception**
- **Real-time Capability**
- **No Active Illumination Required**
- **Cost-Effective**
- **Wide Field of View**

Ibrahim, O., ElGendy, H., & ElShafee, A. M. (2011). Speed Detection Camera System using Image ProcessingTechniques on Video Streams. International journal of computer and electrical engineering, 3(6), 771.

STEREO VISION-BASICS (I)





Two camera setup

- Stereo Vision can infer the depth of objects using two images from two different cameras sitting a known distance apart.
- By computing the distance between the same object in the two images, the depth can be calculated using simple triangulation.

Ibrahim, O., ElGendy, H., & ElShafee, A. M. (2011). Speed Detection Camera System using Image ProcessingTechniques on Video Streams. International journal of computer and electrical engineering, 3(6), 771.

STEREO VISION-BASICS (II)

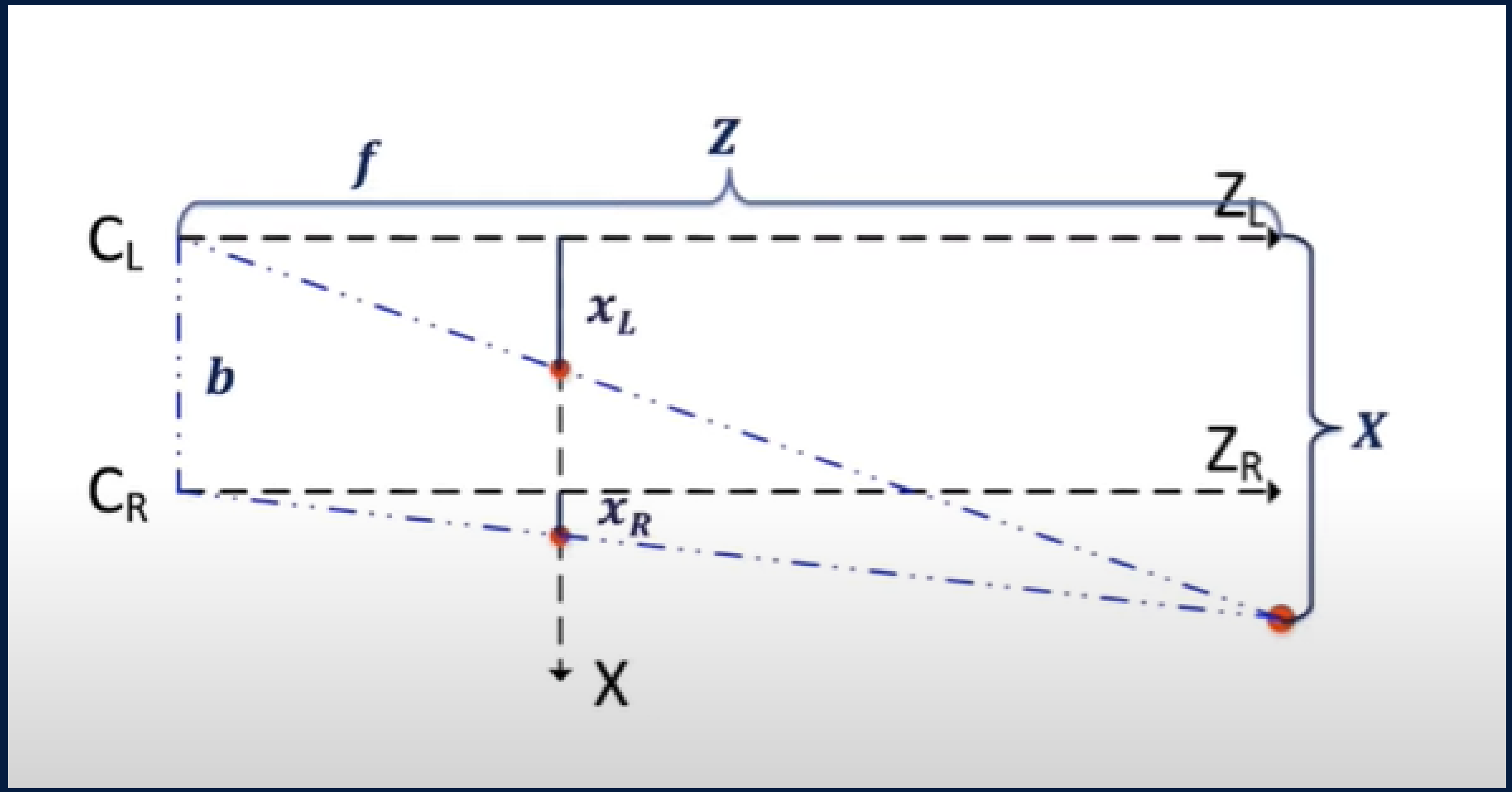
For a simple stereo system, the depth of a point (z) is given by:

$$\boxed{Z = \frac{f * b}{d}}$$

f = focal length

b = baseline, or interocular distance between the cameras

d = disparity between corresponding points.



$$\frac{Z}{f} = \frac{X}{x_L} \qquad \frac{Z}{f} = \frac{X-b}{x_R}$$

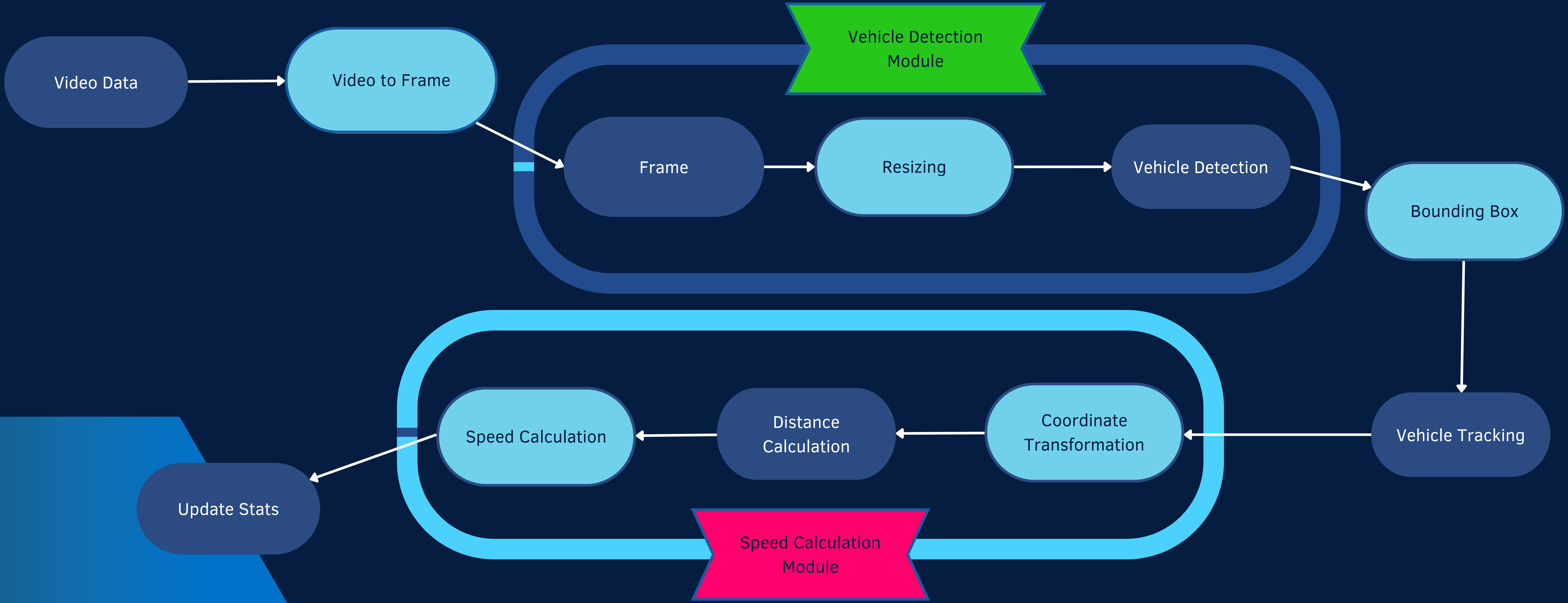
$$x_L = \frac{X}{Z} \cdot f \qquad x_R = \frac{X-b}{Z} \cdot f$$

$$Disparity = x_L - x_R$$

Ibrahim, O., ElGendy, H., & ElShafee, A. M. (2011). Speed Detection Camera System using Image ProcessingTechniques on Video Streams. International journal of computer and electrical engineering, 3(6), 771.

PROPOSED APPROACH

REAL-TIME SPEED ESTIMATION PIPELINE



RELATIVE VELOCITY IMPLEMENTATION

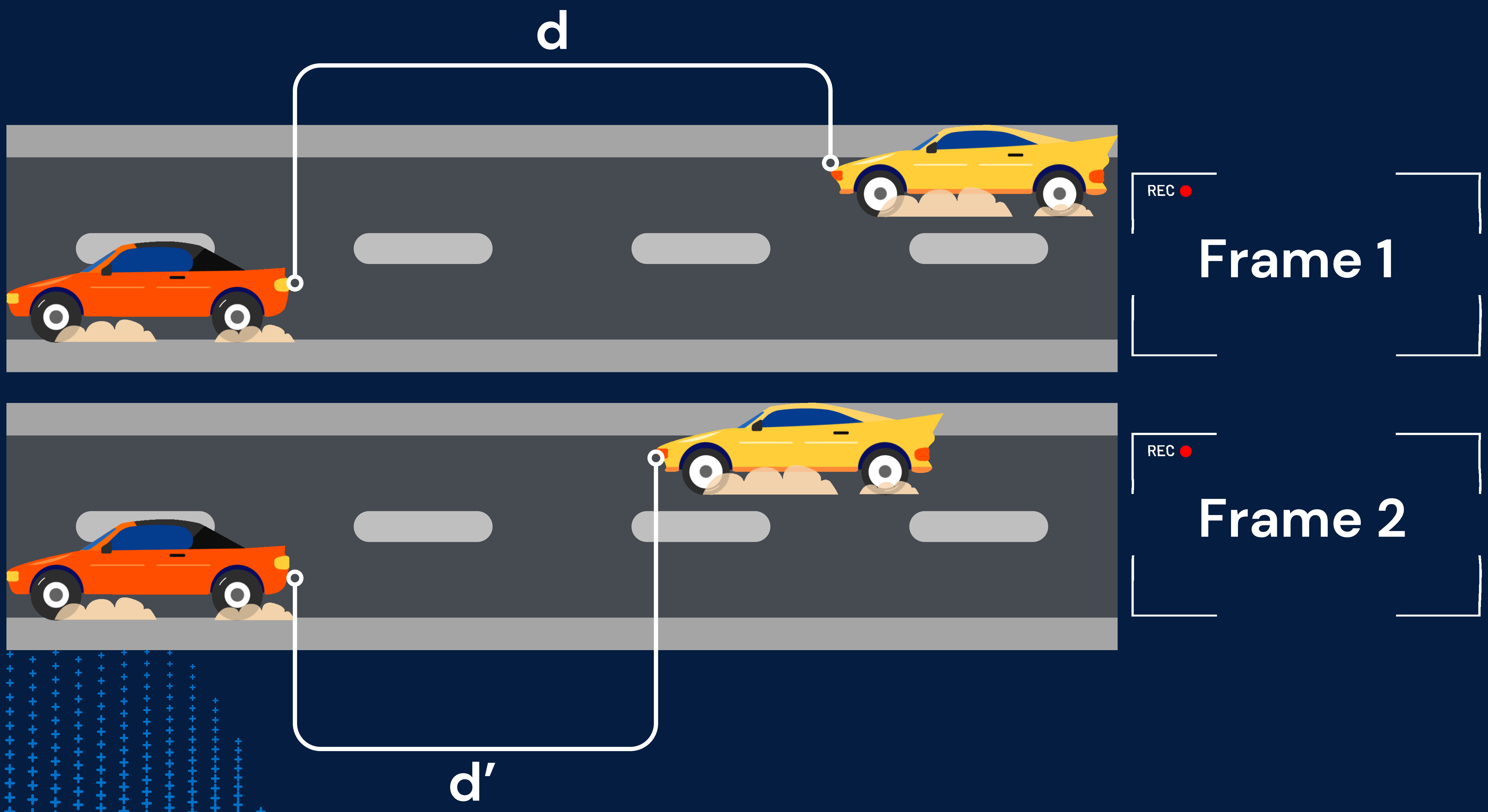
Two Implementations -

Dynamic Camera

Static Camera

DYNAMIC CAMERA APPROACH

- Camera mounted on the car itself.



DYNAMIC CAMERA APPROACH(II)

CALCULATIONS:

- Change in distance between cars $= d - d'$
- Time between every frame $= t$
- Relative velocity for given frame change $= \dfrac{\text{change in distance}}{\text{Time between frames}}$

$$\boxed{\text{Relative velocity} = \frac{d - d'}{t}}$$

CURRENT SCENARIO

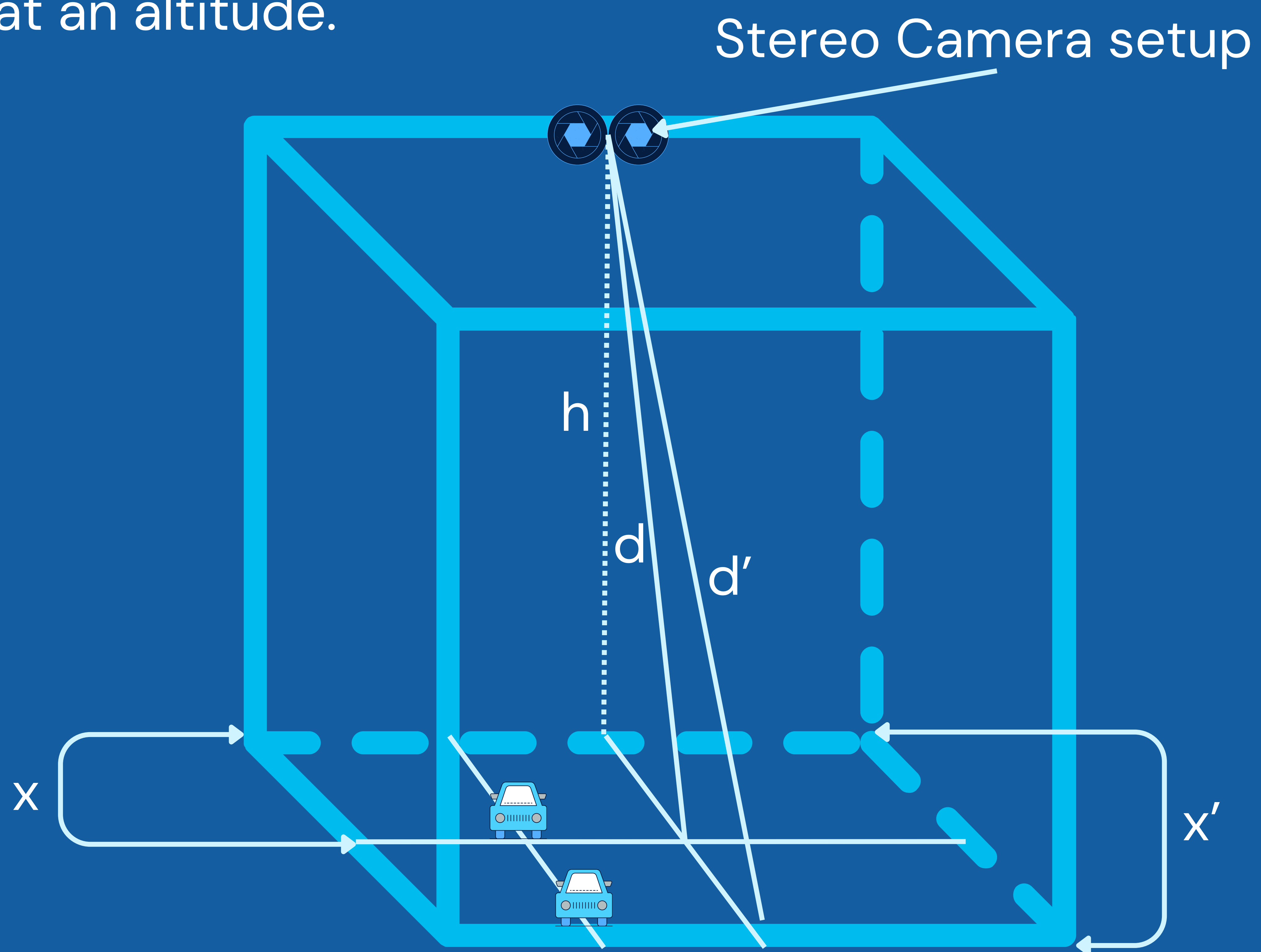
- Speed detected is an average speed from point A to point B.
- Real time relative velocity not calculated.
- Autonomous driving depends upon the instantaneous relative velocity of the neighbouring vehicles.
- Optimal velocity for autonomous driving cannot be calculated!

$$\text{SPEED} = \frac{\text{DISTANCE BETWEEN A AND B}}{\text{TIME TAKEN FROM A TO B}}$$

NOT REAL TIME SPEED!

STATIC CAMERA APPROACH

- Camera mounted at an altitude.



AFTER MID-EVALUATION

DATASETS

KITTI Stereo 2012

KITTI Stereo 2015

Cityscapes dataset

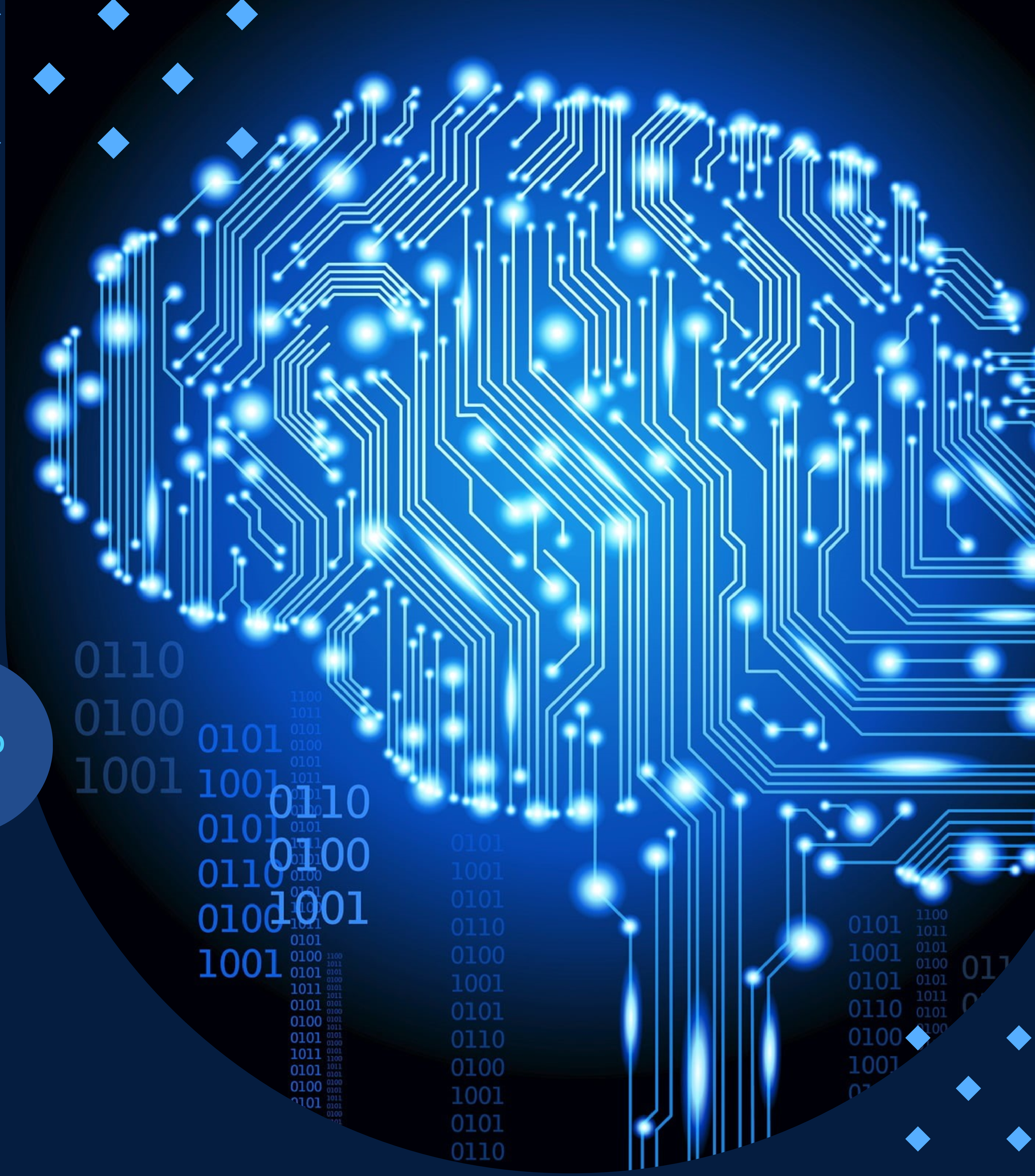
Synthetic Datasets:

- Sintel
- Scene Flow
- Virtual KITTI

DrivingStereo

ETH 3D

Middleburry Stereo



G. Yang, X. Song, C. Huang, Z. Deng, J. Shi and B. Zhou, "DrivingStereo: A Large-Scale Dataset for Stereo Matching in Autonomous Driving Scenarios," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 899-908, doi: 10.1109/CVPR.2019.00099.

Karlsruhe Institute of Technology and
Toyota Technological Institute

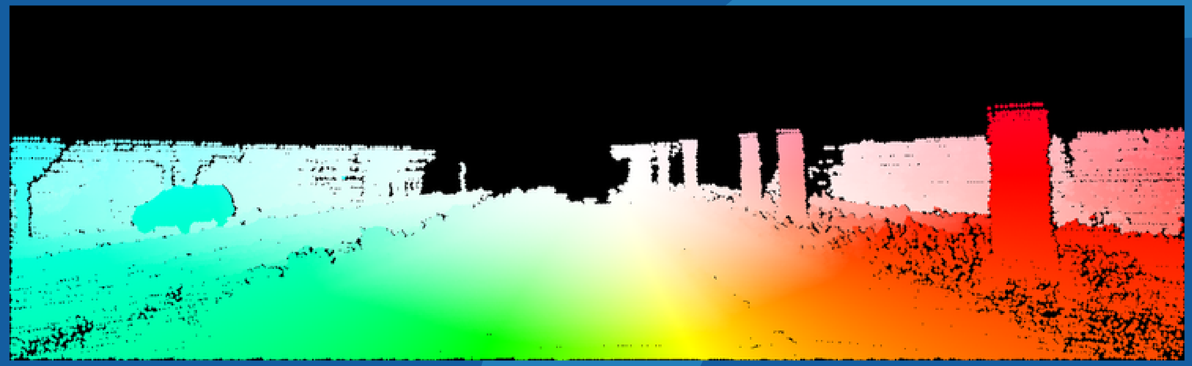
DATASET EXAMPLE

- KITTI Stereo 2015

Left camera

Right camera

Ground truth LiDAR reading



Ground truth Depth reading

ristoph ... Raquel Urtasun. “Vision meets robotics: The KITTI dataset.” The International Journal of Robotics Research 32 (2013): 1231 – 1237.

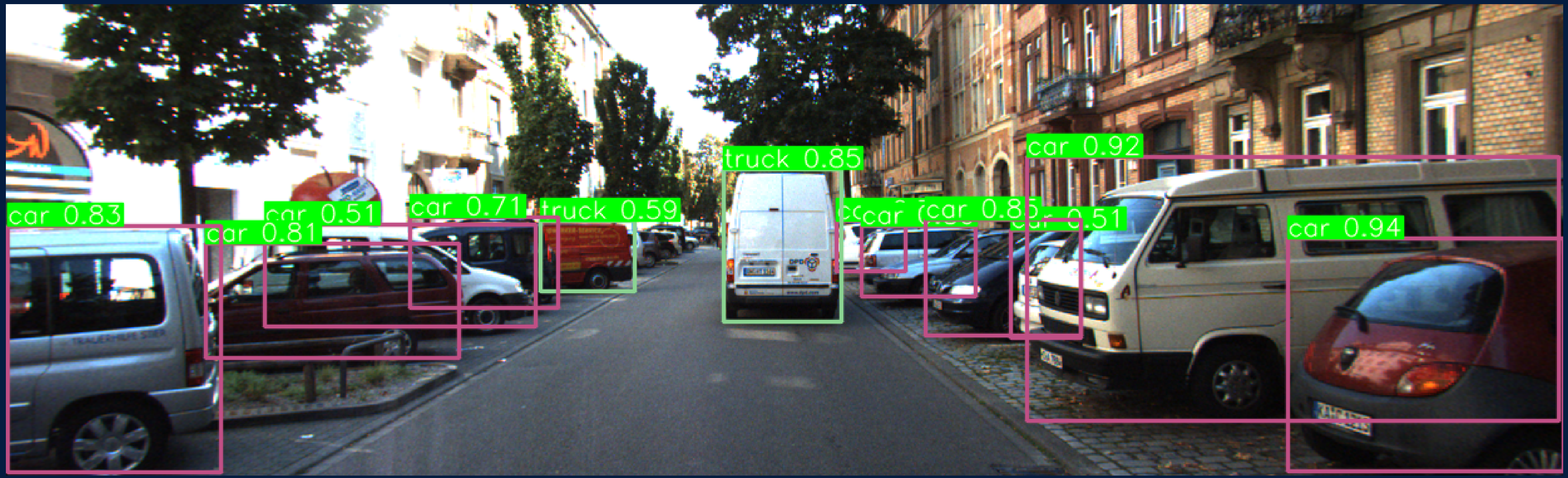
IMPLEMENTATION

OBJECT DETECTION

01 The detection and localisation of the object using any object detector

02 YOLOv7 surpasses all known object detectors in both speed and accuracy

03 YOLOv7 trained on MS COCO dataset having 80 classes



Wang, C., Bochkovskiy, A., & Liao, H. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. ArXiv. /abs/2207.02696

IMPLEMENTATION

OBJECT TRACKING

- Algorithm that tracks objects and assigns a unique ID to each object

- Types
  - Single Object Tracker: CSRT, KCF
  - Multi Object Tracker: DeepSORT, JDE and CenterTrack.

- DeepSORT introduces deep learning into the SORT algorithm by adding an appearance descriptor to reduce the identity switches, making the tracking more efficient.

Wojke, N., Bewley, A., & Paulus, D. (2017). Simple Online and Realtime Tracking with a Deep Association Metric. ArXiv. /abs/1703.07402

IMPLEMENTATION

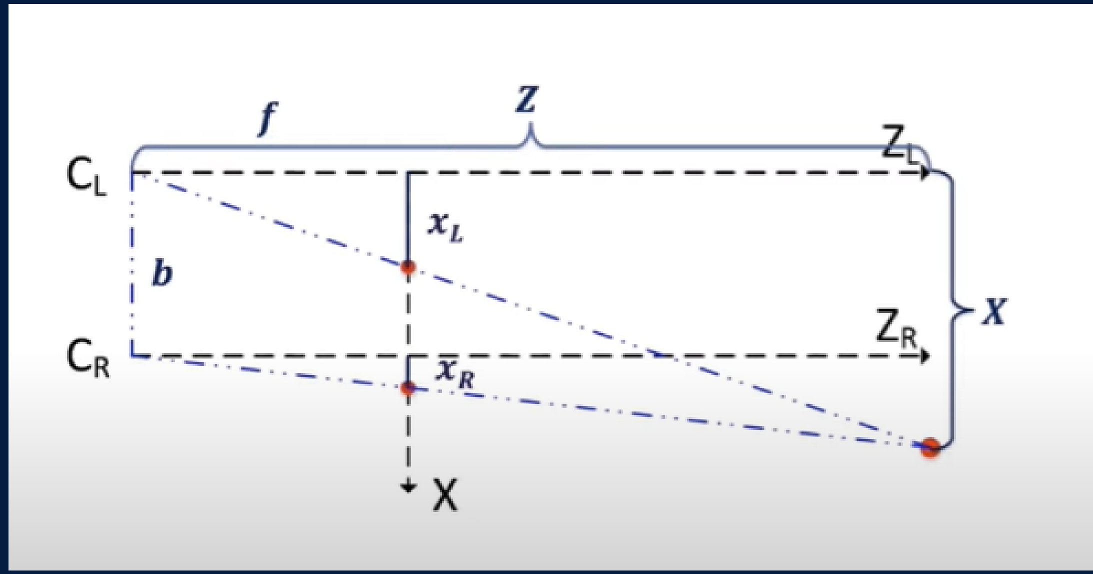
STEREO MATCHING

- Process of finding the pixels in the multiscopic views that correspond to the same 3D point in the scene
- Because the accuracy of depth is dependent on the accuracy of disparity, estimating an accurate disparity map is crucial in stereo vision
- Iterative Geometry Encoding Volume (IGEV-Stereo), a deep network architecture for stereo matching, ranks 1st on KITTI 2015 and 2012 (Reflective) among all published methods

Xu, G., Wang, X., Ding, X., & Yang, X. (2023). Iterative Geometry Encoding Volume for Stereo Matching. ArXiv. /abs/2303.06615

IMPLEMENTATION

Disparity Map Generation

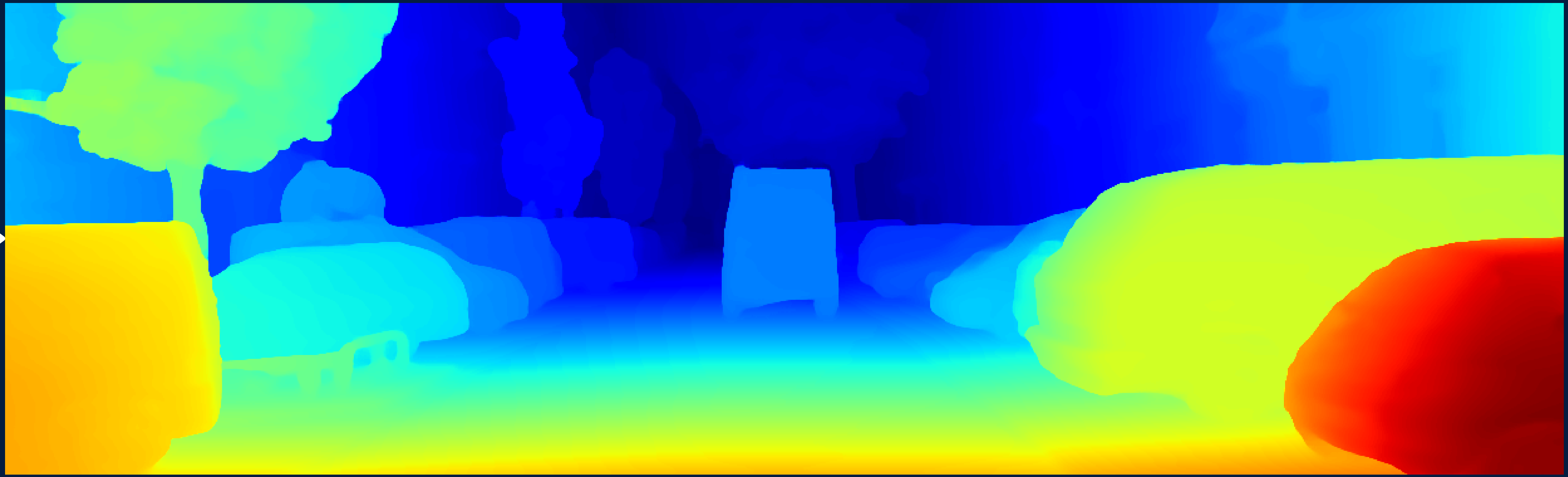


$$\frac{Z}{f} = \frac{X}{x_L} \qquad \frac{Z}{f} = \frac{X-b}{x_R}$$

$$x_L = \frac{X}{Z} \cdot f \qquad x_R = \frac{X-b}{Z} \cdot f$$

$$Disparity = x_L - x_R$$

Generated Disparity Map using the above formulas



Ibrahim, O., ElGendy, H., & ElShafee, A. M. (2011). Speed Detection Camera System using Image ProcessingTechniques on Video Streams. International journal of computer and electrical engineering, 3(6), 771.

IMPLEMENTATION

Depth Map Generation

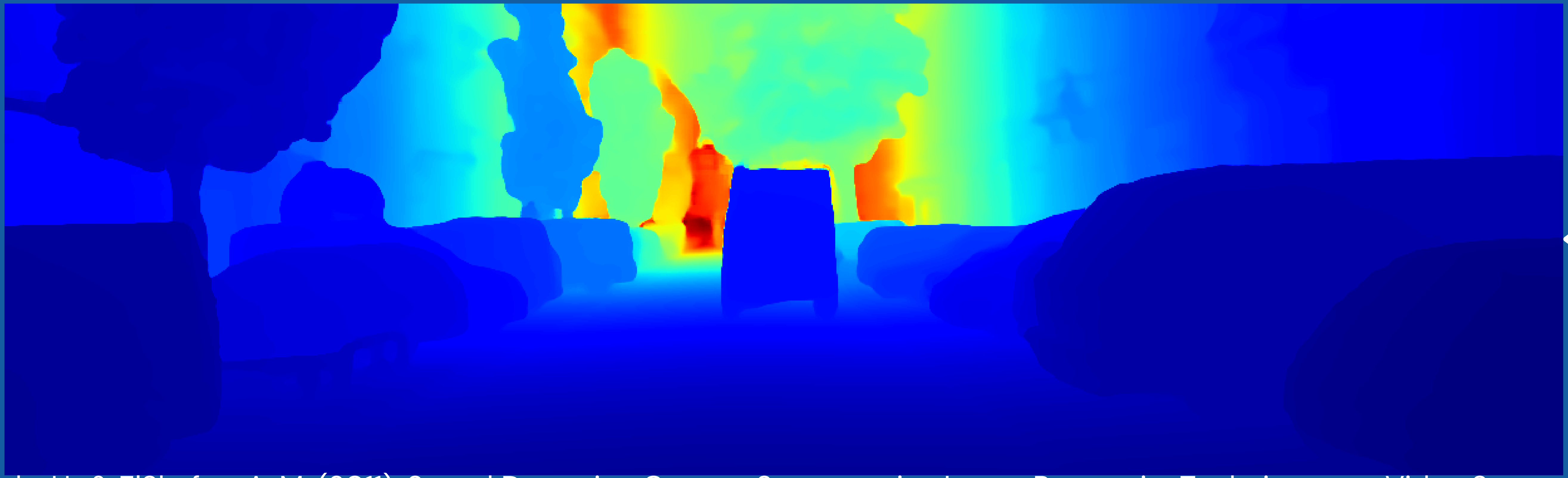
For a simple stereo system, the depth of a point (z) is given by:

$$Z = \frac{f * b}{d}$$

f = focal length

b = baseline, or interocular distance between the cameras

d = disparity between corresponding points.



Generated Depth Map using the above formula

Ibrahim, O., ElGendy, H., & ElShafee, A. M. (2011). Speed Detection Camera System using Image ProcessingTechniques on Video Streams. International journal of computer and electrical engineering, 3(6), 771.

IMPLEMENTATION

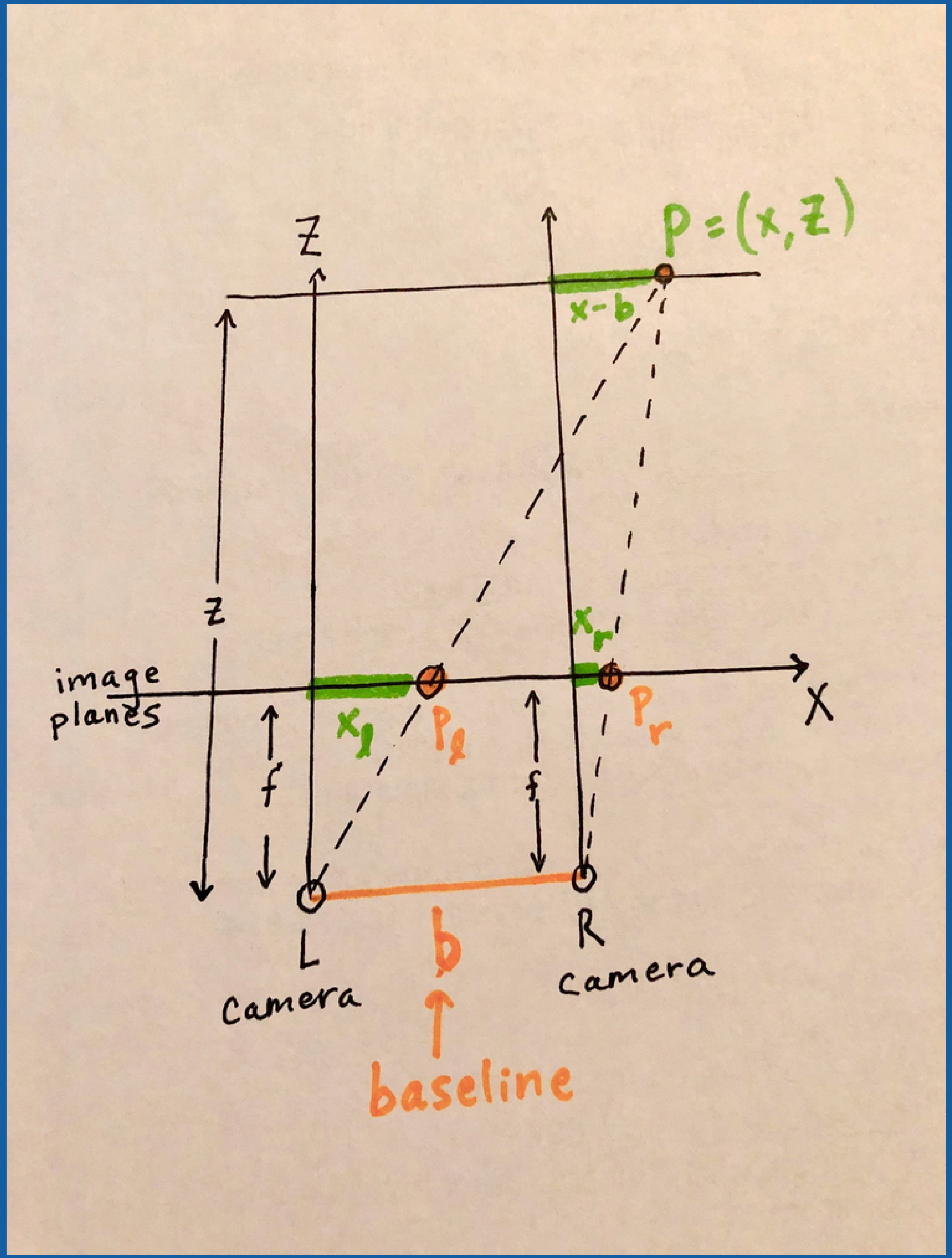
Finding World Coordinates

- By Pinhole Camera Model

$$X = u * Z / f$$

$$Y = v * Z / f$$

u,v are image plane coordinates



Sturm, P. (2014). Pinhole Camera Model. In: Ikeuchi, K. (eds) Computer Vision. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-31439-6_472

IMPLEMENTATION

Finding Realtime speed

For Dynamic camera setup

- As vehicles will be in same Y- plane, hence we'll have to see only in XZ plane

- Distance = $\left((x_{curr} - x_{prev})^2 - (y_{curr} - y_{prev})^2\right)^{\frac{1}{2}}$
- Speed = $Distance * fps \left(\frac{m}{s}\right)$

Ibrahim, O., ElGendy, H., & ElShafee, A. M. (2011). Speed Detection Camera System using Image ProcessingTechniques on Video Streams. International journal of computer and electrical engineering, 3(6), 771.

RESULTS AND DISCUSSION

Result 1 - Highway

Input

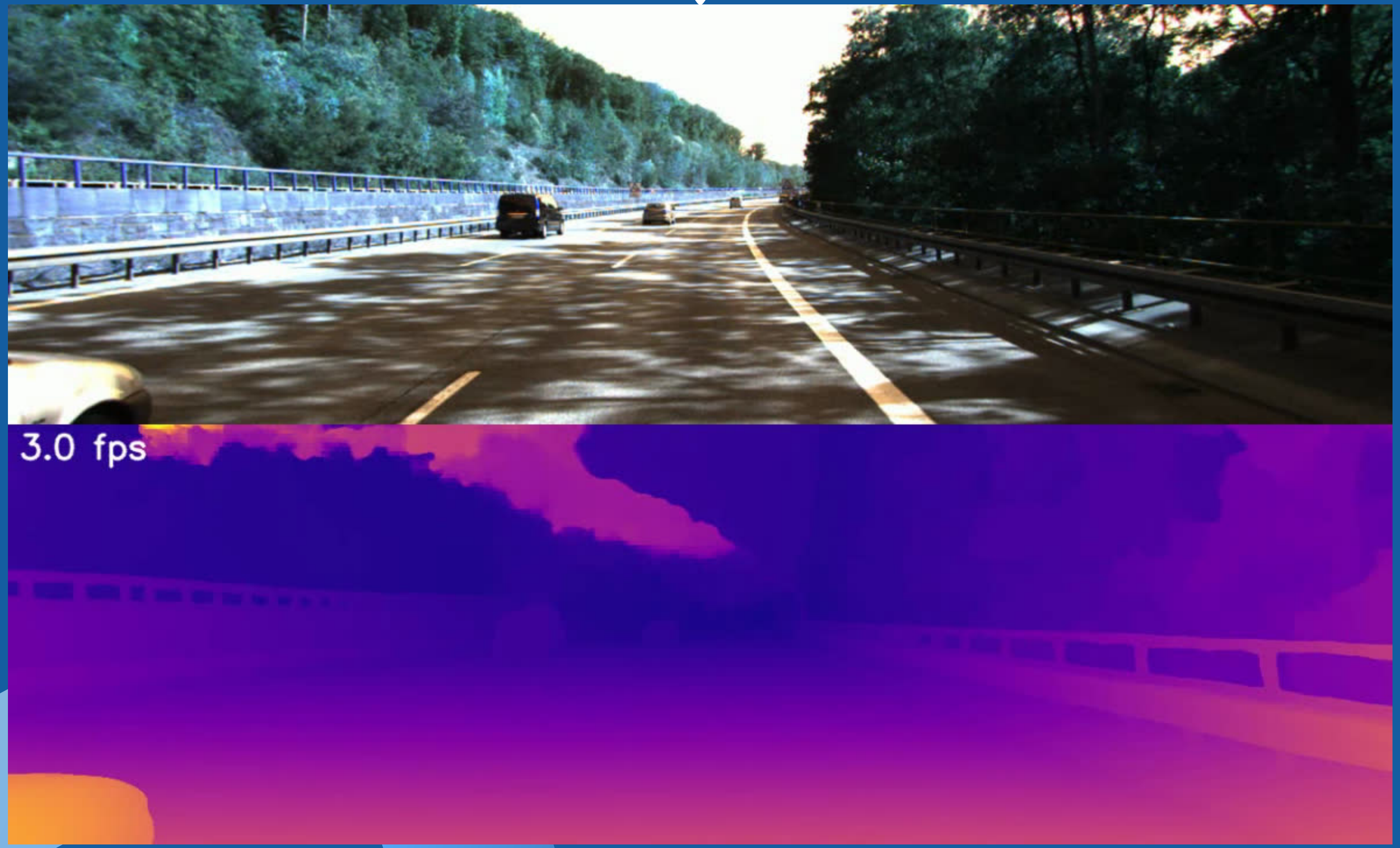
Left

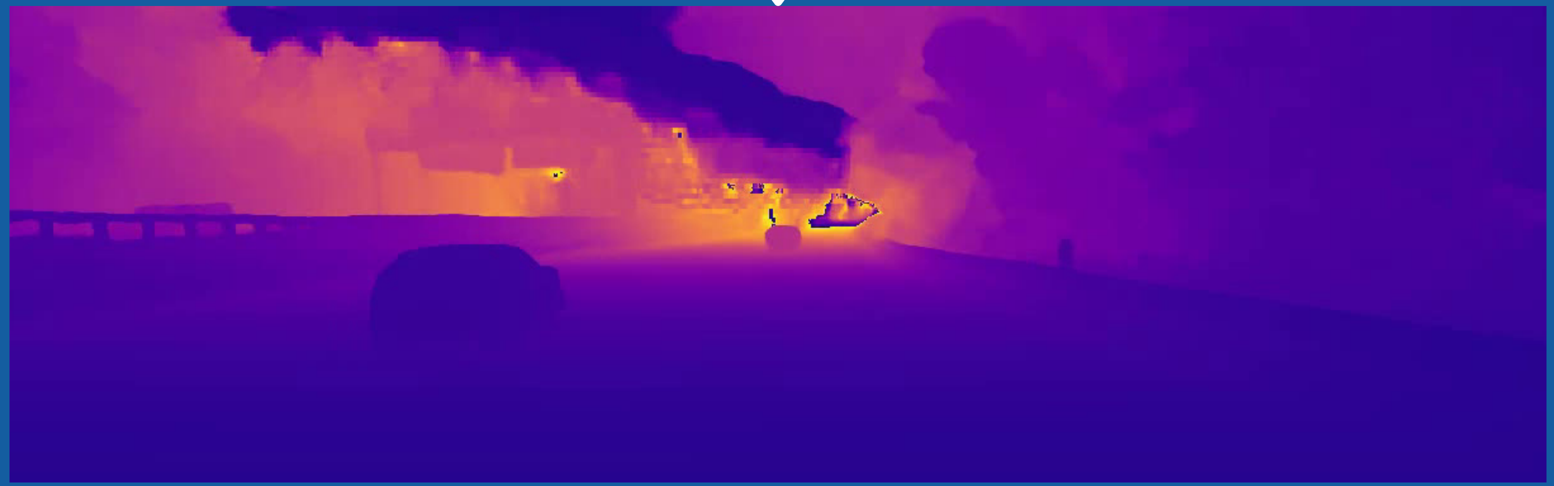
Right

Result 1 - Highway

Object Detection

Result 1 - Highway





Disparity

Depth

Result 1 - Highway

Result 2 - Static

Left

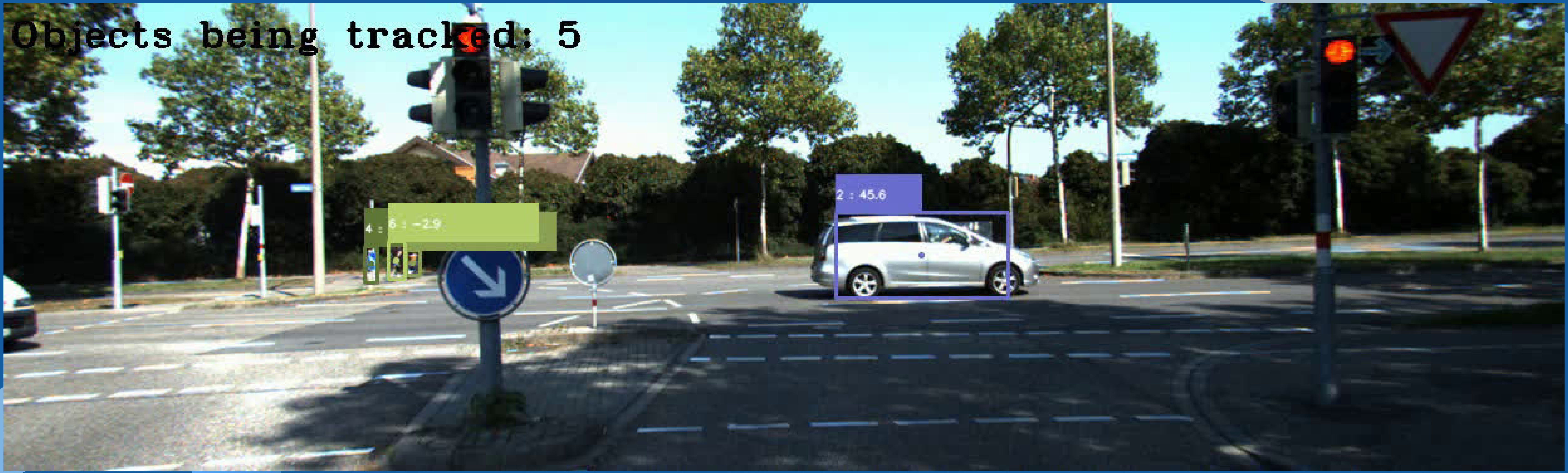
Right

Result 2 - Static

Disparity

Depth

Result 2 - Static



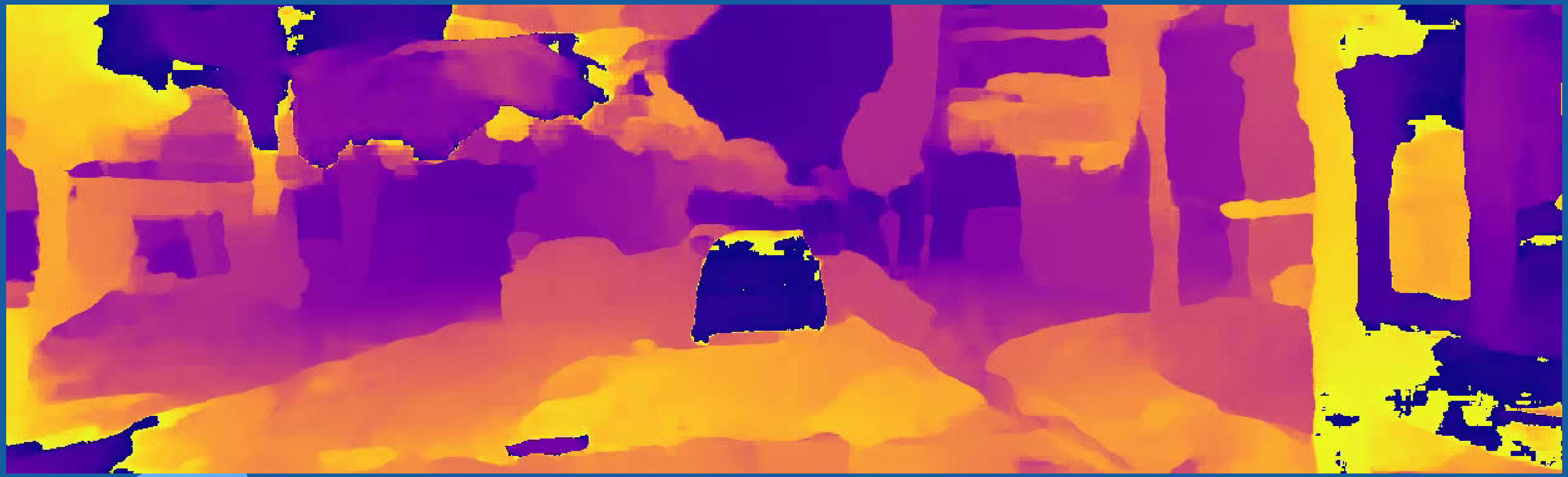
Speed

Result 3 - Gridlocked

Left

Right

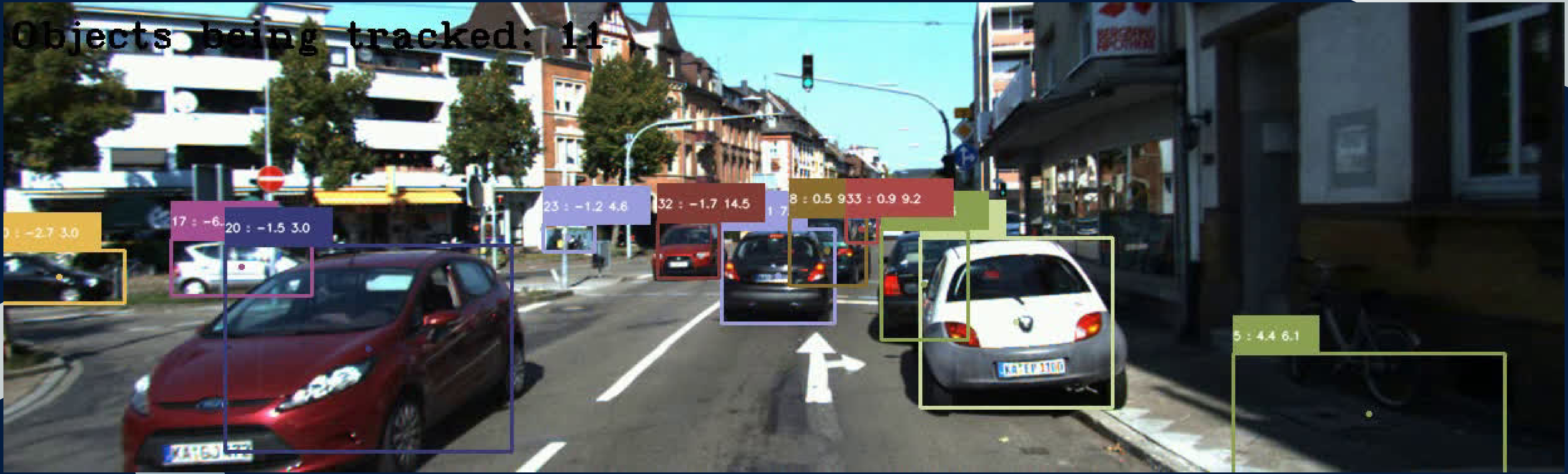
Result 3 - Gridlocked



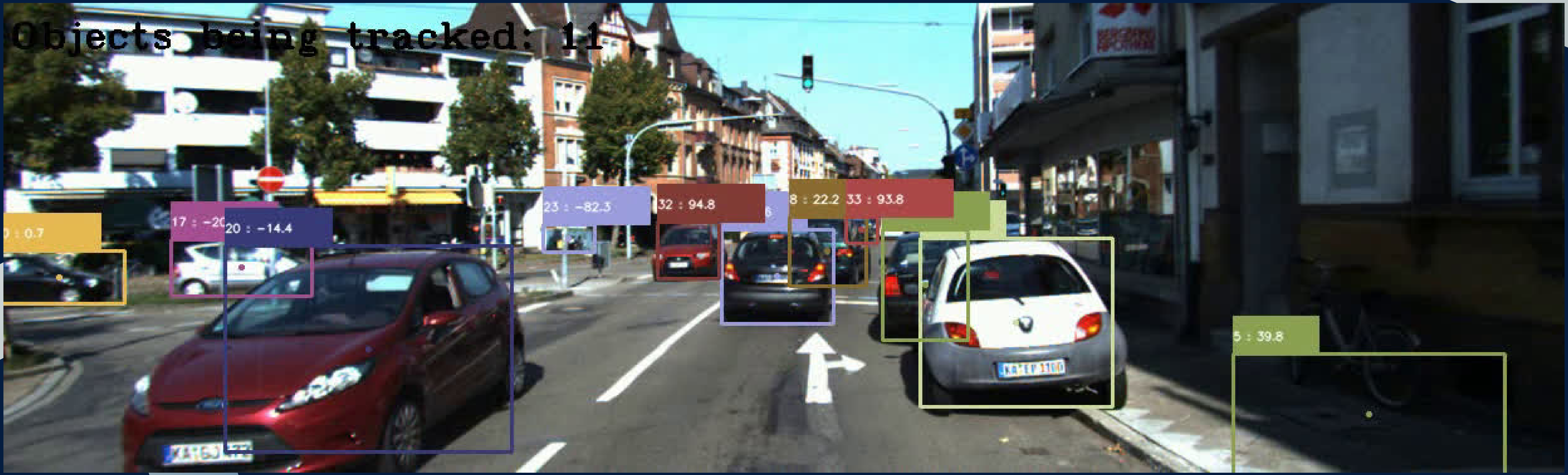
Disparity

Depth

Result 3 - Gridlocked



Result 3 - Gridlocked



APPLICATIONS OF THE PROJECT

Following are some of the applications of this project

Autonomous Vehicles

Obstacle Detection: Implementing our stereo vision method can contribute to obstacle detection, a crucial aspect of autonomous vehicle navigation

Collision Avoidance Systems

Safety Enhancement: The project can be extended to develop collision avoidance systems by providing real-time information about the distance to obstacles and the relative velocity of surrounding vehicles.

Traffic Monitoring and Management

Traffic Flow Analysis: The technology can be applied to monitor and analyze traffic flow, helping in optimizing traffic management strategies

Driver Assistance Systems

Advanced Driver Assistance Systems (ADAS): Integrate the technology into ADAS to provide drivers with real-time warnings about potential collisions or unsafe following distances

Parking Assistance

Parking Guidance Systems: Facilitate parking assistance systems by providing accurate distance measurements to obstacles assisting drivers during parking maneuvers.

Road Safety

Intersection Safety: Enhance safety at intersections by providing real-time information about the distance and velocity of approaching vehicles.

Real-time Applications

Real-time Processing: Optimize the algorithms and processing pipeline for real-time applications, ensuring timely and accurate responses

CONCLUSION

Objective

Introduced an algorithm for real-time speed calculation of moving vehicles captured by a camera.

Utilization of YOLO Algorithm

Employed the YOLO (You Only Look Once) algorithm for precise vehicle detection.

Focus on a Single Vehicle for Methodological Simplicity

The current methodology concentrates on a single vehicle to simplify the speed calculation process, especially when dealing with a moving camera.

Scalability to Multiple Vehicles

The underlying principles of the method are designed to accommodate and guide the handling of multiple vehicles.

Real-Time Implementation

The algorithm is designed for real-time speed calculation, indicating its applicability for scenarios where timely information is critical.

Relevance to Traffic Management and Safety

The proposed method can contribute to effective traffic management and enhance safety measures on the road.

Potential for Future Development

Layed the groundwork for potential future developments, especially in refining the handling of multiple vehicles and advancing of the real-time traffic monitoring.

FUTURE DIRECTIONS

Cross-Dataset Validation

Conduct thorough cross-dataset validation by testing the system on diverse datasets from different geographic locations, traffic conditions and camera setups. This will ensure the model's robustness and generalizability across varied scenarios.

Adversarial Testing

Implement adversarial testing to assess the system's resilience against deliberate attempts to deceive or manipulate its object detection and tracking components. This involves introducing intentional disturbances
or occlusions to evaluate the model's response

Temporal Variability Testing

Explore the impact of temporal variations in traffic patterns by testing the system across different times of day, days of the week, and seasons. This will reveal the system's adaptability to fluctuating traffic conditions over time.

Quantitative and Qualitative Metrics

Employ a comprehensive set of quantitative metrics (accuracy, precision, recall) and qualitative metrics (visual assessments, interpretability) to evaluate the system's performance.

Human-Audit Verification

Introduce a human-audit verification process to assess the accuracy of the system's outputs. This involves comparing the system's results against human-verified ground truth data to ensure high precision and reliability.

REFERENCES

[1]. Volkan Cevher, Rama Chellappa and James H. McClellan, “Vehicle Speed Estimation using Acoustic Wave Patterns”, IEEE Transactions on Signal Processing Volume: 57, Issue: 1, Jan. 2009

[2]. Osman Ibrahim, Hazem ElGendy, and Ahmed M. ElShafee. “Speed Detection Camera System (SDCS) using Image Processing Techniques on Video Streams”, January 2011 International Journal of Computer and Electrical Engineering

[3]. Jozef Gerat, Dominik Sopiak, Milos Oravec, Jarmila Pavlovicova. “Vehicle Speed Detection from Camera Stream Using Image Processing Methods”, International Symposium ELMAR, 18-20 Sept. 2017

[4].Huang, T. (2018). Traffic speed estimation from surveillance video data. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 161-165).

[5]. Khazukov, K., Shepelev, V., Karpeta, T., Shabiev, S., Slobodin, I., Charbadze, I., & Alferova, I. (2020). Real-time monitoring of traffic parameters. Journal of Big data, 7(1), 1-20.

https://github.com/Arjun-s227/Real-Time-Speed-Detection

THANK YOU !