Abstract

*We are dealing with the **problem of position estimation of traffic sign** from **single camera** attached to a **vehicle** for which **two mathematical approaches** have been developed. We have created a **synthetic data set**, which enabled us to test the methods in a controlled environment. **Real data set** has been created as a recording in a town. The position can be estimated within the **error of 1 m** in the laboratory conditions. Frequent and accurate GPS sampling is a necessary condition, which has not been fully satisfied in our real data set resulting in an error range starting from **below 1 m up to 4 m on average**. We have also suggested a **potential application** of our approaches for **improving object tracking** in a sense of a **control mechanism**.*

Author: **Ing. Milan Ondrašovič** (milan.ondrasovic@gmail.com)
Supervisor: **Ing. Peter Tarábek, PhD.** (peter.tarabek@fri.uniza.sk)

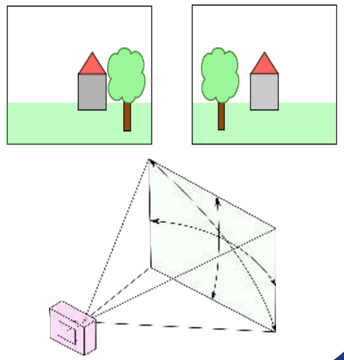
University of Žilina in Žilina,
Univerzitná 8215/1, 010 26 Žilina.

Faculty of management science and informatics, department of mathematical methods and operational analysis.

Position estimation of static object from moving camera

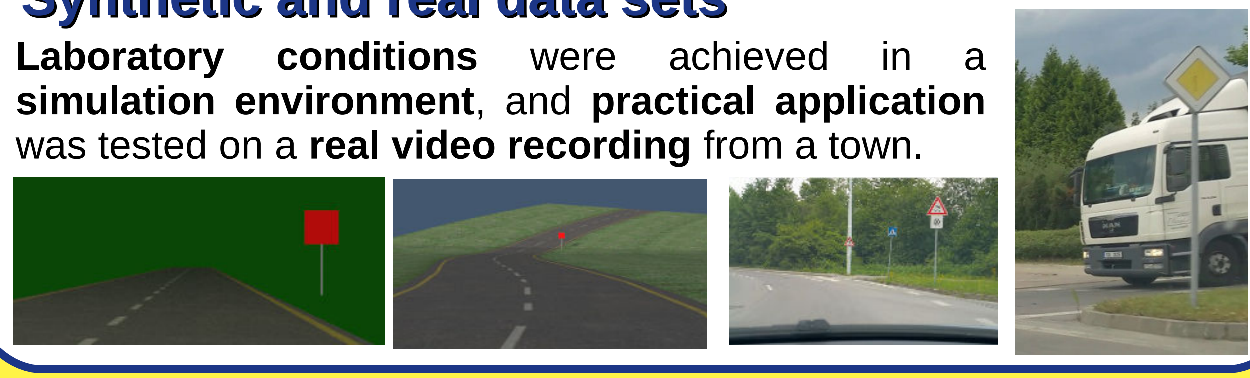
General solution description

The two approaches use **bounding box of the object** of interest and **current position of the camera** in the world to **estimate the object position**. **Two different frames** are always required. Camera calibration is unnecessary, the only needed parameter is the **horizontal angular field of view**.



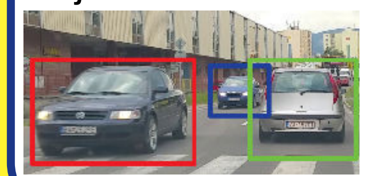
Synthetic and real data sets

Laboratory conditions were achieved in a **simulation environment**, and **practical application** was tested on a **real video recording** from a town.

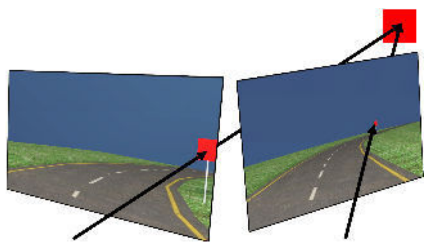


Object tracking

We have **proposed** a way to **utilize our methods** for **improving object tracking** as a **control mechanism** for potentially incorrectly tracked object.



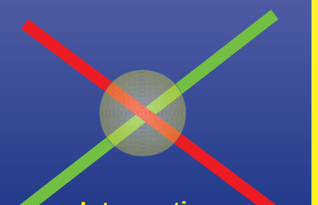
Estimation using intersection of lines



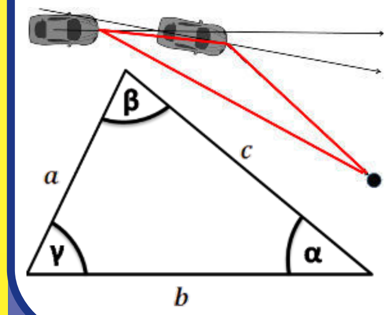
Position is estimated using calculation of **intersection of two lines** in space, in a way similar to **shooting a ray through the optical center of the camera** from two different positions in space.

$$x_t = \frac{y_2 - y_1 + tan\left(p_1\right) x_1 - tan\left(p_2\right) x_2}{tan\left(p_1\right) - tan\left(p_2\right)}$$

$$y_t = tan\left(p_1\right) x_t + y_1 - tan\left(p_1\right) x_1$$



Estimation using triangle properties



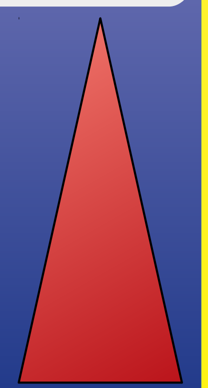
Position is estimated using **properties of the triangle** (fundamentally **law of sines**). Triangle is created from **two different views** of the object with **known distance between positions**.

Object positioning

$$\phi\left(a, b\right) = tan^{-1}\left(\frac{a - b}{1 + ab}\right)$$

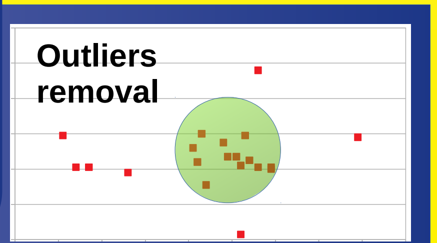
$$x_t = x_1 + sin\left(\pi - \phi\left(m_0, m_2\right)\right)\left(\frac{\sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}}{sin\left(\phi\left(m_0, m_2\right) - \phi\left(m_0, m_1\right)\right)}\right) cos\left(m_1\right)$$

$$y_t = y_1 + sin\left(\pi - \phi\left(m_0, m_2\right)\right)\left(\frac{\sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}}{sin\left(\phi\left(m_0, m_2\right) - \phi\left(m_0, m_1\right)\right)}\right) sin\left(m_1\right)$$



Improving of position estimation

Since the two methods operate **only on two frames** there is a potential to use **all the available frame pairs** where the object is visible to **create a final position from elementary ones**. Two approaches were implemented, first **based upon centroid**, the second (more sophisticated) involving **statistical properties of points** in space and principles from **clustering**. The idea is to **remove outliers**.



Experiments and results

We experimented with **various types of scenarios**, different **road curvature** or **elevation**, object **position** or **distance** from camera, etc. Errors ranged from **4 cm** up to **8 m**. Object distance from camera should be below **30 m**.

Conclusion

Solution can be used **when stereo vision cameras are unavailable**. Laboratory error of **below 1 m** shows the potential, while the real world error of **4 m** provides a useful, yet slightly inaccurate **information about object position**.