

Single-view (2D)

Candidate extraction

•From thousands of *candidate extraction methods*, we use Boolean Linear Programming formulation for obtaining an optimal tradeoff among FN,FP and the complexity of solution

$$T^* = \arg \min_T FP(T) + \kappa_1 \cdot FN(T) + \kappa_2 \cdot \text{card}(T)$$

Color-based methods

$T = (t, a, b, c)^T$

$$I(T) = \begin{cases} 1 & a \cdot I_R + b \cdot I_G + c \cdot I_B \leq t \\ 0 & \text{otherwise} \end{cases}$$

$\bar{T} = \underbrace{(t, a, b, c, s_r, s_c)}_T$

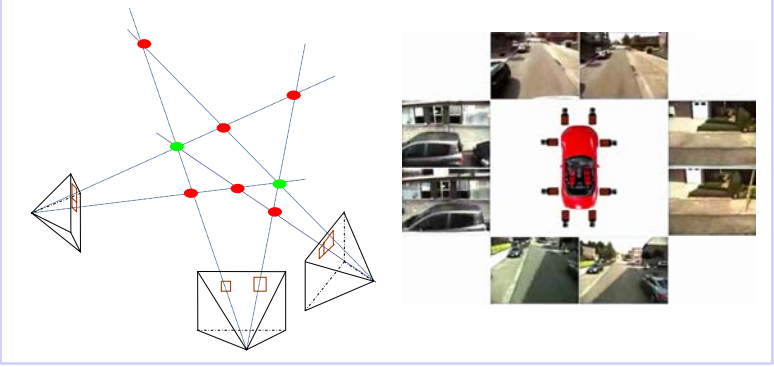
Shape-based methods

Detection

•Cascades of Discrete AdaBoost classifiers trained on Haar-like features computed for each channel of HSI space with I normalized by variance



Multi-view (3D)



3D hypothesis generation

•3D hypotheses (depicted in figure by red and green) generated from visually and geometrically consistent pairs of detected bounding boxes

MDL hypothesis selection

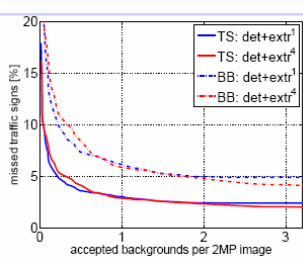
•Global optimization by Minimum Description Length
•MDL selects the **best** subset of 3D hypotheses explaining the overall set of 2D candidates

Sign type recognition

•The supporting 2D candidates for each selected 3D hypothesis contributes to the sign type assignment via voting scheme
•Shape-specific SVM classifiers running on RGB channels
•2D candidates classified into 62 different classes

Results

| | FN-TS | | FN-BB | | FP per 2MP img |
|---------------------------------------|-------|--------|-------|--------|----------------|
| | [%] | #/1274 | [%] | #/3756 | |
| Extr (color) ¹ | 0.5% | 7 | 1.5% | 58 | 3 442.4 |
| Extr (color+TMSER) ² | 0.4% | 5 | 1.4% | 53 | 4 008.5 |
| Extr (color+shape) ³ | 0.2% | 2 | 1.0% | 36 | 6 670.3 |
| Extr (color+TMSER+shape) ⁴ | 0.1% | 1 | 1.0% | 36 | 7 157.3 |
| Det + Extr ¹ | 2.4% | 31 | 4.9% | 184 | 2.5 |
| Det + Extr ⁴ | 2.2% | 28 | 4.3% | 163 | 2.5 |



| # | No.frames/Ts | 3D Localised TS | FP | Recognised TS |
|---|-----------------|-----------------|----|---------------|
| 1 | 8 × 3001 / 78 | 75(96.2%) | 3 | 74(98.7%) |
| 2 | 8 × 6201 / 71 | 68(95.8%) | 7 | 65(95.6%) |
| 3 | 8 × 2001 / 44 | 41(93.2%) | 2 | 41(100%) |
| 4 | 8 × 4001 / 76 | 73(96.1%) | 8 | 71(97.3%) |
| Σ | 8 × 15204 / 269 | 257(95.6%) | 20 | 251(97.7%) |

•Average accuracy of 3D localisation is 24.54 cm

