A Bimodal Sound Source Model for Vehicle Tracking in Traffic Monitoring

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Patrick Marmaroli[1], Jean-Marc Odobez[2], Xavier Falourd[1], Hervé Lissek[1]
1 : Laboratoire d'Electromagnétisme et d'Acoustique (LEMA), EPFL, Switzerland
2 : Idiap Research Institute, Martigny, Switzerland

19th European Signal Processing Conference – August 29 – September 2, 2011, Barcelona (Spain)

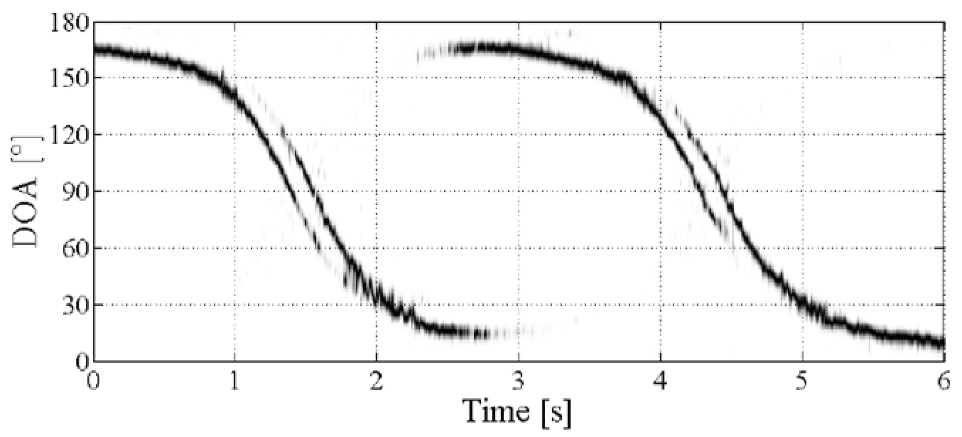
1 - Motivation

BACKGROUND: Road traffic monitoring is usually done using **active and/or intrusive** sensor-based technologies (e.g. radar, wire loop, pressure sensors, coaxial cable).

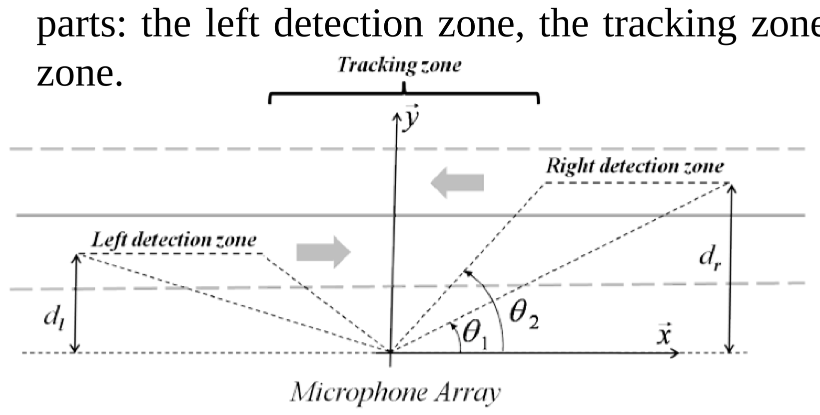
OUR APPROACH: use **one compact array of omnidirectional microphones** disposed near the road.

ADVANTAGES: passive and non-intrusive system.

PRINCIPLE: Exploiting the **Time Difference Of Arrival (TDOA)** of both axles sound sources to estimate the vehicle location, speed and wheelbase length. The TDOA and then the Direction of Arrival (DOA) is given by the acoustic path differences between sensors through the well-known **PHAT weighted Cross-Correlation (GCC-PHAT)** technique [1]. The figure below shows an example of two vehicles pass-by. Notice the change of the dominant axle.

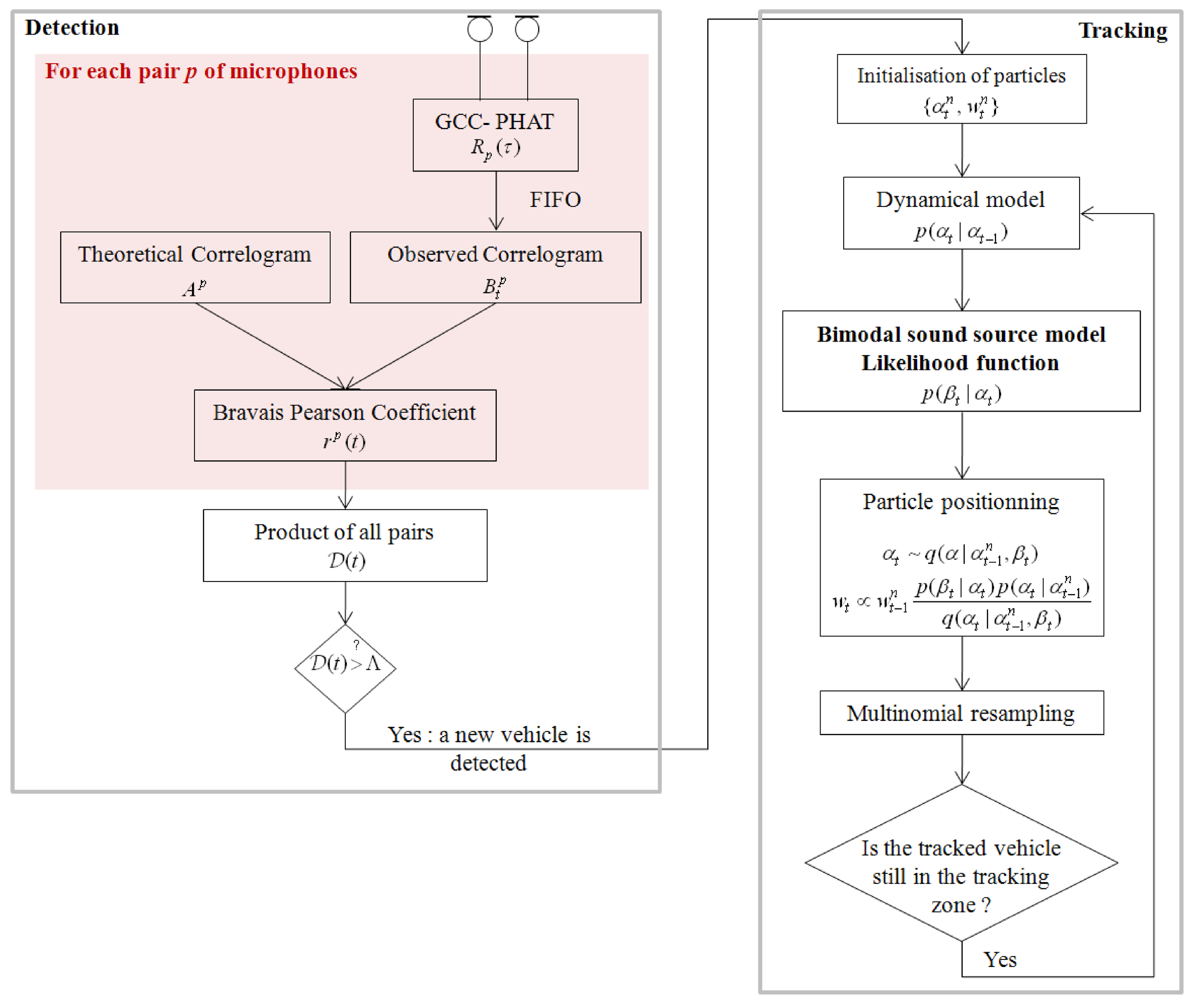


Setup:
A microphone array is disposed near the road. The road is considered as straight and with two lanes of opposite circulation. It is divided in three parts: the left detection zone, the tracking zone and the right detection zone.



2 - Global algorithm synopsis

Detection precedes tracking . The latter estimate speed and distance between axles of each detected vehicle during their pass-by.

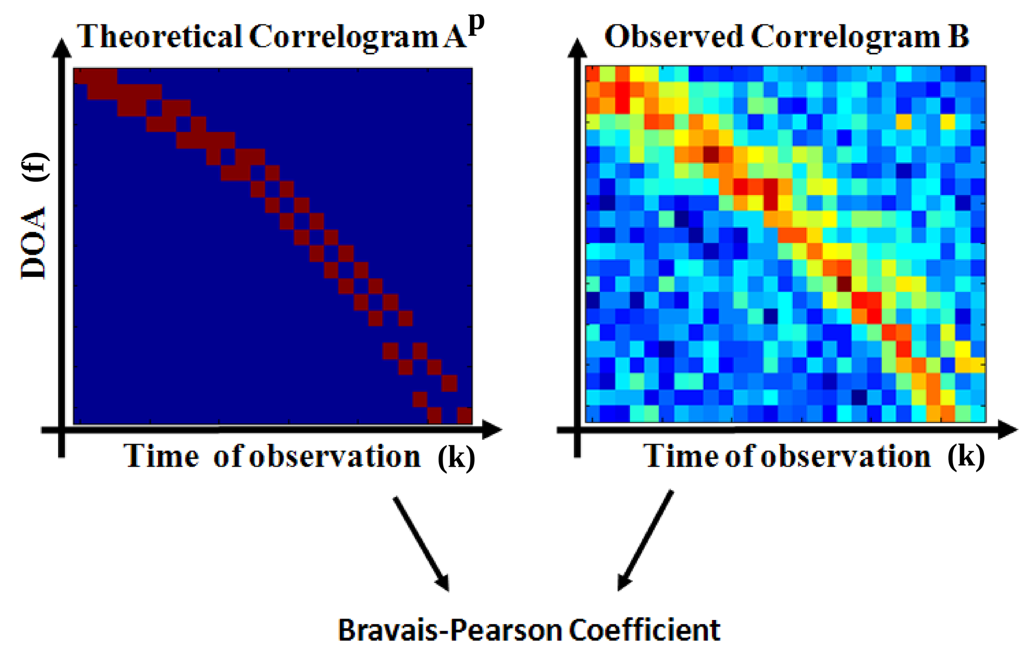


3 - Detection

OBJECTIVE: determine if and when a vehicle is entering the tracking zone.

MAIN ISSUE: in the outdoor traffic monitoring context, a simple sound pressure level based detector (dB) is generally insufficient to detect a vehicle presence because of multiple possible noise sources.

STRATEGY:
- consider the dynamic information of each sound source which can come from each detection zone.
- compare it with one (or several) theoretical vehicle movement(s).
- similarity above a threshold → vehicle is detected and tracking algorithm is launched.



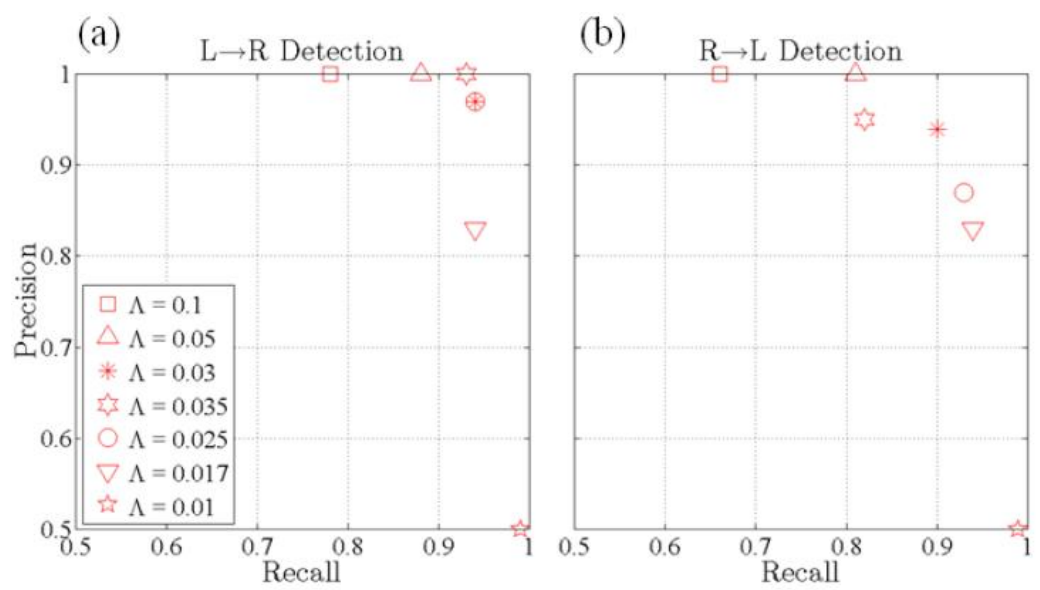
Details:
- one coefficient $r^p(t)$ is computed for each pair p of microphone.

$$r^p(t)=\frac{\sum_{k=1}^{K}\sum_{f=1}^{N_f}\left(A^p(f,k)-\bar{A}^p\right)\left(B_t^p(f,k)-\bar{B}_t^p\right)}{\sqrt{\sum_{k=1}^{K}\sum_{f=1}^{N_f}\left(A^p(f,k)-\bar{A}^p\right)^2\sum_{k=1}^{K}\sum_{f=1}^{N_f}\left(B_t^p(f,k)-\bar{B}_t^p\right)^2}}$$

- The final decision is done according to the product of all coefficients.

$$\mathcal{D}(t)=\prod_{p=1}^{P} r^p(t) \longrightarrow \mathcal{D}(t) \lessgtr_{H_1}^{H_0} \Lambda$$

- Λ is found thanks to a ROC curve analysis [2] (Fig. below): Positive Predictive Value (PPR or Precision) as a function of True Positive Rate (TPR or recall). An optimal threshold gives the point [1,1].



Precision-Recall performances in function of threshold Λ based on real measurements for vehicles comes from left (a) and right (b)

4 - Tracking

OBJECTIVE: estimate speed and wheelbase length of each detected vehicle.

MAIN ISSUES:
- the highest peaks of the GCC-PHAT may not always come from sounds emitted by the vehicle because of high risks of noise.
- the change of dominant axle + vehicle crossing can disrupt the tracking.

STRATEGY:
- use the filtering theory [3] by making distinction between peaks that follow a presumed dynamical movement (vehicle) and the others (noise and vehicles in opposite direction).
- the robustness of the tracking is improved by considering two different sound sources (axles) for each vehicle. This technique allows to estimate the wheelbase length in addition to speed.

Details:

State space:

$$\alpha_t = [x_t, y_t, \dot{x}_t, wb_t]^T$$

abscissa, ordinate, speed, wheelbase

Prediction: the dynamical model is used to propagate the filtering distribution at time *t-1* to provide the predictive distribution.

Update: the predictive distribution is combined with the likelihood to obtain the new filtering distribution at time *t*.

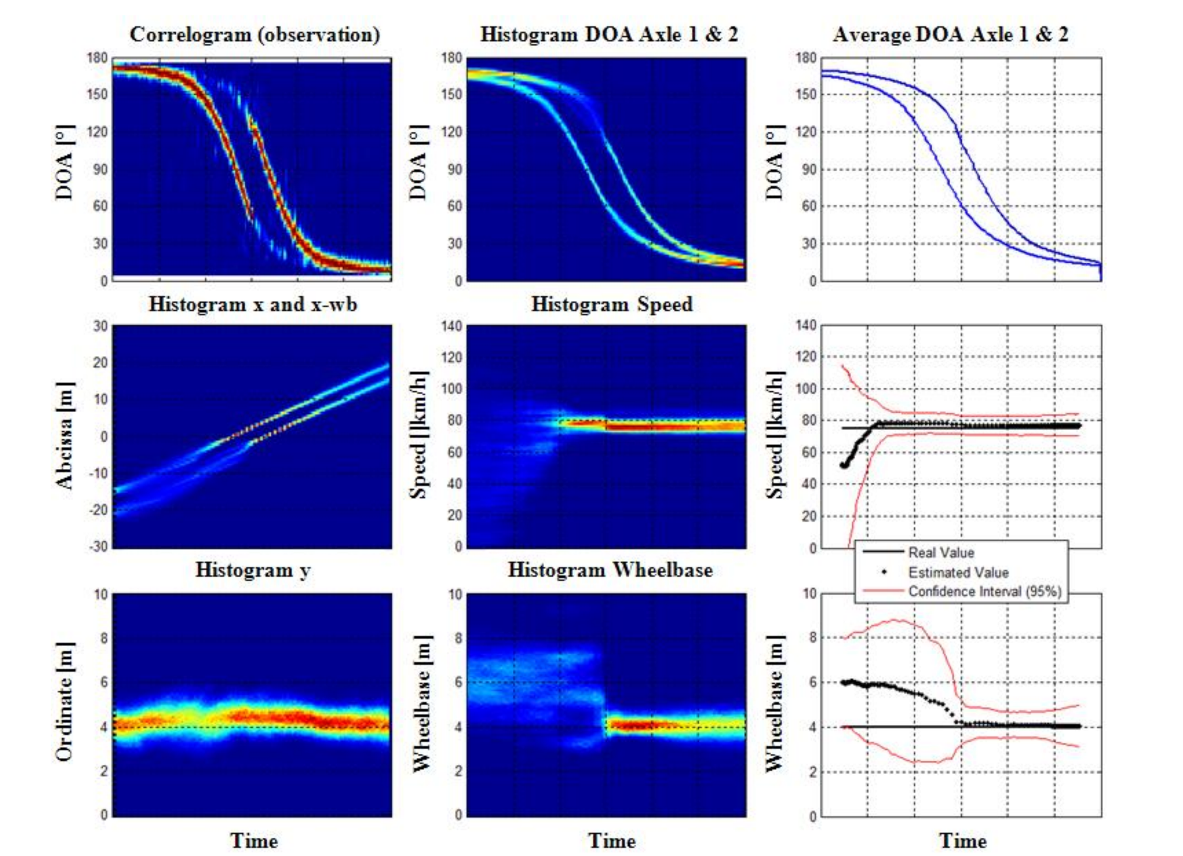
Likelihood: **follows a bimodal sound source model:**

$$p(\beta_t|\alpha_t^{(i)})=\gamma(\mu_\tau)R_t(\underbrace{\tau(x_t^i,y_t^i)}_{1^{st}\text{ Axle TDOA}})+(1-\gamma(\mu_\tau))R_t(\underbrace{\tau(x_t^i-wb_t^i,y_t^i)}_{2^{nd}\text{ Axle TDOA}})$$

GCC-PHAT

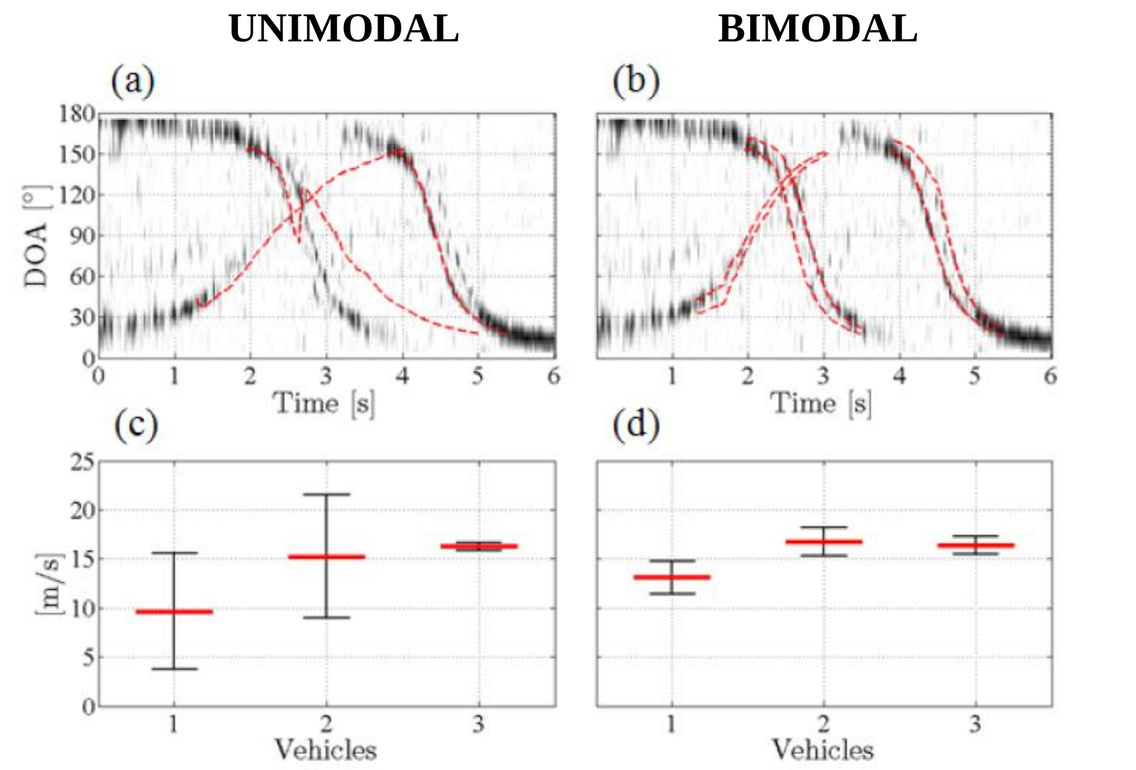
γ : weight function evolving with the average of both axles TDOA's μ_τ by giving more importance to the 1st axle when vehicle approaches and more importance to the 2nd axle when the vehicle goes away.

Example of speed and wheelbase estimation and convergence:



5 - Results

Using a unimodal observation model (no wheelbase taken account), makes the particles follow the most dominant of both axles, and requires to overcome a large gap when the dominant axle is changing, which typically happens when the vehicle is in the broadside situation. Risks of failures during this gap are accentuated when another vehicle is tracked at the same time in the opposite direction. This risk is drastically reduced using the bimodal observation model where no gap is noticed anymore and a wheelbase estimation is provided. More results of robustness are available in the paper.



DOA and speed estimation of three vehicles in a real harsch situation. Raw observation and results are superimposed using (a): a unimodal sound source PF based model, (b): a bimodal sound source PF based model. The same measurement is processed 200 times, mean and IC95 in speed estimates are represented fot both cases on (c) and (d).

6 - Conclusion

Acoustic method for both vehicle detection and characterization.
Detection step: good experimental results. Remaining cases of failures come from situations where several vehicles pass each other inside the end-fire detection zone.
Tracking step: bimodal sound source model allows a much lower variability of the results in the estimation of vehicles that pass each other in front of the microphone array in comparison with a unimodal sound source model. Moreover, the proposed method permits to estimate the wheelbase in addition to speed with a totally passive and non intrusive device. The sensitivity study of the proposed method (not shown here but presented in the paper) ensures effective speed and wheelbase estimations under realistic parameters.
Forthcoming works: extend the algorithm to vehicles with unkwown number of axles.

Aknowledgments:
The authors wish to thank AER-Sàrl (Lausanne, Switzerland) for their help on measurements and the provided data.

Bibliography:

[1] C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *Acoustics, Speech and Signal Processing, IEEE Transactions on 24(4):320-327, Aug 1976*
[2] T. Fawcett. An introduction to roc analysis. *Pattern Recogn. Lett., 27:861-874, June 2006*
[3] F. Gustaffson. Particle filter theory and practice with positionning applications. *Aerospace and Electronic Systems Magazine, IEEE, 25(7):53-82, July 2010*