BROADBAND ANALYSIS OF A MICROPHONE ARRAY BASED ROAD TRAFFIC SPEED ESTIMATOR

R. López-Valcarce

Departamento de Teoría de la Señal y las Comunicaciones
Universidad de Vigo, 36200 Vigo, Spain

ABSTRACT

Recently, a maximum likelihood estimate has been proposed for road vehicle speed based on two omnidirectional microphones. We undertake a broadband analysis in a stochastic setting in order to expose the effect of input SNR, target speed, bandwidth, and observation time in the SNR at the output of the modified crosscorrelator on which the estimate is based. The use of 1-bit quantized signals, which provides an important hardware simplification, is also considered and seen to result in mild performance degradation.

1. INTRODUCTION

Effective traffic management systems require accurate estimation of parameters such as traffic density and flow, for which a sensor infrastructure capable of automatic monitoring of traffic conditions must be deployed. The design of a transit data collecting system must include the choice of sensor type as well as the development of adequate signal processing and parameter estimation methods.

Several sensor technologies are commercially available at present, differing in terms of robustness, cost, safety regulations, etc. A desirable system would be passive, cheap, easy to install and maintain and operational in all-weather day-night conditions. These goals can be achieved with microphone based schemes; however, available systems tend to be expensive since they use arrays based on highly directive microphones. We address the problem of directly estimating vehicle speed from the acoustic signals received at a pair of omnidirectional microphones located next to the traveling path.

Our previous work [6, 8] presented an approximate maximum likelihood (ML) vehicle speed estimate for such a setting, as well as an analysis under a narrowband deterministic framework. One advantage of this estimate is that it requires neither modeling or knowledge of the acoustic source (thus being effectively "blind"), nor intermediate time delay estimation steps, which are potentially troublesome in real applications [3, 7]. In this paper we embrace a broadband stochastic setting for the analysis, and also study the effects of using 1-bit quantized signals in the estimation process. Such drastic quantization is attractive if estimation is done *in situ* at the sensor location, in order to simplify the required sampling and processing hardware. Cost reduction becomes especially significant if a large sensor network is to be deployed.

Work supported by a *Ramón y Cajal* grant from the Spanish Ministry of Science and Technology. Author contact: valcarce@gts.tsc.uvigo.es

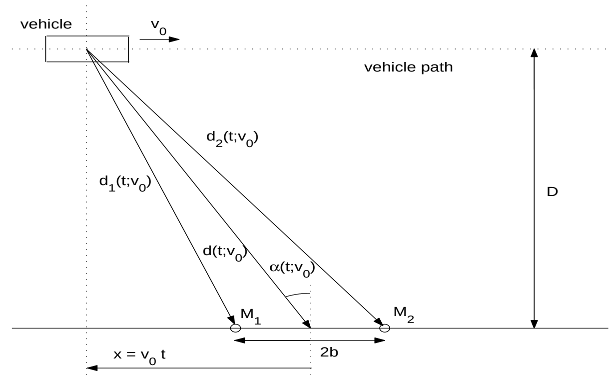

Figure 1: Geometry of the problem

2. APPROXIMATE ML ESTIMATE

Figure 1 illustrates the setting. The microphones M_1, M_2 are separated $2b$ m and placed at D m from the lane center. The acoustic source is assumed to travel at constant speed v_0 on a straight path along the road, parallel to the array. The time reference is set at the closest point of approach (CPA), i.e. $t = 0$ when the source is equidistant to M_1 and M_2. The propagation time $\tau_i(t; v_0)$ from the source to microphone M_i is given by

$$\tau_{1,2}(t; v_0) = \frac{1}{c}\sqrt{D^2 + (v_0 t \pm b)^2},$$

with $c = 340$ m/s, the speed of sound (assumed constant for simplicity). Define also the angle and distance between the source and the array center

$$\alpha(t; v_0) = \arctan\frac{v_0 t}{D}, \qquad d(t; v_0) = \frac{D}{\cos\alpha(t; v_0)}.$$

Let the sound wave emitted by the vehicle be $s(t)$. Taking into account the attenuation of sound with distance, the signal received at M_i within the observation window $[-T/2, T/2]$ is

$$\begin{aligned} r_i(t) &= s_i(t) + w_i(t) \\ &= \frac{s(t-\tau_i(t;v_0))}{(d(t;v_0)/D)} + w_i(t), \quad |t| \leq T/2, \end{aligned} \quad (1)$$

with $w_1(\cdot)$, $w_2(\cdot)$ additive noise processes, assumed stationary, independent, zero-mean jointly Gaussian with psd $N_0/2$ W/Hz in the band $|f| < f_s/2$ (f_s = sampling frequency). The problem is to estimate v_0 given the observed signals $r_i(t)$, and without knowledge of the sound wave $s(t)$ or its power spectrum.

In [6, 8] an approximate ML estimate was derived. It is given by $\hat{v}_0 = \arg\max_v \psi(v)$, where

$$\psi(v) \triangleq \frac{1}{T} \int_{-T/2}^{T/2} r_1(t - \delta\tau(t;v)) r_2(t) dt, \quad (2)$$

with the differential time delay (DTD) $\delta\tau(t;v)$ defined as

$$\delta\tau(t;v) \triangleq -\frac{2b}{c} \frac{\sin\alpha(t;v)}{1 - \frac{v}{c}\sin\alpha(t;v)}. \quad (3)$$

This estimate is based on the fact that, taking into account source motion, $s_2(t) \approx s_1(t - \delta\tau(t;v_0))$ is satisfied. It exploits knowledge of the DTD parametric dependence with v to accordingly time-compand the signal $r_1(t)$ before performing the crosscorrelation (2), which must be computed over the whole observation window for each candidate speed.

3. UNQUANTIZED SIGNALS

We proceed to derive expressions of the mean and variance of the crosscorrelator output for a broadband acoustic signature, assuming no quantization.

3.1. Expected value of $\psi(v)$

Modeling the acoustic wave $s(\cdot)$ as a realization of a WSS stochastic process with autocorrelation $R_s(\tau)$ and psd $G_s(f)$, the expected value of $\psi(v)$ is given by

$$\begin{aligned} E[\psi(v)] &= \frac{1}{T} \int_{-T/2}^{T/2} E[r_1(t - \delta\tau(t;v)) r_2(t)] dt \\ &\approx \frac{1}{T} \int_{-T/2}^{T/2} \frac{R_s[\Delta^2\tau(t;v_0,v)]}{d^2(t;v_0)/D^2} dt \end{aligned} \quad (4)$$

where

$$\Delta^2\tau(t;v_0,v) \triangleq -\frac{2b}{c}[\sin\alpha(t;v_0) - \sin\alpha(t;v)] \quad (5)$$

and a first-order approximation of $\tau_1(t - \delta\tau(t;v);v_0)$ has been used, as in [6]. It is also shown in [6] that (5) can be accurately approximated as

$$\Delta^2\tau(t;v_0,v) \approx q\sin[2\arctan(zt)], \quad (6)$$

where

$$q \triangleq \frac{b(v-v_0)}{c\sqrt{2v_0 v}}, \qquad z \triangleq \frac{\sqrt{2v_0 v}}{D}. \quad (7)$$

Writing $R_s[\Delta^2\tau]$ in terms of $G_s(f)$ in (4), changing the order of integration and using the expansions

$$f(r\sin x) = \sum_k J_k(r) f(kx),$$

where $f(\cdot)$ is either $\sin(\cdot)$ or $\cos(\cdot)$ and J_k is the k-th order Bessel function, then after retaining only the dominant term in the summation one obtains

$$E[\psi(v)] \approx \frac{2\alpha_0}{\tan\alpha_0} \int_0^\infty G_s(f) J_0(2\pi f q) df, \quad (8)$$

where

$$\alpha_0 \triangleq \alpha(T/2; v_0)$$

is the *angular aperture*. Thus $E[\psi(v)]$ can be seen as the scaled *Hankel transform* [5] of $G_s(f)/f$ evaluated at q.

Assume that $s(\cdot)$ is a bandpass process centered in f_c and with bandwidth B, i.e.

$$G_s(f) = \begin{cases} S_0/2, & |f \pm f_c| < B/2, \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

The corresponding integral (8) has no closed-form solution. However, since $R_s(\tau) = S_0 B \operatorname{sinc}(B\tau)\cos(2\pi f_c \tau)$, then from (6),

$$\begin{aligned} & R_s[\Delta^2\tau(t;v_0,v)] \\ &\approx S_0 B \frac{\sin[\pi Bq \sin(2\arctan zt)]}{\pi Bq \sin(2\arctan zt)} \\ &\quad \times \cos[2\pi f_c q \sin(2\arctan zt)] \\ &= S_0 B \sum_{k=-\infty}^{\infty} \frac{J_k(\pi Bq)}{\pi Bq} \frac{\sin(2k\arctan zt)}{\sin(2\arctan zt)} \\ &\quad \times \sum_{n=-\infty}^{\infty} J_n(2\pi f_c q)\cos(2n\arctan zt). \end{aligned} \quad (10)$$

Retaining only the dominant terms in the summations ($n = 0$ and $k = \pm 1$), substituting back in (4) and integrating, one obtains

$$E[\psi(v)] \approx \frac{\alpha_0}{\tan\alpha_0} S_0 B \operatorname{somb}(Bq) J_0(2\pi f_c q), \quad (11)$$

where

$$\operatorname{somb}(x) \triangleq \frac{2J_1(\pi x)}{\pi x}$$

is the sombrero function [5]. (11) depends on v and v_0 via q; it peaks at $v = v_0$ since in that case $q = 0$. The somb and J_0 factors reflect the influence of bandwidth B and central frequency f_c respectively. As f_c increases, the main lobe of the J_0 factor becomes narrower. The somb factor has the effect of attenuating the lateral lobes, more so as B increases.

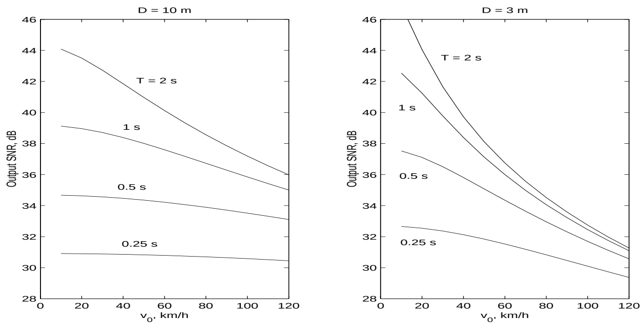

Figure 2: Output SNR vs. target speed. Lowpass signal and noises with $B = 5$ KHz and $\sigma_s^2/\sigma_w^2 = -10$ dB.

3.2. Variance of $\psi(v_0)$

If the signal and noise processes are lowpass Gaussian with bandwidth B and powers σ_s^2, σ_w^2, then the variance of $\psi(v)|_{v=v_0}$ can be determined following the approach in [2]:

$$\mathrm{var}[\psi(v_0)] \approx \frac{\alpha_0}{\tan\alpha_0}\frac{\sigma_s^2(\sigma_s^2+\sigma_w^2)}{BT} + \frac{\sigma_w^4}{2BT} \qquad (12)$$

This expression is more accurate for smaller α_0. The corresponding output SNR of the correlator is given by

$$\begin{aligned} \mathrm{SNR}_o &\triangleq \frac{E^2[\psi(v_0)]}{\mathrm{var}[\psi(v_0)]} \\ &= \frac{\left(\frac{\alpha_0}{\tan\alpha_0}\right)^2 BT}{\left(\frac{\alpha_0}{\tan\alpha_0}\right)\left(1+\frac{\sigma_w^2}{\sigma_s^2}\right)+\frac{1}{2}\left(\frac{\sigma_w^2}{\sigma_s^2}\right)^2} \end{aligned} \qquad (13)$$

SNR_o is plotted in Fig. 2 vs. target speed, for several values of the observation interval T, and in two settings: $D = 10$ m (array far from road) and 3 m (array close to road). A degradation in SNR_o is observed as v_0 increases, for fixed T; this is due to the signal attenuation which becomes more pronounced at a given time separation from the CPA as the target moves faster. Increasing T improves the output SNR, even though the factor $\alpha_0/\tan\alpha_0$ is decreased, since its influence in the time-bandwidth product BT outweighs this effect. It is also seen that placing the array closer to the source trajectory improves the output SNR for low speed values, but decreases it for high speeds.

4. 1-BIT QUANTIZED SIGNALS

If the received signals $r_i(t)$ are quantized to a single bit before performing the crosscorrelation (2), then for Gaussian $s(\cdot)$, following [4] one obtains

$$E[\psi(v)] \approx \frac{2}{\pi T}\int_{-T/2}^{T/2} \arcsin\left(\frac{R_s[\Delta^2\tau(t;v_0,v)]}{\sigma_s^2+\sigma_w^2\frac{d^2(t;v_0)}{D^2}}\right)dt. \qquad (14)$$

Note the dependence of $E[\psi(v)]$ with the input SNR σ_s^2/σ_w^2 in this case. For high SNR, the denominator of the arcsin argument in (14) is close to σ_s^2.

For a bandpass process with $G_s(f)$ as in (9), we can mimick the development in section 3.1 to arrive at

$$E[\psi(v)] \approx \frac{2}{\pi}\arcsin\left[\mathrm{somb}(Bq)J_0(2\pi f_c q)\right]. \qquad (15)$$

On the other hand, for low SNR, the arcsin argument is comfortably less than one, so that one can approximate

$$E[\psi(v)] \approx \frac{2}{\pi\sigma_w^2}\frac{1}{T}\int_{-T/2}^{T/2}\frac{R_s[\Delta^2\tau(t;v_0,v)]}{d^2(t;v_0)/D^2}dt, \qquad (16)$$

a scaled version of the expression for the unquantized case.

Computation of $\mathrm{var}[\psi(v_0)]$ in the 1-bit case involves the computation of a set of 4-variate orthant probabilities for which no closed-form expression is known [1]. For low input SNR, we can approximate this variance by the variance obtained when the signal is absent. Using the fact that, for x, y jointly Gaussian zero-mean random variables with correlation coefficient ρ,

$$E[\mathrm{sign}(x)\,\mathrm{sign}(y)] = \frac{2}{\pi}\arcsin\rho,$$

this variance computes to

$$\begin{aligned} \mathrm{var}[\psi(v)] &\approx \frac{(2/\pi)^2}{BT}\int_0^\infty [\arcsin(\mathrm{sinc}\,x)]^2 dx \\ &= \frac{4}{\pi^2}\frac{0.78}{BT}, \end{aligned} \qquad (17)$$

given a large time-bandwidth product BT. To obtain (17), use has been made of the integral identity

$$\int_{-T/2}^{T/2}\int_{-T/2}^{T/2} f(x-y)dxdy = T\int_{-T}^{T} f(x)dx - \int_{-T}^{T}|x|f(x)dx.$$

In the case $s(\cdot)$ is lowpass with bandwidth B, the output SNR that results is

$$\mathrm{SNR}_o \approx 1.28\left(\frac{\alpha_0}{\tan\alpha_0}\right)^2\left(\frac{\sigma_s^2}{\sigma_w^2}\right)^2 BT. \qquad (18)$$

Comparing (18) to (13) for $\sigma_w^2/\sigma_s^2 \gg 1$, it is seen that the asymptotic loss in output SNR due to 1-bit quantization for low input SNR is $10\log(2/1.28) = 1.9$ dB.

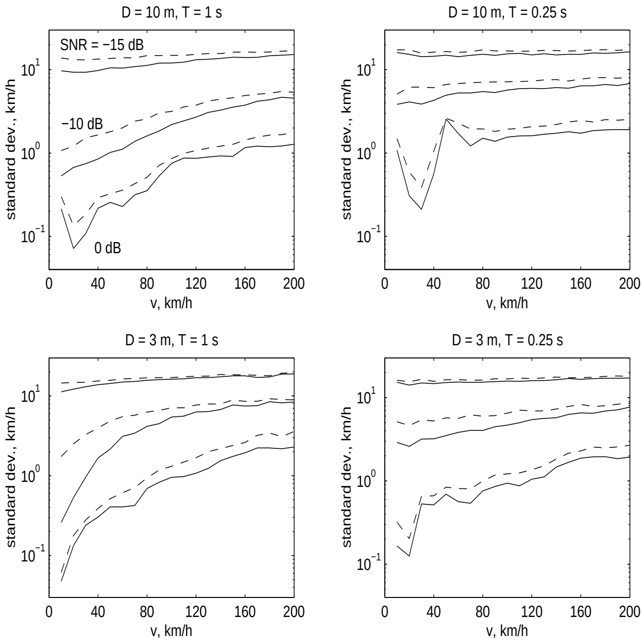

Figure 3: Standard deviation of the speed estimate using double-precision (solid) and 1-bit quantized (dashed) signals. $2b = 1$ m, $B = 5$ KHz.

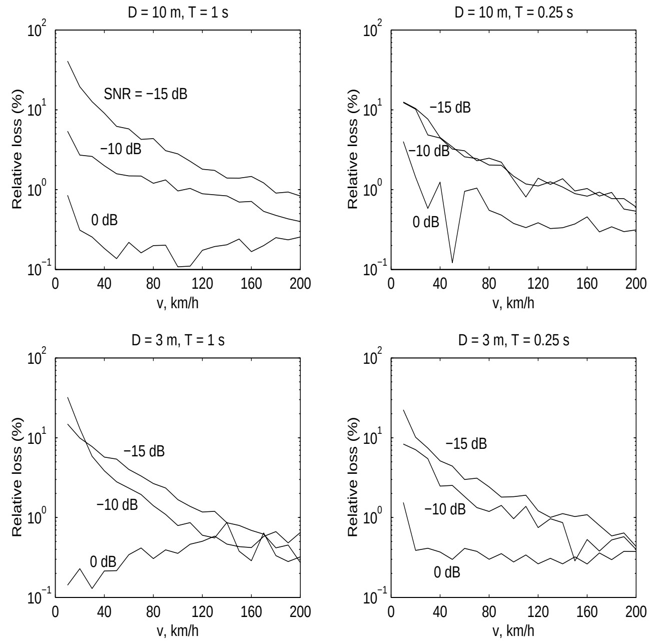

Figure 4: Relative loss of the 1-bit based w.r.t. the double-precision based estimator.

5. NUMERICAL RESULTS

In order to gain more insight into the speed estimator performance, Monte Carlo simulations based on 2000 independent trials were conducted. Fig. 3 shows the standard deviation $\sigma(\hat{v}_0)$ for both double-precision and 1-bit quantized signal cases, assuming an array separation $2b = 1$ m, $c = 340$ m/s, and observation windows of $T = 1$ and 0.25 s, for two settings: $D = 10$ m and $D = 3$ m. The acoustic signature was modeled as a lowpass Gaussian process with bandwidth $B = 5$ KHz, sampled at $f_s = 10$ KHz. The delayed values required to generate the synthetic received signals were computed via interpolation.

The standard deviation of the estimate increases with target speed v_0, as expected in view of the degradation of the output SNR given by (13) with v_0 seen in Fig. 2. The effect of a larger T is also seen to be more beneficial for slower targets.

The degradation in terms of variance incurred when using 1-bit quantization is seen to be mild. Notice from Fig. 4 that the relative loss, defined as

$$\frac{\sigma(\hat{v}_0)|_{\rm 1bit} - \sigma(\hat{v}_0)|_{\rm double}}{v_0}$$

is higher for low target speeds. Also, increasing T from 0.25 s to 1 s perceptibly reduces this loss for lower values of v_0, but not so much for higher ones.

T (s)	1.5	1.0	0.75	0.5	0.4	0.3	0.2	0.1
16-bit	43	44	44	44	48	48	50	31
1-bit	43	43	44	44	44	38	44	47

Table 1: Estimated speed (km/h) with varying T.

Last, we present the results obtained with real traffic signals, sampled at $f_s = 14.7$ KHz and recorded with 16 bit precision with omnidirectional microphones in a setting similar to that of Fig. 1 with $2b = 0.9$ m and $D \approx 14.5$ m. (The signals are available at http://www.gts.tsc.uvigo.es/˜valcarce/traffic.html). Fig. 5 shows the waveform and spectrogram of the acoustic signature of a compact car traveling at ≈ 40 km/h. Observe the presence of time-localized disturbances of impulsive nature as well as lowpass background noise with no significant spectral content beyond 1 KHz.

Fig. 6 shows $\psi(v)$ computed with 16- and 1-bit quantized signals. The CPA was taken at $t = 3.45$ s, determined from the short-time signal variance estimate using a sliding window of length 70 ms after highpass filtering the recorded signals (also shown in Fig. 5). Note how the main lobe widens as T is reduced. Table 1 gives the speed estimates for several T. It is apparent that a window of $T \geq 0.5$ s is required, and that using 1-bit quantization is a viable option.

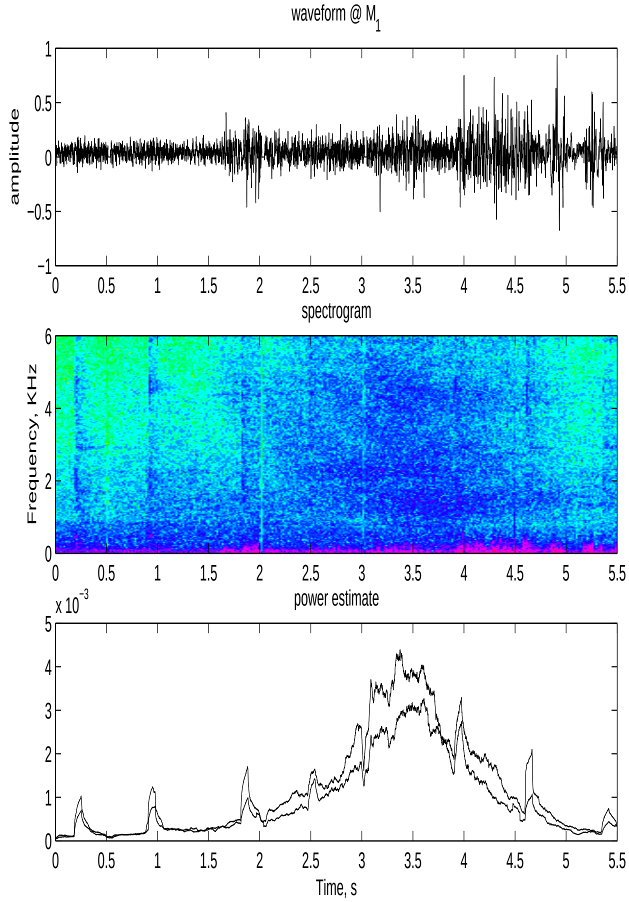

Figure 5: Acoustic signature of a passing car: waveform (top), spectrogram (middle), and short-term power estimate at both channels after highpass filtering (bottom).

6. CONCLUSIONS

Under a broadband Gaussian model for the signal and the noises, we have developed expressions for the expected log likelihood function for the speed estimation problem, as well as for the output SNR of the modified crosscorrelator on which the estimate is based. The influence of parameters such as true speed, time-bandwidth product, and angular aperture has been highlighted. The case of 1-bit quantized estimation and its subsequent output SNR degradation has also been examined. From this analysis, together with numerical results from synthetic signals as well as real traffic data, single bit quantization appears as a very attractive choice in order to reduce hardware complexity at the sensor location with a moderate loss in performance.

The estimate analyzed requires knowledge of the CPA location. In practice this parameter must be estimated as well, jointly or otherwise. Also, the robustness of the estimate to uncertainties in the array parameter values (mostly c and D) has to be assessed. Ongoing work is aimed along these directions.

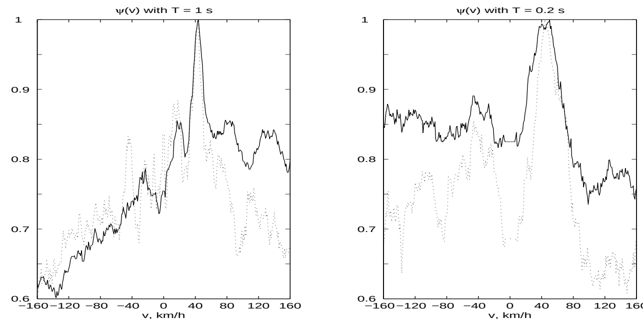

Figure 6: $\psi(v)$ computed with signals quantized to 16 (solid) and 1 bit (dotted): $T = 1$ s (left) and 0.2 s (right).

7. REFERENCES

[1] O. Bar-Shalom and A. J. Weiss, 'DOA estimation using 1-bit quantized measurements', *IEEE Trans. Aerosp. Elect. Syst.*, vol. 38, pp. 868-884, Jul. 2002.

[2] J. Betz, 'Comparison of the deskewed short-time correlator and the ML correlator', *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 32, pp. 285-294, Apr. 1984.

[3] S. Chen, Z. P. Sun and B. Bridge, 'Traffic monitoring using digital sound field mapping', *IEEE Trans. Vehicular Technology*, vol. 50, pp. 1582-1589, Nov. 2001.

[4] R. Cusani, 'Performance of fast time delay estimators', *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 37, pp. 757-759, May 1989.

[5] J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics.* Wiley, 1978.

[6] R. López-Valcarce, D. Hurtado, C. Mosquera and F. Pérez-González, 'Bias analysis and removal of a microphone array based road traffic speed estimator', to be presented at *EUSIPCO 2004*, Vienna.

[7] K. W. Lo and B. Ferguson, 'Broadband passive acoustic technique for target motion parameter estimation', *IEEE Trans. Aerospace and Electronic Systems*, vol. 36, pp. 163-175, Jan. 2000.

[8] F. Pérez-González, R. López-Valcarce, C. Mosquera, 'Road vehicle speed estimation from a two-microphone array', *Proc. ICASSP*, vol. 2, pp. 1321-1324, 2002.