Advances in Radio Science (2004) 2: 169–174

Advances in Radio Science

Entropie analysis of floating car data systems

F. Gössel, E. Michler, and B. Wrase

Dresden University of Technology, Fakultät für Verkehrswissenschaften „Friedrich List“, Lehrstuhl Informationstechnik für Verkehrssysteme, D-01062 Dresden, Germany

Abstract. The knowledge of the actual traffic state is a basic prerequisite of modern traffic telematic systems. Floating Car Data (FCD) systems are becoming more and more important for the provision of actual and reliable traffic data. In these systems the vehicle velocity is the original variable for the evaluation of the current traffic condition. As real FCD-systems are operating under conditions of limited transmission and processing capacity the analysis of the original variable vehicle speed is of special interest. Entropy considerations are especially useful for the deduction of fundamental restrictions and limitations.

The paper analyses velocity-time profiles by means of information entropy. It emphasises in quantification of the information content of velocity-time profiles and the discussion of entropy dynamic in velocity-time profiles. Investigations are based on empirical data derived during field trials. The analysis of entropy dynamic is carried out in two different ways. On one hand velocity differences within a certain interval of time are used, on the other hand the transinformation between velocities in certain time distances was evaluated.

One important result is an optimal sample-rate for the detection of velocity data in FCD-systems. The influence of spatial segmentation and of different states of traffic was discussed.

1 Introduction

The initial point of the historical development and original area of application of the statistical information theory is communications engineering. But already Shannon suggested the application of many conclusions from information theory in other areas of science and technology (Ebeling et al., 1998; Olberg and Rákóczi, 1984). Thereby the objective is not a formal transformation of termini from one area of science into another but to discover new interrelations, to generalise and extend the applicability of known relationships and to find principles of classification and systematisation. In this way the following information-entropical approach on FCD-systems should be understood. The usage of information theory in traffic science is not new. Already Potthoff used the entropy and transinformation to analyse the structure of railway systems (Potthoff, 1972).

The original variable in FCD-systems (Michler, 2001; Gössel et al., 2002) is the vehicle velocity. The vehicle velocity can be measured directly and is suitable for the detection of the actual traffic state. Certainly vehicle velocity can be regarded as a random variable. From this point of view the information source “vehicle velocity” will provide a certain amount of information. As real FCD-systems are operating under conditions of limited transmission and processing capacity the analysis of velocity-time resp. velocity-position profiles is of special interest. Entropy considerations thereby are especially useful for the deduction of fundamental restrictions and limitations.

In the following the amount of information in velocity-time profiles will be quantitatively analysed. Thereby entropy dynamic is of special interest. Results of the analysis are consisting of an estimation of the data rate in FCD-systems, possibilities and limitations in efficient coding of velocity data and the choice of an optimal sample rate.

2 Consideration of velocity-time resp. velocity-position profiles as sources of information

The vehicle velocity is a time and value continuous variable. Entropy analysis of value continuous variables usually require the knowledge of the probability-density function of this variable. Another possibility consists in sampling and discretisation of the velocity profile and consideration of the velocity profile as a sequence of symbols. The entropy of a discrete source X with n independent symbols x_i is given by:

$$H\,(X) = -\sum_{i=1}^{n} p\;(x_i)\;ldp\;(x_i), \tag{1}$$

Correspondence to: F. Gössel
(goessel@vini.vkw.tu-dresden.de)

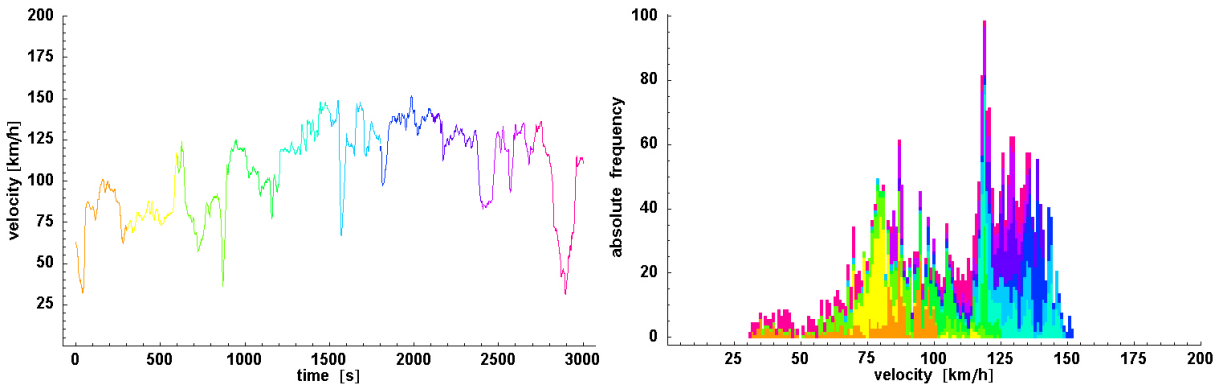

Fig. 1. Exemplary representation of a velocity-time profile (left) and the corresponding absolute frequency distribution, colours indicating segmentation (duration 300 s), quantisation 1 km/h.

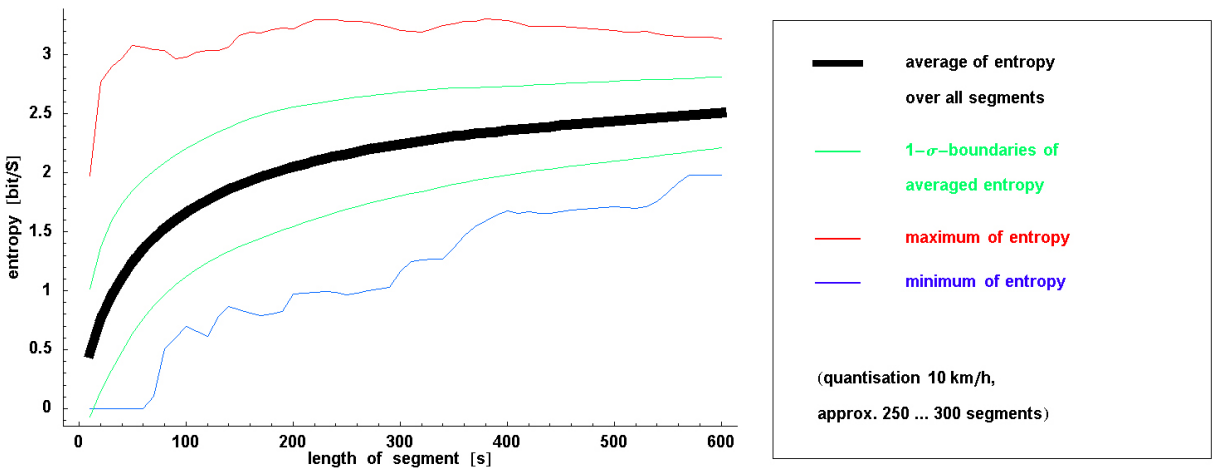

Fig. 2. Relation between entropy and length of segment.

with $p(x_i)$ giving the probability of the symbol x_i. Dividing the entropy of a source X with n independent symbols by its maximum $\hat{H}(X) = ld\, n$ yields the relative entropy:

$$H_r(X) = \frac{H(X)}{\hat{H}(X)}. \tag{2}$$

In the paper the probabilities of symbols are estimated from their relative frequencies in empirical data. The basis for estimating probabilities from the corresponding relative frequencies is the stationarity of the underlying random process. In the case of velocity profiles only stepwise stationarity can be assumed. Therefore empirical investigations are restricted to certain traffic scenarios assumed to be stationary.

The further investigations are restricted to velocity-time profiles based on following two reasons. Firstly velocity-time profiles emphasise parts of the velocity profile driven at low speed which are usually more important for traffic telematic systems. Second velocity profiles are normally measured using the Global Positioning System (GPS) resulting in velocity values equally spaced in time (typical distance one second). Obtaining velocity values equally spaced in position (velocity-position profiles) needs a data conversion meaning additional effort and sources of error.

2.1 Empirical determination of (information) entropy

The empirical calculation of entropy requires the probabilities $p(x_i)$ of symbols x_i of the source "velocity-time profile". Empirical data show very different probability density functions in different parts of a velocity-time profile (see Fig. 1). Using Eq. (1) to determine the entropy shows the dependence of entropy from quantisation, length of segment (segment = space-time area of interest with stationary conditions) and traffic state.

As expected the length of the observation area (segment) has an important influence on the entropy (see Fig. 2). Considering short segments velocity usually does not change very much. This results in a small number of occupied velocity classes leading to a comparatively small entropy. But for some short segments nevertheless the entropy is quite large, especially in case of a noticeable change of speed within the segment. The consequence is a wide range of entropy. One problem in calculating the entropy from empirical data of

small segments is the small amount of data not allowing a stable estimation of the frequency distribution.

On the other hand using long segments often leads to different street classes and traffic states within one segment. In these cases stationarity can not be assumed. As a trade-off the following segment length will be used: segments with duration of 10 min on motorways and segments with duration of 2 min in city traffic.

2.2 Empirical investigation in the dependence of information entropy from street class and traffic state

Street class and traffic state have considerable influence on the information entropy. Driving on motorways usually means little variation in velocity especially in case of speed limits. Assuming free traffic results in a comparatively small entropy. Substantial differences occur in Stop-and-Go situations (which are comparable with city traffic) and on motorways without speed limit (requiring a stronger adaptation of the speed on external conditions). Figure 3 gives examples of a typical velocity-time profile on a motorway and a typical velocity-time profile in city traffic. The corresponding entropies are given in Table 1.

Investigations have shown very little difference of the entropy (approximately 0,5 bit/S) on various German motorways (with different topographic characteristics and different speed limits). The entropy of velocity-time profiles in city traffic is characterised by a wide range of segment-based entropy values. Many segments achieve nearly the theoretical maximum of entropy (e.g. acceleration after stop at traffic lights), others show very little entropy (e.g. slowly approaching traffic lights). Comparing information entropy of motorways and city traffic shows significant differences in relative entropy but quite similar (absolute) entropy values.

Summarising the empirical investigations the information entropy of velocity-time profiles has proven to be quite moderate in the order of 2 bit/S. It should be mentioned that correlations between successive velocity values are yet not considered in the investigation. Consideration of these correlations will cause a further reduction of information entropy of velocity-time profiles.

3 Investigation in entropy dynamic (Ebeling et al., 1998; Fey, 1968)

Entropy dynamic provide an insight into the dependencies and correlations of symbols within a sequence. Special sequences are the independent sequence, the periodic sequence and Markov sequences of order m. Without doubt velocity-time profiles are neither independent nor periodic sequences. Due to physical reasons (inertia) the acceleration/deceleration of a vehicle is limited. This results in strong statistical dependencies between successive velocity values. Of cause these dependencies are decreasing with increasing time shift. Apparently speed-time profiles have some properties of Markov sequences.

The analysis of entropy dynamic is carried out in two different ways. On one hand velocity differences within a certain interval of time are used, on the other hand the transinformation between velocities in certain time distances was evaluated. Analysis of velocity differences is a very simple concept. Its main advantage is the effective use of the available empirical data. But this concept does not consider information of the concrete value of the initial velocity (e.g. a very high velocity is much less probable followed by an even higher velocity than by a lower velocity). This disadvantage can be avoided by the concept of transiformation. The transinformation evaluates the statistical dependencies between velocity values in a certain time distance. Compared to the analysis of velocity differences longer empirical sequences are needed.

3.1 Analysis of velocity differences

In the following all velocity differences $\Delta v = v\ (t_i + \Delta t) - v\ (t_i)$ for the run of a vehicle through a (time) segment $t_{\min} < t_i < t_{\max} - \Delta t$ are considered. Based on the resulting frequencies for certain velocity differences after a given time Δt (see Fig. 4) the associated information (entropy) can be calculated. Small values of information entropy indicate a nearly deterministic relation between the velocities at the considered times whereas a large entropy means a more random relationship.

Figure 5 shows the entropy of velocity differences over the time difference Δt calculated from empirical data of several segments on a motorway. This figure confirms the strong statistical dependencies for small time differences Δt. The given runs show an asymptotic behaviour. Close statistical dependencies can be found up to a time difference of about $\Delta t_{max} = 40\,\mathrm{s}$ ...100 s (depending from the segment). Noticeable differences between the given runs can be found for time differences $\Delta t > 10\,\mathrm{s}$. Thereby the single runs differ not more than about 25% from each other.

3.2 Analysis of transinformation

The transinformation $T\ (v_1, v_2)$ between two single velocities $v_1 = v\,(t_1)$ and $v_2 = v\,(t_2)$ with $t_2 = t_1 + \Delta t$ gives the amount of information which can be obtained from the knowledge of the velocity value v_1 over the velocity value v_2. The transinformation $T\ (v_1, v_2)$ is given by (Ebeling et al., 1998):

$$T\ (v_1, v_2) = H\ (v_1) + H\ (v_2) - H\ (v_1, v_2)$$
$$\text{with } H\ (v_1, v_2) \ldots \text{ compound entropy.} \tag{3}$$

For an independent random process the transinformation between two symbols is identical zero. In case of a deterministic relationship the transinformation between two symbols is identical to the (identical) entropy of both single symbols, i.e. the transinformation represents the entire information of the following symbol.

Figure 6 gives the relationship between transinformation of two velocity values in time distance Δt over this

Table 1. Entropy calculation for Fig. 3 (sliding segments of 10 min duration (motorway) and 2 min duration (city)).

	motorway (Fig. 3)	city (Fig. 3)
Minimum of entropy of segment $\mathbf{H}_{\min}$ [bit/S]	1,36	0,25
Minimum of relative entropy of segment $\mathbf{H}_{r,\min}$	33%	9%
Maximum of entropy of segment $\mathbf{H}_{\max}$ [bit/S]	2,76	2,53
Maximum of relative entropy of segment $\mathbf{H}_{r,\max}$	66%	90%
Entropy [bit/S]	2,04	1,68
Relative entropy	49%	60%

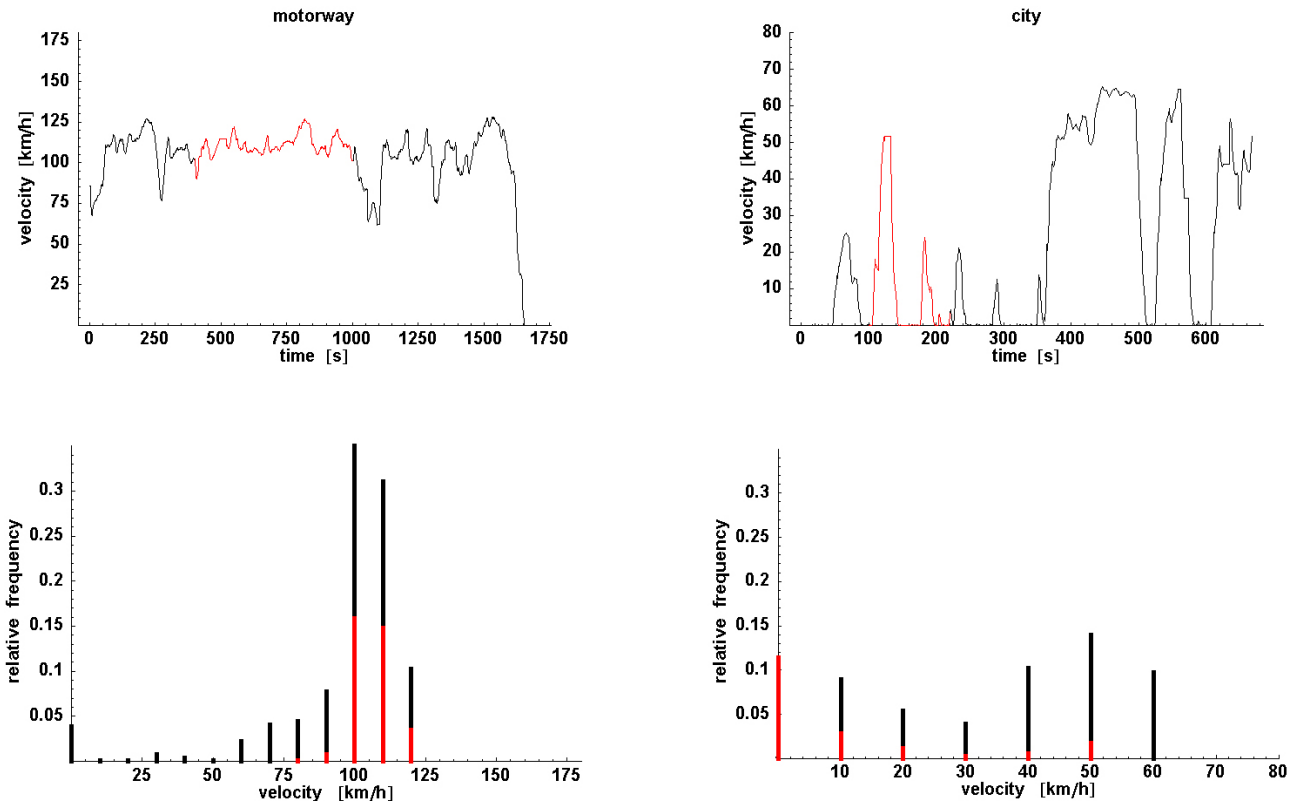

Fig. 3. Exemplary velocity-time profiles; example for single segment highlighted (red).

time distance Δt for empirical data (several segments of a test run). All runs show a nearly asymptotic decreasing behaviour. There can be distinguished between three characteristic sections. For very small time differences ($0<\Delta t<10$ s) the transinformation decreases very steep with increasing time difference. Further increasing time differences ($10\,\text{s}<\Delta t<80$ s) also result in a further but slower decrease of the transinformation. For time differences $\Delta t>80$ s the transinformation is approaching a limit asymptotically.

The comparative large values of the transinformation at small time differences are an expression of inertia, i.e. the limited acceleration/deceleration ability of real vehicles. The range of medium time differences is dominated by the desire of the driver to keep a certain velocity. The not completely vanishing transinformation for larger time differences can be traced back to the choice of stationary velocity-time profiles. Stationarity means unchanging street classes and unchanging traffic condition resulting in preferring certain velocity values.

4 Optimal (acquisition) sample-rate in FCD-systems from an information entropic point of view

The acquisition sample rate is an important parameter of FCD-systems as it significantly influences the necessary effort in data acquisition, data transmission and data processing. Transmission of velocity data in very short intervals means small information content of the single values as they could distinguished with high probability from the preceding values. Contrary detection of velocity values only in very long intervals results in slow reactions of the system on a

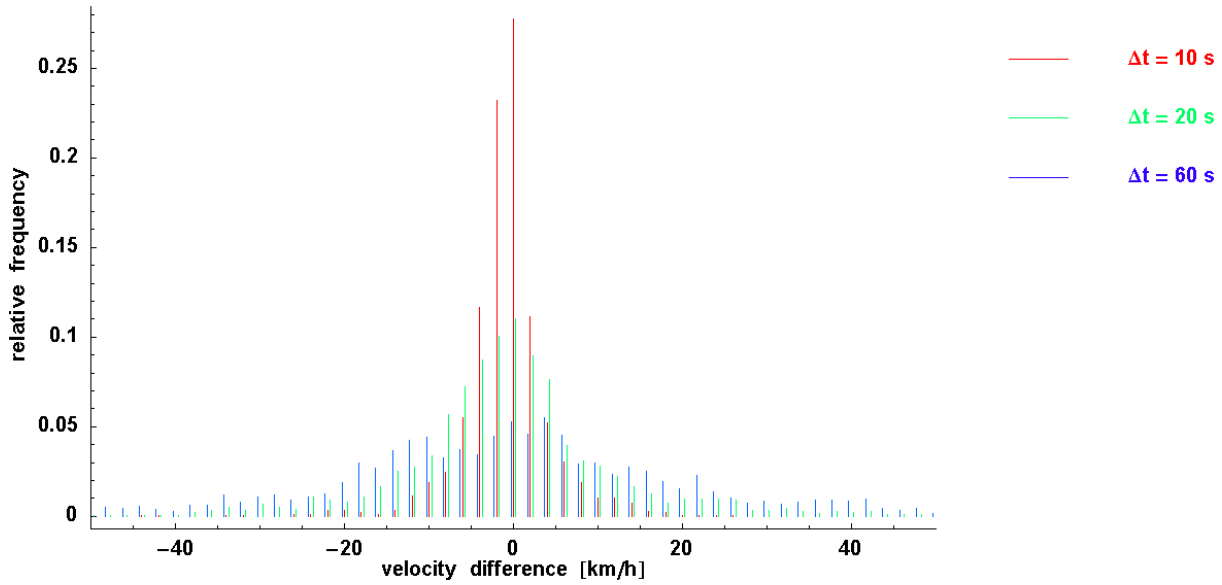

Fig. 4. Relative frequencies of velocity differences in a given time interval (parameter).

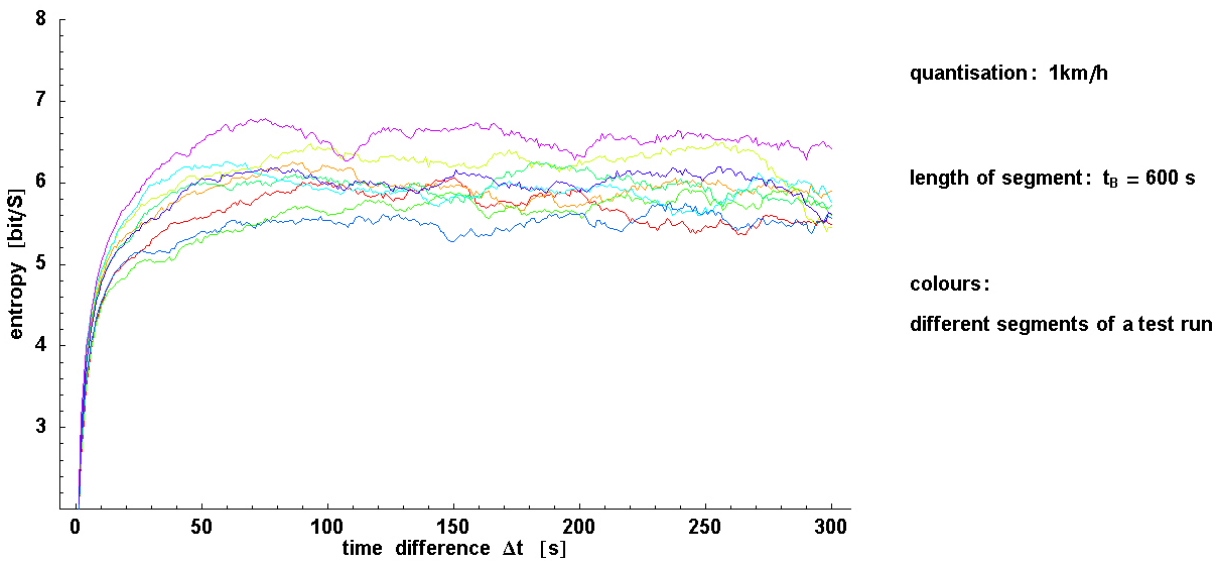

Fig. 5. Entropy of velocity differences over time difference for several segments (parameter) of a test run on a motorway.

changing environment. Considering the entropy of velocity differences (Sect. 3.1) velocity values within an interval of 10 s show very strong statistical dependencies but after about 100 ms velocity differences occur nearly randomly. Correspondingly an acquisition sample rate of $t_E < 10$ s means (considerable) redundant data whereas an $t_E > 100$ s would not completely characterise the velocity-time profile. Analysis of transinformation yields similar results. Statistical dependencies can be found up to time differences of about 80 s. Further proceeding dependencies result from the overall speed level on a certain street class and are therefore for a traffic observation system of little interest.

5 Conclusion

An acquisition sample rate of more than 0,1 Hz (i.e. ≤ 10 s time distance between samples) yields only a marginal information gain. The information gain per sample increases with the reduction of the acquisition sample rate up to approx. 0,01 Hz. At a further reduction of the acquisition sample rate the information gain per sample keeps nearly constant. Thus an acquisition sample rate between 0,01 Hz and 0,1 Hz should be taken.

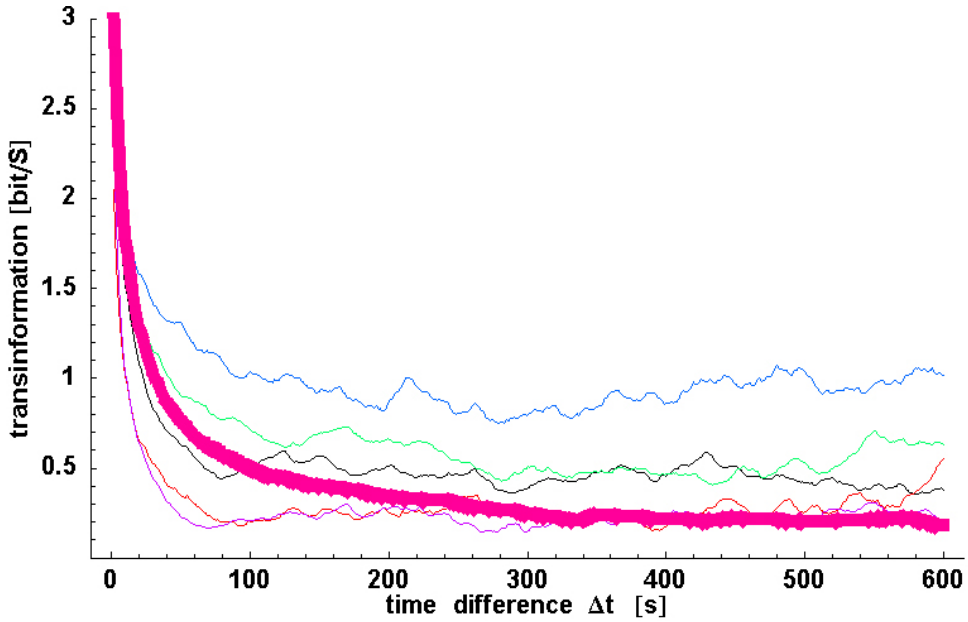

Fig. 6. Transinformation between velocity values in time distance Δt for segments (parameter) of a test run, result for whole test run highlighted (bold).

6 Summary

The paper analysis velocity-time profiles based on the (information) entropy. It could be shown that velocity-time profiles are characterised by a strong redundancy. Empirical investigations point out significant differences of the relative entropy for different street classes and traffic states. For the (absolute) entropy this significant difference could not be found. The resulting entropy of velocity-time profiles is moderate with about 2 bit/S. Investigation in entropy dynamic shows strong statistical dependencies for velocity values within short time distances. Based on entropy dynamic an optimal acquisition sample rate was distinguished.

References

Ebeling, W., Freund, J., and Schweitzer, F.: Komplexe Strukturen: Entropie und Information, Teubner Stuttgart Leipzig, 1998.

Fey, P.: Informationstheorie, Berlin, Akademie-Verlag, 1968.

Gössel, F., Michler, E., and Wrase, B.: Spectral Analysis of Floating Car Data, Kleinheubacher Berichte 2002/Advances in Radio Science, 2003.

Michler, E.: Verkehrsmonitoring/Verkehrsdatengewinnung, Kleinheubacher Berichte, 44, 296–305, 2001.

Olberg, M. and Rákóczi, F.: Informationstheorie in Meteorologie und Geophysik, Berlin, Akademie Verlag, 1984.

Potthoff, G.: Verkehrsströmungslehre, Band 4, Berlin, Transpress Verlag für Verkehrswesen, 1972.