The image resulting from plane rectification will be referred to as the rectified domain or the transformed domain. In the rectified domain, the motion of vehicles can be safely described as a first-order linear equation with an added random noise.

One important limitation regarding the dynamic model in existing methods is the interaction treatment. Most approaches to multiple vehicle tracking involve independent motion models for each vehicle. However, this requires an external method for handling of interaction, and often this is simply disregarded. In contrast, we have designed an MRF-based interaction model that can be easily integrated with the above-mentioned individual vehicle dynamic model.

Finally, a method is necessary to detect new vehicles in the scene so that they can be integrated in the tracking framework. This is addressed in the current work by using a two-step procedure composed of an initial hypothesis generation and a subsequent hypothesis verification. In particular, candidates are verified using a supervised classification strategy over a new descriptor based on HOG features. The proposed feature descriptor and the classification strategy are explained in Section 6.

The explained framework is summarized in the general scheme shown in Fig. 1. The scheme shows the main constituent blocks of the method, i.e., observation model (which in turn relies on appearance and motion analysis), motion model, vehicle tracking algorithm, and new vehicle detection algorithm, as well as the techniques used for their design. These blocks are explained in detail in the following sections.

3 Vehicle tracking algorithm

The designed vehicle tracking algorithm aims at estimating the position of the vehicles existing at each time of the image sequence. Hence, the state vector is defined to comprise the position of all the vehicles \( s_k = \{s_{i,k}\}_{i=1}^M \), where \( s_{i,k} \) denotes the position of vehicle \( i \), and \( M \) is the number of vehicles existing in the image at time \( k \). As stated, the position of a vehicle is defined in the rectified domain given by the transformation \( T \), although back-projection to the original domain is naturally possible via the inverse projective transformation \( T^{-1} \).

An example of the bird’s-eye view obtained through IPM is illustrated in Fig. 2. Observe that the upper part of the vehicles is distorted in the rectified domain. This is due to the fact that IPM calculates the appropriate transformation for a given reference plane (in this case the road plane), which is not valid for all of the elements outside this plane. Therefore, analysis is focused on the road plane image and the position of a vehicle will be defined as the middle point of its lower edge. This is given in pixels, \( s_{i,k} = (x_{i,k}, y_{i,k}) \), where \( x \) indicates the column and \( y \) the row of the corresponding point in the image, while the origin is set at the upper-left corner of the image.

In order to estimate the joint state of all of the vehicles, the MCMC method is applied. As mentioned, in MCMC the approximation to the posterior distribution of the state is given by (2), which, assuming that the likelihood of the different objects is independent, can be rewritten as follows:

\[
p(s_k | z_{1:k}) \approx c \cdot \prod_{i=1}^M p(z_{i,k} | s_{i,k}) \sum_{r=1}^N p(s_k | s_{k-1}^{(r)})
\]

where \( z_{i,k} \) is the observation at time \( k \) for object \( i \). In MCMC, samples are generated sequentially from a proposal distribution that depends on the current state, therefore the sequence of samples forms a Markov chain. The Markov chain of samples at time \( k \) is generated as follows. First, the initial state is obtained as the mean of the samples in \( k-1 \), \( s_k^0 = \sum_r s_k^{(r)} / N \). New samples for the chain are generated from a proposal distribution \( Q(\cdot) \). Specifically, we follow a Gibbs-like approach, in which only one target is changed at each step of the chain. At step \( \tau \) the proposed position \( s_{i,k}^{(\tau)} \) of the randomly selected target \( i \) is thus sampled from the proposal distribution, which in our case is a Gaussian centered at the value of the last sample for that target, \( Q(s_{i,k}^{(\tau)} | s_{i,k}) = N(s_{i,k}^{(\tau)}, \sigma_q) \). The candidate sample is therefore \( s_k^{(\tau)} = (s_{i,k}^{(\tau)}, s_{i,k}^{(\tau)}) \), where \( s_{i,k}^{(\tau)} \) denotes \( s_k \) but with \( s_{i,k} \) omitted. This sample is or is not accepted according to the Metropolis algorithm, which evaluates the posterior probability of the candidate sample in comparison to that of the previous sample and defines the following probability of acceptance [31]:

\[
\begin{align*}
\text{Accept } & \quad \text{if } \frac{p(s_k^{(\tau)} | z_{1:k})}{p(s_k | z_{1:k})} \geq 1 \\
\text{Reject } & \quad \text{otherwise}
\end{align*}
\]