

Learning to Predict Slip for Ground Robots

Anelia Angelova

Computer Science Dept.

California Institute of Technology

Email: anelia@vision.caltech.edu

Larry Matthies, Daniel Helmick, Gabe Sibley

Jet Propulsion Laboratory

California Institute of Technology

lhm,dhelmick,gsibley@jpl.nasa.gov

Pietro Perona

Electrical Engineering Dept.

California Institute of Technology

perona@vision.caltech.edu

Abstract—In this paper we predict the amount of slip an exploration rover would experience using stereo imagery by learning from previous examples of traversing similar terrain. To do that, the information of terrain appearance and geometry regarding some location is correlated to the slip measured by the rover while this location is being traversed. This relationship is learned from previous experience, so slip can be predicted later at a distance from visual information only. The advantages of the approach are: 1) learning from examples allows the system to adapt to unknown terrains rather than using fixed heuristics or predefined rules; 2) the feedback about the observed slip is received from the vehicle’s own sensors which can fully automate the process; 3) learning slip from previous experience can replace complex mechanical modeling of vehicle or terrain, which is time consuming and not necessarily feasible. Predicting slip is motivated by the need to assess the risk of getting trapped before entering a particular terrain. For example, a planning algorithm can utilize slip information by taking into consideration that a slippery terrain is costly or hazardous to traverse. A generic nonlinear regression framework is proposed in which the terrain type is determined from appearance and then a nonlinear model of slip is learned for a particular terrain type. In this paper we focus only on the latter problem and provide slip learning and prediction results for terrain types, such as soil, sand, gravel, and asphalt. The slip prediction error achieved is about 15% which is comparable to the measurement errors for slip itself.

I. INTRODUCTION

The Mars Exploration Rover (MER) mission sent two rovers to the surface of Mars which have been exploring it since January 2004, collecting and returning invaluable scientific information. Future planetary exploration missions are being targeted to cover a larger variety of terrains, to drive faster, and to navigate with a higher level of autonomy. More intelligent planning is required in which various aspects of the terrain and rover-terrain interaction are considered, so that the safest, most efficient path is found.

One major problem, encountered in the ongoing MER mission, is that the rovers experience a significant amount of slip on certain terrains. The rovers can get trapped in regions from which they cannot recover due to a 100% slip, which can lead to mission failure, or can slow progress toward their goals. Being able to automatically predict the expected slip beforehand, using vision, can help prevent the rover from going into regions unsafe to traverse, for example, to avoid getting stuck in a sand dune, or to determine if it would be possible to climb back out of a crater before descending into it. Thus, the rover would be able to access more targets of scientific interest in craters or on hills, which are avoided

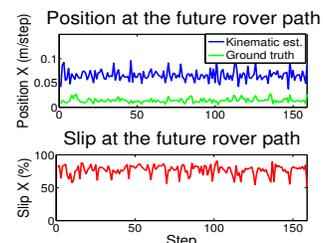


Fig. 1. Learning to correlate how things look to how they feel: the rover collects visual information about the forthcoming terrain from a sequence of stereo pair images, which is later correlated to the experienced slip at the locations traversed by the rover. An image from the left stereo camera displaying the rover’s future steps and the corresponding map cells (left), and the measured slip on the future path of the rover (right). Slip is measured as the difference between the commanded position per step (estimated by the vehicle’s kinematic model) and the actual position (ground truth), normalized by the commanded position.

in the current more conservative planning. The usefulness of slip prediction is not limited to only planetary exploration. Enhancing a path planning algorithm with information about the terrain type to be traversed and the potential slip would considerably improve robot mobility for any autonomous ground vehicle.

In this paper we propose a method to predict the amount of slip a planetary rover would experience using visual information about the terrain. We hypothesize that visual characteristics of the terrain can give clues to its physical properties and the eventual rover-terrain interaction, one aspect of which is slip. The main idea is to extract information about the terrain observed from a distance (using information from a stereo pair only), measure the slip of the rover when it traverses this particular region, and create a mapping between visual information and the resultant slip (Figure 1). We learn this functional relationship using training examples, which makes it possible to predict what the rover would experience at a future location by only observing it at a distance. More importantly, we use the rover’s own sensors as feedback to the learning, which can remove the human-in-the-loop factor for data labeling. To our knowledge, none of the previous work has attempted to predict slip using only information from stereo imagery. This idea may apply to predicting some other properties of the terrain or the vehicle-terrain interaction but here we concentrate only on slip.

Stereo imagery contains information about both the geometry (from range data) and the appearance (from the input

images) of the terrain. This gives us the intuition that visual information can also be used to predict slip, because, from a mechanical point of view, slip depends again on physical and geometrical properties of the terrain. The main issue is how to utilize vision data to infer properties about the terrain or predict slip. We propose a general framework in which the task is subdivided into recognizing terrain type from appearance and learning/predicting the slip behavior when the terrain type is fixed. In this paper we mainly focus on the latter problem, namely, learning the functional relationship between terrain geometry and slippage, when the terrain type is known. Recognizing the terrain type from appearance, which is a problem of visual texture recognition [21], is on-going work of ours and is beyond the scope of this paper.

In this work, we use slip measurements from real-life terrains with all complications, such as uneven and nonhomogeneous terrain, clumps on the ground, or rocks in front of the wheels, which can make the rover’s motion quite unpredictable. The problem is further complicated by uncertainties of predicting the terrain type from imagery and estimating slopes from terrain geometry obtained by the stereo vision.

A. Previous work

Rover-terrain interactions. Slip can be defined as the difference between the vehicle’s position estimated by its kinematics and its actual position. However, it is caused by complicated vehicle-terrain interactions [4], [20], [23], and is very hard to model from a mechanical point of view. Slip depends on the mechanical soil characteristics [20] which can be further modeled [1], or can be crudely described by a few parameters, e.g. cohesion and friction angle [4], [23] which, in turn, need to be estimated. A method for online terrain parameters estimation has been proposed in [11], but it is not predictive, as its measurements have to be done at the traversed location. The same applies to the method proposed for learning of terrain types from vibrations [6]. Outside the realm of mechanical modeling very few authors have addressed the problem of estimating, characterizing, or counteracting slip (with a few exceptions [9], [10], [12]), although slip has been acknowledged as an omnipresent problem in localization, or in rough-terrain mobility. In summary, from a mechanical point of view, slip requires a complex and vehicle-specific modeling and involves additional measurements about the terrain, which might not always be feasible at a distance.

Mars Rover tilt-table experiments. Previous investigations of the behavior of a Mars Rover on slopes have been done by Lindemann [15]. The experiments consist of short traverses on a tilt-table platform set to varying slope angles with the rover approaching the terrain at different orientations. From these tests we learn that slip is a highly nonlinear function of terrain slopes [15], which is important for our choice of learning method later in the paper. We have to note that the measured slip values are applicable only for the particular vehicle model (in this case MER). For example, a small design modification in the pattern of the wheels can change the slip behavior [3], affecting a potential physical model. We believe

that learning slip is a more general approach, namely, the same learning algorithm can be applied to another vehicle to learn its particular behavior on different terrains.

Slip compensation. A significant advancement in navigation on slopes in high slip environments has been achieved with the slip compensation algorithm by Helmick et al. [9], [10], who demonstrated improved goal acquisition and path following. The slip measured at a particular step is taken into account to adjust the next step, compensating for the distance which was not traversed. However, this does not allow for planning at a distance, which our method enables.

Using slip to determine terrain characteristics. The work described above concerns estimating slip from mechanical measurements, or, in our case, visual information. Conversely, slip measurements have been used to infer mechanical terrain parameters on the Mars Pathfinder Mission in a controlled one-wheel soil-mechanics experiment [17]. Similar experiments have been done by [2], [11]. This gives us the assertion that slip characteristics are directly correlated to terrain mechanical properties and the intuition that if the terrain soil type could be correctly recognized (which would entail its mechanical properties) then slip behavior is predictable given the terrain.

II. EXPERIMENTAL SETUP

A. Testbed

This research is motivated mainly by planetary rovers, such as MER, so we used Rocky8, a Mars research rover testbed. For most of our experiments, however, we used an experimental LAGR robot (Figure 2) because it is a more convenient data collection platform. It has two front differential drive wheels and two rear castor wheels. It is equipped with a pair of stereo cameras, wheel encoders, IMU, and GPS (the IMU and GPS are postprocessed into a global pose). It can run in autonomous mode or be manually joysticked using a radio controller. The vehicle can achieve speeds of up to 1.2 m/s.

B. Quantifying slip

The amount of slip s per wheel can be defined as a difference between the velocity measured by the wheel and the actual velocity: $s = wr - v$ [23], or can be normalized $s = \frac{wr - v}{wr}$ [3], [11], where v is the actual linear velocity, w is angular wheel velocity, and r is the wheel radius. Similarly, to define slip for the whole rover we measure the difference between the actual vehicle velocity and its velocity from the kinematic model for each motion DOF of the rover [9]. The actual position (ground truth) can be estimated by visually tracking features, a method called Visual Odometry (VO) [16], [18] or measured with some global position estimation device. VO is a preferred method for ground truth estimation because it is a convenient self-contained sensor on the vehicle, so data collection for training could be done automatically and the whole learning could be done online onboard the vehicle, which coincides with the goals of planetary exploration missions. Furthermore, global positioning devices are not always available, especially regarding planetary missions.



Fig. 2. Rocky8 vehicle in the Mojave desert (left), LAGR vehicle on a sandy terrain (right)

To measure slip, in this paper, we will use the change in position per step, instead of velocity, by normalizing, so that each step takes the same time. A step is defined as the interval between two consecutive stereo pair frames. In our datasets, Rocky 8 stops after each step, while LAGR obtains stereo frames continuously at a particular rate (5Hz). For the kinematic estimate, we use a differential drive model for the LAGR vehicle, and a full rocker-bogie kinematic model for Rocky8 [9]. We consider the slip in the previous rover frame (corresponding to the beginning of the step) which is defined as follows: the X coordinate is along the direction of forward motion, Y is along the wheel axis, and Z is pointing down. We are interested in the longitudinal (slip in X) and lateral (slip in Y) component of slip. We also consider slip in Yaw which is the rotation angle around the Z axis. Note that the LAGR vehicle has only three kinematically observable DOF. Figure 3 visualizes slip behavior of Rocky8 and the corresponding measured slip in X and Y on a tilt-table sand box (at 10°) where the vehicle traverses the platform diagonally. VO estimates are used as ground truth. One should note the significant slip downslope. The kinematic estimate is above the commanded pose because throughout the run the rover automatically adjusts its commands trying to reach its goal waypoint with respect to its current position. We can also notice some randomness in the slip behavior (which is due to uncertainties in the terrain, noise in the vehicle kinematics and measurement errors).

To draw our final conclusions later in the paper we normalize slip by the commanded position (or velocity) in X (thus, slip can also be expressed in %, as in Figure 1). This is more natural because most of the issued commands are towards forward motion, with an exception of a crabbing motion of Rocky8 rover. There are very few steps of this kind in our data so we can ignore them.

C. Visual Odometry validation

As we will be using the VO estimate as ground truth, we need to quantify its precision. We compared VO position estimates to ground truth collected with a surveyor total station (TS), tracking a single point. The TS has ~ 2 mm precision. The whole run was ~ 7.5 m in which the rover took 80 frames, each step was targeted to be of size ~ 0.1 m (the step size is selected in the middle of the range used for our later experiments). The VO error at the end of the run was 0.15m which is 2% of the distance traveled. Considering performance stepwise,

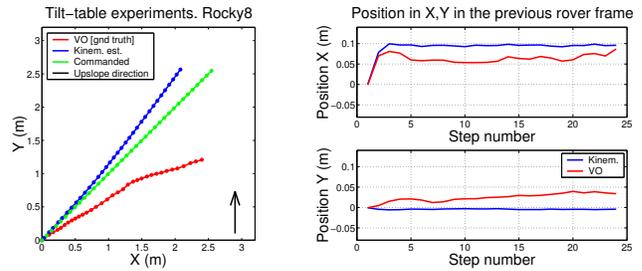


Fig. 3. Traversed path by Rocky8 on a tilt-table platform (left) and the corresponding position measured stepwise in the previous rover frame (right). Slip is the difference between the actual position (measured by VO) and the commanded position (estimated by kinematics). The tilt-table is covered by deep sand. 10° slope. Significant slip occurs in the downslope direction.

we computed the standard deviation of the differences of the distance traveled per step, as measured by the VO and the TS, to be $\sigma = 0.0123$ m. This gives a 1-sigma error of 12% on the ground truth measurement per step. Integration of the step-wise VO errors (both positive and negative) necessitates that the VO error of the entire run is smaller (in percentage) than the stepwise error.

D. Sources of error in measuring slip

Some sources of error in measuring slip come from the kinematic model and its assumptions. Systematic model errors can be due to imprecise measurement of the wheel diameter, different wheel diameters (the wheels might not be at their nominal air pressure), or limited encoder resolution [5]. They give a (usually small) bias in the measurement, but as our slip measurement is done stepwise, it will not be significantly affected by them. Random noise, coming from the terrain and the vehicle-terrain interaction, would also affect slip. It can be due to internal forces (castor wheels), external forces (obstacles in front of any of the wheels), or varying wheel-ground contact area (with the extreme case being when there is no contact with one of the wheels). For example, inconsistent slip measurement will occur when one of the wheels goes over a rock because it will create more or less traction compared to the soil, or can serve as an additional external force to the vehicle. Differential drive makes a flat-terrain assumption which is not always true in our off-road rough-terrain data. Slip also depends on the commanded velocity, although for the relatively small speeds of the rovers, velocity is not a significant factor. We have factored it out by averaging steps or driving at approximately constant velocity, but it will still introduce a certain amount of noise. In summary, other sources of error, which we do not take into consideration, will contribute to noise in slip measurements.

III. SLIP LEARNING AND PREDICTION

In this section we describe a general framework to learn the functional relationship between visual information and the measured slip using training examples. We choose a learning approach to the problem because creating a physical slip model is extremely complicated due to the large number of variables

involved. Also, the mapping between vision information and slip does not take a simple analytical form and the only way to observe it and learn about it is through training examples.

A. Slip modeling

The amount of slip depends both on the type of terrain traversed [20] and its geometry [15] in a nontrivial, nonlinear way. We consider slip learning in a nonlinear regression framework. Our inputs will be terrain geometry (captured by slopes) and terrain type (described by its appearance, such as texture and color); the outputs will be the slip in several independent coordinates. As this is a very complex function, we introduce some structure to make the problem tractable. We can think of slip as a variable S which changes with respect to the slopes of the terrain (let us term the dependence of slip on slope angles as *slip behavior*). As the slip behavior changes on different terrains, we can cast the problem into a Mixture of Experts (MoE) framework [13] in which the input space is partitioned into subregions (corresponding to different terrain types). The appearance information is mostly relevant in doing this partitioning. In each region (which corresponds to a particular terrain type) one (learned) model of slip behavior would be active, i.e. when the terrain type is known, slip will be a function of terrain geometry only. More formally, let I be all the information available from stereo pair images. It can be represented by appearance (A) and geometry (G) components. Specific representations of appearance and geometry will be discussed in the following sections; for now we consider them as abstract terms. Let $f(S|I)$ be the regression function of slip S (slip can be any of the slip in X, Y, or Yaw) on the input variables A, G (used interchangeably with the image information I). Now, considering that we have several options for a terrain type T , each one occurring with probability $P(T|A, G)$, we can write $f(S|I)$ as follows:

$$f(S|I) = f(S|A, G) = \sum_T P(T|A, G) f(S|T, A, G).$$

Naturally $\sum_T P(T|A, G) = 1$. This type of modeling admits one exclusive terrain type to be selected per image, or a soft partitioning of the space, which allows for uncertainty in terrain classification. We assume that the terrain type is independent of terrain geometry $P(T|A, G) = P(T|A)$ and that, given the terrain type, slip is independent of appearance $f(S|T, A, G) = f(S|T, G)$. So we get:

$$f(S|I) = \sum_T P(T|A) f(S|T, G).$$

In other words, to conquer the slip learning problem we divide it into a terrain recognition part, in which we determine the terrain type from appearance information i.e. $P(T|A)$, and a slip prediction part, in which we learn slip behavior with respect to terrain geometry $f(S|T, G)$, given a particular terrain type (Section IV). We take advantage of the independence of appearance and geometry, which is a very reasonable assumption. For example, a sandy terrain in front of the rover,

would appear approximately the same, no matter if the rover is traversing a level or tilted surface. The MoE approach is very adequate for our slip prediction problem because terrain types do not represent a continuum in appearance space and, in general, would form separate regions in the input space (experts). However, several experts might need to be active at the same time, to make smooth transitions in borderline terrain cases. Both cases are naturally incorporated in the MoE framework. The alternative to introducing structure in the problem is pooling appearance and geometry features, which will not only make the problem more complex, because of increased dimensionality, but will also require a formidable amount of training data. Moreover, the MoE framework is quite general and, in principle, allows for different ways of addressing the problems of learning to recognize terrain types from appearance and different algorithms for learning of slip behavior from terrain geometry. The mixing coefficients of the MoE framework can be learned independently and handed down by a terrain texture classifier [21] which can also assign an exclusive winner-take-all terrain class (in this case the MoE would work like a decision tree). They can also be learned in conjunction with the slip behavior per terrain [14]. In the following sections we focus on learning $f(S|T, G)$ as a first important step to realizing the whole MoE framework.

IV. LEARNING SLIP BEHAVIOR ON A FIXED TERRAIN

In this section we learn to predict slip as a function of terrain geometry (called ‘slip behavior’) when the terrain type is fixed. It is an important part in the general slip prediction, as described above. Geometry will be represented by the longitudinal and lateral slopes (see Section V for details). From previous experimental evidence for terrain types like sand [15], we know that slip behavior is a highly nonlinear function of terrain slopes. So, to model the nonlinear dependency $f(S|T, G)$ (i.e. $f(S|G)$) we use receptive field regression [19], [22]. Locally linear functions are learned in small neighborhoods which gives a good tradeoff between memory based nonlinear regression methods [8] and global function approximation methods, such as Neural Networks. We shall note that modeling with local nonlinear regression imposes very little restriction on the functional dependency. It allows for it to be nonlinear but does not assume any particular model, instead, the model is learned from the data. The properties of the receptive field regression approach that we find particularly valuable are: the concept of a receptive field which makes keeping of huge amount of data in memory unnecessary; the adaptability of creating and removing receptive fields as needed; and the possibility to easily extend the approach to online learning. $\hat{S}(\mathbf{x}) = f(S|G = \mathbf{x})$ is estimated as (for convenience, we use \mathbf{x} to denote the two slopes):

$$\hat{S}(\mathbf{x}) = \sum_c^C K(\mathbf{x}, \mathbf{x}_c) (b_0^c + \sum_{i=1}^T b_i^c \langle \mathbf{p}_i^c, \mathbf{x} \rangle),$$

where $K(\mathbf{x}, \mathbf{y}) = \exp(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{\lambda})$ is a smoothing kernel, \mathbf{x}_c is a training example which serves as the receptive field

center, and \mathbf{p}_i^c are several local projections in each receptive field c . In other words, the slip S , corresponding to a query point \mathbf{x} , is computed as a linear combination of C linear functions (one per each receptive field), where the weights are computed according to the distance from \mathbf{x} to the centers of the receptive fields. The centers \mathbf{x}_c could be computed with a clustering type technique (e.g. k-means), or could be allocated wherever needed in the input space as the data arrive in an online fashion [19]. To estimate the parameters b_i^c , \mathbf{p}_i^c , a linear fit is done in each receptive field when the training points are weighted according to their distance to the receptive field center [19], [22]. Weighted linear regression and Partial Least Squares (PLS) methods [8] are used to do the linear fit in [19] and [22], respectively. We use the PLS based linear fit, because it considers the outputs in selecting the most important projections, although, for now, we do not need to do dimensionality reduction because we are currently working with low dimensional inputs. We parameterize the receptive field size by only one parameter λ . More advanced structured kernels could be applied as in [19], but they introduce additional parameters to be learned, which would require a larger sample size, so we use symmetric kernels. We select the parameter λ using a validation set. For example, the best selected λ for our soil dataset (described below) renders a kernel of local activity within about 4° in pitch and roll angles.

V. COLLECTING TRAINING DATA

In this section we describe in detail the information we extract for our training purposes. A 2D map of the environment is built using range information from the stereo pair images. The map has a cell representation with a cell size of 0.2m x 0.2m. The information kept per cell is its extents, average elevation, and association (pointer) to an image (or images) which have observed this cell. This is sufficient to retrieve the required inputs when needed (i.e. when the cell is traversed) and does not overburden the system with keeping huge volume of data per cell, i.e. this is done in a fast and memory efficient way. To collect an example for the training data we do the following: for a particular location in the map (a 6x6 neighborhood of cells centered at the location) which is ‘seen’ by the rover at a distance, we can compute information about appearance and measure the slopes (the input vector); when the rover traverses this location the slip in X, Y, or Yaw is measured (the output value). To be more efficient the data collection goes in the reverse way: at each frame a map is built and in each map cell the average elevation and the pointers to images viewing it are stored, because it is not known which cells are to be traversed. It is only after the rover traverses some region that the particular calculations about slopes and appearance representation are done to be able to use them as training data. To estimate the slope at a particular location, we do a local plane fit to the average elevation in each cell in its neighborhood [7]. The slope is decomposed into a longitudinal (along the forward motion direction) and lateral (perpendicular to the forward motion) component with respect to the current position of the rover. We will call them pitch and roll, since

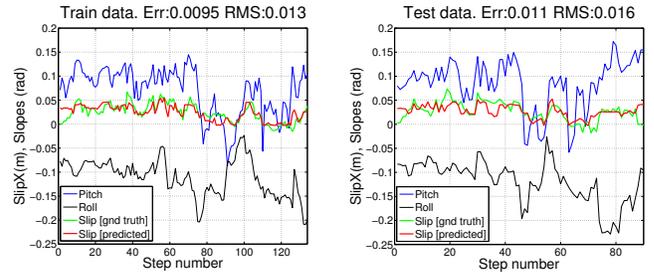


Fig. 4. Prediction of slip in X based on terrain slope angles. Training mode (left), test mode (right). Sandy slopes in the Mojave desert. The step size is $\sim 0.22m$. The test error is 5%-7%. Rocky8.



Fig. 5. Example images from some of the terrains collected by the LAGR vehicle. Sand, soil, gravel.

they correspond to the pitch and roll of the vehicle. Slope angles cannot be perfectly evaluated because of noise in the range data and because the locally planar terrain assumption might be violated. As each location in the map is seen by many frames (as the rover approaches it), we average the roll and pitch estimates to smooth some noise effects. VO is used for the vehicle’s localization and the initial roll and pitch angles of the vehicle (received from a postprocessed IMU) give us initial gravity leveled frame to retrieve correct longitudinal and lateral slope angles from the terrain. Localization is a crucial point in the success of this method.

VI. EXPERIMENTAL RESULTS

In this section we give experimental results of learning and predicting slip from terrain geometry (represented by slope angles). Slope angles are computed using vision information, as described in Section V. This information is immediately available at each point of the local map (wherever there is range data) so slip can be predicted at those locations without the need for the rover to traverse them. Slip prediction error is measured by the average absolute error, $\text{Err} = \sum_{i=1}^n |P_i - T_i| / n$, or by the RMS, $\text{RMS} = \sqrt{\sum_{i=1}^n (P_i - T_i)^2 / n}$, where P_i is the predicted and T_i is the target slip at a particular step i . The latter is more adequate for measuring the error of a regression function but is more prone to outliers and can give an incorrect idea of the error. We report below both training and test error. The training data is used to learn the regression function (Section IV). After learning, the function is tested on the same data (training error) and also on data not used in training (test error). Naturally the training error will be smaller, but the test error is a criterion for the learning method’s *generalization* abilities, i.e. how well it will perform on new, unseen data. We do training and testing pointwise, i.e. not considering potential correlations between consecutive points, which do exist, and can be exploited in a more advanced prediction algorithm.

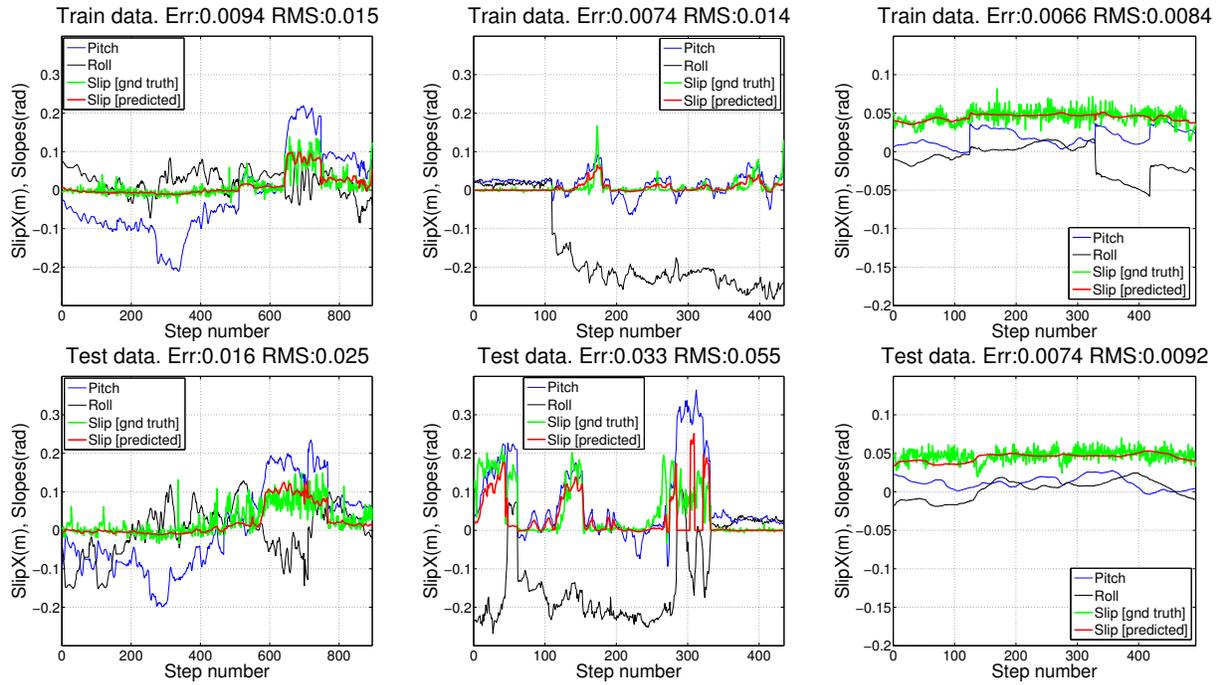


Fig. 6. Prediction of slip in X on soil (left column), gravelly transverse slope (middle column), and flat sandy terrain (right column). Training mode (top), test mode (bottom). The step size is $\sim 0.2\text{m}$ for soil and gravel, and 0.06m for sand. LAGR vehicle.

To be able to measure the test error, we predict slip only on locations traversed by the rover but in principle, as mentioned above, prediction could be done on the whole visible map.

Our first results of learning and prediction of slip behavior are for the Rocky8 rover traversing sandy slopes in the Mojave desert. Figure 2 (left) shows the terrain where the data was collected. The dataset consists of about 200 steps and is taken on slopes which range from -5° to 10° in pitch and up to 12° in roll. Results for learning slip in X are given in Figure 4. Slip is measured in meters per step (a step is $\sim 0.22\text{m}$). As the step size can vary slightly we average the slip across several neighboring steps. The ground truth for this dataset is obtained with a Total Station. Slip prediction captures correctly (with error for the whole data within 5%-7% of the traversed distance per step) a slip of about 20% for high pitch angles. We have to note that this dataset is rather small for learning (200 examples total). We split the data randomly into training and test portions which leaves us with only about 100 examples for training. For all the other experiments in this paper we perform a sequential split of the data into train and test. The latter is a more realistic scenario than the random split because the robot is expected to train on some portion of the terrain first and then continue to traverse the terrain applying what it has learned (testing). It is also more difficult because the distribution of input variables during training might shift to unexplored regions while testing, which makes it much harder to generalize. For example, note that in this dataset there are combinations of roll and pitch angles in the second part of the data (if split consecutively), which have not been seen in the first half to allow us to do a reasonable sequential split.

A second set of experiments was done with the LAGR vehicle in a park. The dataset is composed of the following terrain types: a soil terrain (with small rocks) in which the terrain profile goes up and down, a gravelly terrain, but on a transverse slope, a flat sandy terrain, leafy/woodchip ground, and asphalt terrain (Figure 5). We focused mainly on terrains which are significant for planetary exploration and provided the latter two for comparison only. To do learning, for each terrain type we take the frames up till some time for training, and test on all the frames after that. This is the sequential split of the data, as described above. The first 45% of the data is used for training, the next 10% for validation and the remaining 45% is used for testing. The data is taken by either manually joysticking the rover (soil and gravel datasets) at a speed of about 1 m/s, which can create certain variability in the commanded velocity, or by controlled straight constant velocity drives at 0.3 m/s (all the remaining datasets). The VO estimate is used as ground truth. Slip is measured in meters per step, each step is 0.2 seconds, which gives step sizes of $\sim 0.2\text{m}$ and 0.06m for the above mentioned datasets. The results of slip prediction with the LAGR vehicle on sand, soil and gravel are shown on Figures 6, 7 and 8. Figure 6 shows both training and testing data for slip in X. The soil dataset consists in going up and down a slope twice which helps the testing because similar slope angles have been seen in training. However, this does not happen in the gravel dataset where there is very little variability in the data it was trained on; still, it manages to generalize well. For the gravel dataset we used the vehicle's tilt angles (from the IMU) instead of the ones from the visual information because of localization problems

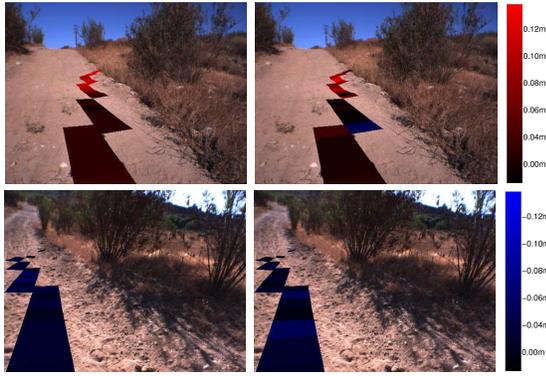


Fig. 7. Predicted slip in X on the future path of the robot (left) and the actually measured slip (right) on soil terrain. Top: robot going upslope, bottom: downslope. Red and blue color coding determine the sign of the slip: red - positive slip (usually occurring upslope), blue - negative slip (downslope). Brighter colors correspond to large amplitude of slip. Black marks zero slip. It has been learned successfully that going downslope on soil terrain there is a small negative slip, whereas going upslope (the same slope) there is much larger positive slip, see also Figure 6 (left). Note the discontinuities in slip ground truth (right column) which are due to random noise or occasional errors in the ground truth. The step size is $\sim 0.2m$.

(due to occasional large rotations between consecutive frames which resulted in incorrect position estimates), but as seen later, with good localization there are no significant differences between the two (Figure 10). Figure 8 shows test data for both slip in Y and Yaw on a transverse gravel terrain. An interesting functional dependence is learned for slip in Yaw: large slip in Yaw, corresponding to large pitch angle, is learned whenever the roll angle is large, but an almost zero slip in Yaw is learned when the roll angle is small, no matter that pitch angle is large. Small but consistent slip in Y is observed whenever the roll angle is large; this dependence could be learned by our algorithm, although it is of the order of a centimeter and is hard to notice on Figure 8. Unfortunately, the LAGR vehicle mobility in deep sand turned out to be extremely poor. On a flat sandy terrain the vehicle experiences a consistent slip of about 80% (see Figure 1, compare to the mobility of Rocky8 on sandy slopes, Figure 4). So it was not possible to collect a dataset on sandy slopes with the LAGR vehicle. The consistent 80% slip in sand forces a constant function to be learned (Figure 6), which is quite natural in this case. Other uninteresting slip behavior from a learning point of view in our dataset is on asphalt and woodchip/leafy ground (Figure 9). Similar to sand, a constant function is learned, because the algorithm could not find any significant correlations with slope angles. On average we get a slip error of about 3%-15% for all the datasets (except for gravel, with 27% MSE, corresponding to 16% absolute error, which is achieved in a hard to generalize learning setup). This is quite a satisfying result in this type of data where a lot of noise is involved (recall that the 1-sigma VO error is 12% per step). In general, our results show very promising prediction of slip in real off-road outdoor environments.

During training we can also use the robot's (postprocessed) IMU to measure the vehicle's tilt angles and compare them

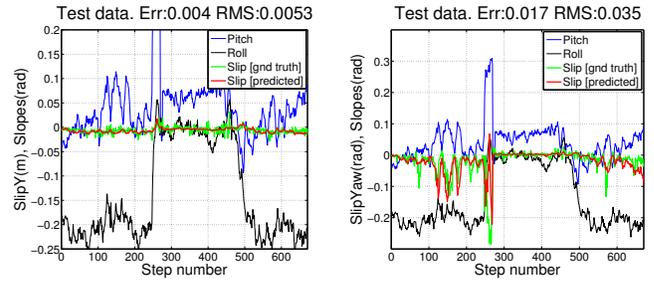


Fig. 8. Predicted slip in Y (left) and in Yaw (right) on a transverse gravelly slope. A small amount of negative slip in Y is learned consistent with the large roll angle, no slip in Y could be detected in any of the other datasets for this vehicle. Test mode. The plots for the training mode are not shown to save space. The step size is $\sim 0.2m$. LAGR vehicle.

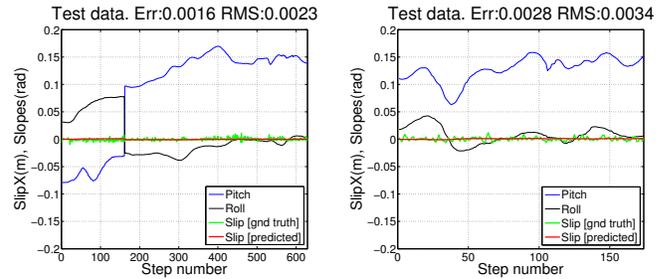


Fig. 9. Slip prediction in X for the trivial cases of asphalt (left) and woodchip/leafy ground (right). The observed slip is close to 0 independent of the slope. Test mode. The plots for the training mode are not shown to save space. The step size is 0.06m. LAGR vehicle.

to the slope angles which were computed from the range data using visual information (Figure 10). Our results show that the estimated slope angles (from vision information) approximate the angles given by the IMU, although we have to be aware that both are, in general, noisy measurements of the actual slope angles: the IMU based measurement gives the tilt of the robot, not of the ground plane, which might be erroneous if the robot traverses a rock, for example; the geometry based slope estimation is susceptible to outliers and can be wrong if there are obstacles in the plane fit area, or a slope estimate can be missing if there are not enough cells under the robot to do a plane fit (this can happen due to missing range data e.g. in sparse vegetation or at the borders of the map). Slope angles estimated from vision data are much smoother (especially for the datasets with smaller speed) because several steps fall in the same cell and because they are averaged measurements. We summarize and compare the slip learning results when learning with respect to the vehicle's tilt angles and with respect to the slope angle estimates which come from visual information only (the latter results have been explored in detail in this section). As seen in Table I, the learning results from both estimates are comparable but the learning from terrain slope angles might potentially benefit from more robust methods for linear fit.

In this section we have shown learning and prediction of slip behavior when the terrain type is fixed. From our results we can conclude that learning of slip is possible and can lead to

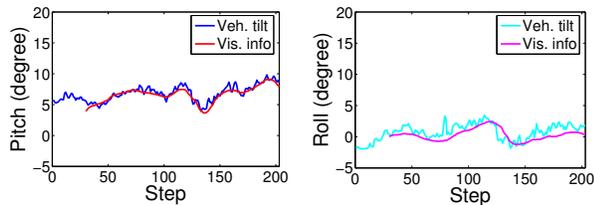


Fig. 10. Estimated slopes from visual information (from stereo imagery), compared to the slopes from the vehicle’s tilt (from the postprocessed IMU), after correction for systematic bias in the vehicle. Some of the vision based estimates might be missing, which happens here at the beginning of the map. Woodchip terrain. LAGR vehicle.

TABLE I

SUMMARY OF THE RESULTS. TEST ERRORS FOR SLIP IN X FROM SLOPES EXTRACTED FROM VISION INFORMATION AND FROM THE VEHICLE’S TILT.

Dataset	Step size(m)	RMS(%) Vision info	RMS(%) Veh. tilt
Soil	~0.2	12.5	12.0
Sand	0.06	15.3	14.8
Gravel	~0.2	N/A	27.8
Asphalt	0.06	3.8	3.6
Woodchip	0.06	5.6	5.8

successful prediction. We want to note that from a mechanical point of view slip behavior of rigid wheels (Rocky8) and of pneumatic wheels (LAGR) is described by different physical models, as well as, slip on off-road terrains, such as soil and sand, is modeled differently from slip on a rigid surface, such as asphalt [23]. Nevertheless, we could learn all those different behaviors from examples because of the flexibility of our learning approach to adjust the learning model according to the data.

VII. CONCLUSIONS AND FUTURE WORK

We have proposed a method to learn to predict slip, which is the result of a complex vehicle-terrain interaction, using vision information only. Our algorithm gives very satisfactory results of slip prediction, given the fact that a lot of noise is involved in measuring it and that the dataset is taken on completely natural off-road terrains. Slip prediction results can be incorporated into a more intelligent path planning algorithm in which regions with potentially high slip would be avoided or more adequate control commands will be issued taking into consideration the expected slip. The importance of this method is that the input information, obtained from stereo imagery, is available at each location in the map (wherever there is range data) so we can potentially predict slip at each seen, but not necessarily traversed, location.

This paper makes a step forward in modeling and learning in complex environments without involvement of detailed mechanical models. We have proposed to map the appearance and geometry of the forthcoming terrain to its mechanical properties, measured when the robot traverses it, and learn this mapping. Our approach is to first recognize terrain type from its appearance and then learn/predict the expected slip

from terrain geometry, the latter being the focus of this paper. Slip prediction from visual information does not undermine the merit of mechanical modeling of terrain, instead, it exploits sensory information which is unavailable or not yet utilized by mechanical models (such as vision) and thus can be complementary to them. An important area of research would be how to combine the two different levels of sensory information.

ACKNOWLEDGMENT

This research is funded by the NASA Mars Technology Program. Thanks also to the JPL LAGR team for giving us access to the LAGR vehicle, to Andrew Howard, Steve Goldberg, Lee Magnone, and Nathan Koenig for helping us with the data collection, to Max Bajracharya for reviewing the paper and providing helpful comments and to Randy Lindemann and Garrett Sohl for discussions.

REFERENCES

- [1] Andrade, C., Ben Amar, F., Bidaud, P., Chatila, R., Modeling robot-soil interaction for planetary rover motion control’, *IROS*, 1998
- [2] Arvidson, R., et al., Localization and physical properties experiments conducted by Opportunity at the Meridiani Planum, *Science*, vol. 306, No. 5685, 2004
- [3] Bauer, R., Leung, W., Barfoot, T., Experimental and simulation results of wheel-soil interaction for planetary rovers’, *IROS* 2005
- [4] Bekker, M., *Introduction to Terrain-vehicle Systems*, Univ. of Michigan Press, 1969
- [5] Borenstein, J., Feng, L., Measurement and correction of systematic odometry errors in mobile robots, *J. of Robotic Systems*, 12(4), 1995
- [6] Brooks, C., Iagnemma, K., Dubowsky, S., Vibration-based terrain analysis for mobile robots, *ICRA*, 2005
- [7] Gennery, D., Traversability analysis and path planning for a planetary rover, *Autonomous Robots*, 6, 131-146, 1999
- [8] Hastie, T., Tibshirani, R., Friedman, J., *The Elements of Statistical Learning*, Springer, 2001
- [9] Helmick, D., Cheng, Y., Roumeliotis, S., Clouse, D., Matthies, L., Path following using visual odometry for a Mars rover in high-slip environments, *IEEE Aerospace Conference*, Big Sky, Montana, 2004
- [10] Helmick, D., Roumeliotis, S., Cheng, Y., Clouse, D., Bajracharya, M., Matthies, L., Slip compensation for a Mars rover, *IROS*, 2005
- [11] Iagnemma, K., Shibly, H., Rzepniewski, A., Dubowsky, S., Planning and control algorithms for enhanced rough-terrain rover mobility, *Int. Symp. on Artificial Intelligence, Robotics, and Automation in Space*, 2001
- [12] Iagnemma, K., Dubowsky, S., Traction control of wheeled robotic vehicles with application to planetary rovers, *Int. Journal of Robotics Research*, Vol. 23, 10, 2004
- [13] Jacobs, R., Jordan, M., Nowlan, S., Hinton, G., Adaptive mixtures of local experts, *Neural Computation*, 3, 79-87, 1991
- [14] Jordan, M., Jacobs, R., Hierarchical mixtures of experts and the EM algorithm, *Neural Computation*, 6, 181-214, 1994
- [15] Lindemann, R., Voorhees, C., Mars Exploration Rover mobility assembly design, test and performance, *Int. Conf. on Systems, Man and Cybernetics*, 2005
- [16] Matthies, L., Schafer, S., Error modeling in stereo navigation, *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 3, June, 1987
- [17] Moore, H., Bickler, D., Crisp, J., Eisen, H., Gensler, J., Haldemann, A., Matijevic, J., Reid, L., Pavlics, F., Soil-like deposits observed by Sojourner, the Pathfinder rover, *J. Geophys. Research*, 104 E4, 1999
- [18] Olson, C., Matthies, L., Shoppers, M., Maimone, M., ‘Stereo ego-motion improvements for robust rover navigation’, *ICRA*, 2001
- [19] Schaal, S., Atkeson, C., Constructive incremental learning from only local information, *Neural Computation*, 10, 8, 2047-2084, 1998
- [20] Terzaghi, K., *Soil Mechanics in Engineering Practice*, John Wiley & Sons, New York, 1948
- [21] Varma, M., Zisserman, A., A statistical approach to texture classification from single images, *Int. Journal of Computer Vision*, Vol. 62, 2005
- [22] Vijayakumar, S., D’Souza, A., Schaal, S., Incremental online learning in high dimensions, *Neural Computation*, 2005
- [23] Wong, J., *Theory of Ground Vehicles*, John Wiley & Sons Inc., 1993