MER-DIMES: A Planetary Landing Application of Computer Vision

CVPR Paper 1206

Abstract

During the Mars Exploration Rovers (MER) landings, the Descent Image Motion Estimation System (DIMES) was used for horizontal velocity estimation. The DIMES algorithm combines measurements from a descent camera, a radar altimeter and an inertial measurement unit. To deal with large changes in scale and orientation between descent images, the algorithm uses altitude and attitude measurements to rectify image data to level ground plane. Feature selection and tracking is employed in the rectified data to compute the horizontal motion between images. Differences of motion estimates are then compared to inertial measurements to verify correct feature tracking. DIMES combines sensor data from multiple sources in a novel way to create a low-cost, robust and computationally efficient velocity estimation solution, and DIMES is the first use of computer vision to control a spacecraft during planetary landing. In this paper, the detailed implementation of the DIMES algorithm and the results from the two landings on Mars are presented.

1 Introduction

During development, the MER Entry, Descent, and Landing (EDL) team determined that wind shear could cause the retro rockets, which are used to slow the lander during final touch down, to fire at an angle and thereby adding a horizontal velocity to the landing system This induced horizontal velocity could drive the platform sideways and down into the surface beyond the safety envelope of the airbag cushions. To deal with this problem, the Transverse Impulse Rockets System (TIRS) was added to reorient the lander system so that the retro rockets would always fire close to vertical.

Another issue surfaced when the EDL team realized that the airbags were also susceptible to wind-induced steady state horizontal velocity. Given a velocity measurement, wind induced horizontal velocity could be taken out by pointing the retro rockets using TIRS. Unfortunately, the lander did not have a sensor for measuring horizontal velocity. Because the spacecraft was already being built, inserting a large and massive traditional Doppler radar velocity sensor into the design was not feasible. At this point, the EDL team realized that, since a rover sun-sensor camera was recently removed from MER design, the MER avionics had an unused camera interface and surplus camera components. This camera hardware could be redirected to development of a downward looking camera which could provide the measurements needed for surface relative velocity estimation with the least impact on the ongoing MER development.

At this point, a group of software, hardware and computer vision engineers was formed to develop this camera based system. Over the following two years, the team designed, implemented, tested and verified DIMES. In the end, DIMES was adopted by the project and used successfully during both spacecraft landings. DIMES determined that horizontal velocity correction was needed for Spirit, but not needed for Opportunity. After-the-fact reconstructions of the landing events showed that without DIMES, the Spirit impact velocity would have been right on the edge of the of the tested airbag performance limits [4].

This paper describes the details of the DIMES algorithm and its implementation. Particular attention is paid to techniques that improved computational efficiency and velocity estimation robustness. The paper finished with a description of the results from the MER-A and MER-B landings.

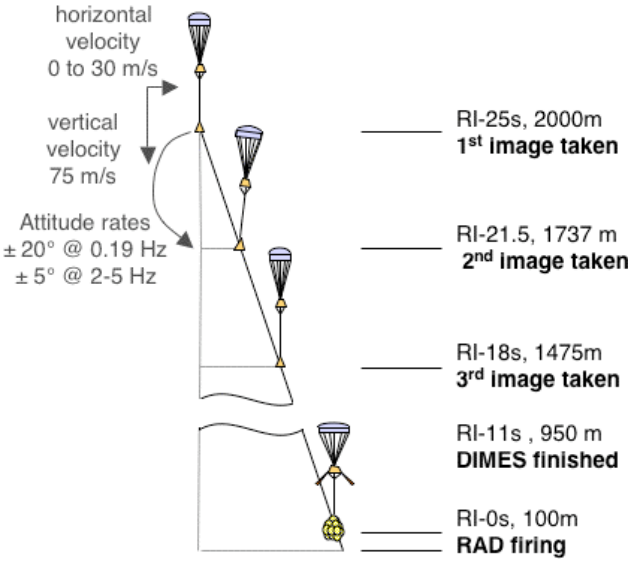

Figure 1 DIMES descent timeline.

2 System Design Constraints

Given the short development time and the novelty of the approach, there was a lot of concern that DIMES could compute an incorrect velocity resulting in mission failure. Consequently, the probability of generating an incorrect velocity estimate was required to be less than 0.1%. To compensate for this strict requirement, DIMES was allowed to not report a velocity 10% of the time. DIMES achieves this level of robustness by tracking multiple features and applying checks on feature correlation and consistency of velocities across image pairs.

The DIMES algorithm was implemented in C under VxWorks on a 20Mhz RAD 6000 flight processor. DIMES was allocated 40% of the processor over 20 seconds which resulted in 160 million instructions total. This severe limitation on available computation drove many of the optimizations in the algorithm.

Since the DIMES camera was created by slightly modifying a camera designed for surface operations, it was not an ideal camera for the highly dynamic descent environment. A fast 5 ms exposure was needed to minimize motion blur, but due to the frame transfer nature of the camera, the images also contained a ramp of useless intensity caused by the open shutter during 5 ms of frame transfer and 5 ms of fast flush. This ramp had to be removed before feature tracking. To ensure enough light made it to the CCD, the 45˚ field-of view MER Navcam optics were used. This field of view was barely wide enough to ensure overlap between images.

The DIMES algorithm takes three images starting at 2000m altitude. During descent, the vertical velocity is around 75 m/s with up to 30 m/s wind induced horizontal velocity. The MER EDL system is essentially a double pendulum composed of a parachute, backshell containing the retro rockets and the lander with the rover inside of it. This double pendulum induced rotation rates up to 60˚/s on top of coning and wrist oscillations. The CCD is 1024x1024, but on chip binning is used to reduce the image to 256x1024. Even with the reduced size, image readout still takes 3.75 seconds resulting in significant motion between images. Figure 1 shows the dynamics and nominal timeline for the DIMES system. The high dynamic, long time between images and limited computing power drove the algorithm design to the use of onboard attitude and altitude measurements for image rectification instead of something more akin to structure from motion.

The DIMES sensors are all located on the lander. The Litton LN200 Inertial Measurement Unit (IMU), containing three gyros and three accelerometers, is inside the rover. The Honeywell RAS radar altimeter and the DIMES camera are attached to the same bracket which is located on the corner of the lander most often pointed at the surface.

The MER landing sites, Gusev Crater and Meridiani Planum, were chosen based on landing safety and science return. These sites are generally smooth and featureless which results in low image contrast (2% to 3% of the image signal) when combined with the camera performance. Low image contrast drove our choice of feature selection and matching algorithms.

3 System Implementation

In the DIMES system, we assume the lander state (attitude and altitude) is correctly measured and that the surface is flat and level. DIMES takes three descent images as input and for each descent image, DIMES also requires elements of the lander state and the sun ray vector (azimuth and elevation) at the time of image exposure. The descent images are scaled and rotated into thesame image coordinate system using the lander state information. Two features are matched for each warped image pair using image correlation. These four matched features are then used to estimate the horizontal velocity, and a consistency check against the delta motion provided by IMU is used to ensure a correct answer.

Under very limited computing power and a very short timeline, the DIMES algorithm was fitted into the flight system using four major strategies

1. Minimize the data volume: It is clear that the slow onboard computer would not be able to accomplish this task in time if each descent image was processed in its entirety. In DIMES's implementation, only two small window of the image, which are just big enough for tracking are used in the actual computation
2. Consolidate processes: To apply 2D correlation between two descent images, radial distortion must be removed and each image must be rectified to the same coordinate system.. By taking advantages of the moderate field of view and very small radial distortion, it is possible to to consolidate these two processes into a single homography based rectification.
3. Simplify the algorithm: Some time consuming computations, such as the Harris operator and image transformation, have been simplified.
4. Use a pyramid scheme: The well known pyramid scheme is not only used for correlation matching but also feature selection. Both implementations yielded a large time saving.

The overall DIMES algorithm is shown in Figure 2.

3.1 Image Binning

The descent camera is programmed to use 4x1 row binning to reduce the time of image readout. Consequently, the images read by DIMES are 256x1024. When DIMES obtains an image, it downsamples it in software from 256x1024 to 256x256 images using 1x4 summing of the 12 bit DN. The purpose of software down sampling is to make square pixels and to reduce computation.

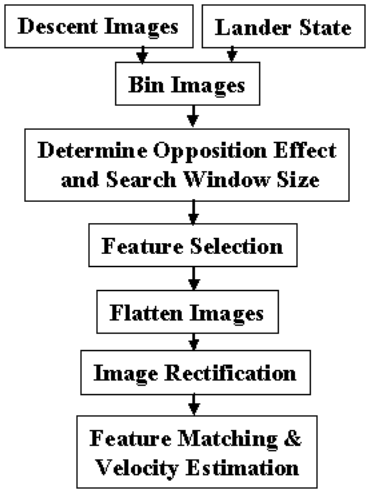

Figure 2 The DIMES system.

3.2 Zero Phase Masking

When viewing the ground from above, there are two non-ideal imaging effects caused by the sun: the shadow of the observer and the zero-phase brightening around the shadow (the opposition effect). The shadow moves with the observer, so it must be avoided so that an incorrect velocity is not reported. The opposition effect applies a non-linear increase to the image intensities near the shadow so it can hinder feature tracking. Fortunately, the location of the zero phase spot and shadow can computed and removed from consideration.

The camera is placed above the surface using the altitude measurement and then the camera is oriented using the surface relative attitude measurement. The sun ray measurement is then projected from the camera and its intersection with the ground is computed using geometry. Finally the ground intersection is projected into the camera to determine the pixel location of the shadow and surrounding opposition effect. A user defined parameter, based on the worst case attitude error and brightening effect, is then used to mask out a circular region around the shadow. This procedure is repeated for each image in the pair.

3.3 Setting the Search Window Size

The correlation window size is determined by considering five factors: spacecraft altitude (H), maximum attitude measurement error (e), maximum horizontal velocity (v_m), time interval between the two descent images (dt), and the template size. Assume the nadir point on the ground for the first image is (0, 0, $Hcos(e)$). If the lander is moving in the y direction, then the nadir point will move to (0, $v_m dt$, $Hcos(e)$). The pixel distance between these two points when projected into the first image plus half of the template size determines the half width of the search window. the search window size and the opposition effect masks are used during feature selection.

3.4 Feature Selection

The features selected for matching must satisfy four criteria (Figure 4)

1. It must be inside the image overlap.
2. It must be away from the opposition effects of both images.
3. It must have enough intensity information, so it can be tracked
4. The two features selected in each image pair must be far apart to ensure independent estimation.

To meet these criteria, the Harris interest operator is applied to the first image inside the overlap region between the images and outside the opposition effect masks. To reduce computation, feature selection is applied on a coarse grid (64x64) in the first image. Overlap is computed as follows. Assume the lander states (attitudes and heights) at the two images are ($\boldsymbol{R}_1$, $\boldsymbol{H}_1$) and ($\boldsymbol{R}_2$, $\boldsymbol{H}_2$). Because feature selection does not require high geometric accuracy radial lens distortion can be ignored. Consequently, if the surface is flat and level, the relation between the two images can be approximated by a homography transform computed from ($\boldsymbol{R}_1$, $\boldsymbol{H}_1$) and ($\boldsymbol{R}_2$, $\boldsymbol{H}_2$) [6][7]:

$$x_1 = \frac{a_1 x_2 + a_2 y_2 + a_3}{a_7 x_2 + a_8 y_2 + 1} \quad Y_1 = \frac{a_4 x_2 + a_5 y_2 + a_6}{a_7 x_2 + a_8 y_2 + 1}$$

A pixel in the first image (X_1, Y_1) is projected into the second image (X_2, Y_2) using this homography and if it is inside the second image bounds (shrunk by the search window size) it is within the overlap. The angle between this pixel and the sun ray in the first image is computed and if it is greater than the zero-phase mask angle it is kept. The inverse homography is used to project the sun ray for the second image into the first image and the zero phase check is repeated. Finally, if the pixel is in the overlap and outsize the zero phase spots, the Harris interest operator is applied.

The pixel with highest interest value is selected as the first feature. The second feature will be the one with highest

interest value outside the search window of the first feature.

To refine the selection made on the coarse grid, the Harris interest operator is applied to a 9 by 9 region centered at each selected feature. The final feature location is set to the pixel with the highest interest value in the region.

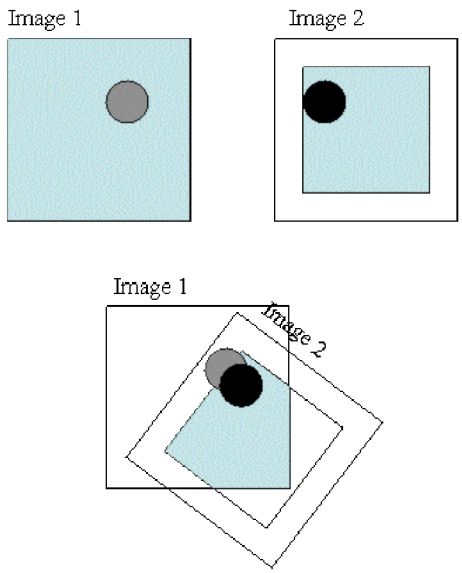

Figure 3 Features have to be selected inside the overlap between two descent images. The dark circles indicate the opposition effects masks. The shaded areas in the top two images are the potential area for feature selection. After overlapping two images, only a small area is left for feature selection (blue area in bottom figure).

3.5 Image Flattening

Before rectification, the templates and window intensities are corrected for image to image variations so that image correlation has the best chance for success in the low contrast Mars environment.

The DIMES camera does not have a shutter, so the CCD is collecting charge during the transfer of the frame off of the CCD after exposure and the flushing of the charge off of the CCD before exposure. Because the exposure time is comparable to the frame transfer and fast flush times, a ramp of intensity is visible in the descent images. However, using the assumption that the camera is not moving during exposure, this intensity ramp can be removed using the image data alone and the formula

$$P'[r][c] = f * P[r][c] + (1 - f) * P'[r-1][c]$$

where P is the pixel intensity, $f = f_t/e_t/r$, f_t is frame transfer plus fast flush time (5ms + 5ms = 10ms), e_t is the exposure time (5 ms), and r is number of rows of the image (256). In order to reduce the computation, only a strip of image, which covers the search window or template is processed. (Figure 4).

After frame transfer removal, radiometric correction is applied to reduce the intensity differences between the images. Radiometric correction reduces the effect of dark current, optical transfer (vignetting) and CCD pixel-to-pixel variation. Radiometric correction cannot eliminate photometric differences between the images (e.g., zero phase angle brightening, lander parachute shadow).

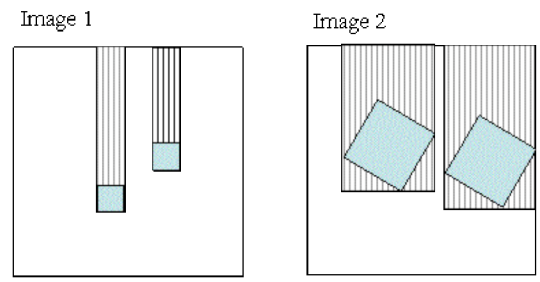

Figure 4 The frame transfer smear removal is only applied to a strip that covers the templates or search windows.

Radiometric correction applies a scale and offset to each pixel. To reduce non-volatile memory requirements, the scale and offset images are represented as biquadric polynomials whose coefficients are stored instead of entire images. Prior to DIMES image acquisition, the scale and offset images are precomputed from the stored coefficients. Since the order of the polynomials is low, they cannot account for pixel to pixel CCD variation. However, testing with representative CCD pixel to pixel variation has shown that not accounting for it with the scale and offset images does not affect DIMES performance. Furthermore, algorithm testing has revealed that the offset images produced by radiometric calibration of the flight cameras have little effect on DIMES results: the velocity error and number of valid solutions are very similar regardless of whether the offset image is used. Consequently, the offset image was not computed or used during the MER landings (turned off by parameter).

3.6 Image Rectification

The templates and corresponding search windows for the two features are rectified to a camera frame parallel to the ground and at the altitude when the first image was taken (Figure 5). Again, a local homography is used for this procedure. The local homography serves two purposes. First, it provides a simple relationship between the rectified plane and the image plane. Second, it compensates for the local radial distortion. Because the DIMES camera lens is close to a perspective pinhole lens, a homography transform is sufficient for approximating the local radial distortion.

The local homography transform is obtained by a numerical solution. The pixel coordinates of four corners of the template or search window between the descent image and level image frame are explicitly calculated through camera models then these four points are used to determine the 8 coefficients of the homography transform.

In order to reduce computation, the homography transform is decomposed into three components.

$$x_1 = \frac{A}{C} \qquad y_1 = \frac{B}{C}$$

$$\text{where} \quad \begin{aligned} A &= a_1 x + a_2 y + a_3 \\ B &= a_4 x + a_5 y + a_6 \\ C &= a_7 x + a_8 y + 1 \end{aligned}$$

If the current pixel's transforms has been calculated, then, the right adjacent pixel's transform can be reduced to three additions and two divisions as.

$$\begin{aligned} A_{i+1} &= A_i + a_1 \\ B_{i+1} &= B_i + a_4 , \\ C_{i+1} &= C_i + a_7 \end{aligned} \qquad \begin{aligned} x_{1i+1} &= \frac{A_{i+1}}{C_{i+1}} \\ y_{1i+1} &= \frac{B_{i+1}}{C_{i+1}} \end{aligned}$$

The pixel's intensity is then calculated using bilinear interpolation.

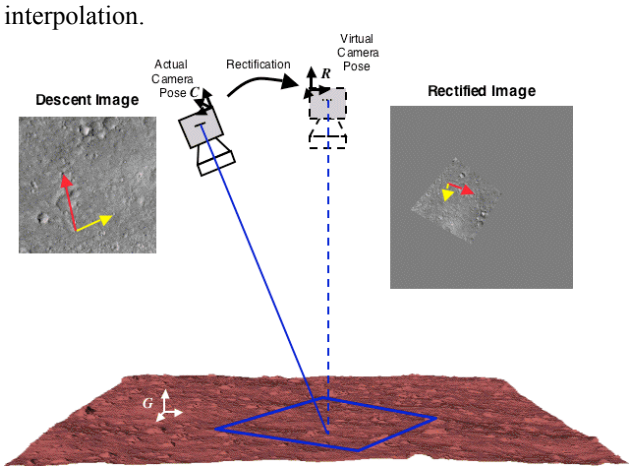

Figure 5 Image rectification concept

3.7 Velocity Estimation

The result of rectification is a pair of templates from the first image and their correspondent search windows from the second image mapped into a single view of the surface. The only difference between templates and windows is a horizontal shift induced by the unknown horizontal motion. A classical Pseudo-normalized correlation [3][4] is used to determine this horizontal motion. The pixel with highest correlation is template's most likely location in the window. A two level pyramid search scheme is used here to speed up the process.

In the final level, a biquadratic interpolation is applied to obtain subpixel matches. The biquadrate form is also used for calculation of peak width and peak ratio, which are used for match validation.

To determine the correlation peak width, the biquadratic used for subpixel interpolation is analyzed. Fitting a biquadratic to a 3x3 neighborhood of correlation scores yields a biquadratic surface

$$C = ax^2 + by^2 + cxy + dx + ey + f$$

The subpixel maximum of the fitted surface is then found by partial differential, which leads to the following solution for the peak pixel (x_p, y_p)

$$x_p = (-2bd + ce)/(4ab - c^2)$$
$$y_p = (-2ae + cd)/(4ab - c^2)$$

Moving the biquadratic form to coordinates around the peak yields

$$C = ax'^2 + cx'y' + by'^2 + f'$$

Let $x' = r\cos t, y' = r\sin t$, then

$$C = ar^2\cos^2 t + cr^2\cos t\sin t + br^2\sin^2 t + f'$$

Given a constant r, the curves maxima is at

$$t_0 = \tfrac{1}{2}\arctan\left(\frac{c}{b-a}\right)$$

Therefore, the maximum peak width can be calculated as

$$w = [(c - f')/(a\cos^2 t_0 + c\cos t_0 \sin t_0 + b\sin^2 t_0)]^{1/2}$$

As described below, this width is used to find incorrect correlations.

After image correlation the center of the template and its peak correlation pixel are projected onto the ground. The horizontal velocity is then

$$v_x = (X_2 - X_1)/dt$$
$$v_y = (Y_2 - Y_1)/dt$$

where (X_1, Y_1) and (X_2, Y_2) are the template center and peak ground points, respectively, and dt is the time interval between them.

3.8 Match Verification

To ensure a correct match, multiple checks are placed on the intermediate and final results. The Harris interest operator value for a feature is the template contrast metric. A low contrast metric indicates little or no contrast and consequently a template that is susceptible to image noise. The correlation peak value is another metric that indicates how similar the template is to its best match location in the search window.

Two new metrics are also implemented. The first one is the correlation peak width, which is a very good indicator of the local frequency content of the window. When low frequency dominates in a template or along certain direction, such as along an edge, the peak width will be large and the match is then not reliable. The correlation peak ratio is the ratio of the highest correlation peak to the second highest correlation peak. This metric is very useful for detecting repetitive terrain or multiple features of similar appearance, such as craters. 0depicts these correlation metrics.

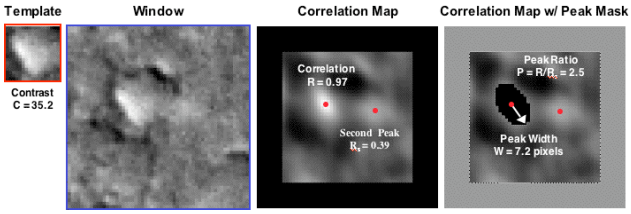

Figure 6 Image correlation metrics.

In addition, the algorithm will not produce any result when it encounters any one of three additional anomalies. The first anomaly is that the matched point is on the border of the window, where the subpixel interpolation is not feasible. The second anomaly is that the peak is too flat, where $(4ab - c^2)$ is close to zero. Finally, if the subpixel correction is greater than 1.5 pixels, the matched result will be rejected.

The DIMES algorithm uses two set of images to determine the horizontal velocity. A total of three images are used with images 1 and 2 for the first pair and 2 and 3 for the second pair. Since, two features are selected in each pair, up to four independent velocity estimates can be generated. For each pair of images, as least one estimate must be obtained. Otherwise, DIMES will not provide a velocity answer. The delta velocity between the two pair can be independently checked by the delta velocity provided by onboard IMU. If at least one delta velocity is consistent with the IMU delta velocity, DIMES is deemed to have measured a correct surface relative velocity, and then it will pass an IMU velocity correction to the control system.

4 Flight Results

DIMES was enabled for both landings. Each time, DIMES computed a valid velocity that was used by the EDL system.

On June 4th, 2004, the MER-A or Spirit spacecraft landed in Gusev Crater, a 160 km diameter crater just south of the Martian Equator. During its decent it took three images of the Martian surface at altitudes of 1983m, 1706m, and 1433m . These images showed a terrain covered with craters, hollows and dust devil streaks. During EDL, DIMES computed a velocity of (4.1, 9.7) m/s at 1570m altitude (all velocities are in the local level (North,East) coordinates.) After propagation down to RAD fire, this steady state velocity was (–1.2,10.7) m/s. Due to a wind gust just before RAD fire, the angle between the lander and backshell was significant at the time of RAD fire, and would have caused an additional horizontal velocity of (-5.7,11.7) m/s making the without -TIRS total velocity (-6.9,22.5) m/s. This velocity was significant enough to cause TIRS to fire in its stronger of two modes resulting in a reduced bridle cut velocity of (0.0,11.0) m/s. The total processing time was 13.75 seconds which was well within the processing margin.

The rad-induced velocity which was less than the threshold to fire TIRS. Had DIMES not been available to measure the steady state velocity, the EDL system would not have fired TIRS and the total velocity would have been just on the threshold of airbag performance. Furthermore, the velocity would have been to the East toward the rockier terrain surrounding the so-called Bonneville Crater.

The DIMES software tracked all of the features correctly, but through out the second feature of the second image pair because the peak width parameter for this feature was larger than the maximum allowable peak width parameter (Table 1). Post flight analysis showed that DIMES did compute the correct velocity.

Figure 7 shows the images taken by the DIMES camera and the associated features that were tracked.

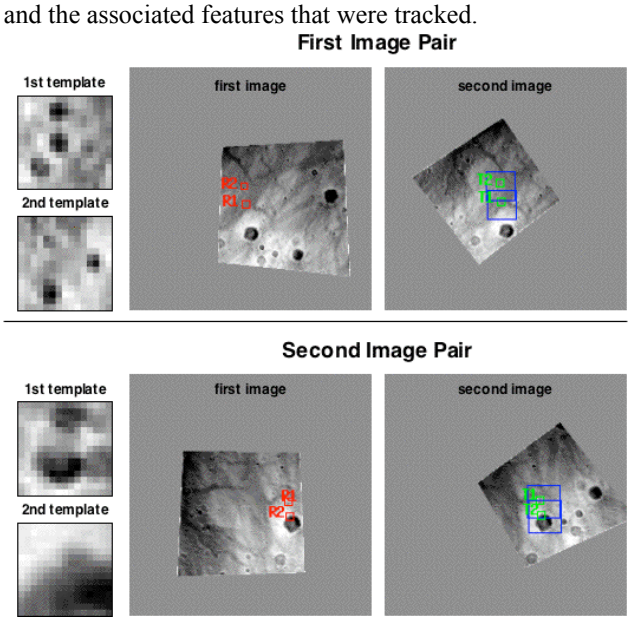

Figure 7 MER-A DIMES results

On June 25th, 2004, the MER-B or Opportunity spacecraft landed on Meridiani Planum, halfway around Mars from Spirit and just north of the Martian Equator. During its decent it took three images of the Martian surface at altitudes of 1986m, 1690m and 1404m. These images

showed a very bland terrain spotted with a few craters. During EDL DIMES computed a mostly northern velocity of (8.0,-0.3) m/s at 1547m altitude. After propagation down to RAD fire, this steady state velocity was still (10.4,-2.8) m/s but in this case the rad induced velocity was (-7.7,3.9) m/s which would effectively negate the effect of the steady state wind. Consequently, TIRS did not fire during the Opportunity lander. The bridle cut velocity was (9.0,-2) m/s, so the lander bounced north ultimately ended up in the so called Eagle crater. The total processing time was 13.875 seconds.

Table 1 MER-A feature outcomes

	Pair 1		Pair 2	
	1st	2nd	1st	2nd
Valid	Y	Y	Y	N
Velocity (m/s)	4.2,10.4	4.1,10.6	4.1,9.7	N/A
Brightness (dn)	4710	3898	4310	4694
Contrast (dn/pix)	47	42	54	56
Correlation	0.88	0.91	0.97	0.997
Peak ratio	1.3	1.5	1.8	1.1
Peak width (pix)	2.4	3.3	4.3	39

Table 2 MER-B feature outcomes

	Pair 1		Pair 2	
	1st	2nd	1st	2nd
Valid	Y	Y	Y	N
Velocity (m/s)	6.3,1.2	6.2,1.1	8.0,-0.3	N/A
Brightness (dn)	3242	3819	3697	4403
Contrast (dn/pix)	31	16	18	17
Correlation	0.82	0.81	0.76	0.68
Peak ratio	1.5	1.1	1.4	2.0
Peak width (pix)	2.1	4.4	3.3	2.8

The DIMES software tracked all of the features correctly, but threw out the second feature of the second image pair because the correlation for this feature was smaller than the minimum allowable correlation parameter (Table 2). Post flight analysis showed that DIMES did compute the correct velocity.

Figure 8 shows the images taken by the DIMES camera and the associated features that were tracked for MER-B.

5 Conclusion

DIMES performed successfully during both rover landings, and post mission reconstruction results shown that DIMES has fully met its mission requirements. DIMES is the first sucessful computer vision system used during planetary landing. Because it is an efficent, low cost and robust velocity estimation approach, DIMES is applicable to future landing mission including Mars Science Laboratory (2009), Phoenix (2007), Mars Sample return or Lunar Sample return. Since a much faster onboard processor is expected to be available for future missions, the performance of DIMES can be inproved by using higher resolution images, adding more features to compensate for low altitude terrain relief, and/or augmenting the IMU for better velocity estimation.

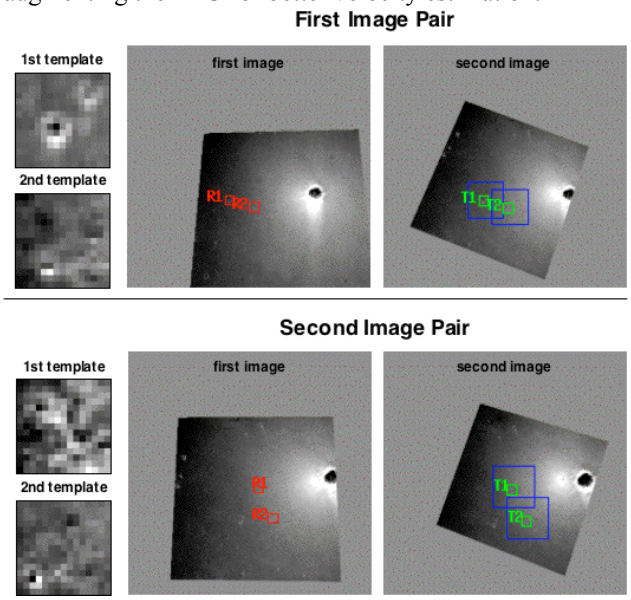

Figure 8 MER-B DIMES results.

References

[1] Anonymous.

[2] C. Harris and M. Stevens, "A Combined Corder and Edge Detector,:, Proc. 4th Alvey Vision Conf. 1988, pp. 147-151.

[3] Anonymous.

[4] Anonymous.

[5] Anonymous.

[6] R. Tsai, and T. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch", IEEE ASSP Vol. 29, No 6, 1981.

[7] Tsai, R., T. Huang, and W. Zhu, "Estimating three-dimensional motion parameters of a rigid planner patch, II: Singular Value Decomposition", IEEE ASSP, Vol. 30, No 4, 1982.