# Autonomous Onboard Surface Feature Detection for Flyby Missions

Thomas J. Fuchs[1], Brian D. Bue[1], Julie Castillo-Rogez[1], Steve A. Chien[1], Kiri Wagstaff[1], David R. Thompson[1]

(1) Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Dr. Pasadena, CA 91109, USA
e-mail: thomas.fuchs, bbue, julie.c.castillo, kiri.l.wagstaff, steve.a.chien, david.r.thompson
@jpl.nasa.gov

## Abstract

To date, most small bodies exploration has involved short timescale flybys that execute pre-scripted data collection sequences. Light time delay means that the spacecraft must operate completely autonomously without direct control from the ground. But in most cases the physical properties and morphologies of prospective targets are unknown before the flyby. Features of interest are highly localized, and successful observations are highly dependent on geometry and illumination constraints. Under these circumstances onboard computer vision can improve science yield by responding immediately to collected imagery, for example by targeting features of opportunity for additional data collection by specialized instruments with a narrow field of view. Among the most challenging targets are those involving specific terrain morphology or differences in surface brightness. Differential illumination and shadowing makes such features extremely difficult to find reliably. Consequently they are a difficult worst-case test for autonomous image analysis. This work presents a framework for the detection and localization of such targets, focusing on the specific problem of high albedo surface features. We evaluate performance using a case study with archival datasets from previous primitive bodies encounters.

## 1 Introduction

Small bodies - the asteroids, comets, and other primitive objects in the solar system - are highly valuable targets for scientific exploration. These objects have undergone little modification since their formation, so they uniquely reveal the processes that shaped our Solar system in the early nebula and as a consequence of large-scale dynamical events. They tell us about our own solar system's history, and inform our interpretation of exoplanet systems and their potential for life. Those objects that we have visited exhibit striking diversity, suggesting that we have only just begun to characterize their populations. Some of these primitive bodies (Near Earth Objects, Phobos and Deimos) have also been identified as possible targets for the extension of Humanity in space and are the focus of reconnaissance missions. All this has given them a high priority in future mission plans by NASA, ESA, and other space agencies.

These targets have difficulties commensurate with their great value. Many lie in remote and challenging orbits that can only be reached by expensive high delta-V maneuvers. Consequently, most encounters of small bodies have to date been flybys that provide just a few minutes or hours to collect data during closest approach. A few rare extended encounters, such as the *Dawn* encounters at Vesta and Ceres and the *Rosetta* mission at 67P/ChuryumovGerasimenko, allow multiple command cycles and opportunities for repeat imaging. But for the most part primitive bodies data collection has occurred autonomously, with light time delays of hours and no direct supervision from the ground. The quality and quantity of information obtained from these pre-scripted flyby sequences so far has been limited. As added difficulty, features of scientific significance to be sought at these objects are faint as their true nature is concealed by space weathering and regolith processes. This makes it difficult to capture the diversity of these bodies in the course of flybys. Nevertheless, flybys provide a low-cost means to characterize many bodies in a single mission and therefore develop a critical population-level understanding of these objects. They figure prominently in many proposed missions such as a Trojan Tour and Rendezvous (TTR) [1] and Main Belt Asteroid Flyby missions [6].

This paper discusses ways that autonomy and onboard intelligence can benefit flyby science. Specifically we demonstrate autonomous onboard data analysis and response to help close the gap in science value of flybys *vis a vis* extended encounters. Missions can command the spacecraft to adapt targeting decisions in real time, migrating rudimentary decisions across the light time gap for more responsive data collection.

Many of these capabilities can be formulated as computer vision tasks. For example, a primary consideration is detection and tracking of the small body itself. This can

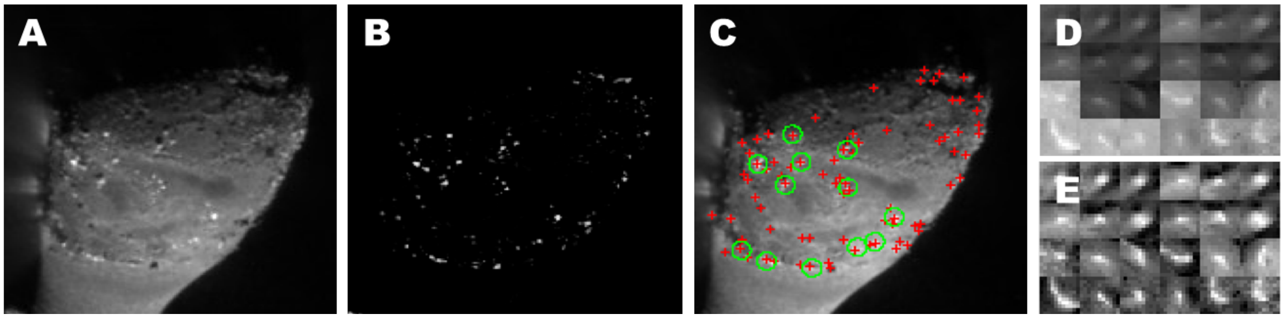


**Figure 1.** Preprocessing pipeline for training a classifier for surface feature detection. **A:** Comet 103P/Hartley on a detail of image mv10110413-5004046 taken by the framing camera of the Deep Impact probe during the EPOXI mission. **B:** Renormalized difference of grayscale and median filtered image. **C:** Set of possible surface features as a result of weighted mean shift clustering (red crosses) and ground truth labels from a domain expert (green circles). **D:** Patches of labeled surface features from 9P/Tempel. **E:** Locally normalized patches, constituting the positive class for training a random forest classifier.

be achieved either by use of a center-of-brightness algorithm [17], by taking connected components into account [19]. Similar strategies were used in the flyby encounters of Deep Space 1 at Borrelly and the Stardust flyby of Annefrank [4]. Plumes or outgassing are also of interest, and a spacecraft that finds them could target them for followup measurements [27]. Plumes have been observed at Jovian and Saturnian moons, and outgassing is also expected at cometary bodies, activated asteroids, and water-rich asteroids [23]. Outgassing is a transient and somewhat unpredictable event and can be faint for main belt asteroids [13]. In addition to offering a compelling science target of opportunity, it is also important that surface feature detectors are invariant against such large-scale changes in brightness. These challenges are aggravated in multinary asteroid systems which occur frequently [16]. One can prevent plumes from confusing subsequent image analysis by intentionally making the later analyses robust to diffuse plumes [26] or by applying automated plume detection as a filter [7, 15, 28]. A third class of targets is made up of spectral outliers indicating distinctive composition. Recent tests have demonstrated autonomous detection of such targets using the Hyperion imaging spectrometer aboard EO-1 [25].

While there has been considerable prior work on autonomy in these areas, there are still compelling phenomena for which no detection methodology currently exists. Among the most challenging are those involving specific terrain morphology or differences in surface brightness. These are scientifically valuable because they can indicate fresh unweathered material with diagnostic mineralogical signatures. On comets, they can point to areas with high concentrations of volatiles. Differential illumination and shadowing makes such features extremely difficult to find reliably. Consequently they are a difficult worst-case test for autonomous image analysis. This work will focus in depth on this challenge, demonstrating a method for detection of high albedo surface features. A series of experiments on archival flyby sequences demonstrates techniques that can reliably achieve this objective. Section 2 describes in detail the proposed surface feature detection framework and Section 3 discusses the results from a performance evaluation on flyby imagery from comets Hartley 2 and Tempel 1.

## 2 Methods

We formulate the surface feature detection task as a classification problem as summarized in Figure 1. Our approach is based on simple image filtering operations and statistical classification. This exploits the strength of statistical object recognition while permitting very fast execution times.

### 2.1 Preprocessing

The aim of preprocessing is to mitigate illumination variations and emphasize high albedo surface features. To this end we apply a cross median filter on the current frame and subtract the result from the original gray-scale image. After renormalizing the difference image highlights high albedo regions independent from local image exposure. As a result, surface features in dark areas and well lit areas give similar responses. In addition, median filters have the desired properties of linear run-time [18] and edge preservation hence suppressing the bright edges at the borders of small bodies. Figure 1B shows the renormalized difference image for a frame from Hartley 2 depicted in 1A. High albedo features are clearly discernible in the illuminated part of the comet as well as in shadow.

### 2.2 Candidate Detection

The next step is the actual detection and localization of surface feature candidates. We use intensity weighted mean shift clustering [8] on the difference image for mode detection. Specifically we use a circular box kernel weighted by the greyscale intensity of the pixels in the difference images. Due to the discrete nature of raster images this step can result in multiple adjacent detections. These clusters are reduced to single locations by running mean shift again but this time on the binary image of detections from the first round. The resulting detections represent the locations of candidate surface features. While this detection procedure covers nearly all true positive samples it yields a large number of false positive detection, which have to be filtered out by a classification algorithms. Figure 1C shows a large number of surface feature candidates depicted as red crosses on a frame from Hartley 2. The ground truth annotations from a domain expert are shown as green circles.

### 2.3 Training Procedure

In order to train a machine learning algorithm to differentiate between true detections and false positives we need a set of positive and negative examples. To this end a planetary scientist annotated all frames from close encounters to the comets Hartley 2 and Tempel 1 by labeling surface features of interest. Figure 2 presents a sequence of 12 frames from Hartley 2 including the expert's annotations shown as red circles. Detected candidates within a predefined distance of 10 pixel to ground truth annotations are labeled as positive while the other detections are labeled as negative. This set of positive and negative training samples can then be used to learn and validate a classifier. Figure 1D depicts image patches of positive examples from comet Hartley 2. We also force the classifier to be rotation invariant by augmenting the training set with rotated and flipped copies of patches from the positive class [12].

### 2.4 Classification

The training set as described in the previous section can be used in a machine learning framework to construct a classifier which is able to discern true surface features from false detections. First we extract images patches of size $11 \times 11$ pixels at the locations of surface features candidate detections from the original greyscale frames as shown in Figure 1D. To achieve robust results each patch is normalized by shifting and scaling the intensity values to the range of $[0, 1]$ (cf. 1E).

The image patches are described by a set of numerical attributes which are subsequently used for classification. These attributes comprise the raw intensity values,

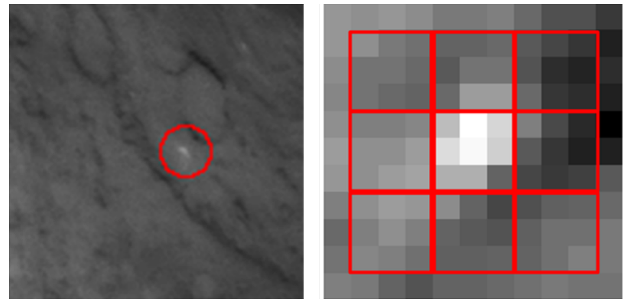

**Figure 3. Left:** Detail of a surface feature of interest on comet 9P/Tempel. **Right:** Enlarged image patch of a surface feature candidate. Overlayed is the grid used for spatial binning of intensity values during attribute extraction.

general image statistics like mean, median and standard deviation as well as the local gray value and gradient histograms. To this end the patches are partitioned spatially as shown in Figure 3 and then the greyscale intensities are histogrammed per bin and for the whole patch. Finally, an attribute vector is constructed containing the raw pixel intensities, the image statistics and the local and global histograms.

Based on the extracted attributes we train a random forest classifier [2, 5] to differentiate actual surface features from false positive candidates. In recent years random forests or decision forests have been extended for clustering [5], online learning [20], interactive learning [10] and density estimation [9] to mention just a few. The applications range from medical imaging [12, 11] over gaming [22] to space exploration [24]. Random forests have a number of properties which make them a suitable choice for autonomous computer vision during flyby missions: (i) They can infer non-linear interactions between attributes and hence are able to construct the complex model necessary for high accuracy in computer vision. (ii) Random forests implicitly perform attribute selection and thus can deal with a large number of attributes while being robust against noisy or non-informative variables. (iii) The ensemble structure favors parallel training of the decision trees in a distributed manner which allows handling of large amounts of training data in a reasonable time frame. When the proposed system is used in several missions the amount of image data demands the use of parallel distributed learning. (iv) Random forests can not only be learned but also tested in parallel which results in fast execution speed. Besides GPU implementations [21], FPGA implementations are already available for space exploration [3] and hence make random forests an ideal choice for onboard computer vision.

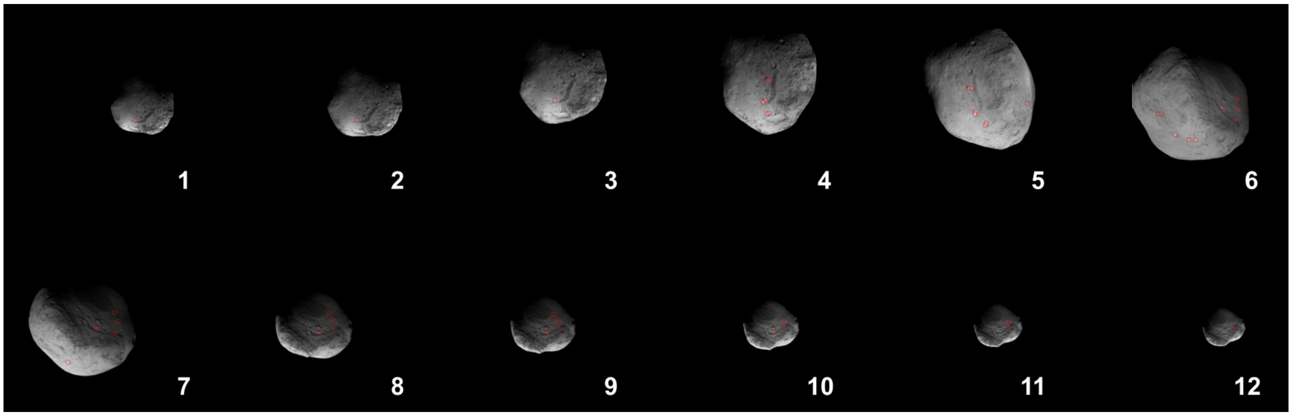


**Figure 2.** Sequence of 12 frames of the closest encounter of deep impact with comet 9P/Tempel. The annotations of the planetary scientist are shown as red circles and represent the surface features of interest. The aim of our system is to detect these features autonomously and reliably with a low false positive rate.

Specifically we train a random forest with 100 trees by bootstrapping the training data for every tree and optimizing over 10 randomly chosen attributes an their threshold at every split node. The trees are grown until completion without pruning and the final prediction is achieved my taking the majority vote of over all trees in the ensemble. The ratio of trees voting for a true surface features versus the ones voting for a false positive can be interpreted as the confidence of the classifier in its overall prediction. We use this confidence estimate to generate the precision/recall plots shown in Section 3 for estimating the generalization power of the classifier.

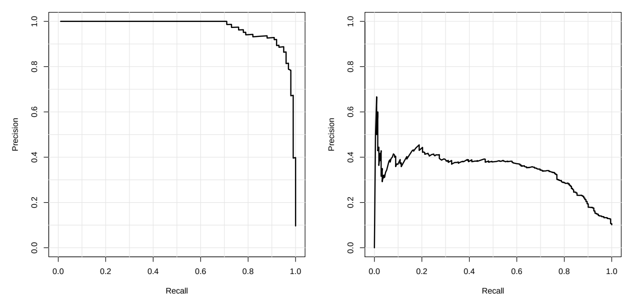


**Figure 4.** Precision and recall curves based on varying classifier confidence. **Left:** Training performance on 103P/Hartley. **Right:** Test performance on 9P/Tempel using the model trained on 103P/Hartley.

## 3 Results & Discussion

### 3.1 Training and Testing Data

The proposed framework is evaluated by conducting cross-validation experiments on public archival data of small bodies. Sequences of navcam images were acquired at regular intervals during the encounters [14]. A domain expert labeled all surface features of interest as described in Section 2.3. In total this amounts to 47 frames from the encounter of deep impact with 103P/Hartley and 72 frames from 9P/Tempel.

### 3.2 Evaluation Procedure

The textbook evaluation procedure for such scenarios in machine learning would be to perform cross-validation on sample level, i.e. training on a subset of surface features from all images and testing on the hold-out set. In our setting this approach leads to nearly perfect classification accuracy and hence vastly overstates the predictive power of such a model. This is mainly due to the fact, that the samples (surface feature candidates) are not i.i.d. but highly correlated. Specifically surface features on one small body look very much alike but can be dramatically different from other small bodies. Similarly, performing cross-validation per frame or using the out-off-bag (OOB) error estimate from the random forest classifier [5] leads to excellent performance, overestimating the power of the model to generalize to a new target body.

To overcome this problem we resorted to a very strict validation scheme by doing leave-one-out cross-validation (LOOCV) on a per body level. In particular we train on all samples from one comet and test the performance on the samples of a complete different, unseen body. This procedure also resembles more closely the scenario we are going to encounter on board in actual flyby missions. We can train a model on all previously seen and labeled as-

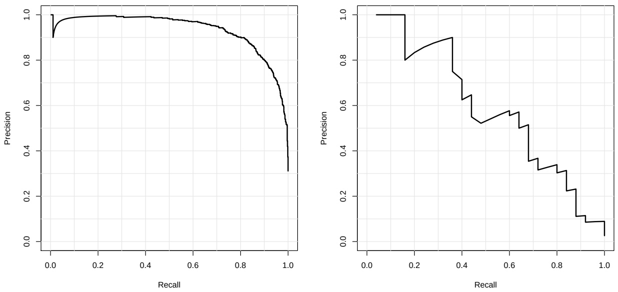


**Figure 5.** Precision and recall curves based on varying classifier confidence. **Left:** Training performance on 9P/Tempel. **Right:** Test performance on 103P/Hartley using the model trained on 9P/Tempel.

teroids and comets but we will not have any knowledge about a small body which we never encounter before. For missions with multiple encounters to the same object one could envision updating the classification model with data from previous flybys. This approach would significantly boost classification accuracy, but is not the focus of this work.

### 3.3 Performance Metrics

Overall detection performance of whole flyby sequences is reported in terms of precision and recall as depicted in Figures 4 and 5. The model is trained on all samples from either Hartley 2 or Tempel 1 and then the training error is calculated on the same data while the test error is estimated on the samples from the unseen small body. We vary the threshold on the confidence estimate of the classifier as described in Section 2.4 to generate precision/recall plots as shown Figures 4-5. Precision is the fraction of detected samples that are true surface features labeled by a planetary scientist, while recall is the fraction of true surface features that are detected. Specifically precision is defined as $TP/(TP + FP)$ and recall as $TP/(TP + FN)$, $TP$ referring to true positive detections, $FP$ to false positive and $FN$ to false negative detection.

### 3.4 Discussion

The performance on the training set presented in Figures 4 (left) and 5 (left) demonstrate that the numerical attributes are expressive enough to accurately model surface features and to train a classifier to differentiate true features from false positive. This is contrasted by a rather poor performance on the test set of samples from previously unseen objects shown in Figures 4 (right) and 5 (right).

In practice we are interested in the high precision regime of the performance curve. During a flyby mission the proposed framework can be used to point a specialised instrument with a narrow field of view at a surface feature of interest. In flyby scenarios the autonomous system will only have time to point once at an interesting feature, making it of lower priority to have a classifier with high recall which would cover all viable features. In that respect the generalization from Tempel 1 to Hartley 2 as shown in Figure 5 (right) would be completely feasible since the spacecraft would target just the features in which the classifier has the highest confidence. The application of the Hartley 2 model in a flyby at Temple 1 on the other hand would likely fail since we have approximately 50% false positive surface features in the high precision regime as illustrated in Figure 4 (right). In this light, the classifier strategy seems most appropriate for encounters involving targets that are well-characterized (i.e. for which representative examples exist) or for which at least one prior flyby has been performed.

The major reason the model generalizes better from Hartley 2 to Tempel 1 than the other way around is most likely the fact that Hartley 2 has significantly more surface features to learn from than Tempel 1. Hence the Hartley 2 model encompasses a wider range of appearances of surface features and can also better discriminate them from the large set of false positive detections. This result demonstrates the need to train such models from the largest variety of comets and asteroid available to guarantee the best generalization performance possible for future flyby missions.

## 4 Conclusion

To the best of our knowledge this is the first demonstration that autonomous surface feature detection is feasible. We described a framework for candidate detection and classification based on median filtering, mean shift clustering and random forest classification. The accuracy of the system was tested in cross validation experiments on data from comets 103P/Hartley and 9P/Tempel which emphasized the need for more and diverse training data to learn robust models with good generalization performance. Future flyby missions to comets and asteroids could benefit from such a framework by allowing precise pointing of narrow field of view instruments like spectrometers and hence significantly increase science return. This same capability is also useful for extended or multiple flyby encounters, where it can precisely target specific features of interest despite positional uncertainty of both spacecraft and target. This can greatly expand the palette of spacecraft commanding options available to the operations team. Future missions visiting more distant targets will introduce increasing light time gaps as well as high encounter speeds resulting in only a few frames for decision making. In this light, autonomous onboard computer vision will become increasingly important for our understanding of the composition and history of the outer solar system.

## Acknowledgments

This research was performed at the Jet Propulsion Laboratory, California Institute of Technology. We acknowledge U.S. Government Support. This study was conducted with sponsorship by the JPL Solar System Exploration Office and the JPL Research and Technology Development Program. Copyright 2014, California Institute of Technology. Government sponsorship acknowledged.

## References

[1] Vision and voyages. *Planetary Science Decadal Survey 2013-2022*, 2011.

[2] Yali Amit and Donald Geman. Shape quantization and recognition with randomized trees. *Neural Computation*, 9:15451588, 1997.

[3] Dmitriy Bekker, David R. Thompson, William J. Abbey, Nathalie A. Cabrol, Raymond Francis, Ken S. Manatt, Kevin Ortega, and Kiri L. Wagstaff. A field demonstration of an instrument performing automatic classification of geologic surfaces. *Astrobiology*, (in press), 2014.

[4] Shyam Bhaskaran, Joseph E Riedel, and Stephen P Synnott. Autonomous nucleus tracking for comet/asteroid encounters: the stardust example. In *Aerospace Conference, 1998 IEEE*, volume 2, pages 353–365. IEEE, 1998.

[5] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[6] Daniel Britt et al. Asteroids. *Community White Paper to the Planetary Science Decadal Survey*, 2010.

[7] B. D. Bue, K. L. Wagstaff, R. Castano, and A. Davies. Automatic plume detection for planetary bodies. *AGU Fall Meeting Abstracts*, page A4, 2006.

[8] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603–619, 2002.

[9] Antonio Criminisi, Jamie Shotton, and Ender Konukoglu. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trends in Computer Graphics and Vision*, 7(2-3):81–227, 2011.

[10] Thomas J. Fuchs and Joachim M. Buhmann. Interactive learning of randomized tree ensembles for object detection. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on Computer Vision*, pages 1370–1377, 2009.

[11] Thomas J. Fuchs and Joachim M. Buhmann. Computational pathology: Challenges and promises for tissue analysis. *Journal of Computerized Medical Imaging and Graphics*, 35(7):515–530, April 2011.

[12] Thomas J. Fuchs, Peter J. Wild, Holger Moch, and Joachim M. Buhmann. Computational pathology analysis of tissue microarrays predicts survival of renal clear cell carcinoma patients. In *MICCAI*, volume 5242, pages 1–8. Springer, 2008.

[13] D. Jewitt and A. Guilbert-Lepoutre. Limits to ice on asteroids (24) themis and (65) cybele. *Astron. J.*, 143:21, 2012.

[14] Kenneth P Klaasen, David Brown, Brian Carcich, Tony Farnham, William Owen, and Peter Thomas. Stardust-next navcam calibration and performance. *Icarus*, 2012.

[15] Yucong Lin, M. Bunte, S. Saripalli, and R. Greeley. Autonomous detection of volcanic plumes on outer planetary bodies. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 3431–3436, 2012.

[16] F. Marchis, P. Descamps, D. Hestroffer, and J. Berthier. Discovery of the triple asteroidal system 87 sylvia. *Nature*, 436:822–824, 2005.

[17] Nikos Mastrodemos, Daniel G Kubitschek, and Stephen P Synnott. Autonomous navigation for the deep impact mission encounter with comet tempel 1. *Space Science Reviews*, 117(1-2):95–121, 2005.

[18] Simon Perreault and Patrick Hebert. Median filtering in constant time. *Image Processing, IEEE Transactions on*, 16(9):2389–2394, 2007.

[19] J.C. Russ. *The Image Processing Handbook*. CRC Press, IEEE Press, 1999.

[20] A. Saffari, C. Leistner, J. Santner, M. Godec, and H. Bischof. On-line random forests. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, page 13931400, 2009.

[21] T. Sharp. Implementing decision trees and forests on a GPU. *Computer VisionECCV 2008*, page 595608, 2008.

[22] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-time human pose recognition in parts from single depth images. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1297 –1304, June 2011.

[23] S. Sonnett, J. Kleyna, R. Jedicke, and J. Masiero. Limits on the number distribution of main belt comets. *EPSC-DPS Joint Meeting*, page 391, 2011.

[24] David R. Thompson, William Abbey, Abigail Allwood, Dmitriy Bekker, Benjamin Bornstein, Nathalie A. Cabrol, Rebecca Castano, Tara Estlin, Thomas J. Fuchs, and Kiri L. Wagstaff. Smart cameras for remote science survey. In *Proceedings of the 10th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS)*, 2012.

[25] David R. Thompson, Benjamin J Bornstein, Steve A Chien, Steven Schaffer, Daniel Tran, Brian D Bue, Rebecca Castaño, Damhnait F Gleeson, and Aaron Noell. Autonomous spectral discovery and mapping onboard the EO-1 spacecraft. *Geoscience and Remote Sensing, IEEE Transactions on*, 51(6):3567–3579, 2013.

[26] David R. Thompson, Melissa Bunte, Rebecca Castaño, Steve Chien, and Ronald Greeley. Image processing onboard spacecraft for autonomous plume detection. *Planetary and Space Science*, 62(1):153–159, 2012.

[27] David R. Thompson, Julie C Castillo-Rogez, Steve A Chien, Richard Doyle, Tara Estlin, and David McLaren. Agile science operations: A new approach for primitive bodies exploration. In *Proc. SpaceOps Conf*, 2012.

[28] Kiri L. Wagstaff, David R. Thompson, Brian D. Bue, and Thomas J. Fuchs. Autonomous real-time detection of plumes and jets from moons and comets. *Submitted Manuscript*, 2014.