An automated method for tracking clouds in planetary atmospheres

D. Luz a,b,*, D.L. Berry c, M. Roos-Serote a

a CAAUL/Observatório Astronómico de Lisboa, Portugal
b Observatoire de Paris, Section de Meudon (T1m), 5, Place Jules Janssen, 92195 Meudon Cedex, France
c Departamento de Física and Centro de Geofísica da Universidade de Évora, Portugal

Received 3 July 2007; accepted 21 September 2007
Available online 29 September 2007
Communicated by W.B. Hubbard

Abstract

We present an automated method for cloud tracking which can be applied to planetary images. The method is based on a digital correlator which compares two or more consecutive images and identifies patterns by maximizing correlations between image blocks. This approach bypasses the problem of feature detection. Four variations of the algorithm are tested on real cloud images of Jupiter’s white ovals from the Galileo mission, previously analyzed in Vasavada et al. [Vasavada, A.R., Ingersoll, A.P., Banfield, D., Bell, M., Gierasch, P.J., Belton, M.J.S., Orton, G.S., Klaasen, K.P., Dejong, E., Breneman, H.H., Jones, T.J., Kaufman, J.M., Magee, K.P., Senske, D.A. 1998. Galileo imaging of Jupiter’s atmosphere: the great red spot, equatorial region, and white ovals. Icarus, 135, 265, doi:10.1006/icar.1998.5984]. Direct correlation, using the sum of squared differences between image radiances as a distance estimator (baseline case), yields displacement vectors very similar to this previous analysis. Combining this distance estimator with the method of order ranks results in a technique which is more robust in the presence of outliers and noise and of better quality. Finally, we introduce a distance metric which, combined with order ranks, provides results of similar quality to the baseline case and is faster. The new approach can be applied to data from a number of space-based imaging instruments with a non-negligible gain in computing time.

© 2007 Elsevier B.V. All rights reserved.
PACS: 93.85.Bc; 92.60.Gn; 96.12.Jg
Keywords: Methods: data analysis; Methods: numerical; Techniques: image processing

1. Introduction

Cloud features have been used successfully in the past as tracers of the winds in the atmospheres of Venus (Rossow et al., 1990), Mars (Wang and Ingersoll, 2003), Jupiter (Vasavada et al., 1998) and Saturn (Sánchez-Lavega et al., 1999), and are used routinely for deriving wind patterns in the terrestrial atmosphere (cf. Schmetz et al., 1993). Imaging systems in planetary missions typically produce data in which the wavelength dimension adds to the two spatial dimensions. Tracking of atmospheric features at given wavelengths, in order to derive the winds, and the construction of global maps from such large data volumes, become particularly time-consuming if done manually. This highlights the importance of automated procedures capable of analyzing sequences of data cubes with minimal user interaction.

We present here an automated method for cloud tracking (CT) which can be applied to planetary orbiter images in general, and in particular to images from the Visible and Infrared Thermal Imaging Spectrometer (VIRTIS) and from the Venus Monitoring Camera (VMC), the two experiments on board Venus Express which possess imaging capabilities (Drossart et al., 2007; Markiewicz et al., 2007). The method implements a digital correlator which
compares two consecutive images and identifies patterns by maximizing correlations between image sections, or blocks. We have tested four variations of the technique, including one, applied here for the first time, that compares favorably with one of the existing methods. The method has been tested on real cloud images by comparing both with cloud motions due to a prescribed wind field and with real motions. Section 2 makes a brief overview of CT methods. Sections 3 and 4 present the method and the numerical tests carried out for validation.

2. Overview

There are two problems associated with the tracking of atmospheric tracers such as clouds. First, an individual, homogeneous, cloud mass needs to be detected for tracking in a pair of images. The second problem is the computation of the velocity vectors, by identifying identical cloud masses in the pair. Some methods do not address direct feature detection, and tracking is based on the complete image instead of individual tracers. Most methods in use for identification are based on linear correlation, using the sum of squared differences, the normalized correlation coefficient or another measure of cross-correlation to find the correspondence between the intensity distributions of certain areas of the images.

One example of an automated tracer-based method is described in Nieman et al. (1997). In the feature selection stage, the gradient of the brightness is computed in a target area, with size defined by the operator, and the maximum is chosen for that target. To exclude undesired tracers, two filters are applied to the targets. Targets with large gradients are selected as tracers, and are then checked by computing their mean pixel brightness to ensure that no data is missing. Targets are divided into boxes of $3 \times 3$ pixels, where means and standard deviations are computed. Those having deviations below a given threshold are analyzed with a two-dimensional clustering scheme in order to determine the mean brightness for the coherent clusters, only two clusters being allowed by target. The automated algorithm then generates wind vectors by following the selected tracers through a sequence of three images. The method compares target boxes from the first image with a search area of the second image. The minimum of the sum of squared differences between the pixel brightnesses is chosen as the best match for each target box, giving the first vector. The procedure is then repeated for the second and third images, yielding a second vector, using the first vector to predict a tracer on the third image. A final quality controlled vector is based on the differences between the two vectors. Mukherjee and Acton (2002) present a scale-sensitive image classification scheme to minimize the problems of detection of cloud mass boundaries in feature selection. The detection of a cloud mass boundary is essentially a problem of image segmentation. The method detects homogeneous cloud segments with minimum intra-segment classification error, using a scheme for cloud detection based on an area segmentation that preserves the morphology of the cloud. This scale space segmentation consists of constructing a set of scaled images, called scale space, using operators that preserve the morphological shape of the area, avoiding the distortion of region boundaries. A filter is then applied to remove all the small areas with high intensities (within a dark region) and small areas with low intensities (within a bright region), allowing objects with similar characteristics and intensities to be clustered, grouping pixels in an image by measuring the distance from a cluster center. The distance is defined as the sum of the squared differences between the intensities at a given position and the cluster mean, through all the scales in the scale space. For the cloud tracking process, the scale space classification is applied to the areas of homogeneous cloud mass determined in the images, extracting the segments that present averaged intensities higher than the average intensity of the background. The selected features are cloud contours that remain almost stable over a sufficiently short time span, and which can then be cross-correlated to find spatial correspondences between the images.

A similar automated method, which tracks the connected zones of infrared images (cloud shields) based on best matches between successive images, is described in Morel et al. (1997). A geographical area containing the same cloud is selected in both images, its velocity is estimated from earlier tracking iterations, and translations are computed of the selected area. Trajectories are found by minimizing the overlapping threshold between the selected and the translated areas. This method has also been used to build a database of trajectories and characteristics of Mesoscale Convective Systems from Meteosat infrared images. Velden et al. (1997) describe the techniques for extracting terrestrial water vapor motion from GOES-8/9 and GMS-5 satellites. Feature selection is based on the brightness temperatures of the clouds, isolating the lowest values within a pixel array, computing local bidirectional gradients for comparison with empirically determined thresholds, and selecting those with maximum value. The tracking metric searches for the minimum of the sum of the squared differences in radiance between the target and the search arrays, in images with a 30-min interval.

In the planetary domain, Rossow et al. (1990) analyzed pairs of ultraviolet image sequences from the Pioneer Venus orbiter cloud photopolarimeter (OCPP), spanning nine years. From images with 30 km nadir resolution, spaced approximately 4 h apart, they determined wind speeds at the cloud tops by measuring the motions of UV cloud features. The smaller cloud features were assumed to move with the bulk flow, and the larger cloud features to represent large scale waves. They describe a fully automated technique with five steps. First the planetary coordinates of each pixel in latitude/longitude are determined and projected onto a standard latitude/longitude mapping. Next a high-pass filter is applied to the images to eliminate
large scale brightness gradients, enhancing the brightness contrast. The third step is the target selection, which systematically samples the first of a pair of images, without considering any specific features. Any location is considered a valid target, and a similar set of features is searched on the second image of the pair. Target dimensions are 32 × 32 pixels, separated by 16 pixels, with an overlap of 50% between adjacent blocks, and the search region is an array with 69 × 152 pixels, starting at the original location of the target in the first image. The identification process is associated with a maximum of correlation of the two numerical arrays representing the target selected, and the displacement of the feature is measured by the distance between the center of the target in the first image and the center of the region with maximum correlation in the search area. A similar planetary application has been the analysis of Galileo imaging data of Jupiter’s atmosphere (Vasavada et al., 1998), which will be discussed in Section 4.

Evans (1999) presents an alternative way to find the correspondence between images through the application of ordinal measures, combined with a correlation-relaxation labeling approach. The method of ordinal measures, which uses the relative order ranks of the image intensities, is known to produce less mismatches than the usual linear labeling approach. The method of ordinal measures, combined with a correlation-relaxation algorithm (Evans, 2000). The method was tested in Meteosat infrared images taken at a time interval of 30 min, yielding a consistent flow suitable for interpretation.

3. Cloud tracking technique

Techniques of cloud tracking can be used for meteorological and dynamical studies. They are used, on the one hand, for deriving winds, using cloud motion vectors as input parameters to weather models and for studying the dynamical behavior of weather systems. On the other hand, CT provides valuable information on growth and dissipation of cloud masses. However, the behavior of cloud structures presents major difficulties to estimation techniques based on automated image processing, since clouds change shape as time passes, which makes it difficult to apply principles of rigid motion. Another problem inherent in CT techniques for planetary atmospheres (other than the Earth’s) is that the available data frequently present high noise and distortion, and are often sparse both spatially and in time. In addition, the data volume from ground-based observations and space missions is very limited compared to terrestrial data.

For a pair of subsequent images, captured at a fixed time interval Δt, the objective is to derive the displacement vectors corresponding to the motion which occurred in the interval between the first image $A(t_0)$ and the second, $B(t_0 + Δt)$. Consider $size(A_0) = size(B_0) = M \times N$, for simplicity. If $M$ and $N$ are large, processing the data becomes quite time-consuming and an automated and optimized tool is necessary.

Defining a distance metric, measuring the distance between the images and finding the best correspondence, is one of the most important aspects of the algorithm. Several metrics can be used. Many CT techniques compute the cross-correlation directly from radiance. In this case the cross correlation coefficient can produce mismatches when in the presence of noise and deformation, since the radiiances of the pixels in an image are neither unique nor normally distributed.

Radiance, however, is not the only possible measure of cloud brightness at a given point. For CT purposes, it can be replaced without loss of generality with any other such measure that preserves the intensity variations across the image. A robust approach is to replace radiances with their order ranks within the images, assigning to each pixel its corresponding order number (Bhat and Nayar, 1998; Evans, 1999). This is less sensitive to noise and represents better the order relation between bright and dark areas within a given area in the presence of feature motion. In the following we shall refer to a value of such a general measure of cloud brightness as a number of digital units (DU).

The methods tested here are based on an algorithm of maximum correlation between blocks of pixels, or sub-windows, selected within the images being compared. Each sub-window (A) in the first image is compared with all possible sub-windows (B) in the subsequent image in order to retrieve displacement vectors. At the end of the comparison stage, a filter (described at the end of this section) is applied to the final vector field in order to remove incoherent vectors. In the following discussion we shall use the terms sub-window and box interchangeably with the same meaning.

We test four variations of the technique, based on combinations of two definitions of the distance between image pairs (which we call 2D and semi-2D, described below) with two definitions of image brightness (pixel radiances and order ranks).

The first distance definition is based on a Euclidean metric. The differences between DUs of the sub-windows are first computed and then the distance, computed as the mean squared difference between the images, is minimized:

$$d(p,q) = \sum_{i=1}^{m \times n} (DU_{B} - DU_{A})^2 / (m \times n)$$

where $m$ and $n$ are the number of rows and columns in the sub-windows, and $p$ and $q$ are labels identifying sub-windows in images A and B, respectively. Point $(p,q)_{\text{min}}$ at the minimum of $d$ is taken as the center of the displaced sub-window in B.

The distance measure defined in Eq. (1) is a 2D, $m \times n$ sum which depends on every pixel in the chosen box, and
its computation is the top time-consuming point in the algorithm. As an alternative approach we introduce a semi-2D definition, which is a combination of Eq. (1) with a 1D distance metric, based on the average values of the box along each of the spatial directions. Consider the distances

\[ d_k = \sum_{i=1}^{m} (|B_i - |A_i|)^2 / m \]  
\[ d_r = \sum_{j=1}^{n} (|B_j - |A_j|)^2 / n \]  

where square brackets denote a mean of the box over the image rows in Eq. (2) and columns in Eq. (3) (with the result being a column vector and a row vector, respectively).

A first trial for the distance is computed as the minimum of

\[ d(p_1, q_1) = d_k^2 + d_r^2. \]

However, since this is a poor estimator of the displacement, we also use a second distance estimation, \( d(p_2, q_2) \), from Eq. (1), but now only from the central \([m/2 \times n/2]\) part of the box. Finally,

\[ d(p, q) = \min(d(p_1, q_1), d(p_2, q_2)) \]  

is taken as the distance between the two sub-windows. This semi-2D method has the advantage of reducing the computing time by about 20% compared with Eq. (1).

Using the measured radiances of the images as DU can sometimes produce poor results, due to the deformation inherent in feature motion and to noise. The post-tracking filter can in this case sometimes remove as many as two thirds of the wind vectors, making the traditional approach of directly comparing radiance values one of limited utility.

In the alternative definition of image brightness, radiance pixel values are ordered and replaced with their relative order ranks, i.e., the darkest pixel value is replaced with 1, and the brightest with the value \( m \times n \). In this case, pixels with equal radiances have to be separated by adding some small value of random noise to them.

The correcting filter applied to the field retrieved at the end of the comparison stage is similar to the one described in Vasavada et al. (1998) A vector may be considered an outlier and be removed if: (a) it has one of its components longer than half the box size; (b) its length is more than 50% the 1st-neighbor average (four closest points); (c) its orientation differs from the neighbor average by more than 30°. The filtered field thus obtained can be somewhat patchy, but if the number of vectors left by the filter is sufficient, it can be interpolated onto a regularly spaced mesh to produce a smooth field (but this is not done here).

4. Numerical experiments

In the tests presented here we use the following combinations of distance and DU definitions: Eq. (1) and image radiances (case A); Eq. (1) and pixel order ranks (case B); Eq. (4) and image radiances (case C); and Eq. (4) and order ranks (case D).

Three batteries of tests have been carried out, using images of Jupiter’s clouds from the Galileo spacecraft’s Solid State Imaging system (SSI). In the Galileo operation period, SSI acquired spatial, spectral and temporal coverage of many features in Jupiter’s atmosphere, such as the Great Red Spot, belts and zones, hot spots and the white ovals. A time series of multispectral image mosaics was collected for each target region during each orbit. Mosaics were obtained in four spectral bands of reflected sunlight, in the violet (410 nm), near infrared continuum (756 nm), and the two CH₄ absorption bands at 727 and 889 nm. The cloud morphology is seen to vary over 1-h intervals.

Vasavada et al. (1998) tracked cloud motions in the Great Red Spot, the equatorial region and two of the white

![Fig. 1. (a) Root-mean-squared error in horizontal (*) and vertical (+) velocity components, versus \( u_0 \) and \( v_0 \), for method A. (b) Same as (a), compared with case C (circles and diamonds). (Note the different scales.)](image-url)
In the manual technique, a human operator takes measurements by matching cloud tracers on a pair of maps, displayed side-by-side or blinking. The automated software maximizes a two-dimensional cross-correlation within a predefined box of pixels. The method is more successful at matching tracers between maps separated by short time intervals (1 h) than long ones (10 h), because the changing morphology of the tracers does not allow to determine their locations accurately in the latter case. Measurements with both the manual and automated techniques are consistent.

For testing with real data, we chose images of the white ovals, which exhibit mostly ordered circular motions, while on the rim of the ovals and cyclonic features more complex motions with shear are apparent. The resolution for the white ovals is on the order of 22–28 km px\(^{-1}\). The images are analyzed in detail in Vasavada et al. (1998).

Our tests are based on a pair of SSI color composite images separated by 1.4 h, showing a white oval and a

---

**Fig. 2.** Motions retrieved using a square box of 40 pixels (shown in red) and the prescribed distorting field (blue arrows). RMS\(U,V\) are both equal to 0.85, but decrease to 0.61 and 0.56 if a 20-pixel box is used.

**Fig. 3.** Motions retrieved for methods A, B and D. White arrows show the retrieved displacements vector field and black arrows show the field after the correcting filter has been applied.
cyclonic feature (Fig. 10 in Vasavada et al. (1998)). In the first set of tests only one image ($A_0$) is used in its original form, with the second image ($B_0$) being generated by applying the motion of a prescribed wind field to the first image. The prescribed wind has been defined as

\begin{align}
U &= U_0 + u_0 \cos(k_x X + k_y Y) \\
V &= V_0 - v_0 \sin(k_x X + k_y Y)
\end{align}

where $U$ and $V$ are the horizontal and vertical components, $X$ and $Y$ are image pixel numbers and $k_x$ and $k_y$ are defined arbitrarily as $k_x = 2\pi/200$ and $k_y = 2\pi/100$. A given pixel, $(X_0, Y_0)$ in $A_0$, is transported to position $(X_0 + U \Delta t, Y_0 + V \Delta t)$ in $B_0$. While the constant terms amount to a simple translation, the wave components allow to introduce distortion in the second image. We have performed a series of tests either with translation or distortion. In the end of the test the retrieved wind field $(U_r, V_r)$ is compared with the prescribed one, using the root-mean-squared error for each component, $\text{RMS}_U$ and $\text{RMS}_V$, with $\text{RMS}_U = \sqrt{\sum(U_r - U)^2/N_p}$ (and similarly for $V$), where $N_p$ is the number of points where motions were computed.

To optimize the use of computer memory and time the tests were restricted to a region $250 \times 250$ px within the white oval. In tests with prescribed motions we use methods A and C as defined above, with $\Delta t = 1$ and square boxes.

The method handles simple translations ($u_0 = v_0 = 0$) successfully, as expected, yielding a null RMS error whenever the displacements are within the box size ($U_0$ and $V_0 \leq m/2$).

Wave motions are also well retrieved, albeit with a loss of accuracy as their amplitude increases. Fig. 1 shows $\text{RMS}_U$ and $\text{RMS}_V$ versus $u_0$ and $v_0$, respectively, for a series of tests made with a box of 40 by 40 pixels. The values used for $(u_0, v_0)$ were $(-1, -1), (1, 1), (-3, -3), (3, 3), (-5, -5), (-5, 5), (5, -5), (5, 5), (-10, 10)$ and $(10, -10)$. Panel (a) shows case A only, while panel (b) shows a comparison with case C.

Fig. 2 shows the retrieved wind field for the case $u_0 = v_0 = 3$. In this case, if the box size is divided by 2, then the RMS error decreases from $\sim 0.85$ to 0.6 in both components, indicating that for motions with distortion the accuracy of the method is sensitive to that parameter. For the 40-px box used in Fig. 1a, case C compares favorably with case A because the distorting motion favors a smaller box. Case C is also faster.

In the second set of tests we use both images in the SSI pair. Instead of a prescribed motion, the method retrieves the physical displacements of the features from the first image.
to the second image. In Fig. 3, methods A, B and D are compared. The selected region is 340 × 320 px, containing the white oval. From a total of 1088 vectors, the correcting filter removes 26% in case A, 13% in case B, 53% in case C (not shown) and 24% in case D. The similar results for methods A and D, with D being 20% faster, make the latter a better choice than the classical approach of case A used in Vasavada et al. (1998). The quality of results for case B make the combination of the classical approach with order ranks the best approach overall.

In the third group of tests we also use both images in the SSI pair, and the retrieved field is compared with the results of Vasavada et al. (1998). The box size was tested with dimensions ranging from 20 × 20 to 50 × 50 px. The number of pixels between successive box positions (box step) is not important for the vectors retrieved, affecting only the spatial density of the field. We tested values between 4 and 10 pixels for the box step. Fig. 4 shows a comparison between case A (red arrows) and Vasavada et al. (1998) (blue arrows) for the full image (500 × 1000 px). From experience, the optimal box size is on the order of 40 px. Smaller boxes produce worse results. Fig. 5, where a 32-px box was used, shows larger patches of blank areas where the filter removed incoherent vectors.

Results for case B are shown in Fig. 6, where it was the only one producing displacement vectors in the low contrast region at the center of the oval.

In order to compare results quantitatively, both vector fields (after filtering) have to be interpolated onto the same

Fig. 5. Same as Fig. 4 for the white oval only and a box size of 32 px. Note the larger patches of blank areas left after correction, due to the smaller box size.

Fig. 6. Same as Fig. 4 for the white oval only and for method B. Note the better results (absence of blank areas) at the center of the oval.
regular grid (Fig. 7). The root-mean-square difference between the retrieved vector field in case D and the results of the digital correlator of Vasavada et al. (1998) is 2.5 pixels. The mean displacement (computing the field for the full image) is 8.4 pixels. Since they do not cite the mean error of their wind retrievals, the similarity of the results of both schemes is difficult to assess.

The four variations of the method presented here impose few requirements, are simple and fast to execute. The consecutive images must present a minimum amount of spatial overlap, at least equal to the box size, and they need to have similar spatial resolutions. The spatial overlap sets an upper limit on the time interval between the images. The minimum overlap required, equal to the box size, leads to a very patchy vector field. It must be greater than half the field of view, in order to ensure the continuity of the vector field.

The image overlap depends on the detector field of view and on its distance to the planet, as well as on the wind velocity. For an instrument such as VIRTIS on Venus Express, with a FOV of 64 mrad, at a typical distance of 10,000 km, Venus cloud tops moving with zonal wind speed \( \sim 100 \text{ m s}^{-1} \) move across one FOV in \( \sim 107 \text{ min} \), making the maximum time interval \( \sim 53 \text{ min} \).

To optimize computer memory and time, large images have to be divided into smaller ones, typically 100 \( \times \) 100 px, which are treated separately. The partial vector fields for the various sub-images are stored and merged together to create the final field.

5. Conclusions

We have presented four variations of a method for retrieving cloud motion vectors, based on the correlation between two consecutive images. We have intentionally avoided the problem of direct feature detection (Morel et al., 1997; Mukherjee and Acton, 2002), which requires extensive computational power. The method produces coherent results for satellite images that respect certain conditions regarding the time interval between them and the resolution. Experiments with different cases have shown that valid information can be retrieved for various types of motion, including pure translation, rotation, and in the presence of more complex cloud deformation.

Our method A is similar to the basic digital correlators described in Vasavada et al. (1998), Rossow et al. (1990), Velden et al. (1997) and Nieman et al. (1997). Method B, with the same distance metric but correlating order ranks (Bhat and Nayar, 1998) rather than radiances, is more robust for handling real images, while keeping a computational load comparable to case A. As long as computer time is not an issue, it is shown to be the best approach to the problem.

Fig. 7. Comparison between the vector fields of Vasavada et al. (1998) (in black) and case D (white arrows), after filtering and interpolation onto a regularly spaced grid.
The distance metric introduced in cases C and D is an interesting alternative to the classical metric. Although it gives poor results for real images when correlating radiiances (case C), combined with order ranks (case D) it yields results qualitatively very similar to the classical method, with a gain of ~20% in computer time.

Our principal objective with this work has been to create a tool to analyze planetary imaging data, in particular from Venus Express (VEx) and its VIRTIS and VMC instruments. In April 2006, VEx was inserted into Venus orbit, and started its mission to unveil the Venusian atmosphere and surface with unprecedented precision. Ground-based observations and atmospheric modeling will complement satellite data in important ways, providing data at different wavelengths and altitude levels, and helping to understand mechanisms which are not apparent from the separate analysis of the various data sets.

Winds and temperatures obtained through VEx data analysis will allow the characterization of the dynamics of the Venus atmosphere in great detail. The present method of cloud tracking can be applied to quickly derive dynamical information from atmospheric feature motions. It will help to understand superrotation and the mechanisms controlling the northern and southern polar vortices, to characterize wave phenomena with sufficient spatial and temporal resolution, and to measure winds for data assimilation purposes, testing and fine tuning general circulation models.

This method is also available to the general scientific community by contacting the authors.

Acknowledgements

The authors thank Ashwin Vasavada for his invaluable help with the Jupiter images and for providing results in a text file. The authors acknowledge support from Fundação para a Ciência e a Tecnologia, Portugal, through Project PDCTE/FNU/49822/2003. D.L. and D.L.B. acknowledge support from FCT Project POCI/CTE-AST/57655/2004. D.L. acknowledges financial support from FCT fellowship PRAXIS XXI/BPD/3630/2000.

References