Reprint

from

Proceedings of the Fifth International Conference on Computer Vision

Cambridge, Massachusetts

June 20-23, 1995

Structure and Semi-Fluid Motion Analysis of Stereoscopic Satellite Images for Cloud Tracking

K. Palaniappan, Chandra Kambhamettu, Fredrick Hasler, and Dmitry Goldgof

IEEE COMPUTER SOCIETY

Washington, DC ◆ Los Alamitos ◆ Brussels ◆ Tokyo

PUBLICATIONS OFFICE, 10662 Los Vaqueros Circle, P.O. Box 3014, Los Alamitos, CA 90720-1264 USA

Structure and Semi-fluid Motion Analysis of Stereoscopic Satellite Images for Cloud Tracking

K. Palaniappan, Chandra Kambhamettu, A. Frederick Hasler† and Dmitry B. Goldgof‡

Universities Space Research Association
† Lab. for Atmospheres, Code 912
NASA/Goddard Space Flight Center
Greenbelt, MD 20771

‡ Dept. of Computer Science and Eng.
University of South Florida
Tampa, FL 33620

Abstract

Time-varying multispectral observations of clouds from meteorological satellites are used to estimate cloud-top heights (structure) and cloud winds (semi-fluid motion). Stereo image pairs over several time steps were acquired by two geostationary satellites with synchronized scanning instruments. Cloud-top height estimation from these image pairs is performed using an improved automatic stereo analysis algorithm on a massively parallel Maspar computer with 16K processors. A new category of motion behavior known as semi-fluid motion is described for modeling cloud motions and an automatic algorithm for extracting semi-fluid motion is developed to track cloud winds. The time sequential dense estimates of cloud-top height depth maps in conjunction with intensity data are used to estimate local semi-fluid motion parameters for cloud tracking. Both stereo disparities and motion correspondences are estimated to sub-pixel accuracy. The Interactive Image SpreadSheet (IISS) is a new versatile visualization tool that was enhanced to analyze and visualize the results of the stereo analysis and semi-fluid motion estimation algorithms. Experimental results using time-varying data of visible channel from two satellites in geosynchronous orbit is presented for the Hurricane Frederic.

1 Introduction

The analysis of remote sensing datasets is an important application area for computational methods developed in computer vision. The estimation of cloud-top structure and motion using multiple satellite views is an extremely challenging problem due to the complex dynamics of the imaging instruments and the underlying non-linear phenomena of cloud formation and weather.

Accurate cloud heights and winds are important for a number of meteorological and climate applications [1, 7] such as cloud model verification, physically-based numerical weather prediction and data assimilation, and radiation balance estimation for Mission to Planet Earth type climate baseline studies. Stereoscopic multispectral measurement of cloud heights and time sequential estimation of cloud winds using geostationary and polar orbiting satellite-based instruments are essential because they provide information which is independent of other meteorological measurements. The height and wind information is derived from direct geometric measurements and does not depend on complicated inversion techniques that have non-unique solutions and also require a knowledge of the cloud emissivity or the local lapse rate when using infrared-based techniques [7]. Stereoscopic-based heights could assist in the accurate estimation of other cloud properties.

The computational approach to cloud structure and semi-fluid motion estimation involves two primary modules. The first module estimates stereo-based cloud structure using a hierarchical, multi-resolution coarse-to-fine search strategy with image warping at each level [8] and has been parallelized for the Maspar. The second module estimates semi-fluid motion using a sequence of dense depth maps and corresponding intensity imagery, in order to capture the time evolution of cloud heights. The IISS is a visualization tool that was developed for rapidly analyzing gigabyte-sized geophysical datasets [2], and was enhanced to visualize the results of the stereo analysis and semi-fluid motion estimation algorithms. A brief overview of stereo and motion analysis for cloud tracking is presented next, followed by a description of the algorithms.

1.1 Stereo Analysis

Stereo analysis of remotely sensed images is useful for a variety of applications including cloud height measurement, digital terrain models for geological and hydrological studies, land use mapping, geometric rectification, topography related radiometric correction, etc. Stereopsis both in human vision and in remote sensing relies on the same principal of parallax (relative displacement). One significant difference is that in human stereopsis, the range of accurate depth information is on the order of several meters whereas in remote sensing, depth information can be recovered over distances that are hundreds to tens of thousands of kilometers due to the extremely long baselines with satellite geometries [8]. The most difficult aspect of developing computational algorithms for stereopsis

0-8186-7042-8/95 $4.00

that match the intrinsic capabilities of human vision is the correspondence problem, i.e, locating the same point in multiple time-varying sensor measurements.

Correspondences are usually determined using feature-based or region-based matching algorithms with bottom-up or top-down implementations. The bottom-up or low-level approach for stereo analysis includes: i) extracting feature points or area measures in both satellite views, ii) matching the feature points or area measures under certain geometric and illumination constraints, and iii) computing a depth or cloud-top height map using the disparity values from correspondences, with the satellite ephemeris, sensor geometry and scanning configuration. A number of stereo algorithms attempt to use aspects of the human visual system including multiresolution or hierarchical filtering and coarse-to-fine constraint propagation. The bottom-up multiscale approach with image warping at each resolution is used in the parallelized stereo analysis algorithm.

1.2 Motion Analysis

Motion perception in humans is a natural, well-developed visual capability with the facility to rapidly categorize the relative motion between non-rigid objects with respect to an arbitrary co-ordinate system. Yet, developing automatic computational motion analysis algorithms, capable of handling non-rigid motion has been a difficult task. Most of the work in motion analysis has been based on the rigidity assumption where the shape of objects (i.e. distances between points) do not change over time. The rigidity assumption fails in numerous real-world examples. Non-rigid motion behavior is prevalent in nature including the dynamics of clouds, aerosols, water waves, currents, sand and soil, human movement, motion of animals, plants, biological organs and cells, deformation of flexible or articulated materials, structures and industrial components etc.

Cloud motion is a special case of non-rigid motion, where there is both partial fluid and partial solid motion, that we describe as being *semi-fluid motion*. The motion of objects exhibit a wide variation in behavior ranging from the simplest ideal frictionless rigid motion to the most complex turbulent fluid motion. Within this broad range various classes of motion can be identified including *articulated motion*, *quasi-rigid motion*, *isometric motion*, *homothetic motion*, *conformal motion* and *elastic motion* [5]. Elastic motion deals with shape changes of continuous surfaces, and fluid motion deals with particle motion in which there is usually no continuity constraint among neighboring particles that are free to move according to the underlying dynamics. Although there has been an increasing amount of research in the individual categories of motion behavior, there have not been many attempts to combine different classes of motion models. In the fluid motion case when the local continuity of a surface is partially preserved, it should be advantageous to use both the characteristics of local continuity required for estimating surface properties, along with the flexibility of larger localized particle motion that would otherwise violate most local continuity assumptions. An ideal example of this category of motion, which we call *semi-fluid motion* is the non-rigid motion of clouds. Other examples may be found in the atmospheric sciences, as well as in hydrospheric modeling of ocean eddies and currents that mix and split thus maintaining some larger scale features identifiable in multi-spectral imagery, or biological cells that undergo fission or fusion processes. Time varying properties of clouds represent a prototypical example of semi-fluid motion behavior and offers a new experimental area for further investigation of non-rigid motion models.

2 Automatic Stereo Analysis

A successful application of the special purpose Goodyear Massively Parallel Processor (MPP) computer was the interactive Automatic Stereo Analysis (ASA) program developed for land stereo analysis [8]. The MPP enabled interactive experimentation with cross correlation techniques using large templates on high resolution grids with the results being displayed on a MacIntosh personal computer. The Goddard ASA algorithm has been reimplemented on the commercially available Maspar MP-2 parallel computer with various improvements including sub-pixel interpolation using neighboring correlations, automatic selection of template sizes, manipulating disparity values using floating point arrays for greater matching precision, handling data sets with greater dynamic range (up to 16 bits per pixel), restricting the epipolar search direction (positive or negative) and displaying the results using a simple X Window System graphical interface. The Maspar MP-2 has a Single Instruction Multiple Data stream architecture with 16,384 processing elements, each with 64KB of memory arranged in a 128 column by 128 row mesh configuration. The MPL data parallel language was used to implement the stereo analysis algorithm on the MP-2.

The primary goal of the Goddard ASA is to find correspondences between the reference (left) and search or test (right) satellite-based stereo datasets in order to estimate the disparity or depth map to subpixel accuracy. The ASA algorithm uses coarse-scale matching with large templates for estimating an initial fusion and then iteratively increases to fine-scale matching in analogy with psychological observations of human stereopsis. In many multiresolution stereo analysis approaches, the cost of generating images and matching over all levels even using an image pyramid is very high. In the Maspar parallel implementation the image size remains constant and the matching template size is reduced. Matching is based on maximizing a normalized cross-correlation measure for a template centered around the pixel of interest in the reference dataset. The search window in the test dataset incorporates epipolar and maximum disparity constraints and specifies the search region. The pixel in the test dataset corresponding to the maximum template correlation value is used for the cumulative disparity estimate at the current level. The updated dense disparity map is used to warp the test image so that searching at the next finer level can use a smaller sized template. The process is repeated until the changes in the disparity estimates are small or the finest level of

the hierarchy is reached.

The main steps in the Goddard ASA algorithm implemented for the Maspar are summarized below.

- *Preprocessing images.* Due to complex and time-varying sensor geometries the stereo pairs are reprojected or rectified so that epipolar lines are horizontal. The sensor data may need to be radiometrically corrected and enhanced so that the reference and test data have histograms of roughly the same shape.
- *Automatic template size search.* Determine the starting resolution (coarsest template size) for initial matching using the local gradient of the weighted entropy and gradient of the standard deviation measures.
- *Initialize disparities.* If available a priori disparity estimates from other sources can be incorporated (i.e. from infrared measurements) to constrain matching.
- *Warp test image.* The original test image is warped using the cumulative disparities estimated up to the current scale (iteration).
- *Determine image matches.* For each reference image pixel, a match score is computed between a template centered at the pixel and neighborhoods within a search area in the test image. At finer-scales the template sizes and search areas are reduced to extract smaller scale features. The region-based measure used for stereo matching is a normalized mean and variance correlation, called match score given by,

$$\frac{\sum_{i,j}(x_{i,j}-\bar{x})(y_{i-k,j}-\bar{y}_k)}{\left[\sum_{i,j}(x_{i,j}-\bar{x})^2\right]^{\frac{1}{2}}\left[\sum_{i,j}(y_{i-k,j}-\bar{y}_k)^2\right]^{\frac{1}{2}}} \quad (1)$$

where $x_{i,j}$ and $y_{i,j}$ correspond to template pixels within the reference and test images respectively, $x_{i,j}$ is the grey level of the (i,j)th pixel within the template neighborhood and $y_{i-k,j}$ is the grey level of the (i,j)th pixel in kth search area of the test image. The values $\bar{x}$ and $\bar{y}_k$ are the corresponding mean values. For each pixel in the reference image, the match scores for all neighborhoods within the search area are computed. The pixel at the center of the search template with the highest correlation match score is selected as corresponding candidate pixel and the vector k gives an estimate of the disparity. The search window size and sub-pixel interpolation window size can be controlled. Vertical disparities are possible unless searching is constrained to (horizontal) epipolar lines and the search direction can be constrained to only positive disparities. Unreliable match values can be filtered using thresholds based on local image variance.
- *Detect illegal disparities.* Local continuity of surfaces is used by constraining the disparity gradient to mark outlier disparities for correction. A disparity gradient threshold is used to constrain the maximum local slope.
- *Interpolate over outlier disparities.* Interpolation using local neighboring disparities is done using an iterative second order diffusion model at positions where the maximum disparity gradient is exceeded.
- *Smooth disparities* Smoothing of the disparity field can be done to mitigate noise before image warping and matching at the next finer-scale of the hierarchy.

The difficulties in cloud stereo analysis include viewing angle dependence of cloud opacities, occlusion, multilayer clouds with differing transparencies, sunglint, illumination differences between views, and small cloud motions between views due to sensor scanning synchronization differences.

3 Automatic Semi-fluid Analysis

A novel approach is presented to estimate the semi-fluid motion parameters for a small local neighborhood experiencing deformable motion. Our emphasis in this work is on the application to cloud motion analysis with the long-term goal that automated methods of estimating the global 3-D wind field will lead to an improved understanding of atmospheric processes and better operational weather forecasts. Both the stereo-based cloud-top height (depth) and intensity information is used in a complementary fashion for estimating the semi-fluid behavior of clouds. The invariant properties used to characterize semi-fluid motion are discussed below.

When an image of a scene is acquired, one important aspect is to define the imaging coordinate system for precise quantitative scene description. It may be easy to define the coordinates w.r.t the camera. However, if we consider more complex satellite imaging geometries where it is difficult to keep track of the camera position and pixel location with respect to an inertial coordinate system, or to keep track of the objects' motion dynamics, then it is advantageous to use characteristic properties that are invariant w.r.t the coordinate system for motion analysis. Surface properties such as curvature and unit-normal can be used for such purposes. We use these properties in particular for the estimation of semi-fluid motion parameters and establish point correspondences for cloud tracking. Note that curvature is invariant w.r.t translation and rotation of an object, whereas the unit-normal is invariant to only translation.

A local small deformation assumption can be used to derive relationships between the unit-normal at a point on the surface before and after motion, and also between the curvature at the point before and after motion. We build on these derivations [9, 4] and extend the approach to include semi-fluids, in order to estimate point correspondences with sub-pixel accuracy. Consider a point defined parametrically as, $z(x,y)$, on a surface that undergoes small deformation changes and is transformed to the point $z'(x',y')$. The following relationship holds between the corresponding unit-normals for the point undergoing a small non-rigid motion [9, 4]:

$$\mathbf{n}' = \mathbf{n} - \mathbf{n} \times rot\ \mathbf{s} \quad (2)$$

where $\mathbf{n}$ corresponds to the unit-normal of a point before motion, and $\mathbf{n}'$ is the unit-normal of the same point after motion. The displacement function, $\mathbf{s}$, models the non-rigid motion of the local neighborhood or small patch around the point of interest and can be expressed in the imaging instrument coordinate system at the initial time step as,

$$\mathbf{s} = (x'-x, y'-y, z'(x',y') - z(x,y)). \quad (3)$$

the hierarchy is reached.

The main steps in the Goddard ASA algorithm implemented for the Maspar are summarized below.

- *Preprocessing images.* Due to complex and time-varying sensor geometries the stereo pairs are reprojected or rectified so that epipolar lines are horizontal. The sensor data may need to be radiometrically corrected and enhanced so that the reference and test data have histograms of roughly the same shape.
- *Automatic template size search.* Determine the starting resolution (coarsest template size) for initial matching using the local gradient of the weighted entropy and gradient of the standard deviation measures.
- *Initialize disparities.* If available a priori disparity estimates from other sources can be incorporated (i.e. from infrared measurements) to constrain matching.
- *Warp test image.* The original test image is warped using the cumulative disparities estimated up to the current scale (iteration).
- *Determine image matches.* For each reference image pixel, a match score is computed between a template centered at the pixel and neighborhoods within a search area in the test image. At finer-scales the template sizes and search areas are reduced to extract smaller scale features. The region-based measure used for stereo matching is a normalized mean and variance correlation, called match score given by,

$$\frac{\sum_{i,j}(x_{i,j}-\bar{x})(y_{i-k,j}-\bar{y}_k)}{\left[\sum_{i,j}(x_{i,j}-\bar{x})^2\right]^{\frac{1}{2}}\left[\sum_{i,j}(y_{i-k,j}-\bar{y}_k)^2\right]^{\frac{1}{2}}} \quad (1)$$

where $x_{i,j}$ and $y_{i,j}$ correspond to template pixels within the reference and test images respectively, $x_{i,j}$ is the grey level of the (i,j)th pixel within the template neighborhood and $y_{i-k,j}$ is the grey level of the (i,j)th pixel in kth search area of the test image. The values $\bar{x}$ and $\bar{y}_k$ are the corresponding mean values. For each pixel in the reference image, the match scores for all neighborhoods within the search area are computed. The pixel at the center of the search template with the highest correlation match score is selected as corresponding candidate pixel and the vector k gives an estimate of the disparity. The search window size and sub-pixel interpolation window size can be controlled. Vertical disparities are possible unless searching is constrained to (horizontal) epipolar lines and the search direction can be constrained to only positive disparities. Unreliable match values can be filtered using thresholds based on local image variance.
- *Detect illegal disparities.* Local continuity of surfaces is used by constraining the disparity gradient to mark outlier disparities for correction. A disparity gradient threshold is used to constrain the maximum local slope.
- *Interpolate over outlier disparities.* Interpolation using local neighboring disparities is done using an iterative second order diffusion model at positions where the maximum disparity gradient is exceeded.
- *Smooth disparities* Smoothing of the disparity field can be done to mitigate noise before image warping and matching at the next finer-scale of the hierarchy.

The difficulties in cloud stereo analysis include viewing angle dependence of cloud opacities, occlusion, multi-layer clouds with differing transparencies, sunglint, illumination differences between views, and small cloud motions between views due to sensor scanning synchronization differences.

3 Automatic Semi-fluid Analysis

A novel approach is presented to estimate the semi-fluid motion parameters for a small local neighborhood experiencing deformable motion. Our emphasis in this work is on the application to cloud motion analysis with the long-term goal that automated methods of estimating the global 3-D wind field will lead to an improved understanding of atmospheric processes and better operational weather forecasts. Both the stereo-based cloud-top height (depth) and intensity information is used in a complementary fashion for estimating the semi-fluid behavior of clouds. The invariant properties used to characterize semi-fluid motion are discussed below.

When an image of a scene is acquired, one important aspect is to define the imaging coordinate system for precise quantitative scene description. It may be easy to define the coordinates w.r.t the camera. However, if we consider more complex satellite imaging geometries where it is difficult to keep track of the camera position and pixel location with respect to an inertial coordinate system, or to keep track of the objects' motion dynamics, then it is advantageous to use characteristic properties that are invariant w.r.t the coordinate system for motion analysis. Surface properties such as curvature and unit-normal can be used for such purposes. We use these properties in particular for the estimation of semi-fluid motion parameters and establish point correspondences for cloud tracking. Note that curvature is invariant w.r.t translation and rotation of an object, whereas the unit-normal is invariant to only translation.

A local small deformation assumption can be used to derive relationships between the unit-normal at a point on the surface before and after motion, and also between the curvature at the point before and after motion. We build on these derivations [9, 4] and extend the approach to include semi-fluids, in order to estimate point correspondences with sub-pixel accuracy. Consider a point defined parametrically as, $z(x,y)$, on a surface that undergoes small deformation changes and is transformed to the point $z'(x',y')$. The following relationship holds between the corresponding unit-normals for the point undergoing a small non-rigid motion [9, 4]:

$$\mathbf{n}' = \mathbf{n} - \mathbf{n} \times rot\ \mathbf{s} \quad (2)$$

where $\mathbf{n}$ corresponds to the unit-normal of a point before motion, and $\mathbf{n}'$ is the unit-normal of the same point after motion. The displacement function, $\mathbf{s}$, models the non-rigid motion of the local neighborhood or small patch around the point of interest and can be expressed in the imaging instrument coordinate system at the initial time step as,

$$\mathbf{s} = (x'-x, y'-y, z'(x',y') - z(x,y)). \quad (3)$$

The rotation or *curl* of the displacement function denoted as, *rot* s , is given by,

$$rot\ \mathbf{s} = \frac{1}{E}\mathbf{z}_1 \times \mathbf{s}_1 + \frac{1}{G}\mathbf{z}_2 \times \mathbf{s}_2. \quad (4)$$

The subscripts 1 and 2 indicate differentiation with respect to the parameters x and y respectively and the denominators, $E\ (= z_1 \cdot z_1)$ and $G\ (= z_2 \cdot z_2)$, are coefficients of the first fundamental form [9]. Polynomial functions are the simplest models for the local displacement function of a small patch undergoing small deformation. In our cloud-motion experiments, and in order to derive simple analytical relationships for differential geometry parameters, a small patch is assumed to undergo a local affine transformation defined as,

$$\mathbf{s}(x,y) = (a_i x + b_i y + c_i, a_j x + b_j y + c_j, a_k x + b_k y + c_k). \quad (5)$$

Let the unit-normal of a point before motion be defined as, $\mathbf{n} = (n_i, n_j, n_k)$ and the unit-normal of the same point after motion as, $\mathbf{n}' = (n_i', n_j', n_k')$. Surface normals and other intrinsic surface properties are estimated by using locally fitted quadratic surfaces for each small patch though the relationship between the normals (2) holds for all surface models. From (2), (3), (4) and (5), the following linearly independent equations can be derived [4, 6]:

$$\left[\frac{-n_k z_1}{E}a_i + \left(\frac{n_i - n_k z_2 - n_j}{G}\right)b_i + \left(\frac{n_j - n_k z_1 - n_i}{E}\right)a_j - \frac{n_k z_2}{G}b_j + \frac{n_k}{E}a_k + \frac{n_k}{G}b_k - n_i + n_i{}' - n_j + n_j{}'\right] = 0, \quad (6)$$

$$\left[\frac{n_i z_1}{E}a_i + \frac{n_i z_2}{G}b_i + \frac{n_j z_1}{E}a_j + \frac{n_j z_2}{G}b_j - \frac{n_i}{E}a_k - \frac{n_j}{G}b_k - n_k + n_k'\right] = 0. \quad (7)$$

The components of the displacement function, $\mathbf{s}(\cdot)$, are the unknowns in the above equations. Although the displacement function has nine unknowns, Eq. (3) can be used to solve for the constant terms. Using (6) and (7) provides two independent constraints for each point correspondence, so three point correspondences are needed to solve for the six unknowns in the affine displacement model (5). Using just three point correspondences is not as reliable as the least squares minimization approach described next.

Since each small patch is assumed to experience small continuous deformations locally, the continuity constraint can be used to hypothesize a set of point correspondences or equivalently a template mapping η_R, between corresponding patches. The hypothesis point set or search area is defined as L_R with the subscript R used to indicate that the hypothesis testing is done using range or disparity data. The best match is the hypothesis which minimizes the total sum of the squared errors, given by (6) and (7) defined as $\varepsilon_{R_1}(x,y)$ and $\varepsilon_{R_2}(x,y)$ respectively,

$$\varepsilon_R(x,y) = \sum_{i \in \eta_R} (\varepsilon_{R_1}(x,y))_i{}^2 + (\varepsilon_{R_2}(x,y))_i{}^2. \quad (8)$$

The error function (8) can be minimized w.r.t each of the six unknowns, (a_i, b_i, a_j, b_j, a_k, b_k), in the affine displacement function, resulting in six equations that can be solved using the hypothesized patch correspondence(s). Once c_i, c_j, c_k are calculated using Eq. (3), the displacement function (5) for a local patch is completely specified. Using the estimated displacement function associated with each hypothesis the errors $\varepsilon_R(x,y)$ can be ranked. The matching hypothesis associated with the smallest error is assumed to represent the deformable motion.

The minimization procedure is suitable for locally continuous surfaces that have small deformations. However, it needs to be modified to include semi-fluid motion where the continuity constraint is relaxed locally. Intensity-based information is used to complement the matching process discussed above which relies primarily on local surface shape characteristics and depth analysis. The algorithm for semi-fluid motion analysis has three steps. The first step deals with determining a set of possible semi-fluid template mappings using intensity information for each local patch. The second step evaluates the error for each hypothesized correspondence provided by the template mapping and selects the most reliable semi-fluid mapping. The third step is to refine the motion parameters of the semi-fluid mapping to sub-pixel accuracy.

The neighborhood template mapping for η_R can be continuous as shown in Fig. 1(a), or quite discontinuous under semi-fluid motion as shown in Fig. 1(b). Evaluating all possible semi-fluid template mappings between corresponding neighborhoods is combinatorially explosive. Determining the most plausible semi-fluid template mapping is made more tractable by using intensity information and tracking intensity features. For the cloud motion experiments, the discriminant was used as an intensity-based differential geometry parameter [6, 9]. The intensity-based algorithm allows small local regions within the template to undergo discontinuous semi-fluid motion. Other techniques such as optic flow-based segmentation that can be adapted to robustly estimate the local semi-fluid template mapping for motion tracking are also being investigated [3]. In order to determine the best semi-fluid mapping, Eq. (8) that measures the error in the surface or range normal constraint, is evaluated for each correspondence hypothesis using the semi-fluid template maps generated in the previous step. The error $\varepsilon_R(x,y)$ is computed within a small patch neighborhood for the pixel of interest and overlapping windows are used in computing the errors (6) and (7). The reliability in selecting the appropriate template mapping is improved by evaluating the error function over the entire neighborhood of points (η_R), rather than at a single feature or individual pixel. The template mapping correspondence hypothesis with the minimum error is assumed to represent the best estimate of the local semi-fluid motion behavior. The total error $\varepsilon_R(x,y)$ can be used as a confidence measure when the underlying motion assumption is valid.

Semi-fluid motion tracking to sub-pixel accuracy is obtained by minimizing an energy function involving Gaussian curvature variations within the pixel. For a

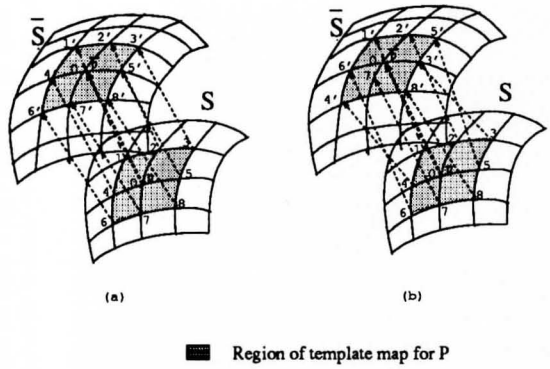

Figure 1: Local non-rigid deformation can be continuous as shown in the template mapping (a), or discontinuous under semi-fluid motion as shown in (b).

small patch defined by the vector function, $\mathbf{r}(x, y, z)$, that undergoes small deformation, the Gaussian curvatures before and after motion, K and K' respectively are related as [6],

$$K' = \frac{-1}{2D(1+\theta)}\left[\left(\frac{G + 2r_2 \cdot s_{21}}{D(1+\theta)}\right)_1 + \left(\frac{E + 2r_1 \cdot s_{12}}{D(1+\theta)}\right)_2\right]$$
$$= \frac{1}{(1+\theta)^2}K + f(x, y) \qquad (9)$$

The discriminant (or infinitesimal area) at the point of interest is D $(= EG)$, and θ is the divergence or dilation at the same point. Dilation can be defined as the unit expansion/contraction of the feature during deformation and is expressed as,

$$\theta = div\ \mathbf{s} = \frac{1}{E}\mathbf{r}_1 \cdot \mathbf{s}_1 + \frac{1}{G}\mathbf{r}_2 \cdot \mathbf{s}_2\ . \qquad (10)$$

4 Visualization Using the Interactive Image SpreadSheet (IISS)

Visualization of high dynamic-range stereoscopic data, evaluation of results from stereo analysis and semi-fluid motion analysis was accomplished using the IISS. The IISS software tool designed at Goddard for manipulating large geophysical and remote sensing datasets produced by the Earth Observing System program, offers interactive browsing, visualization and analysis capabilities for gigabytes of satellite imagery [2]. The IISS organizes (2-D and 3-D) datasets in a multidimensional spreadsheet configuration, containing cells and frames. The datasets are arranged compactly, thus providing quick access to thousands of datasets and rapidly developing complex relationships between the datasets using a formula language. The complex inter-relationships between datasets are maintained by the IISS without the necessity for complicated database systems.

The IISS environment provides a spreadsheet-based visual interface for performing satellite image analysis tasks and enables the user to concentrate on algorithm development and explore the information available in the data. In our work, the capabilities of IISS were enhanced and extended for visualizing the results of stereo and semi-fluid analysis. Extension of the traditional spreadsheet concept (pervasive on personal computers) to the image processing field has recently been possible using the current generation of moderate cost powerful graphics superworkstations with advanced image and graphics hardware architectures and a very large general purpose memory (256 MB or more) with fast access to tens of gigabyte of local disk storage. Highly interactive browsing tools, such as synchronized animation, roam and zoom, navigated data probing, surface rendering and texture mapping applied to arbitrary-sized datasets in each frame of the IISS are versatile methods for inspecting hundreds of datasets. The IISS facilitates intercomparison and remapping between different geophysical coordinate systems such as satellite, aircraft or ground-based radar. Figure 2 shows a four cell image spreadsheet, with each cell containing hurricane cloud data for a different time step overlaid with wind barbs and annotation, for evaluating the results of semi-fluid cloud tracking.

The IISS has been enhanced to display any two frames in stereoscopic mode that can be viewed using the CrystalEyes glasses from Stereographics. In the stereoscopic mode, 24-bit stereo images are refreshed at 60 frames per second for each eye, and can be animated, zoomed and roamed interactively. The capability of examining stereoscopic data using a true 3-D display capability was found to be a valuable tool for evaluating and intercomparing the results of stereo analysis with the input stereo datasets. A feature under development, for example, is the coordinated probing of stereo and motion correspondence data so that matching points from multiple frames are accessed and displayed visually while the user probes data values in only one frame. Extensions to the interactive true color electro-stereoscopic display capabilities of the IISS for examining time-varying stereo datasets include the display of estimated three-dimensional fields using a virtual reality system such as the Fakespace VR-boom system at Goddard.

5 Experiments

Experiments have been performed on both synthetic and geostationary satellite stereo data. Synthetic stereo image pairs were generated using NOAA/AVHRR data of Hurricane Andrew from August 25, 1992. The infrared channel which measures the temperature at each point of the scene can be approximately related (inversely) to the cloud heights and used to generate synthetic stereo data by shifting each pixel horizontally along the epipolar direction in proportion to the corresponding infrared value. The pixel having the coldest temperature will be shifted the maximum which was 25 pixels in the test case. The infrared image scaled to the range 0 to 25 is used as the ground truth. The output of the Goddard ASA, which produces disparities for every pixel can then be compared against the ground truth for error analysis. The error between the estimated disparity and

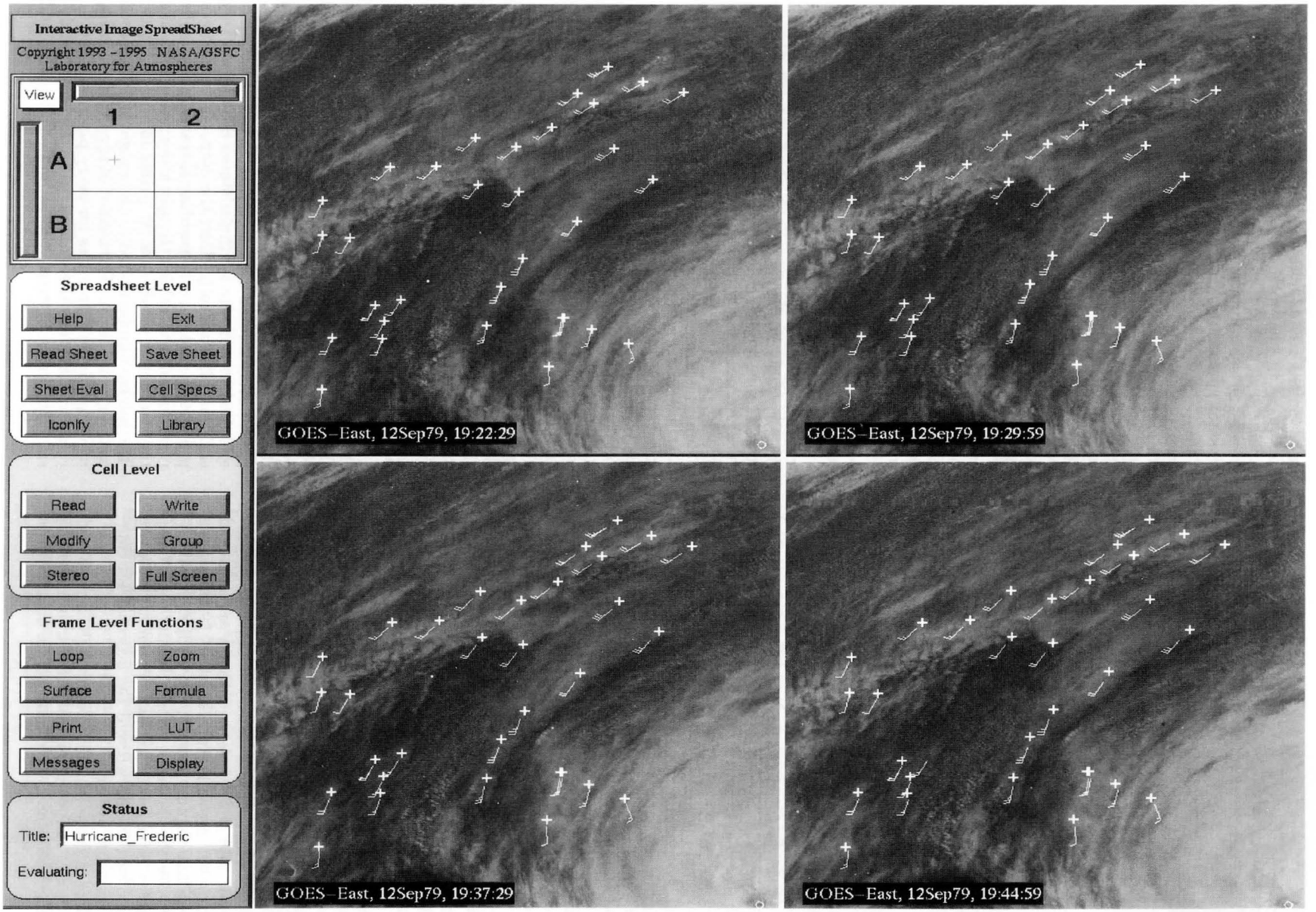

Figure 2 Visualization of semi-fluid motion estimation results for the Hurricane Frederic 22.5 min. time sequence using the IISS. The Options panel is to the left and the 2X2 cell image spreadsheet to the right. Each satellite image is labelled with the nominal UTC time of observation, the wind barbs are from manual tracking which can be compared with the plus symbols from automatic motion tracking.

the ground truth was found to be very low with a mean of -0.274, a variance of 6.31, and a standard deviation of 2.51. We have also tested the algorithm by generating different stereo pairs with different disparity ranges and have observed the mean error to be consistently near zero.

The stereo and semi-fluid motion analysis results for the Hurricane Frederic datasets acquired by two geostationary weather satellites in 1979 are summarized below. Four time sequential stereo pairs with a time interval of 7.5 minutes between each stereo pair and a lag of 15 seconds between the left and right pair were used. The small time lag can be neglected for stereo analysis since there is very little cloud motion within this small time difference. Automatic stereo analysis was first performed on the stereo pairs to estimate the disparity. The estimated depth map time sequence along with corresponding intensity images from the eastern satellite were used to perform semi-fluid motion analysis. A local quadratic surface is fitted to estimate the various differential geometric measures that were used in the algorithm. The automatic results were then compared with the manually tracked winds using cloud tracers done by expert meteorologists. Figure 2 shows the four time frames (arranged in each cell of IISS) along with the manually estimated wind barbs. The branches at one end of the wind barb represent the magnitude of speed, and the barb points in the opposite direction of the cloud motion; the wind barbs indicate the direction of the source of the wind. The relationship between a wind barb and cloud motion can be observed by following the motion of the white cross marks. These marks are produced by the semi-fluid motion algorithm and the cloud wind direction compares well with the manual estimates. One can also see how well the feature points are tagged to the cloud structure and follow them thru time. There is an average error of only 8.89 degrees between the semi-fluid motion algorithm estimates of cloud wind direction and the manually derived estimates. The amount of error is not necessarily a reflection of algorithm performance as some of discrepancies can also be attributed to human error in manual cloud tracking.

6 Conclusions and Future Directions

An integrated set of algorithms for stereo analysis and semi-fluid motion analysis with sub-pixel accuracy, for satellite-based cloud tracking has been developed. The Goddard ASA algorithm can accurately recover the complex structure of clouds given multiple satellite views. However, additional research is needed to develop automatic parameter tuning methods for selecting optimum window sizes and various thresholds for different scenes using possibly machine learning methods. The parallel implementation of the ASA algorithm enables interactive cloud structure determination for even very large satellite scenes. The automatic semi-fluid motion estimation algorithm which can produce a dense motion field produces favorable results compared to manually tracked cloud winds. In addition to affine, higher-order polynomial and nonlinear displacement functions guided by experiment can be used. A semi-fluid motion model that contains rigid motion as a special case is desirable and the automatic classification of non-rigid motion types would improve motion estimation. Related to the computational vision work was the development of new visualization tools and techniques, such as the IISS, for evaluating vision algorithms which offer a new area of study for integrating techniques in computer vision (analysis) and computer graphics (synthesis). Future work involves using illumination and multispectral information, incorporating robustness, motion segmentation and adaptive searching for improving operational cloud-tracking performance.

7 Acknowledgements

The encouragement and support of Glenn Mucklow, program manager for the Applied Information Systems Research Program (NASA NRA-93-OSSA-09) in the Office of Space Science at NASA HQ is greatly appreciated. Hal Pierce in the Mesoscale Atmospheric Processes Branch assisted in the manual tracking and assembly of stereoscopic satellite datasets.

References

[1] A. F. Hasler. Stereoscopic measurements. In P. K. Rao, S. J. Holms, R. K. Anderson, J. Winston, and P. Lehr, editors, *Weather Satellites: Systems, Data and Environmental Applications, Section VII-3*, pages 231–239. Amer. Meteor. Soc., Boston, MA, 1990.

[2] A. F. Hasler, K. Palaniappan, M. Manyin, and J. Dodge. A high performance interactive image spreadsheet (IISS). *Computers in Physics*, 8(3):325–342, 1994.

[3] Yan Huang, K. Palaniappan, Xinhua Zhuang, and Joe Cavanaugh. Optic flow field segmentation and motion estimation using a robust genetic partitioning algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence*, To be published.

[4] Chandra Kambhamettu, Dmitry B. Goldgof, and Matthew He. Determination of motion parameters and estimation of point correspondences in small nonrigid deformations. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 943–946, June 1994.

[5] Chandra Kambhamettu, Dmitry B. Goldgof, Demetri Terzopoulos, and Thomas S. Huang. Nonrigid motion analysis. In Tzay Young, editor, *Handbook of PRIP: Computer vision*, volume II, pages 405–430. Academic Press, San Diego, California, 1994.

[6] ChandraSekhar Kambhamettu. *Nonrigid Motion Analysis Under Small Deformations*. PhD thesis, University of South Florida, December 1994. Department of Computer Science and Engineering.

[7] Robert T. Merrill, W. Paul Menzel, Wayman Baker, James Lynch, and Eugene Legg. A report on the recent demonstration of NOAA's upgraded capability to derive cloud motion satellite winds. *Bull. American Meteorological Society*, 72(3):373–376, March 1991.

[8] H. K. Ramapriyan, J. P. Strong, Y. Hung, and C. W. Murray, Jr. Automated matching of pairs of SIR-B images for elevation mapping. *IEEE Trans. Geosciences and Remote Sensing*, 24(4):462–472, 1986.

[9] C. E. Weatherburn. *Differential Geometry in Three Dimensions*, volume II. Cambridge University Press, Cambridge, 1930.