Statistical Predictive Rendering for Robust Passive Relative Navigation

Michael J. Veth
Veth Research Associates, LLC
Niceville, FL 32578
USA

michael.veth@vethresearch.com

ABSTRACT

Recent advances in nonlinear feature extraction algorithms have led to significant research contributions in the area of feature aided navigation techniques (e.g., simultaneous localization and mapping), most notably in the image-aided navigation application. These feature extraction and tracking algorithms are attractive due to their ability to extract a collection of interest points in an autonomous or semi-autonomous fashion. The location of these extracted features in a sequence of images can easily be exploited to solve for pose using the principles of multiple-view geometry.

One of the limitations inherent in feature tracking approaches is the requirement for the signal-processing algorithm to make hard decisions regarding the information remaining in the observation during the signal pre-processing stages. Because these decisions are made without incorporating a priori knowledge, they result in a loss of information and, as a result, sub-optimal performance. For these reasons, a more holistic approach to extracting navigation information from a sequence of images is motivated.

In this paper, the statistical predictive rendering (SPR) technique is explored. SPR techniques seek to minimize the deleterious effects of sub-optimal pre-processing stages by using the entire observation to improve the navigation state estimate. The paper begins with an overview of the SPR algorithm, highlighting the importance of properly modeling the time-varying error statistics. The algorithm is then applied to the passive relative navigation problem encountered in autonomous air-to-air refueling. Conclusions are presented regarding the accuracy and stability of the SPR algorithm and future research directions.

1.0 STATISTICAL PREDICTIVE RENDERING

The Statistical Predictive Rendering (SPR) algorithm is designed to present an alternative to feature-based image tracking algorithms by leveraging known shape and textural information in a holistic fashion. In addition, the algorithm is designed to use the existing library of high-performance graphical rendering algorithms and associated accelerated hardware (e.g., graphical processing units). The SPR algorithm consists of three main components: predictive rendering, comparison, and statistical back projection.

1.1 Predictive Rendering

In the most general terms, rendering is the process of generating an expected sensor observation using a model of the world and of the sensor itself. In most cases, the sensor produces a multi-dimensional observation. One of the most common forms of rendering is generating images using computer graphics libraries such as OpenGL [1] and DirectX. Because of the natural application of predictive rendering techniques to computer vision applications, the predictive rendering discussion in this paper will focus on the image rendering problem.

In order to render an image, both a model of the world and the camera are required. The world consists of all objects and media that generate or transform energy in the wavelengths detectable by the camera sensor. These combine and interact to create an irradiance pattern, or *scene*, on the surface of the camera lens [2]. The camera uses lenses to focus the *image*, on the surface of the detector array. In a digital camera, the detector array measures the light intensity on each pixel and produces an array of digital numbers known as the sampled image. A typical digital camera imaging configuration is shown in Figure 1.
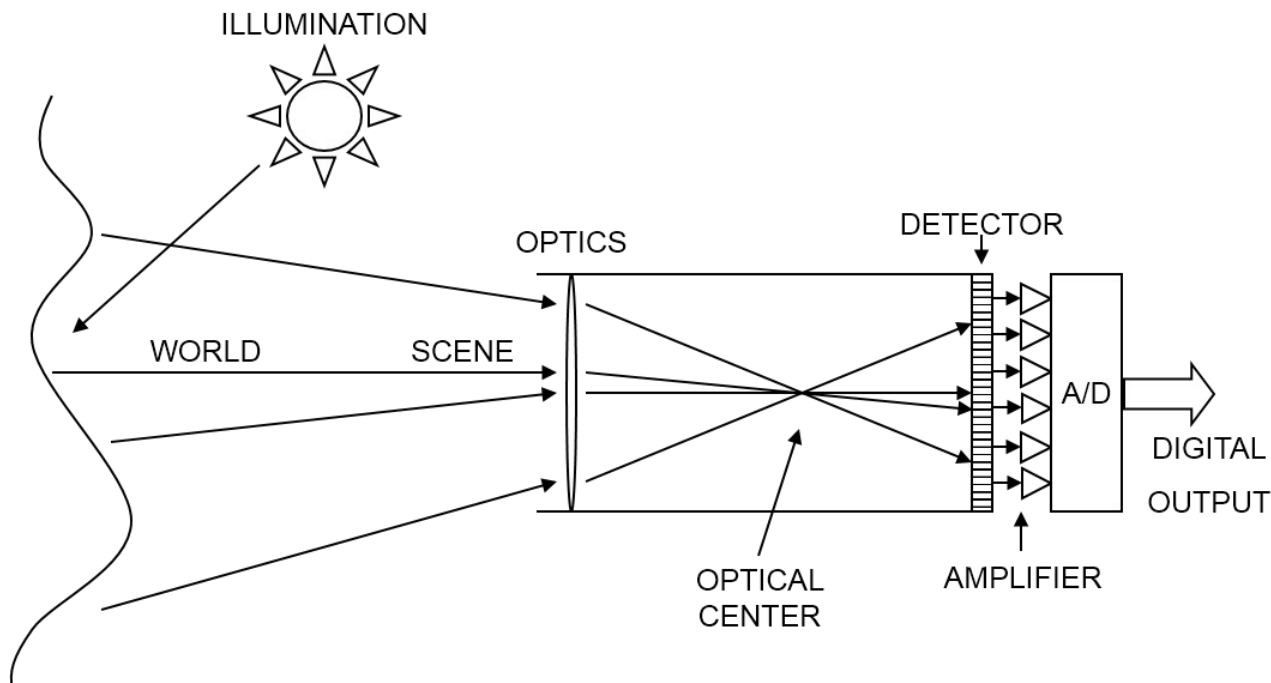

Figure 1: Typical Digital Imaging Model. The pattern of light energy entering the aperture is focused on the detector array to produce a digital image.

Each pixel location is mapped to a direction in the world frame. In general, this relationship is nonlinear, however a calibration can be performed to map the relationship to a projective transformation known as the pinhole camera model. The pinhole camera model consists of both intrinsic (specific to the camera) and extrinsic (specific to the imaging geometry) components. Additional details can be found in [3].

For the purposes of this article, it is sufficient to model the image at time $I_k(\cdot, \cdot)$, as an $M \times N$ array of k, pixels which is mapped via the observation function

$$\mathbf{I}_k = \mathbf{h}\left[\mathbf{x}_k, \mathbf{m}_k\right] \tag{1}$$

where $\mathbf{I}_k$, the image at time k, is $\mathbf{x}_k$ the state vector, is the $\mathbf{m}_k$ model of the scene and camera, and is the $\mathbf{h}\left[\cdot, \cdot\right]$ vation function. The state vector contains the parameters of interest. For navigation applications, the position and orientation of the camera are common. The model vector consists of the parameters that are known *a priori*.

Unfortunately, the model of the scene and camera are not known perfectly. As such, a statistical error component must be added to the model:

$$\mathbf{I}_k = \mathbf{h}\left[\mathbf{x}_k, \mathbf{m}_k, \mathsf{v}_k\right] \tag{2}$$

where v_k represents the random errors present in the scene and camera. These errors manifest themselves as differences in pixel intensity and can be caused by many factors, including: unmodeled lens distortions, unmodeled or random effects in the detection and sampling portion of the camera, unmodeled illumination sources and types, errors in the texture or structure of objects in the scene, and many others.

In the presence of a large number of errors, an additive Gaussian model is justified based on the properties of the Central Limit Theorem. Revising (2) yields

$$\mathbf{I}_k = \mathbf{h}\left[\mathbf{x}_k, \mathbf{m}_k\right] + \mathsf{v}_k \tag{3}$$

where v_k is a Gaussian random vector. Without loss of generality v_k is assumed to be zero-mean.

1.2 Comparison

Once a statistical observation function is established, the likelihood of a given image, conditioned on the state vector and model, can be calculated. Given the additive Gaussian model (3), and assuming each pixel's errors are uncorrelated, the conditional probability of a single pixel location is

$$p(\mathsf{z}_k(i,j)|\mathbf{x}_k, \mathbf{m}_k) \propto \exp\left\{-\frac{1}{2}\left(\frac{\mathsf{z}_k(i,j) - \mathbf{h}[\mathbf{x}_k, \mathbf{m}_k](i,j)}{\sigma(i,j)}\right)^2\right\} \tag{4}$$

where $\sigma_k(i,j)$ is the standard deviation of the pixel rendering and measurement errors at time k. The combined likelihood of all pixels yields

$$p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{m}_k) \propto \exp\left\{ -\frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(\frac{z_k(i,j) - \mathbf{h}[\mathbf{x}_k, \mathbf{m}_k](i,j)}{\sigma(i,j)} \right)^2 \right\} \tag{5}$$

The pixel standard deviation is clearly dependent on the model and must be carefully chosen to properly represent the expected variation at each pixel location. This will be discussed later in the article.

1.3 Statistical Back Projection

The statistical back projection step transforms the information available in the observation back into the state space. Applying Bayes' rule to the measurement likelihood (5) yields

$$p(\mathbf{x}_k|\mathbf{z}_k, \mathbf{m}_k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{m}_k)p(\mathbf{x}_k|\mathbf{m}_k)}{p(\mathbf{z}_k|\mathbf{m}_k)} \tag{6}$$

It can be shown that the *a priori* state density $p(\mathbf{x}_k|\mathbf{m}_k)$ is independent of the model at time *k*, thus:

$$p(\mathbf{x}_k|\mathbf{m}_k) = p(\mathbf{x}_k) \tag{7}$$

In addition, the evidence density $p(\mathbf{z}_k|\mathbf{m}_k)$ can be expressed equivalently using a marginalization or $\mathbf{x}_k$:

$$p(\mathbf{z}_k|\mathbf{m}_k) = \int p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{m}_k)p(\mathbf{x}_k)d\mathbf{x}_k \tag{8}$$

Substituting (7) and (8) into (5) yields the back projection equation:

$$p(\mathbf{x}_k | \mathbf{z}_k, \mathbf{m}_k) = \frac{p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{m}_k) p(\mathbf{x}_k)}{\int p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{m}_k) p(\mathbf{x}_k) d\mathbf{x}_k} \tag{9}$$

In general, the *a posteriori* density is non-Gaussian and multi-modal, thus nonlinear estimation algorithms are usually required to implement the SPR technique. In the next section, an example application of the SPR algorithm is presented and analyzed.

2.0 EXAMPLE APPLICATION – AUTONOMOUS AERIAL REFUELING

Aerial refueling is a critical component of modern airpower and is a key enabler of force projection. Previous image-aided approaches have been demonstrated [4]. The majority of these algorithms utilize pre-selected, highly identifiable portions (or visually augmented locations) of the tanker to drive feature trackers. The location of these features in the image can be used to calculate the relative pose [5].

The SPR algorithm is focused on determining the position and orientation of an imaging sensor relative to a known object of interest. As mentioned previously, the SPR algorithm uses a holistic approach and incorporates each pixel in the image, thus eliminating dependency on local features. This makes SPR well-suited to the autonomous aerial refueling (AAR) problem [6],[7].

Two primary methods are used for aerial refueling systems: probe and drogue systems, and boom / receptacle systems. In each case, the aircraft receiving fuel (receiver) is required to fly and maintain in close trail formation with the tanker aircraft. The relative position accuracy required for drogue systems is on the order of a decimeter. Boom / receptacle systems are less stringent and require relative position accuracy on the order of 60 decimeters [7]. An illustration of the KC-135 refueling boom range of motion is shown in Figure 6.
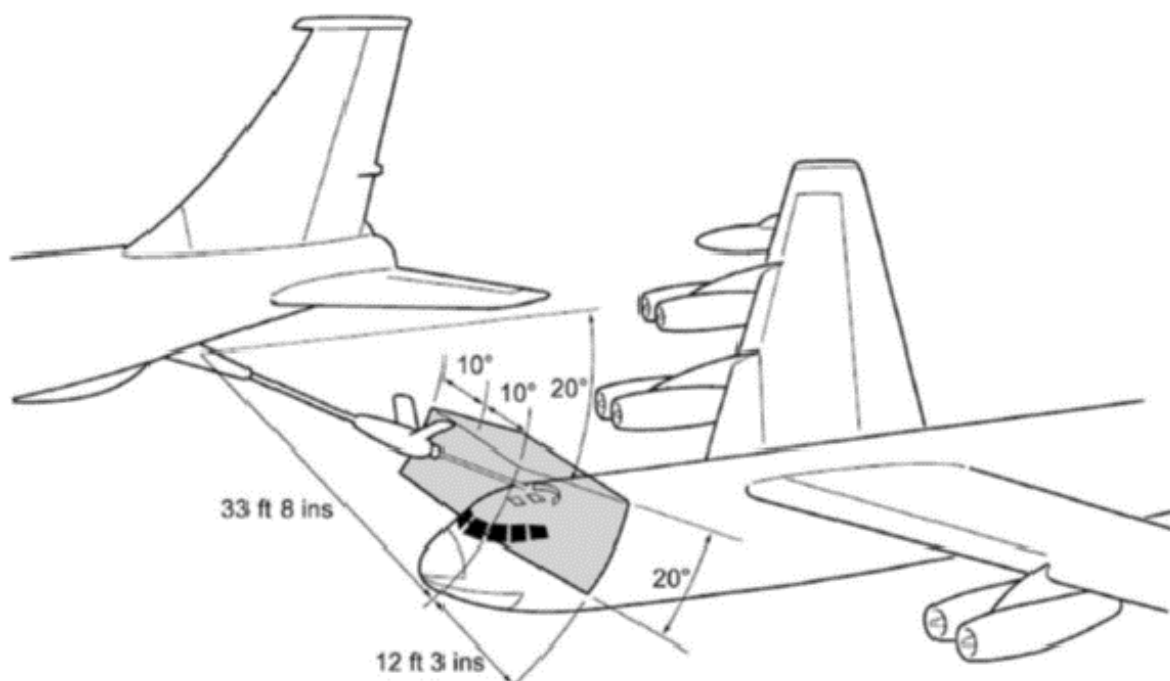
2.1 Simulation and Experiment

The use of the SPR algorithm for the aerial refueling problem has been demonstrated in research conducted by the Air Force Institute of Technology's (AFIT) Advanced Navigation Technology (ANT) Center. Two research efforts from the ANT Center are reviewed in this article.

2.1.1 Weaver and Veth – Observability Predictions and Analysis for the Aerial Refueling Scenario

In [6], the authors present the foundations of the relative navigation problem associated with aerial refueling and conduct a sensitivity analysis over variations of the SPR algorithm using a KC-135 target model. The article demonstrates that the SPR observations for aerial refueling are possible by using experimental image data from the Air Force Research Laboratory's (AFRL) Autonomous Aerial Refueling (AAR) flight test program.

The first step in predicting the observability of the SPR algorithm is defining the aircraft model and the geometry required to relate this model into the camera's reference frame. The three-dimensional KC-135R *Stratotanker* aircraft model was rendered using the Matlab Virtual Reality (VR) Toolbox. The VR Toolbox uses the OpenGL framework to render scenes directly from the Matlab environment. An example rendering of the aircraft model is shown in Figure 3. Note the rendering environment was configured to exclude textural information for this article. In general, textural information should be used if available and reliable.

Figure 3: Sample rendering of the KC-135R from the trailing aircraft's perspective. The model is rendered using the Matlab Virtual Reality Toolbox.

Accurately predicting the camera view requires defining the intrinsic and extrinsic camera parameters. The intrinsic camera parameters are determined using a camera calibration procedure described in [8]. Once nonlinear distortion effects are removed from the images, the intrinsic camera model relates directions in the camera axes to pixel locations in the image using an affine transformation. In computer vision applications, it is common to use a homogeneous coordinate system which effectively remaps affine transformations to the space of linear operators [3]. Given line of sight from the camera to the target is the three-dimensional $\mathbf{s}^c$ vector , the homogeneous pixel location, $\underline{z}$, , is linearly related to the intrinsic camera matrix as follows

$$\underline{z} = \mathbf{P}_i \underline{s}^c \tag{10}$$

where $\mathbf{P}_i \in \mathbb{R}^{3 \times 4}$ is the intrinsic camera matrix. The underline symbol indicates the homogeneous form of vector.
       the

The relationship of the camera reference frame to the position of the lead aircraft is specified by the extrinsic camera matrix. The extrinsic camera matrix is calculated using a physical boresighting procedure once the camera is installed in the trailing aircraft. Given the relative position of the lead aircraft model in the wing aircraft's reference frame, $\mathbf{p}_{wl}^w$, the homogeneous line of sight vector in the camera frame is

$$\underline{s}^c = \mathbf{P}_e \underline{p}_{wl}^w \tag{11}$$

where $\mathbf{P}_e \in \mathbb{R}^{4 \times 4}$. Combining (10) and (11) yields the linear projection function

$$\underline{z} = \mathbf{P}_i \mathbf{P}_e \underline{p}_{wl}^w \tag{12}$$

As mentioned previously, this imaging model is standard in the computer vision and computer rendering fields and is directly supported by off-the-shelf rendering frameworks such as OpenGL. This ease-of-implementation is one of the advantages of the SPR algorithm.

Once the rendering algorithm is initialized, an observability analysis is conducted. The observability analysis is designed to demonstrate the sensitivity of changes in the state vector on the observation. This is accomplished by manually sweeping the state vector about the true state and recording the resulting set of likelihood functions. In equation form, the $\tilde{\mathbf{x}}_k$ perturbed state vector is defined as

$$\tilde{\mathbf{x}}_k = \mathbf{x}_k + \delta\mathbf{x}_k \tag{13}$$

where $\mathbf{x}_k$ is the true state vector and $\delta\mathbf{x}_k$ is the perturbation vector. Substituting (13) into the likelihood equation (5) yields the scaled perturbation likelihood function

$$L(\mathbf{z}_k|\mathbf{x}_k + \delta\mathbf{x}_k, \mathbf{m}_k) = k\exp\left\{-\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{\mathrm{z}_k(i,j) - \mathbf{h}[\mathbf{x}_k + \delta\mathbf{x}_k, \mathbf{m}_k](i,j)}{\sigma(i,j)}\right)^2\right\} \tag{14}$$

The proportionality relation can be changed to an equality relationship without loss of generality as the proportionality constant k is independent of the perturbation $\delta\mathbf{x}_k$.

The sensitivity analysis is conducted for both the pre-contact and contact refueling positions. A typical image from the contact position is shown in Figure 4.
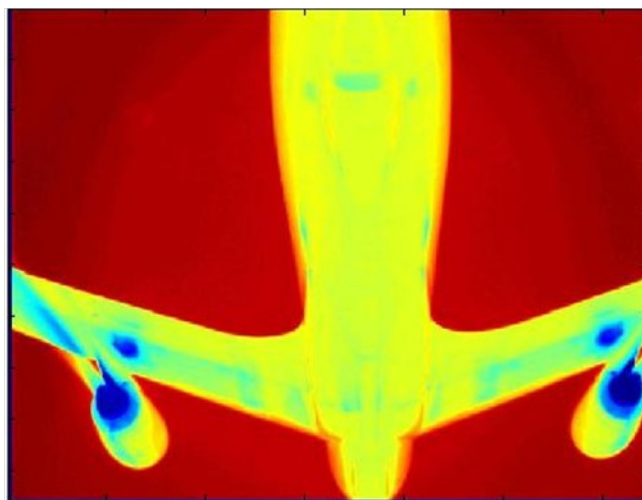

Figure 4: Sample image of a KC-135R from the contact position. The image was captured in flight using an infrared camera.

The predicted image is compared to the observed image by perturbing the state space. An example of the predicted and observed image overlay is shown in Figure 5.
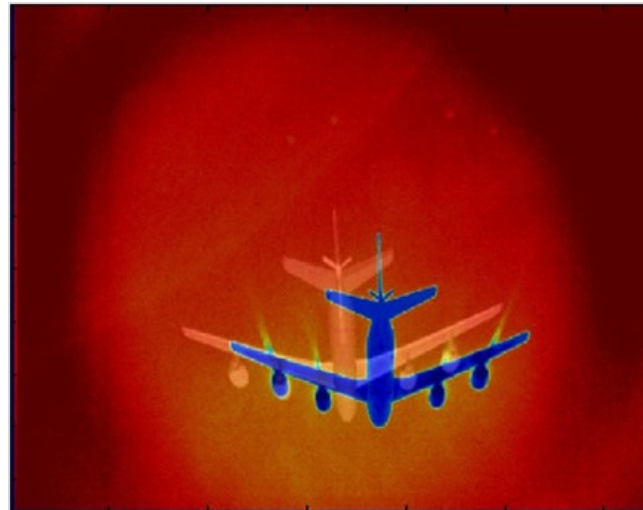
Figure 5: An example of the predicted image overlayed on the observed image showing the effects of a perturbation about the true state vector in the pre-contact position. From [6].

The inverse likelihood is calculated for a representative set of perturbations. In Figure 6, the effects of perturbation in the translation states are shown for the pre-contact position. The inverse likelihood shows a clear global minimum for each axis. This is a desirable characteristic and is indicative of excellent global observability. The apparent bias error between the minimum points on each curve is due to errors in the camera installation and boresighting and errors in the aircraft 3D model.
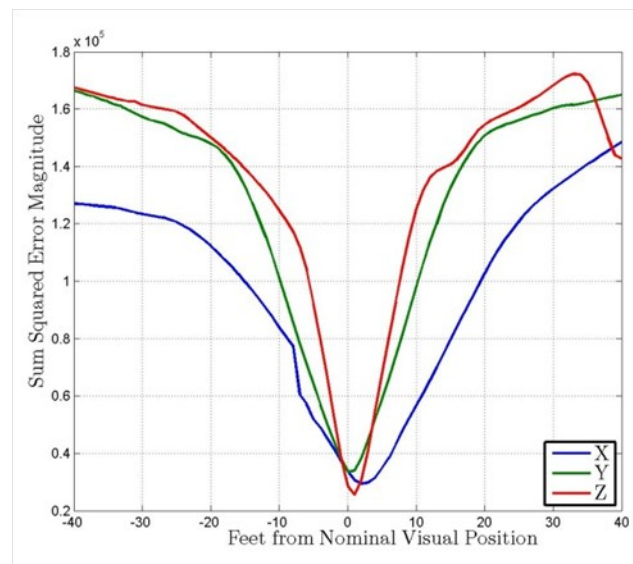

Figure 6: Inverse likelihood of translation perturbations from the pre-contact refueling position.

A similar result is seen for perturbations in the roll, pitch, and yaw states is shown in Figure 7. Again, all axes show a clear minimum error and smooth increases as the states are perturbed about the true.
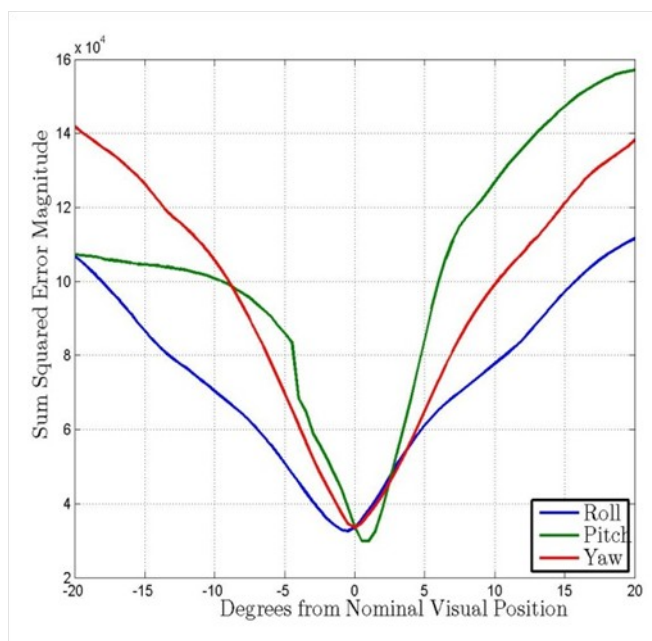

Figure 7: Inverse likelihood of roll, pitch, and yaw perturbations from the pre-contact refueling position.

Further investigation into the sensitivity of the SPR approach illustrates the effects of correlated perturbation states. Two examples are shown. The effects of coupled perturbations in the *x* translation axis (out the nose of the wing aircraft) with roll perturbations are shown in Figure 8. In this example, the inverse likelihood surface shows desirable characteristics with a clear minimum over a large neighborhood.
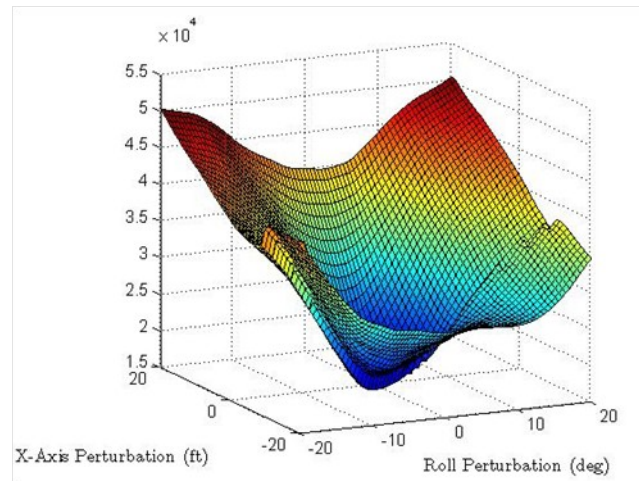
Figure 8: Inverse likelihood of coupled roll and x-translation perturbations. Note the surface shows a clear minimum with smooth characteristics over a large neighborhood. This is indicative of good observability.

Analyzing the inverse likelihood surface for coupled pitch and *x*-axis translation shows a different characteristic (Figure 9). In this case, note the very shallow gradient of the surface about the minimum. This indicates limited observability between the states. This condition makes intuitive sense as small pitch changes are difficult for human pilots to discern from longitudinal changes.
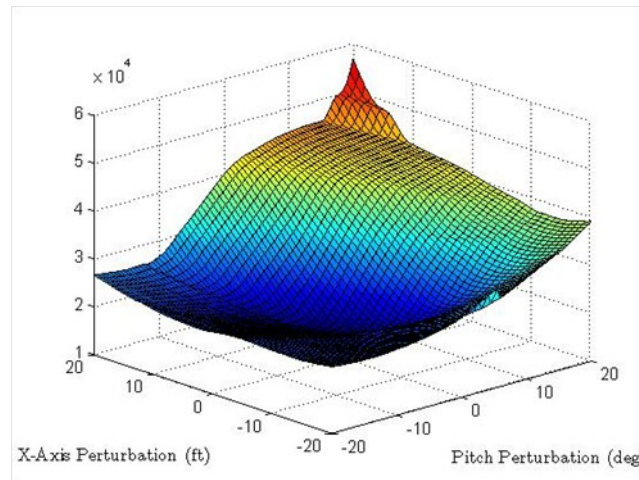

Figure 9: Inverse likelihood of coupled pitch and x-translation perturbations. The surface, although smooth, does not show a distinct minimum. This is indicative of relatively poor observability between pitch and longitudinal excursions.

In summary, Weaver demonstrated the feasibility of the SPR algorithm to support a completely passive, aerial refueling relative navigation system. Using experimental flight test data, his research demonstrated the observability characteristics of the sensor and served as the foundation for future work. One of those resulting efforts (Howard and Veth) will be presented in the next section.

2.1.2 Howard and Veth – Aerial Refueling Experiment Using SPR

In [7], the authors continue the work of Weaver by incorporating the SPR algorithm into an estimator tasked with solving for the relative position and orientation of an aircraft in a refueling position. The integrated algorithm was tested in conjunction with the US Air Force Test Pilot School using instrumented test aircraft. The relative navigation algorithm was implemented using online procedures (causal), however all processing was conducted post-flight.

The algorithm utilizes a two-stage SPR matching algorithm to improve performance. The predicted image is created at the observation time using *a priori* state information from the previous state and measurements from the inertial navigation system. While both stages use a greedy hill-climbing algorithm, the attitude states are perturbed first. The attitude perturbations corresponding to the maximum likelihood are then fixed while the translational states are varied. The resulting state vector corresponding to the maximum likelihood is used as a Kalman Filter measurement or simply used as the *a posteriori* state. An example of the algorithm is shown in Figure 10.
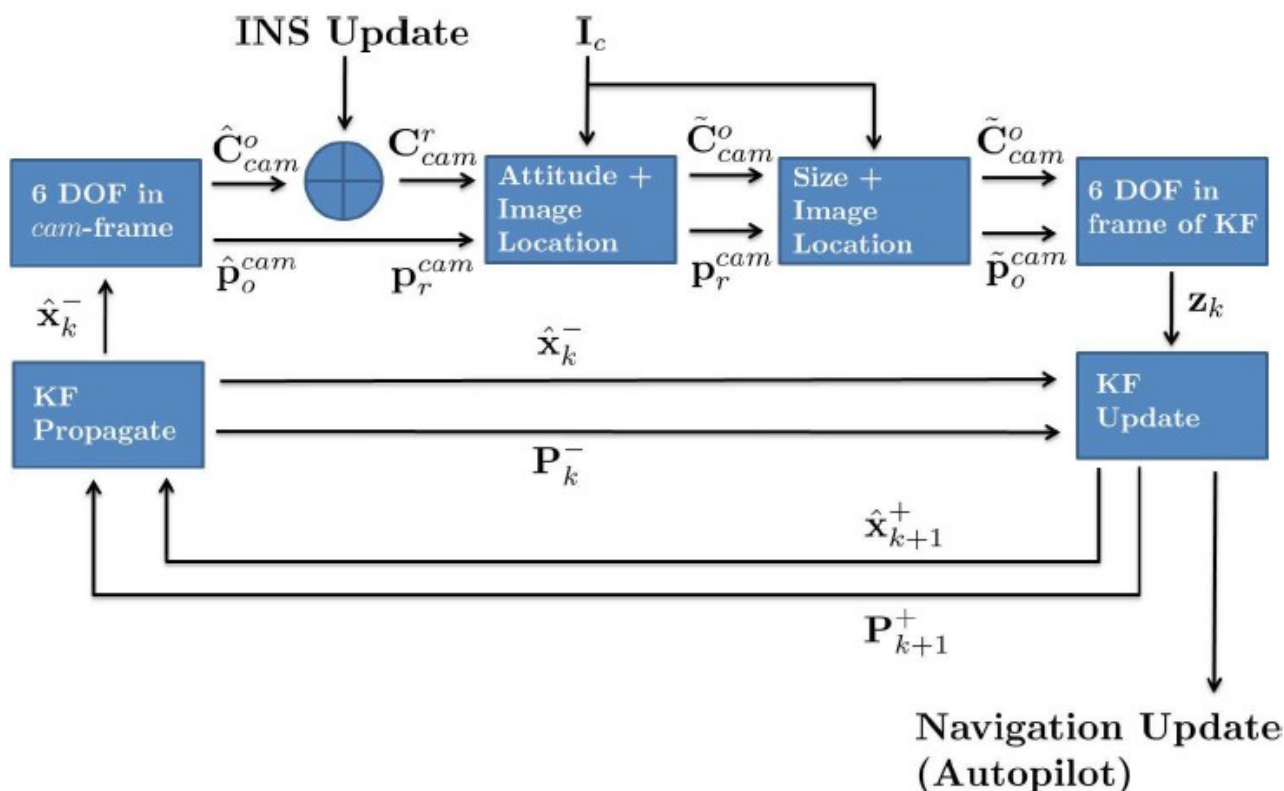

Figure 10: In [7], the SPR update is accomplished using a partitioned likelihood maximization algorithm. The resulting maximal state vector can be used as an update in a Kalman Filter algorithm, if desired.

The flight test was conducted using two aircraft. A Learjet LJ-24 was used as the simulated receiver aircraft and was equipped with a Honeywell HG-1700 inertial measurement unit and a Prosilica machine vision camera. Due to aircraft availability, a T-38 Talon was used as a surrogate for the tanker aircraft. A three dimensional model was developed for the T-38 aircraft using a laser scanner and photos collected from a calibrated camera. A sample rendered view of the T-38 aircraft model is shown in Figure 11.
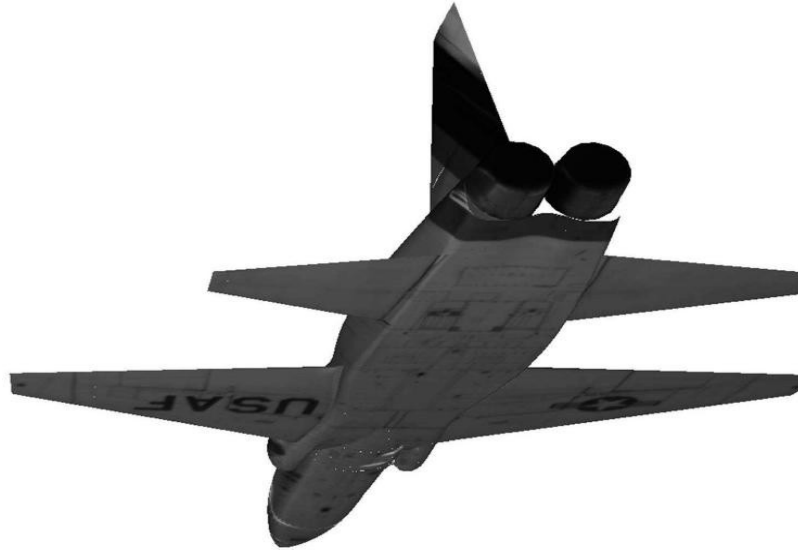

Figure 11: Example rendering of the T-38 aircraft model. The model is stored using the Wavefront Object Model format.

Image and inertial data were recorded in-flight during representative refueling formations. Images were collected at 10 *Hz* from a camera mounted in the cockpit of the Learjet. Following each mission, distortion was removed from the images using the calibrated camera model. The SPR-based relative navigation solution was processed using the algorithm in Figure 10. The results shown in this article do not include the Kalman Filter in order to more purely represent the quality of the raw SPR measurements. The SPR relative navigation solution was compared to the true solution which was calculated using an independent differentially-corrected GPS/INS system located on each aircraft. An example of a successful template match is shown in Figure 12.
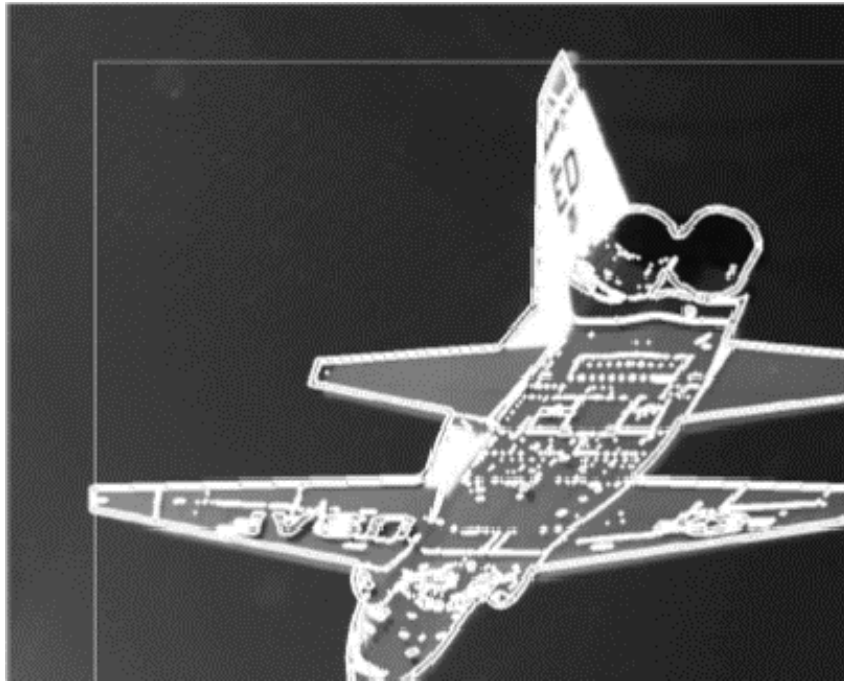
**Figure 12: After the SPR algorithm converges on a
result, an edge-enhanced version of the rendered
image is overlayed on top of the observed image.
This is an example of a successful match.**

Two imaging scenarios are presented. The first scenario represents a typical approach from the pre-contact position to the contact position under normal imaging conditions (i.e., good background contrast, no lens flare or imaging into the sun, etc.). The results are shown in Figure 13. Note that the position errors in all axes are maintained within one meter and roll errors are less than three degrees.
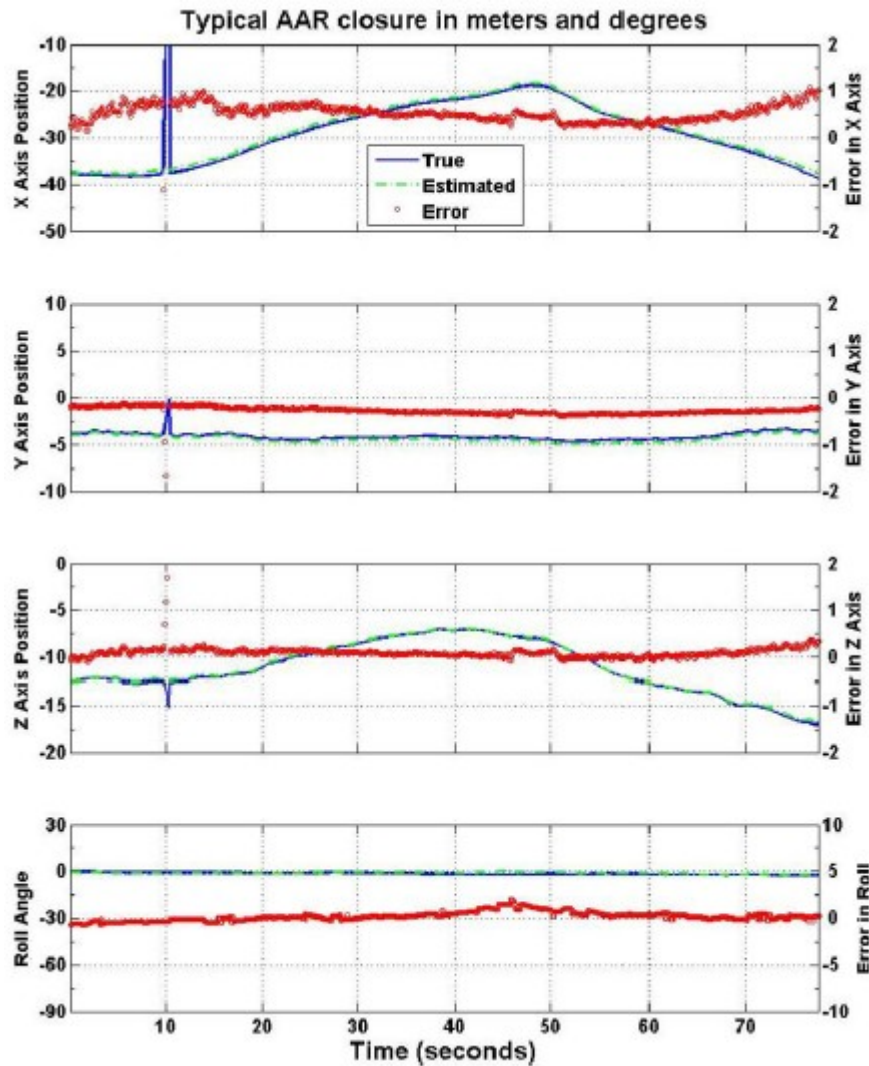
Figure 13: Results of the SPR algorithm for a typical rendezvous from pre-contact to contact position and back out. The spike in the truth data at approximately 10 seconds is due to a momentary data drop out.

The second scenario represents a challenging imaging situation with extreme distortion caused by lens flare due to the sun. An example of the images used during this scenario are shown in Figure 14. As the images degrade in quality toward the end of the run, the errors increase most notably in the *x*-axis. Maximum errors were on the order of three meters. It is noted by the author of the paper (an aerial-refueling qualified test pilot) that the sun conditions during this test were at the extreme edge of what is possible by human pilots. The results are shown in Figure 15.
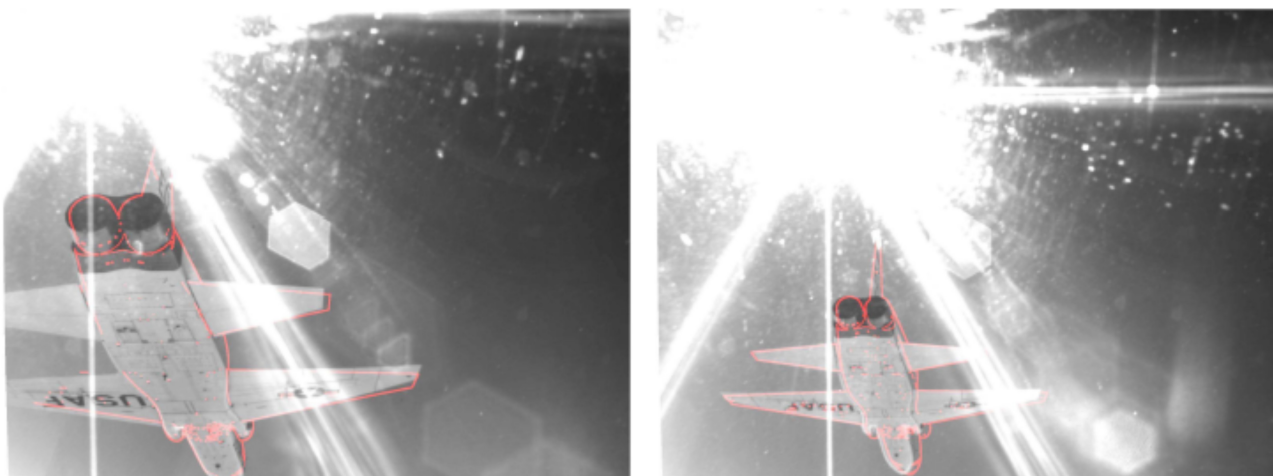
Figure 14: Example images from the challenging rendezvous scenario. The lens flare caused by the sun's position provided a challenge to the SPR algorithm.
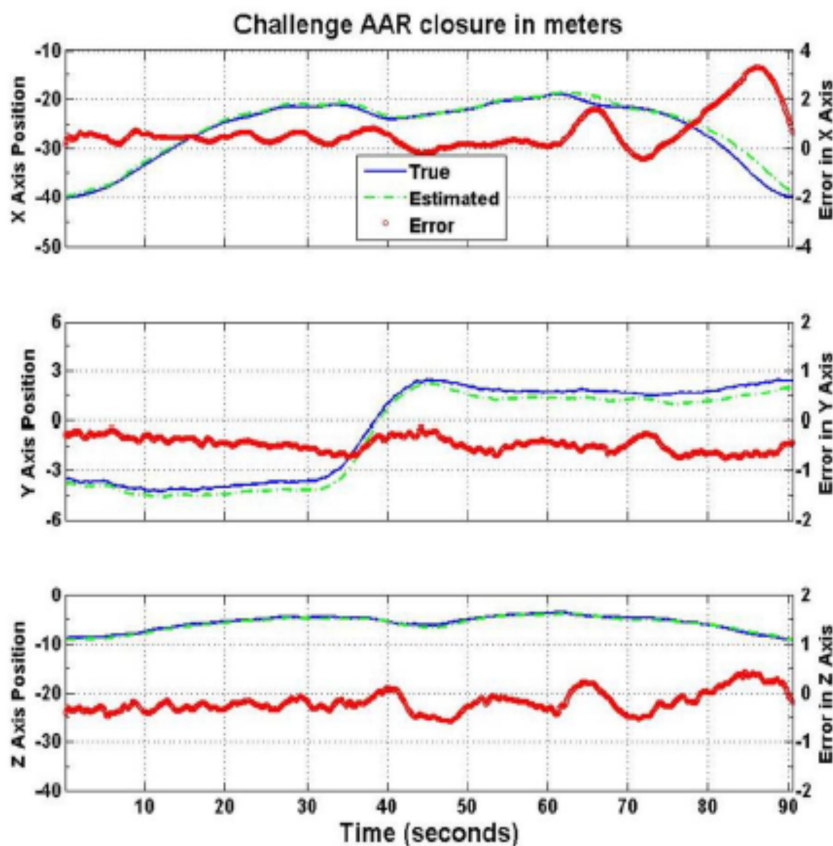

Figure 15: Results from the challenging refueling scenario. In this case, the rendezvous was staged so that the sun was directly in the camera's field of view. The result was an extremely challenging rendezvous.

In summary, the SPR algorithm was successfully demonstrated in a fully-instrumented flight test environment. During the pre-contact and contact portions of the rendezvous, position errors were approximately 3% during typical refueling conditions. This level of performance is at a similar, or better, level than that achievable by human pilots, thus showing the feasibility of this algorithm for autonomous operations.

3.0 CONCLUSIONS AND FUTURE WORK

Autonomous aerial vehicles are rapidly becoming a necessary component of modern airpower. The unparalleled combination of persistence and maneuverability offered by these aircraft will truly change the way air forces are employed.

In this article, the foundations of a completely passive approach to autonomous aerial refueling is presented. The statistical predictive rendering algorithm is presented and the results of two flight test efforts evaluated. In both cases, the algorithm is successful at determining the relative pose of a pre-defined model in real-world conditions. The SPR algorithm demonstrated relative navigation performance at the level of a human pilot, even under very difficult imaging conditions.

While the SPR algorithm shows initial promise, a number of additional research areas are necessary. First, the characterization of the likelihood function should be investigated further. In each of the above examples, only the simplest model (independent, identically distributed Gaussian) was used. This assumption does not properly represent the effects of changes of illumination intensity, reflections from light sources, or unmodeled variations in the vehicle shading. In addition, the uncertainty should be varied across the model to properly represent areas of high versus low confidence.

The second improvement is an effort to properly couple the SPR algorithm into a recursive, particle filer-based estimator. The particle filter shows promise due to its ability to accurately maintain higher-order statistical effects. The main issue with particle filter implementation is the high number of particles necessary to represent a six-dimensional state space.

Finally, additional image pre-processing stages should be considered. Weaver's investigation into a magnitude of gradient pre-filter showed the potential for higher accuracy, although, in his experiments there was a corresponding reduction of the smoothness of the likelihood manifold [6]. This would impact the convergence of the estimation algorithm. Previous computer vision research into image matching has shown benefits when using a Laplacian, Difference of Gaussian, or wavelet filters. These could hold promise for the SPR algorithm as well.

4.0 BIBLIOGRAPHY

1.   Shreiner, Dave, *OpenGL Programming Guide*. Upper Saddle River, NJ: Pearson Education, Inc., 2009.

2.   Jain, Anil K., *Fundamentals of Image Processing*. Upper Saddle River, NJ: Prentice Hall, Inc., 1989.

3.   Hartley, Richard and Andrew Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge: Cambridge University Press, 2003.

4.   Calise, Anthony J., James C. Neidhoefer, Yoko Watanabe Jincheol Ha, and Eric N. Johnson, "Real-Time Vision-Based Relative Aircraft Navigation". *AIAA Journal of Aerospace Computing Information*, 4(4):707-738, 2007.

5.   Drummond, Tom and Roberto Cipolla, "Real-time Visual Tracking of Complex Structures". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):932-946, 2002.

6.   Weaver, Adam and Michael Veth. "Image-Based Relative Navigation for the Autonomous Refueling Problem Using Predictive Rendering". *Proceedings of the IEEE Aerospace Conference*, 2009.

7.   Howard, J. Michael and Michael J. Veth. "Image Aided Relative Navigation for Air Vehicles Using a Passive, Statistical Predictive Rendering Approach". *Proceedings of the ION GNSS 2011*, 2011.

8.   Heikkila, Janne and Olli Silven. "A Four-Step Camera Calibration Procedure With Implicit Image Correction". *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 22(10):1106-1112, 1997.