The Extended-Image Tracking Technique Based on the Maximum Likelihood Estimation

Haiping Tsou and Tsun-Yee Yan
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

ABSTRACT

This paper describes an extended-image tracking technique based on the maximum likelihood estimation. The target image is assumed to have a known profile covering more than one element of a focal plane detector array. It is assumed that the relative position between the imager and the target is changing with time and the received target image has each of its pixels disturbed by independent additive white Gaussian noise. When a rotation-invariant movement between imager and target is considered, the maximum likelihood based image tracking technique described in this paper is a closed-loop structure capable of providing iterative update of the movement estimate by calculating the loop feedback signals from a weighted correlation between the currently received target image and the previously estimated reference image in the transform domain. The movement estimate is then used to direct the imager to closely follow the moving target. This image tracking technique has many potential applications, including free-space optical communications and astronomy where accurate and stabilized optical pointing is essential.

Keywords: Optical pointing; Extended-image tracking; Image stabilization; Maximum likelihood estimation; Optical communications

1. INTRODUCTION

Free-space laser communication is a promising technology to meet the demand of cost-effective broadband support for future NASA missions by allowing spacecraft communication subsystems with lower mass and power consumption.\(^1\)

In order to prepare for development and validation of various deep-space laser communication technologies, JPL has begun construction at its Table Mountain Facility in California of a 1-meter telescope system, known as the Optical Communication Telescope Laboratory and capable of tracking spacecraft from low Earth orbit to deep space. One of these vital technologies is the acquisition and tracking of laser signals to ensure accurate pointing of the receiving telescope to the laser transmitter onboard the spacecraft for an optical space-to-ground link (referred to as an optical downlink) which may be extended as long as several astronomy units. For those deep-space applications, it is not uncommon that the required pointing accuracy is on the order of microradians to ensure the quality of an optical downlink in the presence of many factors that cause pointing error. For example, uncompensated platform motion and/or jitter, sensor noise and/or bias, and atmospheric propagation effects, such as image dancing, blurring, and scintillation, can all contribute to the error in optical pointing.\(^2,3\)

This paper presents an optical downlink acquisition and tracking technique for the fine-pointing control of a ground-based optical terminal. Figure 1 shows the functional block diagram of a typical receiving system of a ground-based optical terminal. It includes a wide-angle acquisition-aid telescope which is aligned with a much larger main telescope used to receive the downlink laser signals. The co-aligned telescopes are driven by a coarse-pointing controller which is capable of keeping the targeted spacecraft within the FOV of the main telescope. The downlink laser beam coming off the mirror assemblies of the main telescope is then directed to the optical detector by a two-axis fast-steering mirror. By controlling this fast-steering mirror through a separate fine-pointing controller, the incoming laser beam can be tracked and the laser spot is centered at the optical detector to ensure optimal reception.

\(^*\) The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology under contract with the National Aeronautic and Space Administration.
The commonly used fine-pointing techniques involve the laser image acquisition and tracking on the focal plane array and can be categorized into two types: the centroid-type algorithm and the correlation-type algorithm. The former calculates the image centroid solely from the detected laser intensity profile, however, the latter performs a detailed feature comparison between the detected intensity profile and the reference profile which may be pre-stored or synthesized. In general, the centroid-type techniques work best if the acquired laser spot is small and has high a signal-to-noise ratio. On laser spot extended several pixels on the focal plane array, the correlation-type techniques perform better, even with low signal-to-noise ratios. The proposed technique is based on a correlation-type extended-source images acquisition and tracking technique developed for optical transceivers onboard the spacecraft. To adopt the extended-source model for the downlink when the laser beam is usually characterized as from a point-source is justified by the fact that, due to atmospheric propagation, diffraction and other system imperfections, the laser spot on the focal plane can cover more than one pixel of the detector array like an image of an extended-source. Furthermore, for data detection, it is a common practice to intentionally defocus the incoming laser beam to ensure that at least a portion of the detector array can capture the signal in the presence of optical turbulence.

The development of this extended-image acquisition and tracking scheme is based on the maximum likelihood estimation, where the uncertainties between the received image and the reference image are modeled as independent additive white Gaussian disturbances. It is also assumed that the relative position between the spacecraft and the ground-based optical terminal is changing with time, but only in a rotation-invariant fashion on a 2-D plane. As suggested by this study, the received laser image obtained from a detector array is first correlated in the transform domain with the reference image derived from the priorly known or synthesized intensity profile of a distant point source. The coordinate of the onboard optical transmitter is then estimated and tracked using the developed open-loop acquisition and closed-loop tracking algorithms, respectively. The optimal acquisition requires solving two nonlinear equations to estimate the coordinate. A suboptimal estimate exists by solving linearized version of the maximum likelihood criterion when computation complexity becomes a concern. A closed-loop tracking algorithm motivated by the maximum likelihood criterion is developed as well for continuously tracking of the translation movement between the spacecraft and ground terminal. The closed-loop tracking structure has its loop feedback signals formulated as weighted transform-domain correlations between the received laser image and the previously
estimated reference image. Only two linear equations are involved in each iteration of a continuous tracking. The closed-loop tracking scheme has many potential applications, including free-space optical communications and astronomy where accurate and stabilized optical pointing is important. This scheme is expected to be able to achieve sub-pixel resolutions in a high disturbance environment.

In this paper, section 2 provides a general description of the mathematical model and the discrete Fourier transform of an image using the lexicographic representation. The effect of rotation-invariant translation movement in the transform domain is also discussed there. Section 3 lays the groundwork for the maximum likelihood estimate of the translation vector, which leads to the open-loop acquisition algorithm. The derivation of the closed-loop tracking scheme is provided in section 4, followed by the summary of this study and a brief description of possible future research efforts in section 5.

2. MATHEMATICAL BACKGROUND

2.1. Representation of an Extended Image

The image of incoming laser beam detected by an \( M \times N \) focal plane array at time \( t_l \), denoted as \( r_l(m, n) \), can be represented by a sum of the source image, \( s_l(m, n) \), and the random disturbance, \( n_l(m, n) \), as follows

\[
r_l(m, n) = s_l(m, n) + n_l(m, n), \quad m = 0, 1, \ldots, M - 1, \quad n = 0, 1, \ldots, N - 1.
\]

(1)

With an additive white Gaussian random disturbance model, \( n_l(m, n) \) is assumed to be an independent zero-mean Gaussian random variable with variance \( \sigma_n^2 \) for all \( m \) and \( n \).

The discrete Fourier transform of the received image at time \( t_l \) becomes

\[
\mathcal{R}_l(m, n) = \mathcal{S}_l(m, n) + \mathcal{N}_l(m, n)
\]

(2)

where the transform-domain source image and random disturbance are

\[
\mathcal{S}_l(m, n) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} s_l(p, q) e^{-i2\pi(p \frac{M}{M} + q \frac{N}{N})}
\]

(3)

and

\[
\mathcal{N}_l(m, n) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} n_l(p, q) e^{-i2\pi(p \frac{M}{M} + q \frac{N}{N})}
\]

(4)

2.2. Translation Movement

When the image of incoming laser beam makes a translation movement within the FOV of the detector array between \( t_l \) and \( t_{l+1} \) by the amount of \( x_l \) and \( y_l \) pixels along the x-axis and y-axis, respectively, the resulting image at \( t_{l+1} \) is related to the previous image at \( t_l \) by

\[
s_{l+1}(m, n) = s_l(m - x_l, n - y_l)
\]

(5)

In the transform domain, it can be easily shown that the relationship becomes

\[
\mathcal{S}_{l+1}(m, n) = \mathcal{S}_l(m, n) e^{i\theta_{m,n,l}}
\]

(6)

where

\[
\theta_{m,n,l} = -2\pi \left( \frac{m}{M} x_l + \frac{n}{N} y_l \right)
\]

(7)

is the phase introduced to the pixel \((m, n)\) of the transform-domain image due to the translation of coordinate from \( t_l \) to \( t_{l+1} \). Note that such a translation movement does not change the magnitude of each pixel of the transform-domain image.
3. MAXIMUM LIKELIHOOD ESTIMATE OF TRANSLATION VECTOR

Based on the Gaussian assumption stated in Eq. (1), the maximum likelihood estimator will declare the estimated translation vector \( \hat{\beta}_t, \hat{\beta}_y \) if

\[
p(\hat{r}_{t+1} | \hat{\beta}_t, \hat{\beta}_y; \tilde{r}_t) = \max_{\{\beta_t, \beta_y\}} p(\hat{r}_{t+1} | \beta_t, \beta_y; \tilde{r}_t)
\]

(8)

where, for notational convenience, \( \tilde{r}_{t+1} \) and \( \tilde{r}_t \) are the vector representation (known as the lexicographic form) of the corresponding received image matrix \( r_{t+1}(m, n) \) and the reference image matrix \( s_t(m, n) \), respectively and

\[
p(\hat{r}_{t+1} | \beta_t, \beta_y; \tilde{r}_t) = \frac{1}{M N} e^{-\frac{1}{2} \| \tilde{r}_{t+1} - \tilde{r}_t \|^2}
\]

(9)

is the conditional probability density function of \( \tilde{r}_{t+1} \) given that the translation vector is \( \beta_t, \beta_y \) and the reference at the beginning of this movement is \( \tilde{r}_t \). The maximum likelihood criterion stated in Eq. (8) is equivalent to

\[
\min_{\{\beta_t, \beta_y\}} \| \tilde{r}_{t+1} - \tilde{r}_t \|^2 = \min_{\{\beta_t, \beta_y\}} \| \tilde{r}_{t+1} - L_{x, y} \{ \tilde{r}_t \} \|^2 = \| \tilde{r}_{t+1} - L_{x, y} \{ \tilde{r}_t \} \|^2
\]

(10)

where \( L_{x, y}\{\} \) is defined as a translation operator which moves the operand by a translation vector \( (x, y) \). It has been shown\(^4\) that the maximum likelihood criterion stated in Eq. (10) can finally be reduced to

\[
\max_{\{\beta_t, \beta_y\}} \left\{ \frac{1}{M N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |w_l(m, n)| \cos(\xi_{m,n} - \theta_{m,n}) \right\} = \| \tilde{r}_{t+1} - L_{x, y} \{ \tilde{r}_t \} \|^2
\]

(11)

where

\[
w_l(m, n) \triangleq R_{t+1}(m, n)S_t^*(m, n) = |w_l(m, n)| e^{i\xi_{m,n}}
\]

(12)

is the pixel-by-pixel product of the transform-domain received image \( R_{t+1}(m, n) \) and the complex conjugate of the transform-domain reference image \( S_t(m, n) \). Note that the likelihood function to be maximized in Eq. (11) can be rewritten as

\[
\text{Re} \left\{ \frac{1}{M N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} R_{t+1}(m, n)S_t^*(m, n)e^{-i\theta_{m,n}} \right\}
\]

(13)

where \( \text{Re}\{\} \) represents the real part of a complex quantity. It is clearly indicated that the likelihood function involves the average over all pixels of the pixel-wise multiplied received and reference images in the transform domain, as well as the phase to be estimated.

Taking the partial derivatives of Eq. (13) with respect to \( x_t \) and \( y_t \) and equating them to zero, we have

\[
\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m |w_l(m, n)| \sin(\xi_{m,n} - \theta_{m,n}) = 0
\]

\[
\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n |w_l(m, n)| \sin(\xi_{m,n} - \theta_{m,n}) = 0
\]

(14)

which form a set of nonlinear equations to be solved for the maximum likelihood estimates of \( x_t \) and \( y_t \).

When the extended-source is close to being acquired, the phase differences, \( (\xi_{m,n} - \theta_{m,n}) \), are small and the approximation of \( \sin(x) \approx x \) can be applied to Eq. (14), rendering the suboptimal linear estimator

\[
\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m |w_l(m, n)| \xi_{m,n} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m |w_l(m, n)| \theta_{m,n}
\]

\[
\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n |w_l(m, n)| \xi_{m,n} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n |w_l(m, n)| \theta_{m,n}
\]

(15)
Calculate $\hat{x}_t, \hat{y}_t$

The estimated vector $(\hat{x}_t, \hat{y}_t)$ obtained from solving this suboptimal linear maximum likelihood criterion must satisfy

$$-\frac{1}{2\pi} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m |w_t(m,n)| \xi_{m,n,t} = \hat{x}_t \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m \frac{m}{M} |w_t(m,n)| + \hat{y}_t \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n \frac{n}{N} |w_t(m,n)|$$

$$-\frac{1}{2\pi} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n |w_t(m,n)| \xi_{m,n,t} = \hat{x}_t \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n \frac{n}{M} |w_t(m,n)| + \hat{y}_t \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n \frac{n}{N} |w_t(m,n)|$$

Figure 2 shows a functional diagram of the open-loop maximum likelihood image acquisition. Note that, for optical downlink, the reference image can be a pre-stored profile of a point spreading function obtained from a point source such as a distant star. Also note that the transform-domain images can be realized by using a single lens.

4. IMAGE TRACKING LOOP

As discussed previously, to maximize the likelihood function of acquiring an image involves a comparison of the received image against the reference image. However, for image tracking, it is the correlation between the transform-domain received image $R_{t+1}(m,n)$ and the translated reference image

$$\hat{R}_{t+1}(m,n) = S_t(m,n)e^{i\hat{\theta}_{m,n,t}}$$

from previous estimation to be continuously monitored. The pixel-wise product of $R_{t+1}(m,n)$ and $\hat{R}_{t+1}(m,n)$ can be expressed by

$$C_{t+1}(m,n) \triangleq R_{t+1}(m,n)S_t^*(m,n)e^{-i\theta_{m,n,t}}$$

$$= |S_t(m,n)|^2 e^{i\phi_{m,n,t}} + N_{t+1}(m,n)S_t^*(m,n)e^{-i\theta_{m,n,t}}$$

where $\hat{\phi}_{m,n,t}$ is the estimate of $\phi_{m,n,t}$ in Eq. (7) and the estimation error is

$$\phi_{m,n,t} = \theta_{m,n,t} - \hat{\phi}_{m,n,t}$$

$$= -2\pi \left[ \frac{m}{M} (x_t - \hat{x}_t) + \frac{n}{N} (y_t - \hat{y}_t) \right]$$

$$\triangleq -2\pi \left( \frac{m}{M} \Delta_x + \frac{n}{N} \Delta_y \right)$$

where $\Delta_x$ and $\Delta_y$ are the associated estimation errors.
According to the maximum likelihood criterion derived in the previous section, the real part of Eq. (18) shall be maximized over the entire detector array, rendering the likelihood function

\[
M^{-1}N^{-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \text{Re}\{C_{i+1}(m, n)\}
\]  

(20)

which is the correlation between \( R_{i+1}(m, n) \) and \( \hat{R}_{i+1}(m, n) \). It turns out without surprise that \( \hat{x}_i \) and \( \hat{y}_i \) can be obtained by solving a set of two nonlinear equations formed by setting the partial derivatives of Eq. (20) with respect to \( \hat{x}_i \) and \( \hat{y}_i \) to be zero. However, in deriving the tracking algorithm to continuously update the estimates, two loop feedback signals are similarly formed as the partial derivatives of Eq. (20) with respect to \( \Delta_x \) and \( \Delta_y \), rendering

\[
\varepsilon_x = \frac{\partial}{\partial \Delta_x} M^{-1}N^{-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \text{Re}\{C_{i+1}(m, n)\}
\]

\[
= \frac{2\pi}{M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m|S_i(m, n)|^2 \sin(\phi_{m,n,i}) + \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \bar{N}_{i,ef}(x)(m, n)
\]  

(21)

\[
\varepsilon_y = \frac{\partial}{\partial \Delta_y} M^{-1}N^{-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \text{Re}\{C_{i+1}(m, n)\}
\]

\[
= \frac{2\pi}{N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n|S_i(m, n)|^2 \sin(\phi_{m,n,i}) + \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \bar{N}_{i,ef}(y)(m, n)
\]  

(22)

where

\[
\bar{N}_{i,ef}(x)(m, n) = \frac{\partial}{\partial \Delta_x} \text{Re}\{N_{i+1}(m, n)S_i^*(m, n)e^{-i\phi_{m,n,i}}\}
\]

\[
\bar{N}_{i,ef}(y)(m, n) = \frac{\partial}{\partial \Delta_y} \text{Re}\{N_{i+1}(m, n)S_i^*(m, n)e^{-i\phi_{m,n,i}}\}
\]

become the effective noises in the loop operation. Equations (21) and (22) characterize the relationship between the estimate errors, \( \Delta_x \) and \( \Delta_y \), and the loop feedback signals, \( \varepsilon_x \) and \( \varepsilon_y \). However, to solve for \( \Delta_x \) and \( \Delta_y \) from these nonlinear equations can be a quite challenging task. With a reasonable assumption valid when the phase error \( \phi_{m,n,i} \) remains small during the tracking mode, one can substitute Eq. (19) for \( \sin(\phi_{m,n,i}) \) in Eqs. (21) and (22). The resulting simultaneous equations are linear for \( \Delta_x \) and \( \Delta_y \) and can be easily solved, yielding

\[
\Delta_x = \frac{C_n E[\varepsilon_x] - C_{mn} E[\varepsilon_y]}{C_m C_n - C_{mn}^2}
\]  

(23)

\[
\Delta_y = \frac{C_{mn} E[\varepsilon_x] - C_m E[\varepsilon_y]}{C_{mn}^2 - C_m C_n}
\]  

(24)

where \( E[\cdot] \) denotes the statistical expectation and

\[
C_m = \frac{4\pi^2}{M^2} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m^2|S_i(m, n)|^2
\]

\[
C_n = \frac{4\pi^2}{N^2} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n^2|S_i(m, n)|^2
\]

\[
C_{mn} = \frac{4\pi^2}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} mn|S_i(m, n)|^2
\]

are coefficients that can be calculated from the transform-domain reference image of the previous iteration at \( t_i \).
An extended-image tracking loop structure can be realized based upon the above discussion and is depicted in Fig. 3. The transform-domain received image \( \{ R_{t+1}(m,n) \} \) is first multiplied pixel-wise with a properly translated transform-domain reference image established according to the estimate \( (\hat{x}_t, \hat{y}_t) \) from the previous iteration at \( t_t \). After being averaged over the entire detector array, the correlation result is used to compute the loop feedback signals, \( \epsilon_x \) and \( \epsilon_y \). It is important to note that, in real implementation, these two loop feedback signals are calculated from \( C_{t+1}(m,n) \) as the following weighted sums

\[
\epsilon_x = \frac{2\pi}{M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m \text{Im}\{C_{t+1}(m,n)\}
\]

\[
\epsilon_y = \frac{2\pi}{N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n \text{Im}\{C_{t+1}(m,n)\}
\]

where \( \text{Im}\{\cdot\} \) represents the imaginary part of a complex quantity. The subsequent calculation of \( \Delta_x \) and \( \Delta_y \) from \( \epsilon_x \) and \( \epsilon_y \) is straightforward as indicated in Eqs. (23) and (24), except that the statistical averages are replaced by time averages performed by low pass filters. The calculated \( \Delta_x \) and \( \Delta_y \) will be used to update the movement estimates through an accumulator, such that

\[
\hat{x}_{t+1} = \hat{x}_t + \Delta_x
\]

\[
\hat{y}_{t+1} = \hat{y}_t + \Delta_y
\]

The updated accumulator contents will be used to calculate the estimate \( \hat{\theta}_{m,n,t+1} \) and prepare the translated reference image for the next loop iteration at \( t_{t+1} \).

5. CONCLUSION

This paper describes an extended-image spatial acquisition and tracking scheme for an optical downlink, which uses transform-domain image correlation between the imaged downlink laser beam and the priorly known or synthesized intensity profile of a distant point source to determine the relative movement of the laser spot on the focal plane and then to drive the fine steering mirror to maintain accurate and stable optical pointing. This maximum-likelihood-based technique provides the optimal solution to spatial acquisition and tracking and is expected to perform better with low signal-to-noise ratios. The direct benefits from this fine-pointing technique include: (1) less burden on the telescope’s coarse tracking mechanism – a very cost-effective way to enhance pointing accuracy, and (2) possible simultaneous beam tracking and telemetry data detection on a single high-speed detector array - minimizes the handover and possible alignment problems between acquisition and tracking operations.
Preliminary numerical simulation using a computer-generated 2-D Gaussian intensity profile have shown that sub-pixel resolutions can be achieved by this correlation-type acquisition and tracking scheme in a high disturbance environment. However, more realistic simulation and demonstration to include small-sized focal plane array and real laser beam profile subject to various optical turbulence effects are required to determine the relationship between system performance and important specifications such as minimum loop update rate.

There are several areas remain for further investigation. For example, the relative movement considered in this paper is limited to be rotation-invariant. The technique to compensate rotational movement encountered in image tracking is currently under study. The need for adaptation of reference image to compensate the line-of-sight optical turbulence is also an important issue to be resolved before successfully applying such a correlation-type technique. Furthermore, the assumption of independent white Gaussian disturbance in each pixel does not address many different types of disturbance. For example, sub-pixel scanning technique for resolution enhancement and various atmospheric effects can cause spatially correlated disturbances which will be an interesting area for future studies.

REFERENCES


