A Survey of State and Disturbance Observers for Practitioners

Aaron Radke and Zhiqiang Gao
Center for Advanced Control Technologies
Fenn College of Engineering, Cleveland State University
Cleveland, Ohio, USA
Email: radke@ieee.org and z.gao@ieee.org

Abstract—This paper gives a unified and historical review of observer design for the benefit of practitioners. It is unified in the sense that all observers are examined in terms of: 1) the assumed dynamic structure of the plant; 2) the required information, including the input signals and modeling information of the plant; and 3) the implementation equation of the observer. This allows a practitioner, with a particular observer design problem in mind, to quickly find a suitable solution. The review is historical in the sense that it follows the evolution of ideas in observer design in the last half century. From the distinction in problem formulation, required modeling information and the observer design goal, we can see two schools of thought: one is developed in the framework of Modern Control Theory; the other is based on disturbance estimation, which has been, to some extent, overlooked.

Index Terms—Estimators, Observers, Kalman Filter, $H_{\infty}$, Robust Estimation, Disturbance Observer, Unknown Input Observer, Extended State Observer, Survey

I. INTRODUCTION

Without a doubt, “observers”, also known as “estimators” or “filters”\(^1\) are indispensable tools for engineering. Their main function is extracting otherwise unmeasurable variables\(^2\) for a vast number of applications including feedback control\(^2\) and system health monitoring\(^3\). In engineering practice, an observer is used for a number of purposes, such as removing phase lag in feedback, reducing the use of costly sensors\(^4\) and estimating disturbances\(^5\),\(^6\).

Over the years, two classes of design methods for observers have emerged. One is concerned with state estimation based on a mathematical plant model; the other is concerned with disturbance estimation based on input output data.

For the first class, sophistication of observer design gradually grew. Initially, it was found that a better estimate could be obtained if more accurate information was incorporated into the observer. This includes knowledge of noise and disturbances characterized by deterministic, differential\(^7\), polynomial\(^8\), bounded\(^9\), and stochastic\(^10\) descriptions. Consequently, many of these enhancements were proposed at the cost of detailed model information. Textbooks have predominately focussed on this class of observers.

Practitioners recognize one can not rely entirely on mathematical models. This leads to the second class of observers developed for practical disturbances\(^8\),\(^11\),\(^12\). Brief surveys can be found in [13]–[16]. This class of observers compliments the first class in practical control problems with significant nonlinearity and uncertainty. They are primarily motivated by the need for effective disturbance rejection in control of mechanical systems.

To give a comprehensive and clear account of observers in each of these classes, a unified framework is proposed in this paper. This survey also shows how the methodology evolved within each school of thought. A primary motivation of this effort is to provide a comprehensive review of observers for practitioners to solve real world problems. One critical question they face is selecting, among many candidates, an appropriate observer for a particular problem. For this purpose, observers are reviewed in terms of the applicable dynamic structure of the plant, required sensors, plant knowledge, and implementation. This unified framework leads to a standard form for users to select a suitable observer.

A unified characterization of observers is first described in Section II. Once this has been established, a clear evolution of observers can be shown in Section III. This starts with a base set of estimators in Section III-A followed by the modern branch in Section III-B and the less known disturbance estimation based observers in Section III-C. Finally, concluding remarks are made in Section IV.

II. A UNIFIED CHARACTERIZATION OF OBSERVERS

In the context of practical applications, each observer is characterized in terms of:

\[\text{Plant Description} \quad (1)\]

\[\text{Input} \rightarrow \text{Estimate} \quad (2)\]

\[\text{Implementation} \quad (3)\]

(1) provides the mathematical description of a physical process, (2) shows the information required by the observer and what estimates it produces, and (3) gives the observer equation as it is implemented.

To simplify notation, the following guidelines are followed. 1) Where possible, each estimator is described in

\(^1\)The terms “observer”, “estimator” and “filter” have loosely described tools that extract information. For this reason, the terms are used interchangeably in this paper. “Filters” are often viewed as single input single output systems from a classical engineering standpoint where the Kalman filter is an exception. The common view of an “observer” is an information extraction tool that uses inputs from the input and output of a plant.

\(^2\)An observer of a dynamic system is formally referred to as another dynamic system whose states converge to the observed system states [1].
a standard form. For example, when both discrete and continuous version are available a continuous version is used to maintain a clear comparison. 2) The variables \(x, y\) and \(u\) are, respectively, vectors of the plant state, output and input. 3) Unless noted, lowercase letters are considered as time varying vectors. For example, \(y\) represents multiple outputs varying with time \([y_1(t), y_2(t), \ldots, y_n(t)]^T\). 4) Uppercase letters such as \(A\) denote constant matrices unless noted as a function of time, for example \(A(t)\). 5) \(y\) is assumed to be measurable. 6) Instead of assuming the goal is to estimate \(y\), the goal is \(x\). Once a state is estimated, any static mappings are trivial and have been left out for simplicity.

III. EVOLUTION OF OBSERVER DESIGN

The following sections use the Plant Description, Input and Estimate, and Implementation characteristics, defined previously, to survey the early, modern, and disturbance estimation based observers, respectively.

A. Early Estimators

Early on, engineers discovered internal values could be extracted from input output data. The mechanism used for this purpose is known as a state estimator. Unmeasured internal values can be extracted from input, output and plant dynamic information. The following discusses the development of early estimators as a popular and important base set.

1) Plant Output Based Estimator (OBE): This estimator simply extracts information from the output of a plant or signal; for this reason, it is called an Output Based Estimator (OBE). Some common types of OBE’s are the low pass noise filter, approximate differentiator [17] and \(\alpha\beta\gamma\) filter [18]. The OBE represented in terms of (1), (2) and (3) is:

\[
\begin{align*}
\dot{x} &= Ax \\
y &= Cx
\end{align*}
\]

\[
\text{OBE} : \{y, A, C\} \rightarrow \{\dot{x}\}
\]

\[
\dot{x} = A\dot{x} + L(y - C\dot{x})
\]

where \(L\) is chosen such that the estimation error is driven to zero.\(^3\) This filter is useful for common applications that only have an output. Although simple, the information is often delayed and corrupted by disturbances and sensor noise.

2) Alpha Beta Gamma Filter: A special case of the OBE is the Alpha Beta Gamma (\(\alpha\beta\gamma\)) filter since the output is the only information used for estimation. The \(\alpha\beta\gamma\) filter [18], [19] was a very early sampled data filter used as a practical radar estimation algorithm for velocity and acceleration when only position is available.

\[
\begin{align*}
y(n) &= f(x, t, w_f) \\
x &= [y \ y' \ \cdots \ y^{(n-1)}]^T
\end{align*}
\]

\[
\alpha\beta\gamma : \{y, u\} \rightarrow \{\dot{x}\}
\]

\[
\Phi_{ij} = \begin{cases} 
\frac{\tau_{ji}}{1 - \tau_{ji}}, & i \leq j; \\
0, & \text{else.}
\end{cases}
\]

\[
\dot{x}_{k+1} = \Phi(\dot{x}_k + L(y_k - \dot{x}_k))
\]

Here \(T\) is the discrete time sampling period. Design is simplified since (7) requires a specified structure that is a special case of the famous Kalman filter [20] and other equivalent forms [21]. Although design is simplified, the problem of all OBE’s still exists for plants with excessive noise, delay, and output disturbances.

3) Plant Input Based Estimator (IBE): One way to get around sensor noise and output disturbances is not using them.

\[
\begin{align*}
\dot{x} &= Ax + Bu \\
y &= Cx
\end{align*}
\]

\[
\text{IBE} : \{u, A, B, x_0\} \rightarrow \{\dot{x}\}
\]

\[
\dot{x} = A\dot{x} + Bu
\]

If the plant model in the observer is accurate, inputs, and initial conditions, \(x_0\), are available then internal system states can be determined from inputs alone. This can be thought as attempting to estimate internal plant information by running a simulated plant in parallel. However, initial conditions must be given. For example, to estimate velocity a noisy position output could be differentiated using the OBE or an acceleration input could be integrated with the IBE. This method is also applicable if \(y\) is not measurable.

4) Input and Output Based Observer (IOBO): Luenberger Observer: The real workhorse began with the IOBO, popularly known as the Luenberger Observer [22].

\[
\begin{align*}
\dot{x} &= Ax + Bu \\
y &= Cx
\end{align*}
\]

\[
\text{IOBO} : \{u, y, A, B, C\} \rightarrow \{\dot{x}\}
\]

\[
\dot{x} = A\dot{x} + Bu + L(y - C\dot{x})
\]

This is a simple combination of the OBE (6) and IBE (12). By feeding back the estimated state along with measured data, it eliminates the IBE requirement for accurate initial conditions [2]. Since the estimate is fed back through the estimator it is also often called a Closed Loop Observer. The key advantage of the IOBO is the ability to use both the input and output data along with plant information to reduce noise and phase lag without the knowledge of initial conditions.

The Luenberger observer established the structure that most estimators are based today. The difference lies in the method of choosing \(L\).
5) Proportional Integral Observer (PIO): The PIO or PI Observer [23]–[25] is an extension of the IOBO aimed at removing steady state error.

\[
\dot{x} = Ax + Bu \\
y = Cx
\]

PIO : \{u, y, A, B, C\} \rightarrow \{\dot{x}\}

\[
\dot{x} = Ax + Bu + L(y - C\hat{x}) + L_i \int (y - C\hat{x}) \tag{18}
\]

The main idea is to use an integral gain, \(L_i\), in addition to the common proportional gain, \(L\), for the estimation error, \(y - C\hat{x}\), in the Luenberger observer. The extra integral state enhances the correction term by accumulating error over time.

6) Basic Nonlinear Observer (NLO): Nonlinear Luenberger: A common initial enhancement to an established linear algorithm is modifying it for nonlinear systems. The NLO is another simple variation of the IOBO for nonlinear functions of states and inputs.

\[
\dot{x} = f(x, u) \\
y = h(x)
\]

NLO : \{u, y, f, h\} \rightarrow \{\dot{x}\}

\[
\dot{x} = f(\hat{x}, u) + L(y - h(\hat{x})) \tag{21}
\]

The applicability is limited by the requirement that nonlinear plant knowledge is known. Furthermore, like many other plant structures, it is not explicitly designed to handle disturbances.

B. Modern Estimators

From this base set of observers, there have been a few key advances. The advance in modern control theory has been made by formulating the problem with disturbances in mind. These methods minimize a cost function based on mathematical assumptions about disturbances [20]. However, the design complexity has substantially increased.

1) Kalman Filter (KF): The Kalman filter [10], [26] was one of the first estimators to include the formulation of disturbances and provide optimal solutions.

\[
\dot{x} = Ax + Bu + w_{N(0,Q)} \\
y = Cx + v_{N(0,R)}
\]

KF : \{u, y, A, B, C, \text{cov}(w_N), \text{cov}(v_N)\} \rightarrow \{\hat{x}\} \text{ s.t. } \min \|x - \hat{x}\|_2

\[
\dot{S}(t) = S(t)A^T + AS(t) + Q - S(t)C^TR^{-1}CS(t) \\
L(t) = S(t)C^TR^{-1} \\
\dot{x} = Ax + Bu + L(t)(y - C\hat{x}) \tag{24}
\]

At each point in time, \(\dot{f}\) and \(h\) are linearized to \(A(t)\) and \(C(t)\) to then be used in the standard Kalman filter.

One of the most recent Kalman filter modifications for nonlinear systems is the Unscented Kalman Filter (UKF) [32], [33]. It moves beyond the EKF by additionally passing intermediate values through the known nonlinear equations [20], [29]. Although the UKF is derived for an algorithmic implementation, the complexity and model information required make it impractical for the majority of practical applications.

\[
\dot{x} = f(x, u, t) + w_{N(0,Q)} \\
y = h(x, t) + v_{N(0,R)}
\]

UKF: \{u, y, f, h, \text{cov}(w_N), \text{cov}(v_N)\} \rightarrow \{\hat{x}\} \tag{29}

The UKF implementation is too involved to include in this survey, however it is an important recent modern extension to include for historical reference.}

3) \(H_\infty\) Estimator: Another significant tool in the modern direction is the \(H_\infty\) estimator.

\[
\dot{x} = Ax + Bu + B_tw_f \\
y = Cx + D_ww_f
\]

\[
H_\infty : \{u, y, A, B, C, \text{B}_w, D_w, \gamma\} \rightarrow \{\hat{x}\} \text{ s.t. } \left\| \frac{x - \hat{x}}{w_f} \right\|_\infty < \gamma \tag{31}
\]
\[ \dot{Q}(t) = Q(t)A^T + AQ(t) + B_w B_w^T - Q(t) (C^TC - \gamma^{-2}C^TC) Q(t) \]
\[ L(t) = Q(t)C^T \]
\[ \dot{x} = Ax + Bu + L(t)(y - C\hat{x}) \]

This also optimizes a cost function based on an assumption about the disturbance. The formulation is significant because it uses a unique characterization of the disturbance. Kalman minimizes the minimum squared error because it is a mathematically manageble optimization problem. Using infinity norms, the \( H_\infty \) estimator is able to minimize the maximum or worst case disturbance \([9], [28], [34]\). In \((32)\), \( w_f \) is unknown but not necessarily random or stochastic. The estimator is guaranteed to be optimal under a user defined upper bound \( \gamma \).

### C. Disturbance Estimators

Although great strides have been made in modern estimation, it has moved beyond the reach of most practicing engineers. For this reason, the following section describes another school of thought not illuminated in mainstream research.

Estimating an unknown disturbance in addition to states is the key idea in this second school of thought. It is powerful in conjunction with feedback control \([7],[35]\). By appropriately including a disturbance estimate, \( \hat{w} \), in a control law, \( u \), disturbance effects can be approximately removed.

\[ \dot{x} = Ax + Bu + w \]
\[ = Ax + B(u + \hat{B}^+ \hat{w}) + w \]
\[ \approx Ax + Bu \]  

(33)

Here \( B^+ \) signifies the Penrose pseudo matrix inverse. Many estimators employ this scheme to handle slight perturbations for a modeled plant. However, few go as far as removing the requirement of a modeled plant by rejecting any unmodeled dynamics. The disturbance rejection concept has been analyzed in several papers. Systems with unknown and known nonlinear dynamics with linear modeled disturbances \([36]\) and un-modeled disturbances \([37]\) have been studied. Discrete \([16]\), non-linear, reduced-order, and robust \([37]\) forms have been designed. A few key disturbance estimation tools will be outlined in the following sections.

1) **Disturbance Observer (DOB):** A common tool to estimate disturbances is a Disturbance Observer (DOB) \([5],[6],[13]\).

\[ y(s) = P_n(s)(u(s) + d(s)) \]  

(34)
\[ DOB: \{u,y,P_n\} \rightarrow \{\hat{d}\} \]  

(35)
\[ \hat{d}(s) = [P_n^{-1}(s)\tilde{g}(s) - u(s)] Q(s) \]  

(36)

To make \( P_n^{-1}(s) \) proper, \( Q(s) \) is frequently a low pass filter. The DOB is different from state estimators because, instead of states, it estimates external disturbances and observer model discrepancies that effectively appear at the plant input. An effective plant input disturbance instead of states. It is usually written in transfer function instead of state-space form. Similar to \((33)\), the estimate \( \hat{d}(s) \) is important in closed loop operation to cancel \( d(s) \).

2) **Unknown Input Observer (UIO):** The UIO \([38]\) uses the DOB concept in state space representation.

\[ \dot{z} = Afz \]
\[ \hat{w}_e = Cfz \]
\[ \dot{\hat{x}} = Ax + Bu + w_e \]
\[ y = Cx \]  

(37)

\[ UIO: \{u,y,A,B,C,A_f,C_f\} \rightarrow \{\hat{x},\hat{w}_e\} \]  

(38)
\[ \begin{bmatrix} \dot{x} \\ \dot{\hat{z}} \end{bmatrix} = \begin{bmatrix} A & BC_f \\ 0 & A_f \end{bmatrix} \begin{bmatrix} x \\ \hat{z} \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + L(y - Cx) \]
\[ \hat{w}_e = C_f \hat{z} \]  

(39)

It is an IOBO with an augmented disturbance model to estimate both states and disturbances. With state space equations, the UIO defines assumptions about the rate of disturbance changes. The disturbance input, \( w_e \), is made to satisfy a differential equation. The most common assumption is a constant disturbance \( w_e(z) = 0 \) where \( A_f = 0 \) and \( C_f = 1 \).

Originally, the UIO focused on unknown external inputs for linear systems \([39],[40]\); later this included nonlinear plants \([41]\) and fault estimation \([42]\). The ability to estimate states and disturbances simultaneously is a practical advantage of the UIO over the DOB.

3) **Perturbation Observer (POB):** With a slight change in notation from the DOB, the POB makes a significant step to include estimation of unmodeled plant variations in addition to external disturbances.

\[ x_{k+1} = Ax_k + Bu_k + w_f \]
\[ y = Cx_k \]  

(40)
\[ POB: \{u,y,A,B,C,A_f,B_f,C_f\} \rightarrow \{\hat{x},\hat{w}_f\} \]  

(41)
\[ z_k = A_f z_{k-1} + B_f (B^+ (\hat{x}_k - Ax_{k-1} - u_{k-1})) \]
\[ \hat{w}_f = C_f z_k \]
\[ \hat{x}_{k+1} = Ax_k + B(u_k + \hat{w}_f k) + L(y_k - C\hat{x}_k) \]  

(42)

By defining \( w_e \) in \((39)\) to be \( w_f \), a bounded, \( L_\infty \) and admissible function \([43]\), the unknown input can represent traditional external disturbances and model variations \( w_f = w_e + \Delta Ax_k + \Delta Bx_k \) \([16]\).

4) **Extended State Observer (ESO):** Most estimators are made to handle slight perturbations for a modeled plant, however the ESO was designed \([44]\) to remove the requirement of a modeled plant by rejecting un-modeled dynamics. The ESO uses a simple canonical form so the un-modeled dynamics appear at the disturbance estimation portion. This decisively captures the subtle but important design methodology shift between modern estimators and disturbance estimators. It

\[ ^4 \text{This is reasonable assuming the disturbance is constant during a short sampling period.} \]


encompasses realistic disturbances and un-modeled plant variations while remaining simple.

\[
y^{(n)} = f(x, t, u, w_f) + b_m u
\]

and

\[
x = \begin{bmatrix} y & \dot{y} & \cdots & y^{(n-1)} \end{bmatrix}^T
\]

\[
\dot{x}_2 = \begin{bmatrix} \dot{x}_1 & \dot{x}_{n-1} & \dot{x}_n & \dot{f} + b_m u & 0 \end{bmatrix}^T + L(y - \hat{x}_1)
\]

(45)

Originally, \( L \) was a set of nonlinear gains (NESO) [14], [44]–[46] but was greatly simplified with a single tuning parameter [47]. Although the ESO can be structurally equivalent to the UIO when \( C(sI - A)^{-1}B = b_m/s^n, A_f = 0 \) and \( C_f = 1 \), it solves a different problem. The ESO (43) is focused around a large class of nonlinear systems with simple system knowledge and the UIO (37) and DOB (34) require modeling information. Due to the ESO’s practical design, there have been many successful applications [48] including: power converters [49], servo motion control [35], web tension [50], bio-mechanics [51] and multivariable jet engines [52].

A recent modification to the ESO is the GESO [53] to include derivative estimates of the disturbance.

\[
y^{(n)} = f(x, t, u, w_f) + b_m u
\]

\[
x = \begin{bmatrix} y & \dot{y} & \cdots & y^{(n-1)} \end{bmatrix}^T
\]

\[
z = \begin{bmatrix} f & \dot{f} & \cdots & f^{(h-1)} \end{bmatrix}^T
\]

\[
\begin{bmatrix} \dot{x}_1 & \cdots & \dot{x}_{n-1} & \dot{x}_n & \dot{z}_1 & \cdots & \dot{z}_{h-1} & \dot{z}_h \end{bmatrix}^T =
\begin{bmatrix} \dot{x}_2 & \cdots & \dot{x}_n & \dot{f} + b_m u & \dot{z}_2 & \cdots & \dot{z}_h & 0 \end{bmatrix}^T + L(y - \hat{x}_1)
\]

(48)

This extension, using \( h \), provides extra information and increases the ability to track different types of disturbances. For example, \( h = 1 \) allows convergence to a constant disturbance and \( h = 2 \) allows convergence to a disturbance with a constant derivative.

IV. CONCLUDING REMARKS

After a half century of continuous research and development, as reviewed in this paper and shown in Figure 1, observers have become an integral part of control theory and practice. Starting from the early estimators, the evolution of observers proceeded with two distinct schools of thought: One, modern estimation, relies on a detailed mathematical model of the plant and seeks optimal solutions. The other, disturbance estimation, acknowledges the limit of available partial plant dynamic information, and seeks to estimate the disturbance, i.e. the discrepancy between the model and the real system. In some cases, the disturbance observers provide both the state and disturbance estimation.

The model-based methods provide rigorous and, in many cases, optimal solutions. The disturbance estimation strategy is less known but addresses the uncertain nature of physical processes; and it seems to offer a more practical design framework to deal with real world control problems. To facilitate the assessment for practical applications, each observer is presented in terms of 1) the mathematical model required, 2) its inputs and outputs, and 3) its implementation equations. This makes it easy for a practitioner to quickly determine whether the assumptions are met, whether the observer provides what is needed for a particular application, and the computational requirement for implementation.

REFERENCES


