CORE
Metadata, citation and similar papers at core.ac.uk
Provided by NASA Technical Reports Server

FINAL TECHNICAL REPORT

on

NASA Grant, NAG 1-1074

APPLICATION OF OPTICAL DISTRIBUTED SENSING AND COMPUTATION TO CONTROL OF LARGE SPACE STRUCTURES

Submitted: May 1992

A. V. Balakrishnan, Principal Investigator

(NASA-CR-190293) APPLICATION OF OPTICAL DISTRIBUTED SENSING AND COMPUTATION TO CONTROL OF LARGE SPACE STRUCTURES Final Technical Report (California Univ.) 35 p CSCL 20F G3/74

N92-23567

Unclas 0086919

Optical Distributed Sensing for Controlling Large Space Structures

Weijian Zhang

1 Introduction

Rapid developments in optical materials and devices in the last two decades have made optical sensing a very attractive sensing technique in control engineering. Compared with the traditional electronic and electro-mechanical sensors, the optical sensing and related information processing has the following advantages:

1. Interconnections: Optical connections are much easier to make than electronic ones, because optical signals do not affect each other if their paths cross.

2. Immunity to electromagnetic interference (EMI): Electric conductors, unlike light beams, when placed too close to each other have EMI upon each other, because the signals they carry generate magnetic fields which can induce currents in nearby conductors. For example, a prime attraction of optical fibre sensors in aerospace is their immunity to EMI, including some electronic countermeasures.

3. Optical measurements are made without making contact, at high speed and with high precision.

4. Reduced size and weight of devices, higher processing capacity for low electrical power consumption, are attractive factors to space industry.

Although optical fibre sensors have been widely used in industry, their applications are limited to the measurements of the following field variations:

1. strain distribution in large structures such as bridges, dams, and pressure vessels;
2. temperature distributions in equipment such as power transformers, electrical generators, boilers and high-voltage cables;
3. electric and magnetic field anomalies in electrical networks.

In this report a real time holographic sensing technique is introduced and its advantages are investigated from the filtering and control point of view. The feature of holographic sensing is its capability to make distributed measurements of position and velocity of moving objects, such as a vibrating flexible space structure. This study is based upon the distributed parameter models of linear time-invariant systems, particularly including the linear oscillator equations describing the vibration of large flexible space structures. The general conclusion is that application of optical distributed sensors brings gain in the situation where Kalman filtering is necessary for state estimation. In this case, both transient and steady state filtering error covariance become smaller. This in turn results in smaller cost in the LQG problem.

Instead of measuring field variations like optical fibre sensors, holographic sensors are used to measure the position and velocity of objects in motion. Rate measurement is made possible by latching two holographic memory devices, where the images are retained, at different times.

The organization of this report is as follows:

In Section 2, the princple of a real time holographic sensing technique—Degenerate Four Wave Mixing (DFWM)—is reviewed briefly.

In Section 3, we first study noise disturbances in holographic sensing and propose white noise as a reasonable formulation of observation noise in holographic sensing.

Then the results on the gains of optical distributed sensing are stated and proved. Finally, by studying the active damping problem of flexible structures, we show that it is important to use distributed sensing in order to reduce the unwanted energy input caused by filtering errors.

Conclusions are in Section 4.

2 The Principles of Real Time Holography

The principle of holography

(1) *Recording of hologram.* For simplicity, let us consider only the one dimensional case. The arrangement of hologram and coordinate is as illustrated in Figure 1.

Obviously, when $|x| \ll |z_i|$, $i = 1, 2$, we have

$$r(x, z_i) = (z_i^2 + x^2)^{1/2} = |z_i| + \frac{1}{2}\frac{x^2}{|z_i|} + O\left(\left(\frac{x}{z_i}\right)^2\right).$$

The reference wave $W_{ref}(x, z)$, is a plane wave and is coherent with the object wave $W_{obj}(x, z)$. In phasor form, we have

$$\begin{aligned} W_{ref}(x, z) &= K \exp\left[ik\left(x\cos\left(\tfrac{\pi}{2} - \theta\right) + y\cos\tfrac{\pi}{2} + z\cos\theta\right)\right] \\ &= K \exp[ik(x\sin\theta + z\cos\theta)] \end{aligned}$$

where $k = 2\pi/\lambda$ is wave number and λ is wave length.

For simplicity, we consider an object consisting of only two points with spacing $|z_1 - z_2|$. Then the object beam is the inteference of two spherical waves produced by the two points. Therefore the object wave in phasor form is given by

$$\begin{aligned} W_{obj}(x, z)\,|_{z=0} &= a\frac{e^{ikr(z_1,x)}}{r(z_1,x)} + a\frac{e^{ikr(z_2,x)}}{r(z_2,x)} \\ &\approx a\frac{e^{ik|z_1|}}{|z_1|}e^{ikx^2/2|z_1|} + a\frac{e^{ik|z_2|}}{|z_2|}e^{ikx^2/2|z_2|} \\ &= aA_1e^{ikx^2/(2|z_1|)} + aA_2e^{ikx^2/(2|z_2|)} \end{aligned}$$

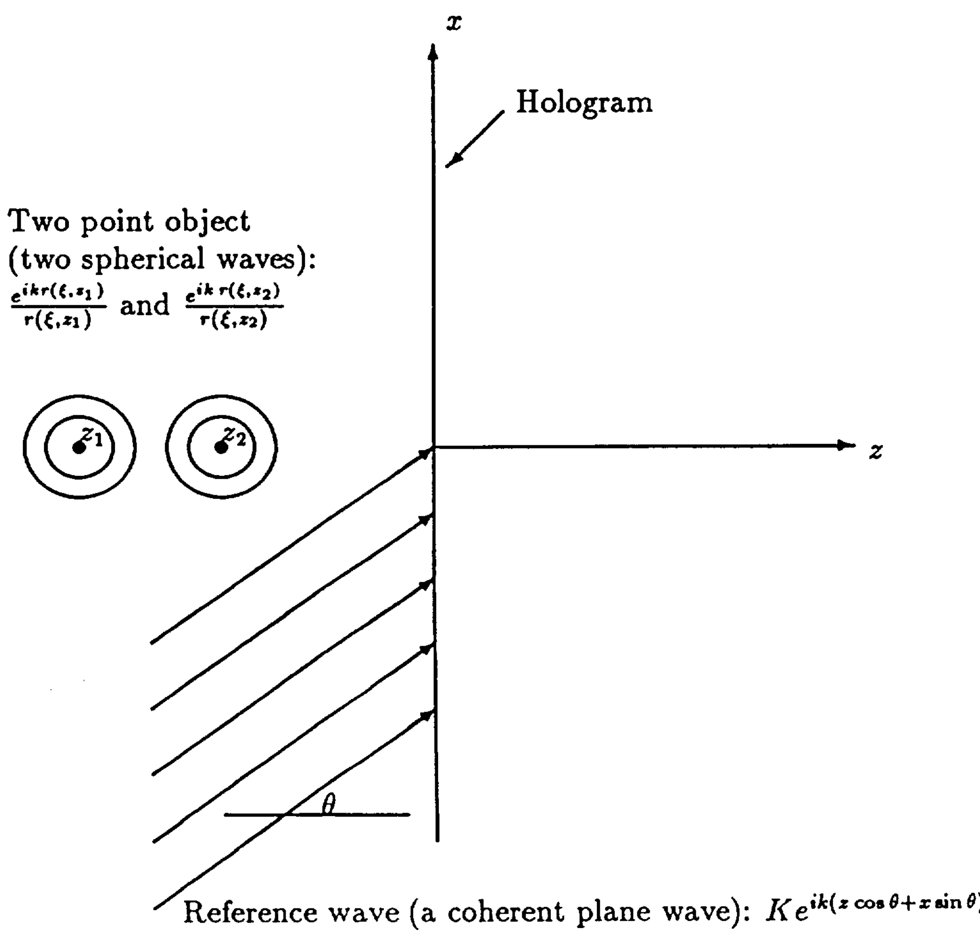

Figure 1: The Schematic for Recording a Hologram

for $|x| \ll |z_2|$.

If we assume the exposure E is in the linear range of the transmittance-exposure relation [3], then what is recorded on the hologram (x-axis) is the intensity of the interference pattern between the object wave $W_{obj}(x, z)$ and the reference wave $W_{ref}(x, z)$ at $z = 0$.

$$\begin{aligned} I(x) &= |W_{ref}(x,0) + W_{obj}(x,0)|^2 \\ &= |Ke^{ikx\sin\theta} + aA_1e^{ikx^2/(2|z_1|)} + aA_2e^{ikx^2/(2|z_2|)}|^2 \\ &= K^2 + aKe^{ikx\sin\theta}(\bar{A}_1e^{-ikx^2/(2|z_1|)} + \bar{A}_2e^{-ikx^2/(2|z_2|)}) \\ &\quad +aKe^{-ikx\sin\theta}(A_1e^{ikx^2/(2|z_1|)} + A_2e^{ikx^2/(2|z_2|)}) \\ &\quad +O(a^2) \end{aligned} \tag{1}$$

for $K \gg a$.

(2)*Reconstruction of the recorded holographic image.* Those two cross product terms in (1) play crucial role in holography. After developing the film, put it in the reconstruction system as illustrated in Figure 2.

Let us consider the contribution of the first cross product term. The spatial domain solution of the diffraction problem is expressed in the Fresnel-Kirchhoff integral as follows: If a plane wave of amplitude a_1, traveling in the direction of the positive z axis, is incident on an object with amplitude transmittance $t(x_1, y_1)$ in the plane normal to the z axis at $z = 0$, the light complex amplitude $a_2(x_2, y_2, d)$ in the plane $z = d$ is

$$\begin{aligned} a_2(x_2, y_2, d) &= \frac{ia_1}{\lambda}\iint_{R^2} t(x_1, y_1)\cos\theta \\ &\quad \times \frac{\exp\{-i(2\pi/\lambda)[d^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2}\}}{[d^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2}} dx_1 dy_1 \end{aligned}$$

The geometry is illustrated in Figure 3. The angle θ is formed by the positive z axis and the straight line connecting the points $(x_1, y_1, 0)$ and (x_2, y_2, d). Cosθ is called

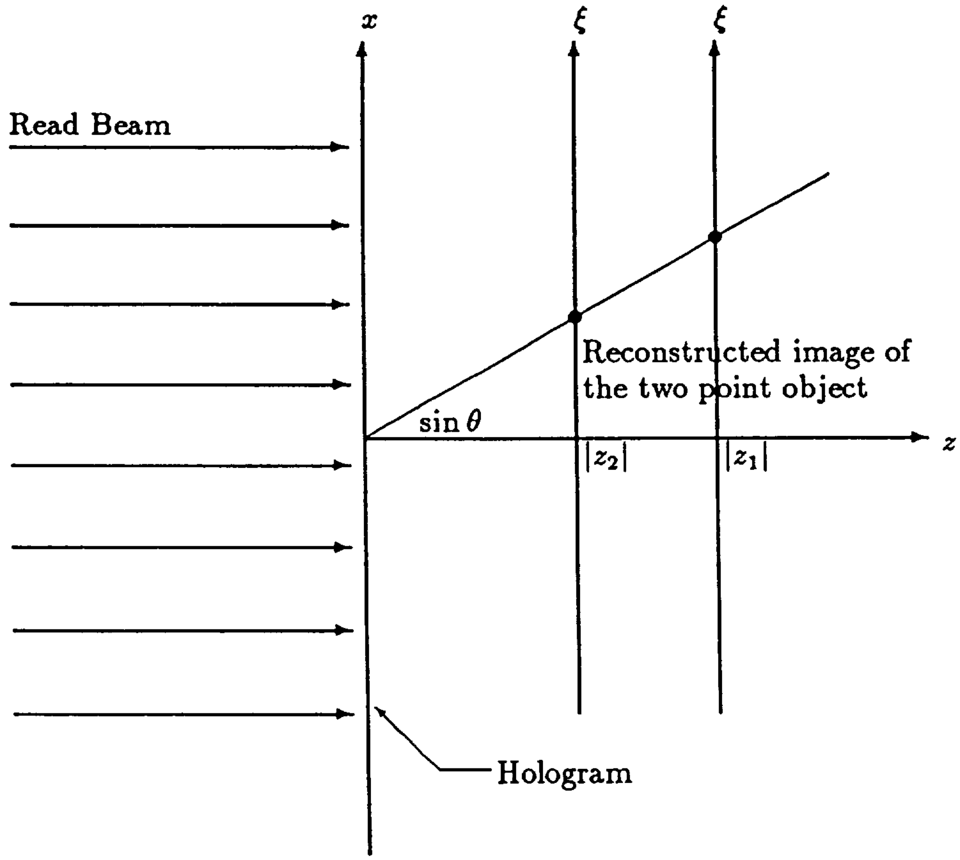

Figure 2: The Schematic for Reconstructing the Image

the ***obliquity factor***, which, in what follows, is considered to be close to 1 since the angle θ is generally close to 0 when d is large.

By Fresnel-Kirchhoff integral [3, p106], and neglecting the multiplicative constant, we obtain the diffraction pattern of the read beam upon the recorded hologram as

$$
\begin{aligned}
g(\xi) &= \int e^{ikx\sin\theta}(\bar{A}_1 e^{-ikx^2/(2|z_1|)} + \bar{A}_2 e^{-ikx^2/(2|z_2|)}) e^{ik\sqrt{(\xi-x)^2+z^2}} dx \\
&= e^{ik(z+\xi^2/(2z))} \int (\bar{a}_1 e^{-ikx^2/(2|z_1|)} + \bar{A}_2 e^{-ikx^2/(2|z_2|)}) \\
&\qquad \times \exp[ikx^2/(2z) - ikx/z(\xi - z\sin\theta)]dx
\end{aligned}
$$

Therefore, at position $z = |z_1|$, the first term in the above integral evaluates to $\delta(\xi - |z_1|\sin\theta)$, while the second term becomes $\delta(\xi - |z_2|\sin\theta)$ at $z = |z_2|$. In other words, the two point object is reconstructed at positions $(|z_1|, |z_1|\sin\theta)$ and $(|z_2|, |z_2|\sin\theta)$.

A Real-time Holography Technology - DFWM

In present days, real-time holography can be accomplished in photorefractive materials, a technique called *degenerate four wave mixing* (DFWM) [4]. Photorefractive materials include the following type of crystals: lithium niobate ($LiNbO_3$), potassium niobate ($KNbO_3$), barium titanate ($BaTiO_3$), strontium barium niobate (SBN) and bismuth silicon oxide ($Bi_{12}SiO_{20}$). Photorefractive crystals provide a unique way of recording light intensity which allows multiple beams to be mixed.

To understand the operation of DFWM, let us first recall that the refractive index n, or index of refraction, of a medium is defined to be c/v, where c is the speed of light in free space and v the speed of light in the medium [7]. For photorefractive crystals, their refractive index is nonlinearly dependent upon the local intensity I of light incident upon it. The total refractive index is given by [9]

$$n = n_0 + n_2 I^p \tag{2}$$

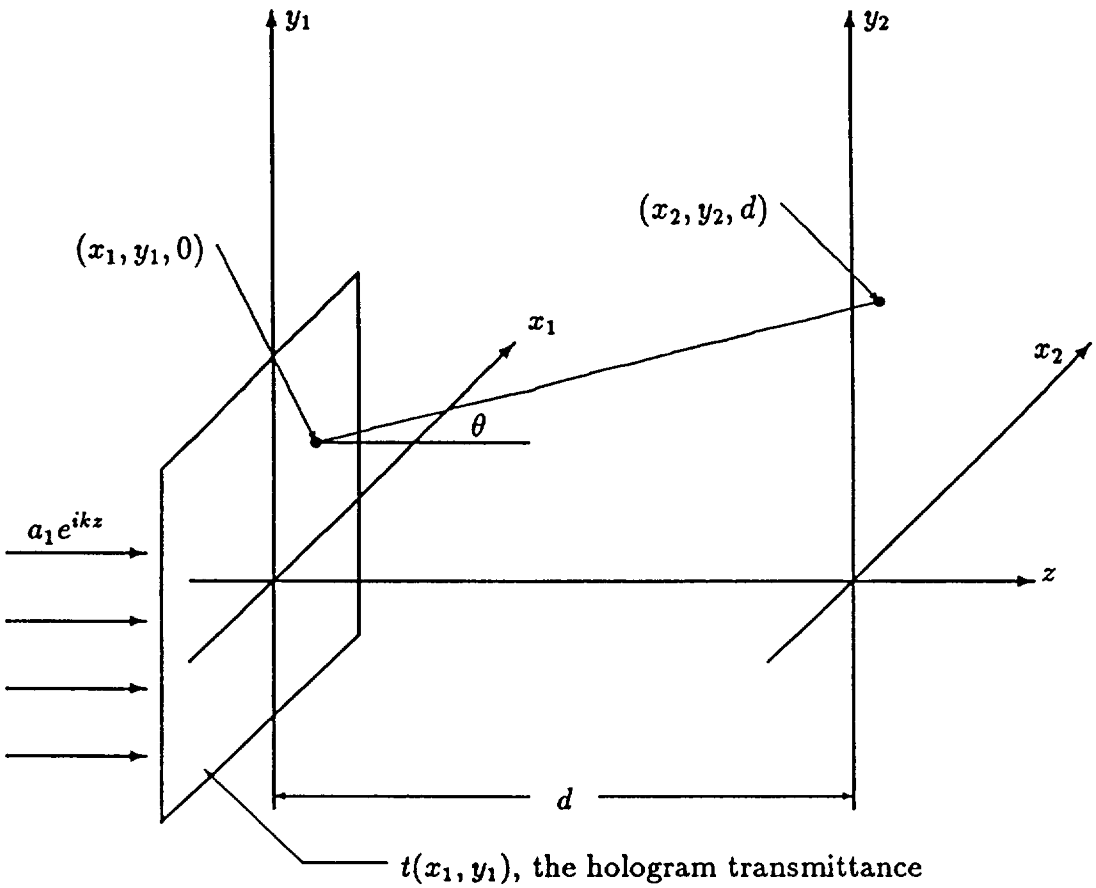

Figure 3: Geometry for Fresnel-Kirchhoff integral

where n_0, n_2, p are constants and it is common that $p \neq 1$.

When light is incident upon a photorefractive crystal, the refractive index within the crystal changes with the local light intensity, by (2). In this way, the distribution of light intensity over the volume of the crystal is recorded as a change in refractive index.

When performing DFWM using photorefractive crystal, as in conventional holography, both the object beam and the reference beam, which are coherent, are incident upon a photorefractive crystal. Al the same time, a third beam, called read beam, is also incident upon the crystal, see Figure 4. The interference pattern generated by the interference of these three beams produces a spatially varying refractive index within the crystal. The variation in refractive index within the crystal is similar to a recorded hologram, and causes diffraction of the beams. If the read beam is identical but phase conjugate to the reference beam, then the diffracted beam is the phase conjuate of the object beam.

DFWM in a photorefractive crystal is similar to conventional holography where the hologram is illuminated with the phase conjugate of the reference beam, producing a phase conjugate of the object. The difference between DFWM and conventional holography is that in the DFWM process, the hologram is made and read simultaneously, avoiding the chemical process of developing the hologram. Because of the fast response times of many photorefractive crystals [5], the holography can be made real-time using DFWM in a photorefractive crystal.

Currently, there are two major factors that limit widespread application of photorefractive crystals:

1. Some of the most promising crystals (e.g., $BaTiO_3$, SBN and $KNbO_3$) are not widely available in large samples with high optical quality.

2. The crystals that are available do not perform well enough in all respects. For instance, in order to demonstrate high speed, low write energy, long memory or large gain, it is necessary to use several different types of materials. Reference [11] gives the performance evaluations of some photorefractive crystals.

3 Results

Noise disturbances in holographic sensing

Noise disturbances are always present in the processes of making a hologram and then reading the hologram. Therefore we have to consider sensor noise in optical holographic sensing. Generally, the noise is attributed the following sources:

1. Random scattering of both signal and reference beams during exposure due to the granularity of the recording material.

2. Random scattering of the reading beam and the reconstructed wave due to the granularity of the material of the hologram.

3. Spatial modulations of the reference and reading beams.

4. Inhomogeneities and surface deformations of the recording material.

5. Nonlinear recording of the signal wave.

Sources 1, 2 and 3 are usually not significant except in certain special cases such as making multiple recording of many holograms in the same material, or using spatially modulated reference beams. Source 4 is related to the preparation of the recording material. Next we study the adverse effects of source 5, the nonlinearity in holographic recording.

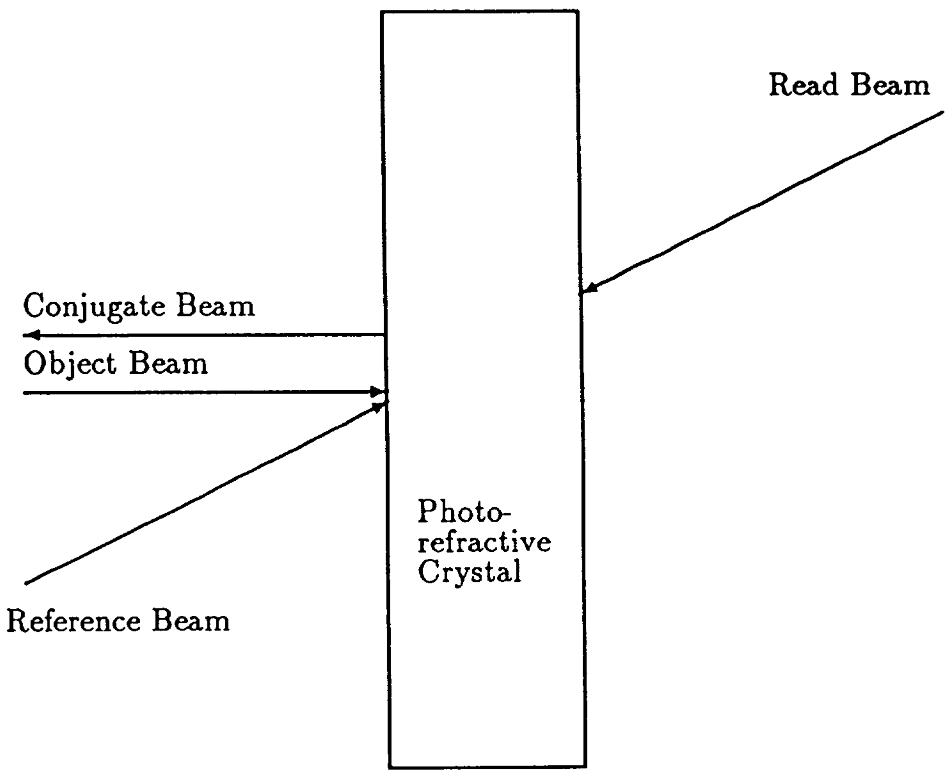

Figure 4: The scheme of Degenerate Four Wave Mixing (DFWM)

The transmittance t of a hologram made is a function of the exposure E, which is defined by

$$E = k_1 \tau_e I$$

where $k_1 > 0$ is a constant, τ_e is the exposure time. In the previous analysis we have assumed that the relation between t and E is linear. However, $t - E$ curve is always nonlinear to some degree. The following form of cubic approximation

$$t = c_0 + c_1 E + c_2 E^2 + c_3 E^3 \tag{3}$$

fits experimental curves quite well over a considerable range for plane absorption hologram.

For the usual range of exposure, nonlinear effects produced by the quadratic term are the most important. Its nonlinear effects are particularly well illustrated for a two point object. If we adopt the reconstruction arrangement as in Figure 5, then the wave at $z = d$ (on the hologram) is given by

$$\begin{aligned} a &= W_{obj}(x_2, y_2, d) \\ &= exp\left(-ik\sqrt{x_2^2 + (y_2 - b_1)^2 + d^2}\right) + exp\left(-ik\sqrt{x_2^2 + (y_2 - b_2)^2 + d^2}\right) \\ &\approx exp\left[-ik/(2d)(x_2^2 + y_2^2 - 2y_2 b_1)\right] + exp\left[-ik/(2d)(x_2^2 + y_2^2 - 2y_2 b_2)\right] \\ &= e^{i\phi_1} + e^{i\phi_2} \end{aligned}$$

For simplicity of analysis, let us assume that the reference wave is unmodulated, i.e., $rr* =$ constant. Then the quadratic term in the transmittance of the recorded hologram is given by

$$c_2 E^2 = c_2 k_1^2 \tau_e^2 (|a + r|^2)^2$$

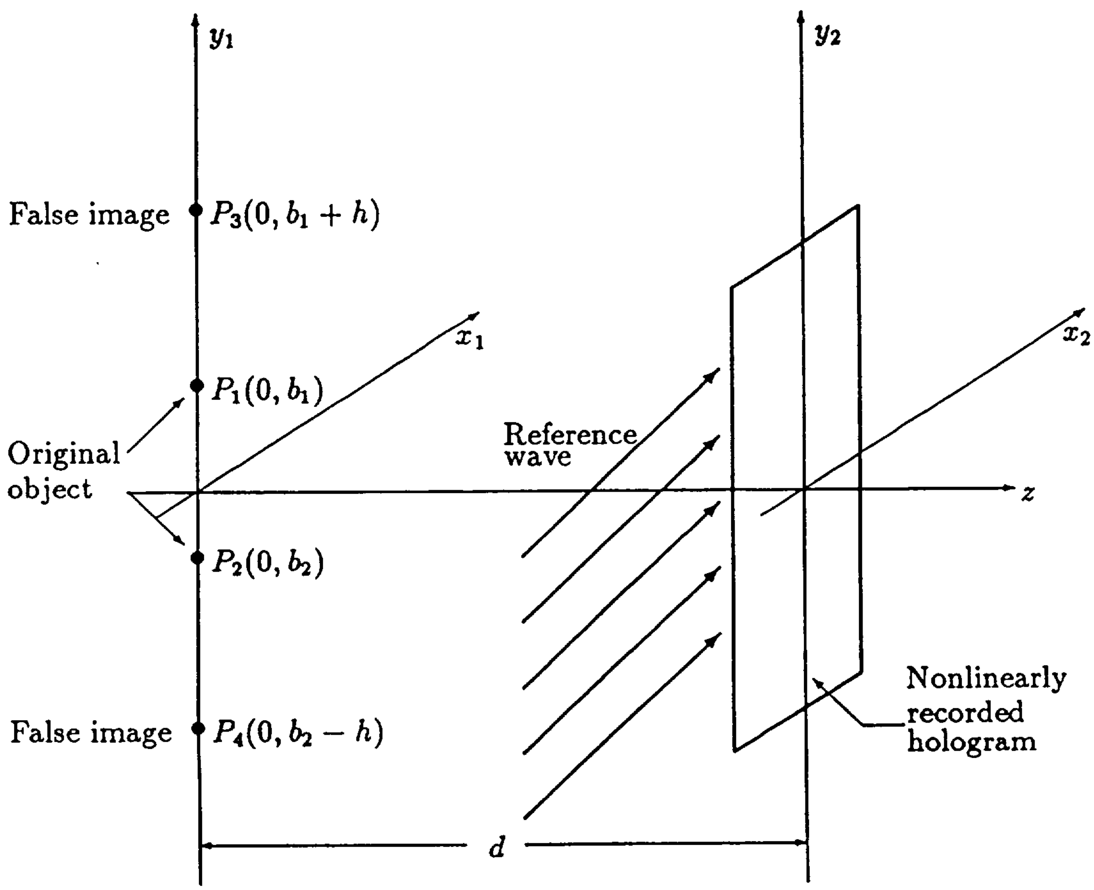

Figure 5: The generation of false images of a two point object by a nonlinearly recorded hologram ($h = b_1 - b_2$).

$$
\begin{aligned}
&= c_2 k_1^2 \tau_e^2 (aa^* + rr^* + ar^* + a^* r)^2 \\
&= c_2 k_1^2 \tau_e^2 [(aa^*)^2 + (ar^*)^2 + (a^* r)^2 + 2aa^* ar^* + 2aa^* a^* r] \\
&\quad + L(aa^*, ar^*, a^* r)
\end{aligned}
$$

where $L(\cdot,\cdot,\cdot)$ is a linear function of its three variables.

When the nonlinearly recorded hologram is illuminated by the original reference wave r, the fourth term in the above equation produces the following diffracted wavefront

$$
\begin{aligned}
r(2aa^* ar^*) &= 2|r|^2 (a^2 a^*) \\
&= 2|r|^2 [e^{i\phi_1} + e^{i\phi_2}]^2 [e^{-i\phi_1} + e^{-i\phi_2}] \\
&= 2|r|^2 [3e^{i\phi_1} + 3e^{i\phi_2} + e^{i(2\phi_1 - \phi_2)} + e^{i(2\phi_2 - \phi_1)}] \\
&= 6|r|^2 (e^{i\phi_1} + e^{i\phi_2}) + 2|r|^2 [\exp(-ik/(2d)(x_2^2 + y_2^2 - 2y_2(b_1 + h))) \\
&\quad + \exp(-ik/(2d)(x_2^2 + y_2^2 - 2y_2(b_2 - h)))\ ]
\end{aligned}
$$

where $h = b_1 - b_2 > 0$.

It can be seen that the third term in the above represents a spherical wave diverging from a false virtual image of a point source $P_3(0, b_1 + h)$. And the fourth term in the above represents a spherical wave diverging from another false virtual image of a point source $P_4(0, b_2 - h)$. Therefore, two false virtual images are produced by the effect of nonlinearity in $t - E$ relation.

The cubic nonlinearity term in (3) can be similarly analyzed and additional false images are introduced in and around the true image.

Therefore, there are numerous sources of noise in holographic sensing and they are in and around the true image. We propose the white noise formulation to represent the observation noise.

The gain of optical distributed sensing

Let us consider the following stochastic system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + F\omega(t) \\ y(t) &= Cx(t) + G\omega(t) \end{aligned}$$

where $GG^* = I$, $FG^* = 0$ (noise independent), and B is Hilbert-Schmidt, and $\omega(\cdot)$ is a Gaussian white noise process.

In the associated LQG problem, the cost functional is given by

$$\begin{aligned} J(u) &= \lim_{T\to\infty}\{\frac{1}{T}\int_0^T Tr.QE[x(t,\omega)x^*(t,\omega)]Q^* dt \\ &\quad + \frac{1}{T}\int_0^T Tr.E[u(t,\omega)u^*(t,\omega)]dt\} \end{aligned} \tag{4}$$

where Q is a Hilbert-Schmidt operator.

By [1], the optimal feedback control for the LQG problem is given by

$$u^*(t) = -B^* P_c \hat{x}_a(t)$$

where P_c satisfies the following SSRE

$$A^* P_c + P_c A + Q^* Q - P_c B B^* P_c = 0 \tag{5}$$

and $\hat{x}_a(t)$ is defined by the asymptotic Kalman filter

$$\begin{aligned} \dot{\hat{x}}_a(t) &= (A - BB^* P_c)\hat{x}_a(t) + P_f C^*(y - \hat{x}_a(t)) \\ &= (A - BB^* P_c - P_f C^* C)\hat{x}_a(t) + P_f C^* y(t) \\ \hat{x}_a(0) &= 0 \end{aligned}$$

where P_f solves the following SSRE

$$AP_f + P_f A^* + FF^* - P_f C^* C P_f = 0 \tag{6}$$

And the minimal final cost is given by

$$J(u^*) = Tr.QP_fQ^* + Tr.CP_fP_cP_fC^* \tag{7}$$

Sufficient conditions on the existence of P_c, P_f such that $(A - BB^*P_c)$ and $(A - P_fC^*C)$ are strongly stable

1. For our particular
$$\mathcal{A} = \begin{pmatrix} 0 & I \\ -A & 0 \end{pmatrix}; \quad \mathcal{B} = \begin{pmatrix} 0 \\ B \end{pmatrix}$$
by [2], if $(\mathcal{A}, \mathcal{B})$ is controllable and
$$||Qw||^2 \leq M||\mathcal{B}^*w||^2 \quad \forall w \in \mathcal{D}(\sqrt{A}) \times H$$
and also $(\mathcal{A}^*, Q^*)$ is controllable, then (5) has unique solution such that $(\mathcal{A} - \mathcal{B}\mathcal{B}^*\mathcal{P}_c)$ is strongly stable.

2. By [2], if $(\mathcal{A}^*, \mathcal{C}^*)$ and $(\mathcal{A}, \mathcal{F})$ are controllable and
$$||\mathcal{F}^*w||^2 \leq M||\mathcal{C}w||^2$$
then
$$\lim_{t\to\infty} \mathcal{P}_f(t) = \mathcal{P}_f \qquad \text{(in strong sense);}$$
and $\mathcal{P}_f$ satisfies (6) and is the unique solution such that both $(\mathcal{A} - \mathcal{P}_f\mathcal{C}^*\mathcal{C})$ and $(\mathcal{A} - \mathcal{P}_f\mathcal{C}^*\mathcal{C})^*$ are strongly stable.

Zabczyk [12] obtained more demanding sufficient conditions which require the exponential stabilizability of $(\mathcal{A}, \mathcal{B})$, $(\mathcal{A}, \mathcal{F})$, $(\mathcal{A}^*, \mathcal{C}^*)$, $(\mathcal{A}^*, Q^*)$. Unfortunately all these conditions are not satisfied in view of Gibson's result [6] and the fact that $\mathcal{B}$ and Q are Hilbert-Schmidt.

From the above well established results, we know that no matter what kind of choice of the observation operator C we make, the optimal control law is given by

$$u^*(t) = -\mathcal{B}^*\mathcal{P}_c\hat{w}_a(t)$$

where $\mathcal{P}_c$ is independent of the observation operator. Therefore the gain of distributed sensing will only happen in filtering, in the sense that $\mathcal{P}_f = lim_{t\to\infty}\mathcal{P}_f(t)$ will be smaller. Furthermore, as a consequence of "better" recovery of state information by Kalman filtering through distributed sensing, the optimal feedback control should give a smaller final cost.

Next we first consider the following two SSRE's in Kalman filtering.

$$AP + PA^* + FF^* - PC^*CP = 0 \tag{8}$$

$$AP_d + P_dA^* + FF^* - P_dC_d^*C_dP_d = 0 \tag{9}$$

where $C : H \to I\!R^n, C_d : H \to H_o$ are linear and bounded.

We are going to show

Proposition 1 *Under the assumptions*

$$CC_d^*C_d = CC^*C \tag{10}$$

$$CC^* > 0 \tag{11}$$

there holds

$$P_d \leq P$$

provided that both P and P_d exist.

Before the proof of the proposition, we notice that the assumption (10) has the clear physical interpretation that the information yielded by the observation operator C is just a subset of that of C_d.

Proof: Consider the following two LQR problems on the infinite horizon,

$$\left\{ \begin{array}{rcl} \dot{x}(t) & = & A^* x(t) + C^* u(t) \\ x(0) & = & x \end{array} \right. \tag{12}$$

with cost functional

$$J(u;x) = \int_0^\infty ||F^* x(t)||^2 + ||u(t)||^2_{R^n} \, dt$$

and

$$\left\{ \begin{array}{rcl} \dot{x}(t) & = & A^* x(t) + C_d^* u(t) \\ x(0) & = & x \end{array} \right. \tag{13}$$

with cost functional

$$J_d(u;x) = \int_0^\infty ||F^* x(t)||^2 + ||u(t)||^2_{H_o} \, dt.$$

For P and P_d satisfying (8) and (9) respectively, we know that

$$[Px, x] = \min_{u \in L^2[(0,\infty), I\!R^n]} J(u;x) \tag{14}$$

$$[P_d x, x] = \min_{u \in L^2[(0,\infty), H_o]} J_d(u;x). \tag{15}$$

Next, in the second LQR problem (13) we perform optimization with respect to only a special class of control.

$$u = C_d C^* (CC^*)^{-1} \tilde{u} \qquad \text{with} \quad \tilde{u} \in L^2[(0,\infty), I\!R^n].$$

Since

$$\begin{array}{rcl} C_d^* u & = & C_d^* C_d C^* (CC^*)^{-1} \tilde{u} \\ & = & (C C_d^* C_d)^* (CC^*)^{-1} \tilde{u} \\ & = & (CC^* C)^* (CC^*)^{-1} \tilde{u} \\ & = & C^* \tilde{u} \end{array}$$

and

$$\begin{aligned} ||u||^2_{H_o} &= [CC_d^* C_d C^*(CC^*)^{-1}\tilde{u}, (CC^*)^{-1}\tilde{u}] \\ &= ||\tilde{u}||^2_{\mathbb{R}^n} \end{aligned}$$

then the second LQR problem becomes

$$\begin{cases} \dot{x}(t) &= A^* x(t) + C_d^* u(t) = A^* x(t) + C^* \tilde{u}(t) \\ x(0) &= x \end{cases}$$

with cost functional

$$\begin{aligned} J_d(C_d C^*(CC^*)^{-1}\tilde{u}; x) &= \int_0^\infty ||F^* x(t)||^2 + ||C_d C^*(CC^*)^{-1}\tilde{u}||^2_{H_o} dt \\ &= \int_0^\infty ||F^* x(t)||^2 + ||\tilde{u}(t)||^2_{\mathbb{R}^n} dt \\ &= J(\tilde{u}; x) \end{aligned} \quad (16)$$

Therefore, in the light of (14)-(16), we obtain

$$\begin{aligned} [P_d x, x] &= \min_{u \in L^2[(0,\infty), H_o]} J_d(u; x) \\ &\le \min_{\tilde{u} \in L^2[(0,\infty), \mathbb{R}^n]} J_d(C_d C^* \tilde{u}; x) \\ &= \min_{\tilde{u} \in L^2[(0,\infty), \mathbb{R}^n]} J(\tilde{u}; x) \\ &= [Px, x] \end{aligned}$$

□

Remark: Notice that in the proof of *Proposition 1*, we only need the assumptions (10) and (11), and C_d is not necessarily a distributed sensing operator. In addition, the requirement that Range(C) = $\mathbb{R}^n$ - a finite dimensional space, is superficial. Therefore, we have proved the conclusion:

$$C_1 C_2^* C_2 = C_1 C_1^* C_1 \quad \& \quad C_1 C_1^* > 0 \quad \Longrightarrow \quad P_{f,2} \le P_{f,1}$$

i.e., more observation implies smaller steady state filtering error.

It is also interesting to apply the above conclusion to finite dimensional Kalman filtering.

Next, we study the influence of distributed sensing to the value of final cost.

Proposition 2 *In addition to the assumptions (10) and (11), suppose* (A, B) *and* (A^*, Q^*) *are exponentially stabilizable. Then,*

$$\min_{u(\cdot,\omega)\in L^2[(0,\infty),H_u]} J_d(u) - \min_{u(\cdot,\omega)\in L^2[(0,\infty),H_u]} J(u) = \text{Tr. } B^* P_c(P_d - P_f)P_c B \leq 0 \tag{17}$$

where $J(u)$ *and* $J_d(u)$*, as given in (4), are the quadratic cost functionals corresponding to the systems with observation operators* C*,* C_d*, respectively.*

Here we impose the much stronger conditions of exponential stabilizability of (A, B) and (A^*, Q^*) as a sufficiency condition such that $(A - BB^*P_c)$ is exponentially stable. And hence

$$P_c x = \int_0^\infty S_c^*(t)Q^*QS(t)x \, dt \qquad x \in H$$

which implies P_c is nuclear, where $S(t)$ and $S_c(t)$ are the C_0-semigroups generated by A and $(A - BB^*P_c)$, respectively.

In general, P_c need not be nuclear without the above assumptions. For example, in the usual notations, let

$$\mathcal{A} = \begin{pmatrix} 0 & I \\ -A & 0 \end{pmatrix}; \qquad Q^* = \mathcal{B} = \begin{pmatrix} 0 \\ B \end{pmatrix}$$

where $B : H_u \to H$ is Hilbert-Schmidt. Then,

$$\mathcal{A} + \mathcal{A}^* = 0 \Rightarrow \mathcal{P}_c = I$$

which is not even compact. Obviously, in this particular case, $(\mathcal{A}, \mathcal{B})$ and $(\mathcal{A}^*, Q^*)$ are not exponentially stabilizable in view of Gibson's results [6], because $\mathcal{B}$ and Q^* are compact.

Proof: Recall that the minimum cost of (4) is given by (7) and P_f and P_c satisfy (6) and (5), respectively.

Let $\{e_n\}$ be an orthonormal basis in H. Then the minimal cost can be rewritten as

$$\begin{aligned}
J(u^*) &= Tr.QP_fQ^* + Tr.CP_fP_cP_fC^* \\
&= Tr.P_fQ^*Q + Tr.P_fC^*CP_fP_c \\
&= \sum_{n=1}^{\infty}\{[Q^*Qe_n, P_fe_n] + [P_fC^*CP_fP_ce_n, e_n]\} \\
&= \sum_{n=1}^{\infty}\{[P_f(P_cBB^*P_c - A^*P_c - P_cA)e_n, e_n] \\
&\qquad +[(AP_f + P_fA^* + FF^*)P_ce_n, e_n]\} \\
&= \sum_{n=1}^{\infty}\{[P_fP_cBB^*P_ce_n, e_n] + [FF^*P_ce_n, e_n]\} \\
&= Tr.P_fP_cBB^*P_c + Tr.FF^*P_c \\
&= Tr.B^*P_cP_fP_cB + Tr.F^*P_cF
\end{aligned}$$

Similarly, we have

$$J_d(u^*) = Tr.B^*P_cP_dP_cB + Tr.F^*P_cF$$

Therefore, subtraction of the two gives (17). □

The conclusions of the above propositions can be extended to the finite time horizon LQG problem:

$$\begin{aligned}
\dot{x}(t) &= Ax(t) + Bu(t) + F\omega(t) \qquad 0 \le t \le T \\
y(t) &= Cx(t) + G\omega(t) \\
Ex(0)x(0)^* &= \Lambda
\end{aligned} \tag{18}$$

where $GG^* = I, FG^* = 0$, and $B : H_u \longrightarrow H$ is Hilbert-Schmidt. The cost functional is given by

$$J(u) = \int_0^T Tr.E[(Qx(t,\omega)(Qx(t,\omega)^*]dt + \int_0^T Tr.E[u(t,\omega)u(t,\omega)^*]dt$$

where $Q : H \longrightarrow H$ is again Hilbert-Schmidt.

Let us consider the differential Riccati equation corresponding to the Kalman filtering

$$\begin{aligned} \dot{P}_f(t) &= AP_f(t) + P_f(t)A^* + FF^* - P_f(t)C^*CP_f(t) \\ P_f(0) &= \Lambda \end{aligned} \tag{19}$$

and the differential Riccati equation corresponding to the optimal feedback gain

$$\begin{aligned} \dot{P}_c(t) &= -A^*P_c(t) - P_c(t)A - Q^*Q + P_c(t)BB^*P_c(t) \\ P_c(T) &= 0 \end{aligned} \tag{20}$$

In what follows, we use $P_d(\cdot)$ to denote the solution of (19) with C being replaced by C_d, and $J_d(u)$ is similarly defined. Then we have

Proposition 3 *Under the assumptions*

$$\begin{aligned} CC_d^*C_d &= CC^*C \\ CC^* &> 0 \end{aligned}$$

there hold

$$P_d(t) \le P_f(t) \quad \text{for} \quad 0 \le t \le T$$

$$\begin{aligned} &\min_{u \in L^2[(0,T);H_u]} J_d(u) - \min_{u \in L^2[(0,T);H_u]} J(u) \\ &= \int_0^T Tr.B^*P_c(t)[P_d(t) - P_f(t)]P_c(t)Bdt \le 0 \end{aligned}$$

Proof: First of all, By [1] we know $P_f(\cdot), P_d(\cdot)$ and $P_c(\cdot)$ all exist and are unique.

Let $K(t) = P_f(T-t)$ and $K_d(t) = P_d(T-t)$, then $K(t)$ satisfies

$$\begin{aligned} \dot{K}(t) &= -AK(t) - K(t)A^* - FF^* + K(t)C^*CK(t) \\ K(T) &= \Lambda \end{aligned}$$

while $K_d(t)$ satisfies the same equation with C being replaced by C_d.

Consider the following finite horizon LQR problems

$$\begin{aligned} \dot{x}(t) &= A^*x(t) + C^*\tilde{u}(t) \\ x(t_0) &= x \qquad 0 \le t_0 \le T \\ J(t_0;\tilde{u},x) &= [\Lambda x(T), x(T)] + \int_{t_0}^{T} [FF^*x(t), x(t)] + \|\tilde{u}(t)\|^2 dt \end{aligned} \tag{21}$$

and

$$\begin{aligned} \dot{x}(t) &= A^*x(t) + C_d^*u(t) \\ x(t_0) &= x \qquad 0 \le t_0 \le T \\ J_d(t_0;u,x) &= [\Lambda x(T), x(T)] + \int_{t_0}^{T} [FF^*x(t), x(t)] + \|u(t)\|^2 dt \end{aligned} \tag{22}$$

It is well known that [1]

$$\begin{aligned} \min_{\tilde{u}\in L^2[(0,T),R^n]} J(t_0;\tilde{u},x) &= [K(t_0)x, x] = [P_f(T-t_0)x, x] \\ \min_{u\in L^2[(0,T),H_o]} J_d(t_0;u,x) &= [P_d(T-t_0)x, x] \end{aligned}$$

Then, similarly as in the proof of *Proposition 1*, by taking

$$u = C_dC^*(CC^*)^{-1}\tilde{u}$$

we can establish

$$\begin{aligned} [P_d(T-t_0)x, x] &= \min_{u\in L^2[(0,T),H_o]} J_d(t_0;u,x) \\ &\le \min_{\tilde{u}\in L^2[(0,T),R^n]} J_d(t_0; C_dC^*(CC^*)^{-1}\tilde{u}, x) \\ &= [P_f(T-t_0)x, x] \end{aligned}$$

For the second half of the conclusion, we can show in a similar manner as in the proof of *Proposition 2*, that

$$\min_{u\in L^2[(0,T),H_u]} J(u)$$
$$= \int_0^T Tr.B^* P_c(t) P_f(t) P_c(t) B + Tr.F^* P_c(t) F dt + Tr.\Lambda P_c(0)$$

from which the result follows by subtraction, noticing that $P_c(\cdot)$ is independent of the observation operator C or C_d. □

Next, let us consider the effect of distributed sensing on active damping of large flexible space structures. Since we are concerned with the design of stablizing control law instead of numerical solution of the structure response, we use the following continuuam model in order to avoid losing the insight of the infinite dimensional nature of flexible structures,

$$\ddot{x}(t) + \gamma D(x(t), \dot{x}(t)) + Ax(t) = Bu + FN(t) \tag{23}$$

where $A : \mathcal{D}(A) \longrightarrow H$ is self-adjoint, nonnegative, has compact resolvent and dense domain. We further assume that 0 belongs to $\rho(A)$, the resolvent set of A, i.e., there is no rigid body mode.

$B : \mathbb{R}^n \longrightarrow H$ is linear and bounded.

$F : H_n \longrightarrow H$ is a Hilbert-Schmidet operator.

$D(x, \dot{x})$ is the (generally nonlinear) damping term, such that $[D(x, \dot{x}), \dot{x}] \geq 0$.

$N(t)$ represents a white noise in the Hilbert space H_n.

Since the damping coefficient $\gamma > 0$ is typically very small in flexible space structures, we set $\gamma = 0$ in the following studies for the sake of simplicity.

As usual, we may define

$$\mathcal{A} = \begin{pmatrix} 0 & I \\ -A & 0 \end{pmatrix}; \quad \mathcal{B} = \begin{pmatrix} 0 \\ B \end{pmatrix}; \quad \mathcal{F} = \begin{pmatrix} 0 \\ F \end{pmatrix}$$

so that (23) can be rewritten as

$$\frac{dY(t)}{dt} = \mathcal{A}Y(t) + \mathcal{B}u + \mathcal{F}N(t) \quad \text{where } Y = \begin{pmatrix} x \\ \dot{x} \end{pmatrix} \tag{24}$$

with $\mathcal{D}(\mathcal{A}) = \mathcal{D}(A) \otimes H$ and equipped with the energy norm

$$||Y||_E^2 = [A^{1/2}x, A^{1/2}x] + [\dot{x}, \dot{x}]$$

on the Hilbert space $\mathcal{H}_E = \mathcal{D}(A^{1/2}) \otimes H$.

From practical point view, only finite number of controllers can be present and they are of saturation type due to limited control forces and moments [10]. Therefore,

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \qquad \text{and } |u_i| \leq U_i, \; i = 1, 2, \cdots, n$$

where U_i are certain positive constants derived from maximum force/moments available.

The total energy of the flexible structure governed by the continuum model (23) is defined by

$$E(t) = 1/2||Y(t)||_E^2 = 1/2([Ax(t), x(t)] + [\dot{x}(t), \dot{x}(t)])$$

Since the energy decay rate is given by

$$\frac{dE(t)}{dt} = [u, B^*\dot{x}(t)] + [FN(t), \dot{x}(t)]$$

then obviously the optimal feedback control for maximum energy decay rate is given by

$$u_j(t) = -U_j \text{sign}[\frac{db_j(t)}{dt}] \quad j = 1, 2, \cdots, n$$

where

$$b(t) = \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_n(t) \end{pmatrix} = B^* x(t)$$

Therefore the control law only requires the rate information at the locations of the controllers $\dot{b}(t) = B^* \dot{x}(t)$. In other words, the rate measurements made by colocated sensors

$$y(t) = B^* \dot{x}(t)$$

are necessary and also sufficient for active damping purpose. More observation than that made by colocated sensors does not make any difference in terms of energy decay. Therefore, in this particular case, optical distributed sensing does not bring any gain.

However, if we take sensor noise into consideration, then the situation will be quite different.

Next, let us consider the case in which the observation of structure deflection rate is corrupted by additive sensor noise, i.e.

$$v(t) = C\dot{x}(t) + GN(t)$$

where $FG^* = 0$ (observation noise and state noise are independent), and C could be colocated sensor B^* or optical distributed sensor C_d for later comparison.

For simplicity of analysis, let us assume that linear feedback is allowed, which is reasonable for small deflection magnitude. Then a reasonable choice for stablizing feedback control would be

$$u(t) = -B^* \hat{\dot{x}}(t) = -\mathcal{B}^* \hat{Y}(t)$$

Then the closed loop system dynamics becomes

$$\frac{dY(t)}{dt} = \mathcal{A}Y(t) - \mathcal{B}\mathcal{B}^* \hat{Y}(t) + \mathcal{F}N(t)$$

$$= (\mathcal{A} - \mathcal{B}\mathcal{B}^*)Y(t) + \mathcal{B}\mathcal{B}^*(Y(t) - \hat{Y}(t)) + \mathcal{F}N(t)$$

We therefore see that due to filtering error, we thereby introduce a random input to the system which may excite higher order modes. And in this case, it is easy to verify that the energy decay rate becomes

$$\begin{aligned} \frac{dE(t)}{dt} &= -\|B^*\dot{x}(t)\|^2 + \|B^*(\dot{x}(t) - \hat{\dot{x}}(t))\|^2 \\ &\quad + [B^*(\dot{x} - \hat{\dot{x}}), B^*\hat{\dot{x}}] + [FN(t), \dot{x}(t)] \end{aligned} \tag{25}$$

The second term in (25) represents a positive energy input, which discounts the active damping provided by the controllers. The third and the fourth terms can be considered as zero mean disturbance. Therefore, we are facing two options to further enhance the structure stability:

1. Increase the active damping by putting more controllers and increasing the force/moment upper bounds, i.e., to increase the stability margin of $(\mathcal{A} - \mathcal{B}\mathcal{B}^*)$ by changing B. But this option generally results in heavier weight and bigger cost.

2. Decrease the positive energy input

   $$\|B^*(\dot{x}(t) - \hat{\dot{x}}(t))\|^2$$

   which is caused by filtering error, i.e. to decrease P_4, the error covariance operator of $\dot{x}(t) - \hat{\dot{x}}(t)$, by employing optical distributed sensor.

Let

$$\mathcal{P} = \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}$$

be the steady state Kalman filtering error covariance operator. It is generally a formidable problem to solve the corresponding SSRE

$$\mathcal{A}\mathcal{P} + \mathcal{P}\mathcal{A}^* + \mathcal{F}\mathcal{F}^* - \mathcal{P}\mathcal{C}^*\mathcal{C}\mathcal{P} = 0 \tag{26}$$

to find $\mathcal{P}$.

Next, a very wild attempt is made to find an "approximate" solution of (26).

First of all, since we are interested in a self-adjoint solution, it is necessary to find the expression of $\mathcal{P}^*$ with respect to the $\mathcal{H}_E$ inner product. It is easy to verify that

$$\mathcal{P}^* = \begin{pmatrix} A^{-1}P_1^*A & A^{-1}P_3^* \\ P_2^*A & P_4^* \end{pmatrix}$$

Then $\mathcal{P}$ is self-adjoint if and only if $P_j,\ j = 1,2,3,4$ satisfy the following equations

$$AP_1 = (AP_1)^* \tag{27}$$

$$P_3 = (AP_2)^* \tag{28}$$

$$P_4 = P_4^* \tag{29}$$

As above, suppose only rate measurement is made, i.e. $\mathcal{C} = (0, C)$, then direct computation using (26) gives

$$P_2 + P_2^* - P_2C^*CP_2^* = 0 \tag{30}$$

$$P_4 - P_1 - P_2C^*CP_4 = 0 \tag{31}$$

$$AP_2 + P_2^*A - FF^* + P_4C^*CP_4 = 0 \tag{32}$$

Therefore, to solve (26) for a self-adjoint solution is equivalent to solving equations (27) - (32). For this purpose, we first eliminate P_1 and P_2 from (27) - (32) by noticing

$$P_2 = A^{-1}P_3^* \tag{33}$$

$$P_1 = P_4 - A^{-1}P_3^*C^*CP_4 \tag{34}$$

from (28) and (31).

Left- and right-multiplication of A to (30) gives

$$AP_3 + P_3^*A - P_3^*C^*CP_3 = 0 \tag{35}$$

And (32) can be rewritten as

$$P_3 + P_3^* - FF^* + P_4C^*CP_4 = 0 \tag{36}$$

Substituting (34) into (27) gives

$$AP_4 + P_4C^*CP_3 = P_4A + P_3^*C^*CP_4 \tag{37}$$

Therefore, the problem becomes solving P_3 and P_4 from (35) - (37) and (29). Once P_3 and P_4 are found, P_1 and P_2 are immediately given by (33) and (34).

Now we attempt to find an "approximate" solution of the form $P_3 = 0$, in which case, $P_2 = 0$ and $P_1 = P_4$.

By the assumption on A, we know there is a sequence of normalized modes ϕ_n's and corresponding frequencies ω_n's such that

$$A\phi_n = \omega_n^2\phi_n \quad n = 1, 2, \cdots$$

with $\lim_{n\to\infty}\omega_n^2 = \infty$ and $\{\ \phi_n, n = 1, 2, \cdots\}$ being an orthonormal basis in H.

Since $P_3 = 0$, (35) is trivially satisfied. We only need to determine P_4 from (36) and (37), which become

$$FF^* = P_4C^*CP_4 \tag{38}$$

$$AP_4 = P_4A \tag{39}$$

and (29).

If we assume ω_n's are distinct, then it is easy to show that

$$AP_4 = P_4A \text{ and } P_4 = P_4^*$$

if and only if there is a sequence of nonnegative numbers $\{\ a_n,\ n = 1, 2, \cdots\}$ such that

$$P_4\phi_n = a_n\phi_n \quad n = 1, 2, \cdots$$

Now a_n's can be determined from (38), i.e.

$$[FF^*\phi_n, \phi_n] = [P_4C^*CP_4\phi_n, \phi_n] = a_n^2||C\phi_n||^2$$

which immediately gives

$$a_n = \frac{||F^*\phi_n||}{||C\phi_n||} \quad n = 1, 2, \cdots$$

Therefore, we have found an "approximate" solution $\mathcal{P}$ of (26) given by

$$\mathcal{P} = \begin{pmatrix} P_4 & 0 \\ 0 & P_4 \end{pmatrix} \tag{40}$$

where P_4 is a self-adjoint, nonegative operator defined by

$$P_4\phi_n = \frac{||F^*\phi_n||}{||C\phi_n||}\phi_n \quad n = 1, 2, \cdots$$

with domain

$$\mathcal{D}(P_4) = \{x \in H \mid \sum_{n=1}^{\infty}(\frac{||F^*\phi_n||}{||C\phi_n||})^2[x, \phi_n]^2 < \infty \}$$

This is an "approximate" solution in the sense that it satisfies only equations (35), (37) and (29), but not (36). In fact, if we consider the left hand side of (36) as an infinite dimensional matrix, then the above solution only makes all the diagonal elements zero, while the off diagonal elements are not necessarily zero.

It is easy to see that a sufficient condition for (40) to be an exact solution of (26) is

$$[FF^*\phi_n, \phi_m] = [P_4C^*CP_4\phi_n, \phi_m] = a_n a_m[C^*C\phi_n, \phi_m] \quad \forall n \neq m$$

or, equivalently,

$$\frac{[F^*\phi_n, F^*\phi_m]}{||F^*\phi_n||\,||F^*\phi_m||} = \frac{[C\phi_n, C\phi_m]}{||C\phi_n||\,||C\phi_m||} \quad \forall n \neq m$$

which is a very strong requirement.

4 Conclusions

In summary, optical distributed sensing brings gain in the situation where Kalman filtering is necessary for state estimation. In this case, both the transient and the stationary Kalman filtering error covariance become smaller. Due to smaller filtering error, the same control law results in smaller cost in LQG problem. These conclusions, of course, also apply to finite dimensional Kalman filtering and LQG problems as special cases.

By considering the stabilization problem of flexible space structures, we realize that it is important to use distributed sensors in order to reduce the positive energy input caused by filtering error.

References

[1] A. V. Balakrishnan, *Applied Functional Analysis*, 2nd edition, Springer-Verlag, 1981.

[2] A. V. Balakrishnan, Strong stabilizability and the steady state Riccati equation, *App. Math. & Optim.* Vol 7, pp 335-345, 1981.

[3] R. J. Collier, C. B. Burckhardt and L. H. Lin, *Optical Holography*, Academic Press, 1971.

[4] J. Feinberg, Optical phase conjugation in photorefractive materials, *Optical Phase Conjugation*, R. Fisher, Editor. Academic Press, 1983.

[5] A. Feldman, R. M. Waxler, Properties of crystalline materials for optics, *Physical Properties of Optical Materials*, Roy F. Potter, Editor. Proc SPIE, Vol. 204, pp68-76, 1979.

[6] J. S. Gibson, A note on stabilization of infinite dimensional linear oscillators by compact linear feedback, *SIAM J. Control* . Vol 18, pp311-316, 1980.

[7] M. Herzberger, *Modern Geometrical Optics.* Interscience Publishers, Inc., New York, 1958.

[8] G. O. Reynolds, J. B. DeVelis, G. B. Parrent, Jr. and B. J. Thompson, *The New Physical Optics Notebook: Tutorials in Fourier Optics.* SPIE Optical Engineering Press, 1989.

[9] C. T. Seaton, G. I. Stegeman, S. S. Yao, C. Karagulett, R. Fortenberry, A. Gabel, W. Banyai and G. Assanto, High speed intensity-dependent refractive index in semiconductor-doped glasses, *Advances in Materials for Active Optics*, Solomon Musikant, Editor, Proc SPIE, Vol. 567, pp132-136, 1986.

[10] L. W. Taylor and A. V. Balakrishnan, A mathematical problem and a spacecraft control laboratory experiment (SCOLE) used to evaluate control laws for flexible spacecraft - NASA/IEEE Design Chanllenge. *Proc. of the 1st NASA SCOLE Workshop, NASA Langley Research Center, Hampton, Virginia*, December 6-7, 1984.

[11] G. C. Valley and M. B. Klein, Optical properties of photorefractive materials for optical data processing, *Advances in Optical Information Processing*, G. Michael Morris, Editor, Proc. SPIE, Vol. 388, pp112-121, 1983.

[12] J. Zabczyk, Remarks on the algebraic Riccati equation in Hilbert space, *App. Math. & Optim.* Vol 2, pp251-258, 1976.

[13] R. C. Montgomery and S. S. Welch, Application of optical distributed sensing and computation to control of a large flexible spacecraft, Seventh VPI&SU/AIAA Symposium on Dynamics and Control of Large Structures. To be published.