JPRS-CST-90-023
31 AUGUST 1990

FOREIGN BROADCAST INFORMATION SERVICE

JPRS Report

Science & Technology

China

JPRS-CST-90-023
31 AUGUST 1990

SCIENCE & TECHNOLOGY

CHINA

CONTENTS

A Multifunction Space-Borne Scanner

90FE0179A Beijing YUHANG XUEBAO [JOURNAL OF CHINESE SOCIETY OF ASTRONAUTICS] in Chinese No 2, 30 Apr 90 pp 52-56

[Article by Kong Huixing [7895 1920 5281], Shanghai Institute of Technical Physics]

[Text] Abstract

A proposal is made to add five detection bands to an existing very-high-resolution (VHR) scanning radiometer to enhance its capability for the observation of meteorological phenomena, ocean environment, and natural disasters.

(I) Performance of the Existing Very-High-Resolution Scanning Radiometer

Test results show that the performance indices of China's existing VHR scanning radiometer in the visible-light and infrared band meet all the design requirements, as shown in Table 1. Both visible-light and infrared channels produce clear images, and have good observation capability for cloud maps, ground covers, lakes and ocean water.

(II) Satellite and Radiometer Design Requirements for Monitoring the Environment and Natural Disasters

Because natural disasters occur abruptly in time and in space, real-time monitoring is required in order to provide timely reporting and assistance. An effective approach of monitoring the entire Chinese landmass is to use a satellite-borne VHR scanning radiometer operating at high repetition frequency.

The results of recent studies in earth science and environmental science show that in order to explain the evolution of certain important environmental phenomena, it is necessary to consider the earth as a single entity, and to study the global effects between the atmosphere, the ocean, and the land, the gain and loss of earth radiation, and the combustion of fossil fuels, etc. To accomplish this requires the accumulation of over 10 years of observations with high spatial and time resolution, as indicated in Table 2.

Table 1. On-Orbit Performance of Very-High-Resolution Scanning Radiometer

Parameter	Design specification	Test results
Detection band	0.58-0.68 μm	Visible cloud maps, ground cover, cloud and snow separation, ocean sediment
	0.725-1.1 μm	Visible cloud maps, ground cover, cloud and snow separation, water and land border separation
	0.48-0.53 μm	Ocean water color, chlorophyll distribution
	0.53-0.58 μm	Ocean water color, chlorophyll distribution
	10.5-12.5 μm	Infrared cloud maps, water and land border separation, temperature measurement
Spatial resolution	1.08 km	1.08 km
Detection sensitivity	Ch_1, Ch_2, $NE\Delta\rho \leq 0.16\%$	$NE\Delta\rho \leq 0.11\%$
	Ch_3, Ch_4, $NE\Delta\rho \leq 0.25\%$	$NE\Delta\rho \leq 0.15\%$
	Ch_5 $NE\Delta T \leq 0.8$ K	$NE\Delta T \leq 0.5$ K
Detection accuracy	Ch_3, Ch_4 ±10%	±10%
	Ch_5 1.0 K	approximately 1.0 K

Table 2. Time and Spatial Resolution Requirements for Environmental Satellite Observations

Object of observation	Repetition period	Ground resolution
Atmosphere	At least 0.5 day/observation	1 km
Ocean	0.5-1 day/observation (for ocean conditions)	0.5-1 km
	3-5 days/observation (for medium and large-scale ocean phenomena)	1-4 km
Ground cover	7-10 days/observation	0.5-1 km

Although earth resource satellites have very high spatial resolution, it cannot satisfy the repetition-frequency requirement for environmental monitoring because it takes 16-26 days to cover the whole earth. In addition, the multi-band scanner of an environmental satellite has 5-10 times higher sensitivity than that of an earth resource satellite (see Table 3) and therefore can detect minute changes in the environmental parameters and predict its future development. Because of these unique features of the environmental satellite, it should be further developed to provide the capability for monitoring the environment and natural disasters.

Table 3. Comparison of Multi-Band Scanners of an Environmental Satellite and of an Earth Resource Satellite

Parameter	MSS	TM	AVHRR	C_2CS	SeawiFS	Multi-band scanner
Repetition period	18 days	16 days	1 day	2 days	1 day	1 day
IFOV visible near-infrared	86 μr	42 μr	1.3 mr	0.865 mr	1.6 mr	1.2 mr
thermal infra-red	172 μr	168 μr	1.3 mr	0.865 mr	1.6 mr	1.2 mr
Ground resolution visible near-infrared	79 m	30 m	1.08 km	0.826 km	1.13 km	1.08 km
thermal infra-red	158 m	120 m	1.08 km	0.826 km	1.13 km	1.08 km
Spectral resolution visible near-infrared	0.1-0.2 μm	0.06-0.27 μm	0.1-0.24 μm	0.02-0.1 μm	0.02-0.044 μm	0.05-0.24 μm
thermal infra-red	2.0 μm	2.0 μm	1.0 μm	2.0 μm	1.0 μm	2.0 μm
Sensitivity visible near-infrared NE$\Delta\rho$	0.50-0.70%	0.50-0.24%	approx. 0.1%	approx. 0.05%	0.5-2.4%	0.08-0.15%
thermal infra-red NEΔT	1.4 K	0.5 K	0.15 K	0.25 K	0.29 K	0.20 K

(III) 10-Band Scanning Radiometer

In order to enhance the monitoring capability of the existing VHR scanning radiometer, it is necessary to add five more detection bands to the existing bands to provide a 10-band observation capability. Specifically, the enhancements include the following:

(1) Adding a 3.55-3.93 μm medium-infrared band for monitoring forest fires and grassland fires; the low water absorption in this band will provide improved temperature measurement accuracy over the ocean surface in the mid-low latitude region.

This band has very high sensitivity with respect to fire because it is located at the radiation peak of a 780 K target. The following table shows the calculated results of radiation luminance for targets at different temperatures.

Table 4. Radiation Luminance of Targets at Different Temperatures in the 10.5-12.5 μm Band and the 3.55-3.93 μm Band

Target temperature / Target luminance	T=1100 K	T=800 K	T=300 K	N(1100 K) / N(300 K)	N(800 K) / N(300 K)
3.55-3.93 μm Luminance ($W/cm^2 \cdot ster$)	0.1924	5.08×10^{-2}	1.7×10^{-5}	1.13×10^{4}	3.9×10^{3}
10.5-12.5 μm Luminance ($W/cm^2 \cdot ster$)	5.59×10^{-2}	3.13×10^{-2}	1.86×10^{-3}	30.0	16.8

The radiated power received by a scanner from a target which does not fill the instantaneous field of view is proportional to the target area. Thus, the radiated power of a 20x20 m^2 fire at 800°K is equivalent to that of a 1.1x1.1 Km^2 target at 300°K. Taking into account atmospheric attenuation and margin threshold setting, it is estimated that the scanner can detect a 30x30 m^2 open fire and issue a fire alarm.

(2) Adding a 1.58-1.64 μm shortwave infrared band for measuring the water content in soil and crops and for monitoring drought conditions.

The soil reflectivity drops sharply in the 1.3-2.5 μm band when it contains water; in particular, there are water absorption peaks at 1.44 μm, 1.90 μm and 2.2 μm. When the soil moisture is in the 0.8 percent to 20.2 percent range, an increase of 1 percent in water content will reduce the reflectivity by 1 percent (see Figure 1). The water content in the leaves of ground vegetation will clearly affect the reflectivity in the 1.3-2.5 μm band (see Figure 2). The sensitivity of the scanning radiometer is estimated to $NE\Delta\rho \leq 0.1\%$, which is sufficient to detect a 1 percent change in water content; therefore, the scanner is effective in monitoring drought conditions. Considering the fact that there are H_2O and CO_2 absorption in the atmosphere, the detection band is chosen to be 1.58-1.64 μm.

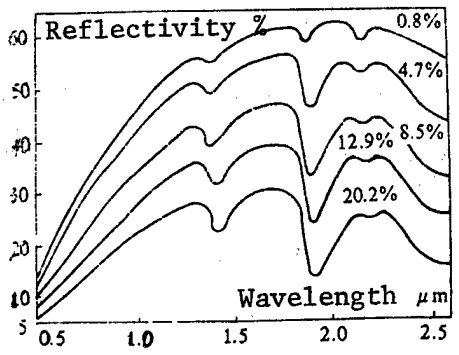


Figure 1. Spectral Characteristics of Soil Reflectivity With Different Moisture Content

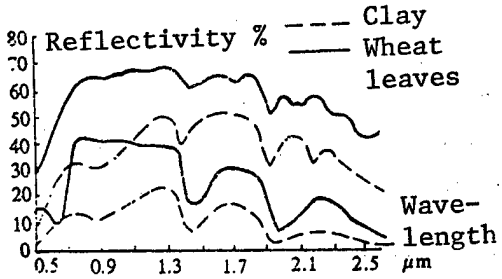


Figure 2. Reflectivity Spectrum of Clay and Wheat Leaves With Different Moisture Content

(3) Adding a 0.433-0.453 μm band for measuring the chlorophyll content in the ocean. This band supplements the existing 0.48-0.53 μm band and the 0.53-0.58 μm band to measure the water color which reflects a wide range of chlorophyll content. The collected data can not only be used for preliminary survey of ocean productivity but also be used in conjunction with surface temperature data for fishing-ground forecast.

(4) Because of scattering by molecules and colloidal particles in the atmosphere in the 0.433-0.453 μm band, the sky background is one order of magnitude brighter than the ocean surface. In order to isolate the water-color signal of the ocean from the signal of the bright sky background, it is necessary to add a 0.745-0.785 μm band to correct for the effect of the colloidal particles in the atmosphere.

(5) The existing 10.5-12.5 μm thermal infrared band of the scanning radiometer is split into two bands: 10.5-11.5 μm and 11.5-12.5 μm, in order to partially cancel the atmospheric effect so that higher measurement accuracy of the ocean temperature (approximately 1.5 K) can be achieved.

(6) Because of the large differences between the reflectivities of soil, ground cover, and water in the 0.725-1.1 μm band, it is an important band for monitoring land-water separation, river flow rate, and flood damage.

In order to improve the capability of separating water and land, the ground resolution in this band is increased to 540 m.

(7) Currently, the 0.725-1.1 μm band and the 0.58-0.68 μm band are combined to monitor the ground cover index, which is a measure of ground cover distribution, growth trend of crops, and estimated wheat production. In order to match the spatial resolution between these two bands, the resolution in the 0.58-0.68 μm band is also increased to 540 m.

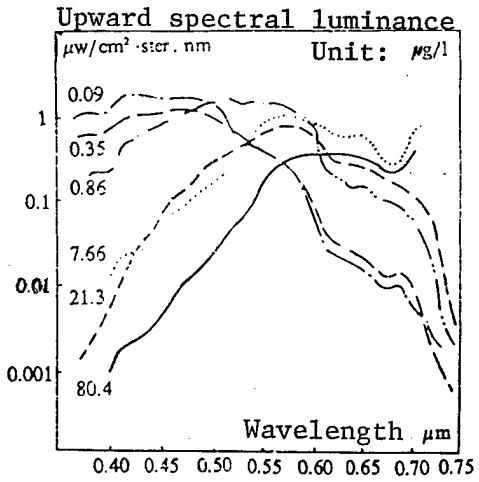


Figure 3. Upward Spectral Luminance of Ocean Water With Different Chlorophyll Content

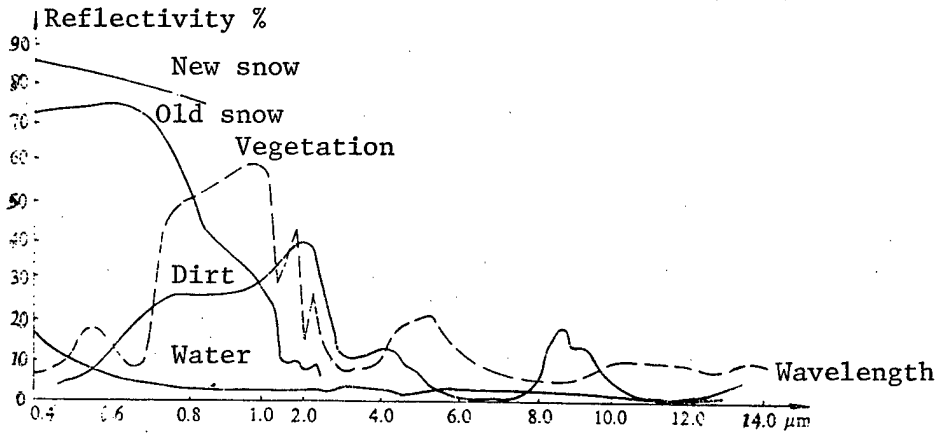


Figure 4. Reflectivity Spectrum of Typical Targets

The modified scanning radiometer has 10 detection bands. It includes the existing five bands of the AVHRR of the NOAA weather satellite and the 1.58-1.64 μm band which is to be added on the NOAA-K, L, M satellites. The addition of the 0.433-0.453 μm band and the 0.745-0.785 μm band to the existing two ocean channels of the scanning radiometer provides the quantitative measurement capability of chlorophyll content and sediment level in

the ocean. Therefore, the improved 10-band scanning radiometer will have an integral capability of monitoring meteorological phenomena, water color as well as natural disasters such as floods, fires, and drought.

The author wishes to express his thanks to Mr Kuang Dingpo and Mr Zheng Qinpo for their assistance in this work.

Table 5. Performance of the 10-Band Scanning Radiometer

Channel No	Band	Spatial resolution	Coverage period	Application
1	0.433-0.453 μm	1.08x1.08 km^2	24 hours	Ocean chlorophyll
2	0.48-0.53 μm	1.08x1.08 km^2	24 hours	Ocean chlorophyll
3	0.53-0.58 μm	1.08x1.08 km^2	24 hours	Ocean chlorophyll, yellow substance
4	0.58-0.68 μm	0.54x0.54 km^2	24 hours	Cloud maps, ground cover, ocean sediment
5	0.725-1.1 μm	0.54x0.54 km^2	24 hours	Cloud maps, ground cover, water and land separation
6	0.745-0.785 μm	1.08x1.08 km^2	24 hours	Atmospheric correction
7	1.58-1.64 μm	1.08x1.08 km^2	24 hours	Drought detection, cloud and snow separation
8	3.55-3.93 μm	1.08x1.08 km^2	12 hours	Fire detection, ocean surface temperature
9	10.5-11.5 μm	1.08x1.08 km^2	12 hours	Infrared cloud maps, temperature measurement
10	11.5-12.5 μm	1.08x1.08 km^2	12 hours	Infrared cloud maps, temperature measurement

References

1. Zheng Qinpo, et al., "Weather Satellite Remote Sensor and Evaluation of the Quality of Cloud Maps," 1988, 10.

2. Fang Zhongyi, "Application of Weather Satellite Data," 1989, 8.

3. Gordon, H. R., Clark, D. K., Muller, J. L., and Hovis, W. A., "Phytoplankton Pigments From the Nioubus-7 Coastal Zone Color Scanner: Comparisons With Surface Measurements," SCIENCE, 210, 63, 1980.

4. Esaias, W., et al., "Moderate Resolution Imaging Spectrometer," NASA EOS. 1986.

5. "Multispectral Scanning System and Their Potential Application to Earth Resource," SURVEY, 1972.

Dynamics of Simplified Model of Solar Panel and Supporting Components

90FE0179B Beijing YUHANG XUEBAO [JOURNAL OF CHINESE SOCIETY OF ASTRONAUTICS] in Chinese No 2, 30 Apr 90 pp 57-64

[Article by Zhang Jiazhong [1728 0857 6988] and Shao Chengxun [6730 2052 8133], Harbin Polytechnical University]

[Text] Abstract

Experimental results and theoretical analysis of a simplified model of a solar panel and its supporting components used on a space station are presented. Specifically, the dynamic characteristics of the system when it is subject to disturbances such as attitude adjustment and orbit maneuvers are investigated. The results show that because of the inertial non-linear coupling effect of a large flexible space structure, the presence of internal resonance will produce abrupt changes in the dynamic response of the structure when the amplitude of the disturbance torque exceeds a threshold value. These changes exhibit non-linear characteristics of simultaneous excitation of the coupled modes. The abrupt increases in the amplitudes are expected to cause difficulties in maintaining the shape and controlling vibration of the flexible structure.

I. Introduction

For a large space structure such as the space station which operates in near-zero load conditions, the strength requirements are very low. To reduce the weight of the structure, many of its components are extremely flexible. However, these components are subject to large overload during orbit injection; therefore they are compactly stowed during launch, then deployed or re-assembled after achieving orbit. A typical example of such flexible component is the foldable solar panel discussed in reference 1. Figure 1 shows the configuration of a postulated space station, which consists of the extremely flexible solar arrays, the radiation panels and a large antenna.

Because a space station generally undergoes a series of on-orbit operations such as deployment and re-assembly of its components, docking with the spacecraft, attitude control and orbit maneuvers, it is inevitably subject to different types of disturbances. A large flexible structure such as the

space station has a high concentration of low-frequency natural modes. An analysis of a postulated space station[2] shows that there are 14 natural modes in the frequency range 0.21-0.75 Hz, and 20 natural modes within a 2-Hz range. These include the modes of the flexible structure as well as the aggregate modes of the space station. Furthermore, these flexible components which are connected through the main body of the space station are mutually coupled. The non-linear coupling has a significant effect on the dynamic characteristics of the system. Reference 3 pointed out that these non-linear coupling effects will cause local vibrations in the structure. References 4 and 5 have analyzed the effect of inertial non-linear coupling on the dynamic characteristics of a flexible structure, and pointed out that the presence of internal resonance in the system will cause abrupt changes in the system dynamic response even if the coupling is weak. For a large flexible structure (Figure 1) with a large number of natural modes, the presence of internal resonance is almost unavoidable. Therefore, this effect must be carefully studied.

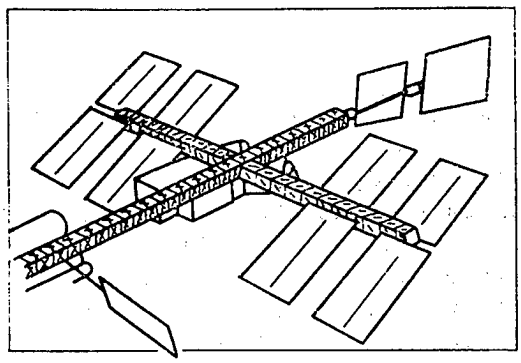


Figure 1. Proposed Space Station

The dynamic equations of a large space structure expressed in modal coordinates are of the following form:

$$\ddot{q}_r + \omega_r^2 q_r = \frac{Q_r}{M_r} - \varepsilon\left[\sum_i \sum_j (A_{rij}\ddot{q}_i\dot{q}_j + B_{rij}\dot{q}_i\dot{q}_j) + 2\omega_r\xi_r\dot{q}_r\right] + \varepsilon^2[\text{3rd order terms}] \tag{1}$$

where ε is a small parameter, and the second term on the right side of the equation is the second-order inertial coupling term. Q_r is the generalized force, M_r is the generalized mass, ω_r is the frequency of the r-th order linear mode, and A_{rij} and B_{rij} are the corresponding coupling coefficients. Under normal conditions, the effect of these non-linear coupling terms can be neglected, which leads to the conventional linear dynamic equations:

$$\ddot{q}_r + 2\omega_r\xi_r\dot{q}_r + \omega_r^2 q_r = \frac{Q_r}{M_r} \tag{2}$$

Equation (2) provides an accurate description of the system characteristics and yields the correct system response. But when internal resonance exists between the natural modes of the system, i.e., when the relationship

$$\omega_i \doteq \omega_j \tag{3}$$

exists between two natural modes, or when the relationship

$$\omega_k \doteq \omega_i + \omega_j \tag{4}$$

exists between three natural modes, or when both relationships are present, then as the amplitude of the disturbance increases, abrupt changes of the system dynamic response may occur. In this case, the effect of non-linear coupling can no longer be neglected. This article presents the experimental results and theoretical analysis of the phenomenon of internal resonance under non-linear coupling for a simplified model of solar panel and its supporting components.

II. Experiments

In order to study the effect of inertial non-linear coupling on system dynamics, we have conducted an experimental investigation and performed theoretical analysis of a simple model of solar panel and its supporting system as shown in Figure 2. The solar panel is modeled as a beam under bending and torsional loads, the supporting system is modeled as another beam. We only consider the case where the solar panel is perpendicular to the supporting beam; the physical parameters of the two components are adjusted so that the natural bending mode in the plane of the solar panel and the supporting components is ω_1 = 6.1 Hz, the natural torsional mode about the axis of the supporting beam is ω_T = 56 Hz, and the natural torsional mode of the solar panel about its own axis is ω_φ = 52 Hz. Clearly, these three frequencies satisfy the following internal resonance relationship:

$$\omega_T \doteq \omega_1 + \omega_\varphi \tag{5}$$

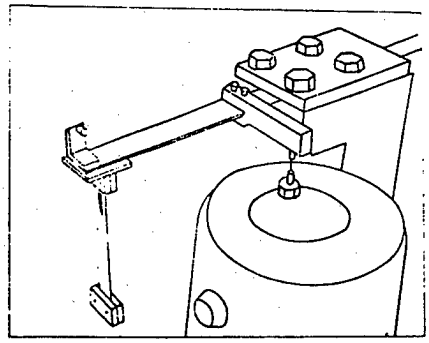
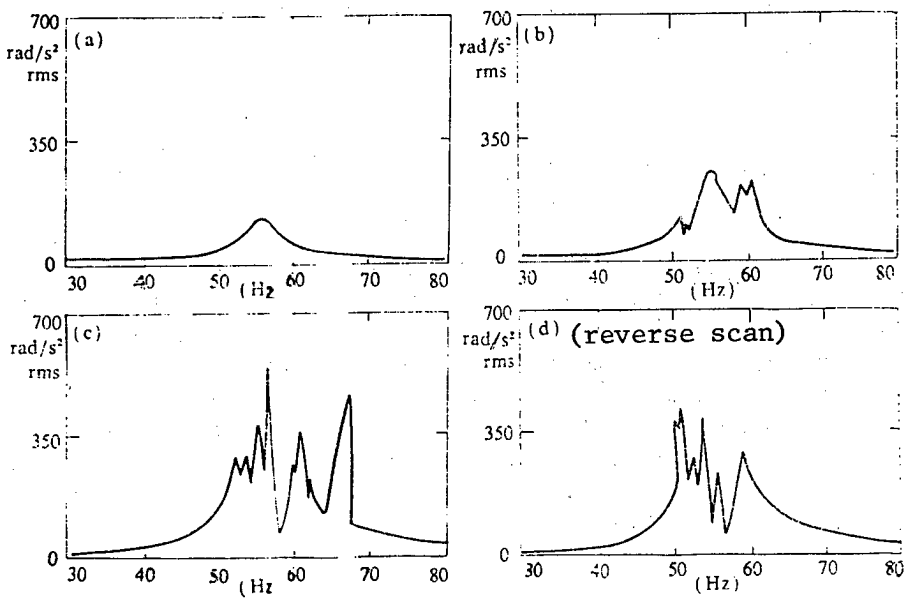


Figure 2

A sine-wave excitation signal is generated using a SC0200 scan-controlled signal generator through a LDS resonator; its spectrum and time history are recorded using a spectrum analyzer and a X-Y recorder. The torsional vibration of the supporting beam is measured by a micro-accelerometer located at the end of the supporting beam. The torsion and bending of the solar panel are measured by the strain gauge located on both sides of the solar panel.

Figure 3 shows the torsional frequency response of the supporting beam under a sine-wave scan. Figure (a) shows the response under low excitation level, which is clearly linear and has a peak response at 56 Hz. Figures (b), (c) show that as the excitation level increases, the non-linear characteristics become increasingly apparent, and multiple peaks appear on the frequency response curve. Figure (d) shows the typical non-linear behavior of the frequency response as the scan direction is reversed.



Excitation level (a) 0.05 Nm (b) 0.1 Nm (c) (d) 0.2 Nm

Figure 3

Figure 4 shows the torsional spectrum of the supporting beam, and the bending spectrum and torsional spectrum of the solar panel under a 56-Hz sine-wave excitation. It can be seen that at low excitation levels (Figure (a)), only torsional vibration is produced; at high excitation levels (Figure (b)), a drive frequency of ω_T will excite not only torsional vibration of supporting beam, but also coupled bending and torsional vibrations of the solar panel. In other words, at high excitation levels, the non-linear effects become very pronounced; the torsional disturbances acting on the supporting beam are transmitted to the solar panel through internal resonance, producing bending and torsional vibrations in the panel.

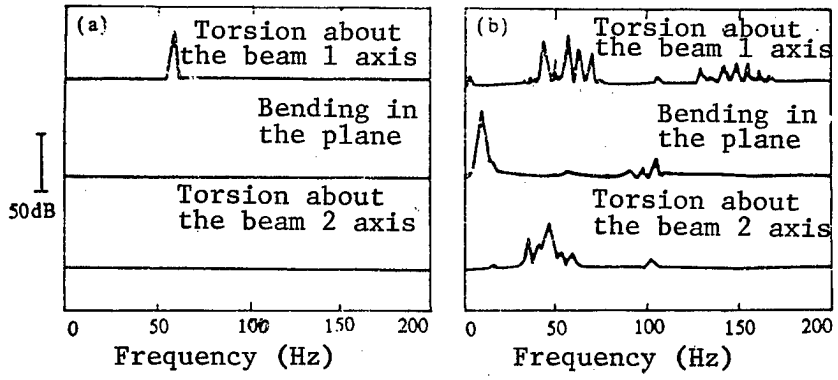
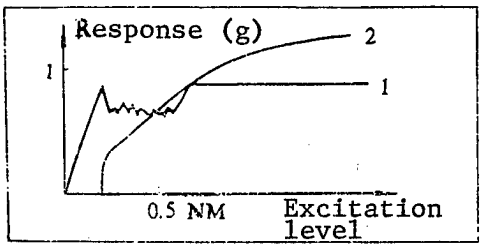


Figure 4

Figure 5 illustrates the effect of the disturbance level: curve 1 is the torsional vibration of the supporting beam, and curve 2 is the vibration of the solar panel; the horizontal axis denotes the excitation level. These curves clearly show that when the excitation level is less than a certain critical value, the system dynamics remain in the linear domain, where torsional vibration is produced only in the supporting beam, the solar panel is free of vibration. When the excitation level is greater than this critical value, the system dynamics become non-linear; the disturbance energy acting on the supporting beam is transmitted to the solar panel through internal resonance, and produces vibration in the panel. As the excitation level further increases, the amplitude of the torsional vibration of the supporting system reaches a saturation level, whereas the amplitude of vibration of the solar panel continues to increase.



Curve 1 denotes the torsional vibration about the beam 1 axis; Curve 2 denotes the bending vibration in the plane

Figure 5

For a large flexible structure with numerous low-frequency natural modes, this internal resonance phenomenon clearly is a potentially dangerous problem.

III. Theoretical Analysis

The solar panel and its supporting system are represented by the simplified model shown in Figure 6. Assume that the disturbance torque is about the axis of the supporting beam and has a sine-wave excitation with frequency Ω. Let the supporting system and the solar panel be represented by beam 1 and beam 2 respectively, both of which are subject to bending and torsion.

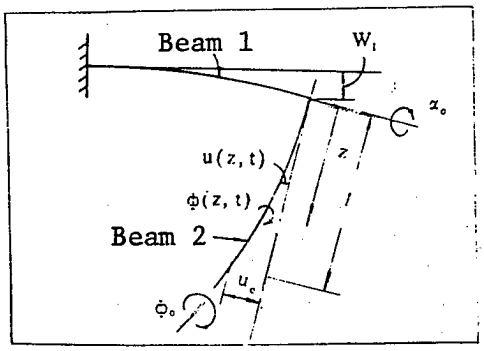


Figure 6

The relative deformation of beam 2, denoted by u(z,t) and $\varphi(z,t)$, can be expressed as follows:

$$u(z,t) = u_0 f_1(z) \tag{6}$$

$$\varphi(z,t) = \varphi_0 g(z) \tag{7}$$

where

$$u_0 = W_1 \psi_{01} \tag{8}$$

ψ_{01} is a component of the eigenvector of the first-order bending mode of the system, and W_1 is the generalized coordinate of the bending mode in the plane. ψ_0 is the coordinate of the torsional mode about the beam 2 axis. Let α_0 be the generalized coordinate of the system torsional mode about the beam 1 axis, then the system dynamic equations expressed in modal coordinates are:

$$\ddot{\alpha}_0 - \omega_T^2 \alpha_0 = \nu_0 \frac{T_0}{J_0} \cos\Omega t - 2\delta_T \omega_T \dot{\alpha}_0 - \frac{M_0}{J_0} l \psi_{01} B_{1\varphi} [W_1 \ddot{\varphi}_0 + 2\dot{W}_1 \dot{\varphi}_0 + \ddot{W}_1 \varphi_0] \tag{9}$$

$$\ddot{\varphi}_0 + \omega_\varphi^2 \varphi_0 = -2\delta_\varphi \omega_\varphi \dot{\varphi}_0 - \frac{M_0}{J_0} l \psi_{01} B_{1\varphi} \ddot{\alpha}_0 W_1 \tag{10}$$

$$\ddot{W}_1 + \omega_1^2 W_1 = -2\delta_1 \omega_1 \dot{W}_1 - \frac{M_0}{M_1} l \psi_{01} B_{1\varphi} \ddot{\alpha}_0 \varphi_0 \tag{11}$$

where $B_{1\varphi}$ is the second-order coupling coefficient:

$$B_{1\varphi} = \int_0^l (l - z) f_1(z) g(z) dz \tag{12}$$

It can be seen from equations (9)-(11) that the coupling term $[W_1\ddot{\varphi}_0 + \dot{W}_1\dot{\varphi}_0 + \ddot{W}_1\varphi_0]$ on the right side of equation (9) absorbs part of the direct excitation of the α_0 response; the coupling terms $\ddot{\alpha}_0 W_1$ and $\ddot{\alpha}_0\varphi_0$ on the right side of equation (10) and equation (11) cause indirect excitation of the torsional deformation about the beam 2 axis φ_0 and of the bending deformation W_1 respectively. These coupling terms are important if the following internal resonance condition exists:

$$\omega_T = \omega_1 + \omega_\varphi \tag{13}$$

Otherwise they can be neglected. This conclusion and the experimental results presented in the previous section can be verified by theoretical analysis of this simplified model using the multiple scale method.

To proceed, we introduce an arbitrary small parameter ε and re-write equations (9)-(11) in the form:

$$\ddot{Y}_1 + \omega_T^2 Y_1 = \varepsilon[\tau\cos\Omega t - 2\xi_T\omega_T\dot{Y}_1 - \mu k_{12}(X_1\ddot{X}_2 + 2\dot{X}_1\dot{X}_2 + \ddot{X}_1 X_2)] \tag{14}$$

$$\ddot{X}_1 + \omega_1^2 X_1 = \varepsilon[-2\xi_1\omega_1\dot{X}_1 - K_{12}\ddot{Y}_1 X_1] \tag{15}$$

$$\ddot{X}_2 + \omega_\varphi^2 X_2 = \varepsilon[-2\xi_\varphi\omega_\varphi\dot{X}_2 - K_{12}\ddot{Y}_1 X_1] \tag{16}$$

where

$$\begin{aligned} &\varepsilon Y_1 = \alpha_0, \qquad \varepsilon l X_1 = W_1, \qquad \varepsilon l X_2 = r\varphi_0 \\ &M_1 r^2 = I_0, \qquad K_{12} = \frac{M_0}{M_1}\psi_{01}B_1 \\ &\mu = \frac{M_1}{J_0}l^2, \qquad \varepsilon^2\tau = \frac{T_0}{J_0}\gamma_0 \qquad \varepsilon\xi = \delta \end{aligned} \tag{17}$$

r is the radius of gyration, M_1 is the generalized mass corresponding to the system bending mode, I_0 is the polar moment of inertia of beam 2, M_0 is the mass at the end of beam 2, T_0 is the amplitude of the applied torque, J_0 is the generalized rotational inertia, γ_0 is a component of the torsional mode at the point of the applied torque, and l is the length of beam 2.

By introducing a slowly time-varying parameter $T_1 = \varepsilon t$, the solution of equations (14)-(16) can be expressed in the form:

$$Y_1(t) = b_1(T_1)\cos[\omega_T t + \beta_1(T_1)] + \varepsilon Y_{11}(t, T_1) \tag{18}$$

$$X_1(t) = a_1(T_1)\cos[\omega_1 t + \alpha_1(T_1)] + \varepsilon X_{11}(t, T_1) \tag{19}$$

$$X_2(t) = a_2(T_1)\cos[\varphi_\varphi t + \alpha_2(T_1)] + \varepsilon X_{21}(t, T_1) \tag{20}$$

Equations (18)-(20) represent the torsion of beam 1, the bending of the system in the plane, and the torsion of beam 2 respectively. It can be seen from these equations that each solution corresponds to a domain of its own natural frequency. From the bounded conditions of the disturbance term $Y_{11}(t, T_1)$, one can derive a first-order differential equation for each of the response amplitudes b_1, a_1 and a_2 with respect to the slowly varying time T_1. The degree of resonance can be defined by two parameters σ and σ_2:

$$\Omega=\omega_T+\varepsilon\sigma \tag{21}$$

$$\omega_T=\omega_1+\varphi_\varphi+\varepsilon\sigma_1 \tag{22}$$

Expanding the differential equations for the modal responses a_1, a_2, and b_1, one obtains:

$$\omega_1 a_1' = -\zeta_1\omega_1^2 a_1+\frac{1}{4}K_{12}\omega_T^2 a_2 b_2\sin\gamma_2 \tag{23}$$

$$\omega_1 a_1 a_1' = -\frac{1}{4}K_{12}\omega_T^2 a_2 b_2\cos\gamma_2 \tag{24}$$

$$\omega_\varphi a_2' = -\zeta_\varphi\omega_\varphi^2 a_2+\frac{1}{4}K_{12}\omega_T^2 a_1 b_2\sin\gamma_2 \tag{25}$$

$$\omega_\varphi a_2 a_2' = -\frac{1}{4}K_{12}\omega_T^2 a_1 b_2\cos\gamma_2 \tag{26}$$

$$\omega_T b_1' = -\zeta_1\omega_T^2 b-\frac{1}{4}\mu K_{12}(\omega_1-\omega_\varphi)^2 a_1 a_2\sin\gamma_2+\frac{\tau}{2}\sin\gamma_1 \tag{27}$$

$$\omega_T b_1\beta_1' = -\frac{1}{4}\mu K_{12}(\omega_1+\omega_\varphi)^2 a_1 a_2\cos\gamma_2+\frac{\tau}{2}\cos\gamma_2 \tag{28}$$

where

$$\gamma_1=\sigma T_1-\beta_1 \tag{29}$$

$$\gamma_2=\sigma_1 T_1-\alpha_1-\alpha_2+\beta_1 \tag{30}$$

In the above equation, the superscript "'" denotes derivative with respect to T_1. The system equations (23)-(28) have two steady-state solutions. One of the solutions is:

$$a_1=0, \qquad a_2=0 \tag{31}$$

$$b_1=\frac{\tau}{2\omega_T\sqrt{\sigma^2+\zeta_T^2\omega_T^2}} \tag{32}$$

which indicates that the indirectly excited motions are stationary, and the response of the directly excited motions is linear.

The second steady-state solution is:

$$b_1=\frac{4\omega_1\omega_\varphi}{\omega_T^2}\sqrt{\frac{\zeta_1\varepsilon_\varphi}{K_{12}}\left[1+\frac{(\sigma_1+\sigma)^2}{(\zeta_1\omega_1+\zeta_\varphi\omega_\varphi)_2}\right]} \tag{33}$$

$$a_2^2=\frac{\zeta_1\omega_2^1}{\zeta_\varphi\omega_\varphi^2}a_1^2 \tag{34}$$

$$a_1^2=\frac{2}{\mu K_{12}(\omega_1+\omega_\varphi)^2}\sqrt{\frac{\zeta_\varphi\omega_\varphi^2}{\zeta_1\omega_1^2}}\left[-\Gamma_1\pm\sqrt{\tau^2-\Gamma_2^2}\right] \tag{35}$$

where

$$\Gamma_1=\Gamma_3[\xi_T\omega_T(\xi_1\omega_1+\xi_\varphi\omega_\varphi)-\sigma(\sigma+\sigma_1)] \tag{36}$$

$$\Gamma_2=\Gamma_3[(\xi_1\omega_1+\xi_\varphi\omega_\varphi)\sigma+\xi_T\omega_T(\sigma+\sigma_1)] \tag{37}$$

$$\Gamma_3=\frac{8\omega_1\omega_\varphi\sqrt{\xi_1\xi_\varphi}}{\omega_T K_{12}(\xi_1\omega_1+\xi_\varphi\omega_\varphi)} \tag{38}$$

This solution represents the non-linear response of the system. The three types of motions are excited simultaneously; also, when $\sigma = \sigma_1$, the torsional amplitude b_1 has a minimum value which is independent of the level of excitation. This is the so-called saturation phenomenon which is clearly evident in the experimental results of Figure 5. The amplitudes of a_1 and a_2 increase with increasing excitation level. This phenomenon can also be observed in Figure 5.

Equation (35) also shows that whether or not a stable solution exists depends on the system parameters.

The τ_0 value of the critical point can be obtained from equations (32) and (33):

$$\tau_0=\frac{8\omega_1\omega_\varphi}{\omega_T}\sqrt{(\sigma^2+\xi_T^2\omega_T^2)\frac{\xi_1\xi_\varphi}{K_{12}}\left[1+\frac{(\sigma_1+\sigma)^2}{(\xi_1\omega_1+\xi_\varphi\omega_\varphi)^2}\right]} \tag{39}$$

Clearly, τ_0 is a minimum when $\sigma_1 + \sigma = 0$ and $\sigma = 0$:

$$\tau_{0min}=8\omega_1\omega_\varphi\xi_T\sqrt{\frac{\xi_1\xi_\varphi}{K_{12}}} \tag{40}$$

However, $\sigma_1 + \sigma = 0$ and $\sigma = 0$ imply that the internal resonance relationship

$$\omega_T=\omega_1+\omega_\varphi$$

and the resonance relationship $\Omega = \omega_T$ exist.

Figure 7 shows the theoretical curves of the amplitudes of the modal responses as a function of the excitation parameter τ. When τ exceeds the critical value, the amplitudes of the coupled modal responses vary differently: one approaches saturation, and the other two undergo abrupt changes, then continue to rise sharply.

IV. Conclusion

In this article, the effect of inertial non-linear coupling on the dynamic characteristics of a flexible system is studied using a simplified model of solar panel and its supporting system. Experimental results and theoretical analysis show that when the external excitation is in resonance with one of the non-linear coupled modes, and internal resonance also exists between the coupled modes, the non-linear coupling can produce abrupt changes in the dynamic characteristics of the system. The effect of internal resonance

produces indirect excitation of the coupled modes which are independent of the external excitation. For a large, flexible structure such as the space station, there are numerous sources of excitation and a high concentration of natural modes which may cause internal resonance. This type of internal resonance is a potential threat to vibration suppression, configuration maintenance, pointing accuracy and control accuracy for a large flexible structure. Therefore, detailed investigation of this effect of non-linear coupled vibration is essential to the design of space station.

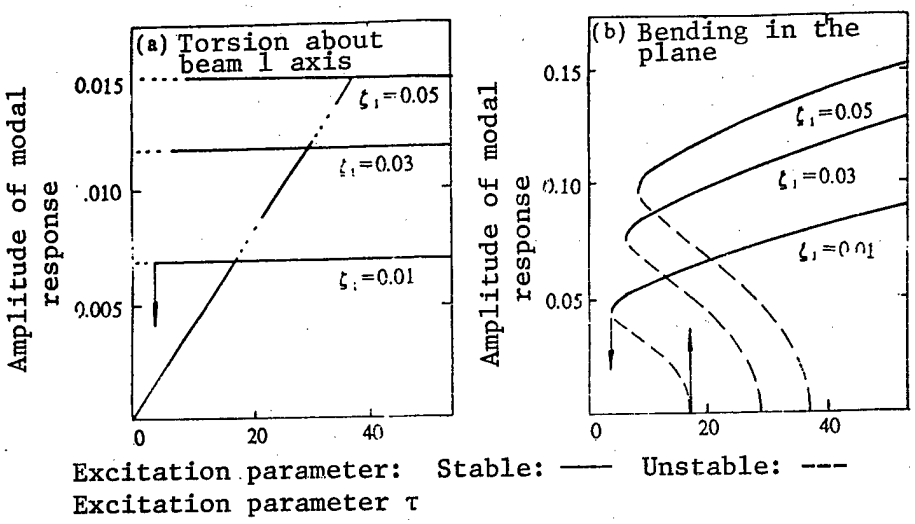


Figure 7

References

1. Weeks, G. E., "Dynamic Analysis of a Deployable Space Structure," J. SPACECRAFT, Vol 23, No 1, 1986, pp 102-107.

2. Ayers, J. K., et al., "Structural Dynamics and Attitude Control Study of Early Manned Capability Space Station Configurations," NASA-TM-89078, Jan. 1987.

3. Pieere, C., Tang, D. M., and Dowell, E. H., "Localized Vibration of Disordered Multispan Beams: Theory and Experiments," AIAA J., Vol 25, No 9, 1987.

4. Haddow, A. G., Barr, A. D. S., and Mook, D. T., "Theoretical and Experimental Study of Modal Interaction in a Two Degree of Freedom Structure," J. SOUND AND VIBRATION, Vol 97, No 3, 1984, pp 451-473.

5. Bux, S. L., and Roberts, J. W., "Nonlinear Vibratory Interactions in Systems of Coupled Beams," J. SOUND AND VIBRATION, Vol 104, No 3, 1986, pp 497-520.

Digital Geometric Rectification of Satellite Panoramic Images

90FE0179C Beijing YUHANG XUEBAO [JOURNAL OF CHINESE SOCIETY OF ASTRONAUTICS] in Chinese No 2, 30 Apr 90 pp 65-70

[Article by He Zhiqin [0149 1807 0530] and Meng Bo [1322 3134], Beijing Institute of Satellite Information Engineering]

[Text] Abstract

A method of coarse geometric rectification of full-size satellite panoramic images is proposed. The method is based on certain intrinsic geometric relationships between the satellite image and the ground, and takes into account the physical process of satellite imaging. The rectification of satellite images is accomplished by applying parameter optimization technique which does not depend solely on the satellite parameters, but also relies on a small number of carefully selected and accurately measured ground control points as well as a geometric model. Simulation results show that the proposed method is very effective for the coarse processing of satellite panoramic images.

I. Introduction

Ground observation and imaging by a retrievable remote-sensing satellite is accomplished by panoramic scanning. Because of the low satellite orbit and the large scan angle, the images suffer severe geometric distortions, particularly at the edges where errors can be as large as several kilometers. In order to promote and apply the satellite panoramic imaging technique, it is essential to find a practical method of geometric rectification.

The geometric methods which are currently used for satellite image processing include the model method and the polynomial method.[1] The model method involves establishing a mathematical model to describe the satellite imaging process and requires accurate knowledge of the satellite orbit, satellite attitude and camera parameters; the accuracy of these parameters will have a significant effect on the accuracy of geometric processing. The polynomial method does not require knowledge of satellite parameters, but it does require a large number of accurately measured and uniformly distributed ground control points; this method produces rather poor results at the edges of the images.

The method proposed in this paper is based on the geometric relationships of satellite imaging and tries to avoid the deficiencies of the above two methods. Specifically, it takes into account the physical process of satellite imaging, but does not rely solely on the satellite parameters; rather, it uses the intrinsic geometric relationship between the satellite image and a small number of accurately measured ground control points and performs geometric rectification of the satellite image by applying optimal parameter estimation technique.

II. Mathematical Model[2]

To simplify the analysis, the earth is assumed to be a sphere of radius R, and the altitude of the satellite is denoted by H. Since the effective area of each film is approximately 200x1400 mm^2, its width corresponds to a size of only several tens of kilometers on the ground. Hence, the projected scan lines can be approximated by straight lines, and the ground segment corresponding to each film can be approximated by a curved surface taken from a cylinder of radius R, as indicated by the surface EFPB in Figure 1. Clearly, for a scan angle β, one can apply the law of sine to the triangle ΔOSP_o to obtain the following equation:

$$2\alpha = \sin^{-1}\left[\sin\beta \cdot \left(1 + \frac{H}{R}\right)\right] - \beta \tag{1}$$

$$l = \frac{R \cdot \sin 2\alpha}{\sin\beta} \tag{2}$$

Recognizing that ΔOS_1P_o is an isosceles triangle, one can determine the corresponding angles, and apply the law of sine to the triangle ΔSS_1P_o to obtain:

$$l = H\cos\alpha/\cos(\alpha+\beta) \tag{3}$$

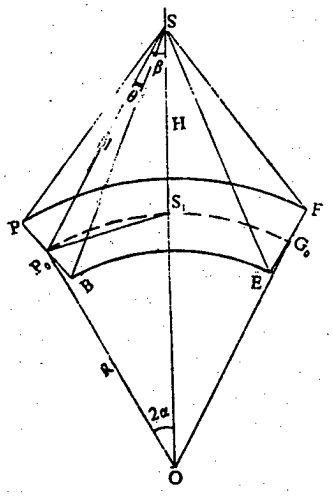


Figure 1. Geometric Relationship of Satellite Imaging

1. Geometric Relationship Along the Scan Direction

As shown in Figure 2, in the ground coordinate system O-XY, L_o is the ground projection of the panoramic scan slit when the scan angle $\beta = 0$; it is represented by the equation Y = KX+E. At time t_i, the ground projection of the slit is L_i, and an image of the straight line L_i' which passes through the point (X_i, Y_i) is taken. The distance between L_i and L_o is vt_i, where v is the translational velocity of L_o due to earth rotation, which is to be determined. It is a function of the inclination of the satellite orbit, the satellite velocity, the earth rotational speed and the latitude and longitude of the satellite subpoint. D_i is the distance between L_i' and L_i.

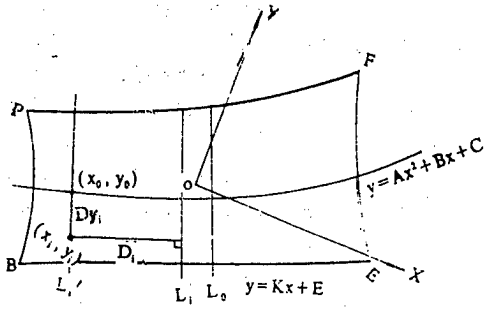


Figure 2. Geometric Relationship in the Ground Coordinate System

In the original film coordinate system O-xy shown in Figure 3, it is assumed that the satellite attitude remains unchanged during the imaging process; in other words, during this period the equation of the theoretical center edge-contact line ($\beta = 0$) remains constant ℓ_o: y = kx+e. By "theoretical" it is meant that ℓ_o passes through the primary image point, and the image line at any time is parallel to ℓ_o. d_i is the distance between ℓ_i and ℓ_o, and corresponds to the ground distance D_i.

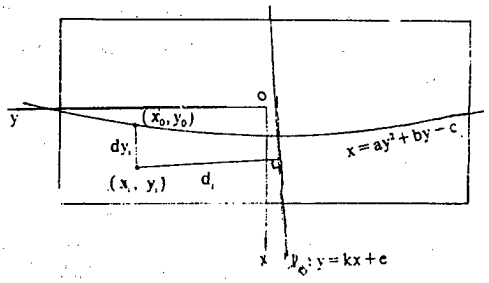


Figure 3. Geometric Relationship in the Film Coordinate System

From the above analysis, one can readily derive:

$$D_i = \frac{|KX_i + E - Y_i|}{\sqrt{1+K^2}} - vt_i \qquad (4)$$

$$d_i = \frac{|kx_i + e - y_i|}{\sqrt{1+k^2}} \tag{5}$$

where K, E, k, e, v are to be determined, $t_i = d_i/(v_o \sqrt{1+k^2})$, v_o is the film speed of the camera.

Substituting $\beta_i = d_i/f$ (f is the focal length of the camera) into equation (1), one can derive the following relationship between D_i and d_i:

$$D_i = R \cdot \left\{ \sin^{-1}\left[\sin \frac{d_i}{f} \left(1 + \frac{H}{R}\right)\right] - \frac{d_i}{f} \right\} \tag{6}$$

It is clear that even if errors are present in the satellite attitude and orbit parameters, as long as they remain unchanged during the imaging process, this geometric relationship also remains unchanged. In other words, equation (6) shows that the accuracy of geometric processing is independent of errors in the satellite attitude parameters, which is a major advantage of this method. Since each film only takes 3 seconds to complete during the imaging process, it is reasonable to assume that the effective parameters remain unchanged during this period.

2. Geometric Relationship Perpendicular to the Scan Direction

To account for small variations in satellite motion, earth rotation and the direction of film advancement, we use a second-order curve to represent the scan line of the slit-center and its ground track; these equations are of the form: $x = ay^2+by+c$ (film) and $Y = AX^2+BX+C$ (ground), as indicated in Figure 2 and Figure 3. In the ground coordinate system, let us draw a straight line from the ground control point (X_i, Y_i): $(Y-Y_i) = K(X-X_i)$, which intersects the center scan line $Y = AX^2+BX+C$ at (X_o, Y_o), then we have:

$$DY_i = \sqrt{(X_i - X_0)^2 + (Y_i - Y_i)^2} \tag{7}$$

Similarly, in the original film coordinate system, we draw a straight line from (x_i, y_i): $y-y_i = k(x-x_i)$, which intersects the center scan line $x = ay^2+by+c$ at (x_o, y_o), then we have:

$$dy_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \tag{8}$$

Now, we shall derive a geometric relationship between DY_i and dy_i. Figure 4 shows the geometry between DY_i and dy_i where a pitch angle φ is present. θ is the angle of the scan beam. It should be pointed out that since φ and dy can be either positive or negative, they can be combined to form several different geometries; however, the derivation in this paper is based only on the geometry shown in Figure 4; it can be readily extended to other geometries. From Figure 4, it is clear that

$$DY = l \cdot [\mathrm{tg}(\varphi+\theta) - \mathrm{tg}\varphi]$$

$$= R \cdot \frac{\sin 2\alpha}{\sin\beta}\left[\frac{\mathrm{tg}\varphi + \mathrm{tg}\theta}{1-\mathrm{tg}\varphi\mathrm{tg}\theta} - \mathrm{tg}\varphi\right]$$

$$= H \cdot \frac{\cos\alpha}{\cos\left(\alpha + \frac{d}{f}\right)}\left[\frac{f\mathrm{tg}\varphi + dy}{f - \mathrm{tg}\varphi dy} - \mathrm{tg}\varphi\right] \tag{9}$$

Equation (9) is affected only by the pitch angle φ; it is independent of other attitude parameters.

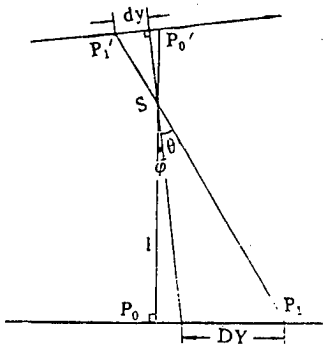


Figure 4. Geometric Relationship Between DY and dy

3. Objective Function

In this paper, the physical parameters are determined by solving the parameter optimization problem using the Powell method. The objective functions to be optimized are:

$$f_1(k,e,K,E,v) = \sum_{i=1}^{n}\left| D_i - R\left\{\sin^{-1}\left[\sin\frac{d_i}{f}\left(1+\frac{H}{R}\right)\right] - \frac{d_i}{f}\right\}\right|^2 \tag{10}$$

$$f_2(a,b,c,A,B,C,\varphi) = \sum_{i=1}^{n}\left| DY_i - H \cdot \frac{\cos\alpha}{\cos\left(\alpha + \frac{d_i}{f}\right)} \cdot \left[\frac{f\mathrm{tg}\varphi - dy_i}{f - \mathrm{tg}\varphi \cdot dy_i} - \mathrm{tg}\varphi\right]\right|^2 \tag{11}$$

4. Geometric Rectification

Having determined the parameters {K, E, k, e, v, A, B, C, a, b, c, φ} using the Powell method, the process can be reversed to establish the relationships between corresponding points on the original film and the new film after rectification. A difficult problem with this process is establishing the quadrant relationship between corresponding points in the new picture and the original picture.

Consider a point p(x,y) on the original film and its corresponding point on the new film P(X,Y). The computational procedure and the key formulas are given below:

1) Determine x_o, y_o

$$\begin{cases} (y-y_0)/(x-x_0) = k \\ x_0 = ay_0^2 - by_0 + c \end{cases} \Rightarrow \quad (x_0, y_0)$$

2) Calculate d, dy

$$d = \frac{|kx + c - y|}{\sqrt{1+k^2}}$$

$$dy = \sqrt{(x-x_0)^2 + (y-y_0)^2}$$

3) Calculate D, DY from equations (6) and (9)

4) Solve the simultaneous equations

$$Y - Y_0 = K(X - X_0)$$

$$Y_0 = AX_0^2 + BX_0 + C$$

$$DY = \sqrt{(X-X_0)^2 + (Y-Y_0)^2} \Rightarrow (X_0, Y_0, X, Y)$$

$$D = \frac{|KX + E - Y|}{\sqrt{1+K^2}} \Rightarrow D$$

This completes the transformation between the points p(x,y) and P(X,Y). It should be noted that the positional relationship between corresponding points on the original film and the new film is controlled by the four "quadrants" formed by the center edge contact and the center scan line; i.e., the "quadrant" relationship between corresponding points remains unchanged.

To save computer time in the actual rectification process, we only use the above model to calculate the coarse lattice points; then we apply the finite element method (first-order approximation inside a triangle) to calculate points within a lattice.

5. Other Considerations

Since the panoramic scan camera uses mechanical drive mechanisms for film advancement, jitters will inevitably appear, causing errors in the images. To solve this problem, we introduce in the objective function an additional parameter δ, which is a function of d_i. The form of this function is shown in Figure 5, where Δ is the residue at d_i, and L is the range of influence of jitter, which is the parameter to be optimized. For a control point whose residue is larger than a certain value, optimization is performed on an objective function with the added parameter $\delta(d_i)$, which determines a model function that reflects the jitter of the entire frame.

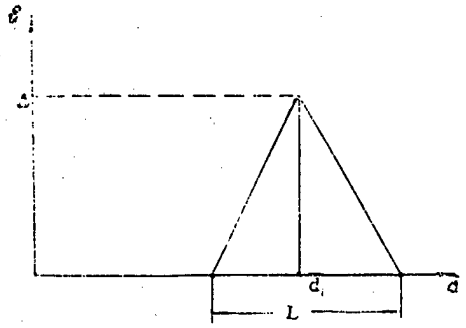


Figure 5. Simulated Jitter Function of Film Advancement

In addition, because of variations in satellite attitude parameters during the scan process and because of dimensional considerations, the film digitization process must be carried out frame by frame (due to the limitation of C-4500). Therefore, the center scan line or its ground track cannot be modeled by a single second-order curve; it must be modeled within each frame to achieve the required processing accuracy.

Another factor that should be considered is the effect of difference in surface altitude. This effect is particularly pronounced for domestic satellites as they are in low-earth orbits and have large scan angles. In this article, the effect of difference in surface altitude is taken into account by introducing a correction factor to the orbit altitude H in the geometric relationship.

III. Test Results

In order to demonstrate the feasibility of the proposed method, we have conducted test on a full-size satellite panoramic image. The dimensions of the image are approximately 200x1400 mm^2. The image is divided into six frames, each of which is digitized using the C-4500; the individual frames are joined together in the computer using identical points to form a complete digitized image. The control points are selected by performing a correlation analysis between the image and the terrain map. However, because a large portion of the area covered by this image is desert, selection of control points is very difficult, and they are not uniformly distributed. As a result, the RMS processing errors using the polynomial method or the traditional model method are of the order of 800-1000 m, which clearly do not meet the accuracy requirements. In this paper, a total of 50 ground target points are selected, 30 of which are control points, and the remaining 20 are check points. After processing the data using the method described in this paper, the RMS error is σ = 206 m, which fully meets the accuracy requirement of coarse geometric processing for satellite images.

This method can also be used for precision rectification of sub-images of a region with a large number of uniformly distributed control points. For example, a test was performed on a 110x110 mm^2 sub-region using 30 control points; the RMS error of the rectified image is σ = 45 m.

Higher processing accuracy can be achieved for a sub-image because: 1) there are larger number of control points; 2) the various parameters remain essentially constant within a shorter time interval; and 3) it is easier to model the center scan line and the center edge contact. However, precision processing the full-scale image all at once would be difficult.

IV. Conclusion

The above test results show that the proposed method is effective for coarse geometric processing of satellite panoramic images. This method is particularly useful in cases where the model method is not applicable because accurate satellite parameters or orbit parameters are not available, or where the polynomial method is not applicable because uniformly distributed control points are difficult to obtain.

Another advantage of the proposed method is that it can perform simultaneous rectification on a full-scale 200x1400 mm^2 image without having to process the individual sub-images. This not only improves processing efficiency but also eliminates possible errors introduced by the joining process of rectified sub-images.

It is generally not difficult to select more than 30 ground target points on a full-scale image. A sufficiently large number of uniformly distributed ground target points should be chosen at the two sides of a satellite panoramic image where large distortions occur. Simulation results show that this method can meet the accuracy requirements for geometric processing satellite images.

The authors wish to express thanks to Comrades Ju Tao, Guo Zhaozeng, and Li Lingzhi for their support in developing this method.

References

1. Guo Zhaozeng, He Zhiqin, Li Lingzhi, Ju Tao, Meng Bo, "Mathematical Models for Domestic Satellite Imaging Process and Recovery of Geometric Distortion," ENVIRONMENTAL REMOTE SENSING, Vol 4, No 3, 1989.

2. He Zhiqin, Meng Bo, "Error Analysis of Geometric Characteristics for Domestic Satellite Imaging," CHINESE SPACE SCIENCE AND TECHNOLOGY, No 5, 1988.

Extracting the Poles of Underground Targets

90FE0179D Beijing YUHANG XUEBAO [JOURNAL OF CHINESE SOCIETY OF ASTRONAUTICS] in Chinese No 2, 30 Apr 90 pp 71-75, 70

[Article by Zheng Jianqing [6774 1696 1987] and Dong Naihan [5516 0035 3211], Beijing Institute of Remote Sensing Equipment]

[Text] Abstract

In this article, an improved method of extracting the poles of underground targets is presented. It is pointed out that in a noise environment, the eigenvalue method is more accurate in determining the number of poles; it is also shown that the eigenvalue method always yields smaller error of approximation than classical methods. Experiments of extracting the poles of buried targets have been performed, from which several significant conclusions can be drawn.

I. Introduction

An effective approach for detecting underground targets is the transient electro-magnetic field method. This method has applications in many different areas such as hydrology, geology, resource exploration, detection of buried articles, mine detection, and anti-stealth technology. However, to use this method, it is necessary to solve the problem of target identification using the transient target response. Since the poles are invariant properties of a target, they can be used as effective signatures for target identification.[1,2,3] In 1975, Rai Mittra et al. proposed the Prony method of extracting poles from transient pulse response. This method attracted a great deal of interest, but further study showed that in the presence of noise and measurement errors, the accuracy of this method is so poor that it may be of little practical use. Therefore, efforts are being made to explore new methods of pole extraction.

II. Comparison of Square Approximation Error

In 1978, Van Blaricum et al. proposed a new method of pole extraction[4] where the problem is formulated as one of finding the smallest eigenvalue (λ_{min}) and the corresponding eigenvector α of the matrix H^TH:

$$\alpha^T(H)^T(H)\alpha=\lambda \tag{1}$$

where

$$(H)=\begin{pmatrix} I_0 & I_1 & \cdots & I_N \\ I_1 & I_2 & \cdots & I_{N+1} \\ \cdots & & \cdots & \cdots \\ I_m & I_{m+1} & \cdots & I_{m+N} \end{pmatrix}$$

I_i is the sample value at time $i\Delta t$, and $\alpha^T = (\alpha_N, \cdots, \alpha_0)$; the vector α satisfies the condition:

$$\sum_{p=0}^{N} \alpha_p Z^p = 0 \tag{2}$$

where N is the number of poles. Once α_i is determined, the poles in the Z-plane can be obtained from equation (2).

Let $\varepsilon = (\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_m)^T$ be the error vector, which can be expressed by the following relation:

$$(H)\alpha = \varepsilon \tag{3}$$

If $W(\varepsilon)$ is the square approximation error, then we have:

$$\alpha^T (H)^T (H)\alpha = \sum_{j=0}^{m} \varepsilon_j^2 = W(\varepsilon) \tag{4}$$

Different α vectors may yield different values of square approximation error; the α vector determined by the eigenvalue method is the eigenvector that corresponds to the smallest eigenvalue λ_{min} of the matrix H^TH, i.e.,

$$\alpha^T (H)^T (H)\alpha = \lambda_{min} \tag{5}$$

Therefore, the problem that needs to be investigated is the relative magnitude between λ_{min} and $W(\varepsilon)$.

By arranging the eigenvalues of H^TH in ascending order:

$$0 \leqslant \lambda_{min} (= \lambda_0) \leqslant \lambda_1 \leqslant \lambda_2 \cdots \leqslant \lambda_N \tag{6}$$

then it is always possible to find a series of (N+1) orthogonal unit vectors E_i (i = 0, 1, $\cdots$, N) such that (u) = (E_0, E_1, $\cdots$, E_N) is an orthogonal matrix, and satisfies the condition:

$$(u)^T (H)^T (H)(u) = \begin{pmatrix} \lambda_{min} & & 0 \\ & \lambda_1 & \\ 0 & & \ddots \\ & & & \lambda_N \end{pmatrix} \tag{7}$$

Consider a normalized vector $a = (a_0, a_1, \cdots, a_N)^T$; from matrix theory we can write:

$$a = \sum_{i=0}^{N} a_i E_i \text{ or } a = (u)a \tag{8}$$

Therefore,

$$\begin{aligned} W(\varepsilon) &= a^T (H)^T (H) a \\ &= a^T (u) a^T (H)^T (H) (u) a \\ &= a^T \begin{pmatrix} \lambda_{\min} & & 0 \\ & \lambda_i & \\ 0 & & \lambda_N \end{pmatrix} a \\ &= \sum_{j=0}^{N} \lambda_j a_j^2 \end{aligned} \tag{9}$$

It can be seen from equation (9) that when all the eigenvalues are equal to λ_{min}, we have the following inequality:

$$W(\varepsilon) \geq \sum_{i=0}^{N} \lambda_{\min} a_i^2 = \lambda_{\min} \sum_{i=0}^{N} a_i^2 \tag{10}$$

Since a is a normalized vector, it follows that

$$W(\varepsilon) \geq \lambda_{\min} \tag{11}$$

Thus, we have proved that the square approximation error of the eigenvalue method is always smaller than that of classical methods.

III. Determining the Number of Poles

The number of poles can be easily determined using the orthogonality property of the eigenvectors.

1. In the noise-free case, let

$$(\phi) = (H)^T (H)$$

then the following procedure can be used to determine the number of poles:

a. Construct a γ x γ matrix (ϕ) from the sample values.

b. Find the eigenvalues of the matrix (ϕ) and determine the number of zero eigenvalues; if there are L zero eigenvalues, then the number of poles is $V = \gamma - L$; if there is no zero eigenvalue, then the dimension of (ϕ) must be increased to form a new matrix (ϕ).

c. When there is only one zero eigenvalue, i.e., when L = 1, then the corresponding eigenvector is the Prony vector α, and the poles are determined along with V.

2. In the case where noise is present, this method can still be applied. Let the noise sequence be zero mean with variance σ^2; since the noise

sequence is uncorrelated in time and also uncorrelated with the signal sequence, the orthogonality property of the eigenvectors is preserved. Analysis has shown that the noise-free matrix (ϕ) and the noise-present matrix (ϕ') are related by the following expression:

$$(\phi') = (\phi) + \gamma\sigma^2 I \tag{12}$$

where I is the unit matrix. Therefore, the noise-free procedure can still be used, except that in this case, the eigenvalue sought is $\gamma\sigma^2$, not zero. Once an eigenvalue of $\gamma\sigma^2$ is found, the corresponding eigenvector is the α vector.

The above analysis shows that this method is very effective in treating the noise problem; once the two basic statistical parameters (mean and variance) of the noise sequence are known, the number of poles N and the α vector can be readily determined.

IV. Experimental Procedure and Data Processing

1. Experimental Set-Up and Procedure

A block diagram of the experimental set-up is shown in Figure 1. The pulse width is τ = 1.2 ns, the amplitude is 40 v, and the repetition frequency is 10 KHz; the number of sample points is 256, the time duration is 20 ns, and the sample interval is 0.078 ns. The environment is a 100 cm x 100 cm x 60 cm sand box. The transient pulses are transmitted by a narrow-pulse generator; the transmitting and receiving antennas are quasi-TEM horn antennas. The two receiving antennas have identical characteristics; the received signal is amplified and stored in a micro-computer. The synchronization of the transmitted pulse and the sampling circuit is also controlled by the micro-computer.

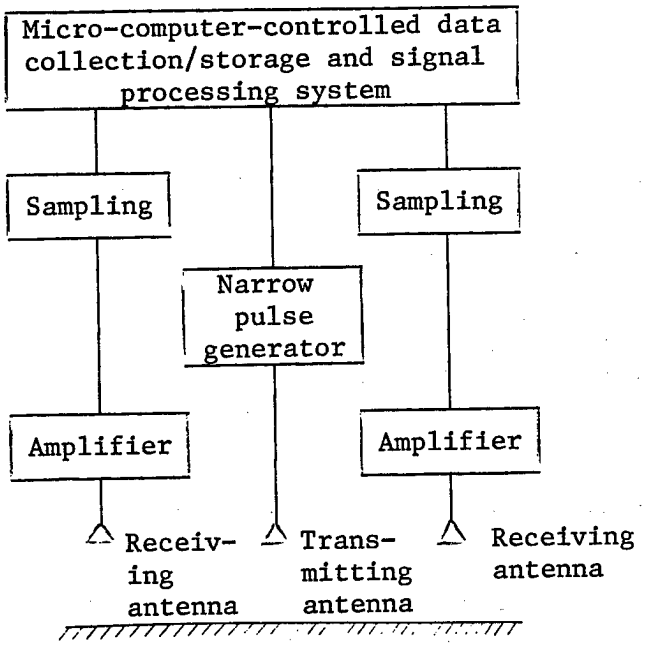


Figure 1. Detection of Underground Targets Using Transient Pulses

2. Processing of the Initial Waveform

a. Elimination of the effect of ground reflection. In this experiment, a dual receiving antenna configuration is used so that the effect of ground reflection can be eliminated by differencing the two received signals.

b. Noise processing. Noise reduction is accomplished by averaging several waveforms measured at the same antenna position. Any d.c. component is eliminated by adjusting the baseline.

c. Gate switching. Gate switching technique is used to eliminate the direct coupling effects of the antenna and premature target response.

d. Filtering of high-frequency components. Because of inconsistencies in sampling times between the two channels of the received signal, high-frequency "burrs" often appear. To solve this problem, a low-pass filter is used to process the waveform.

The above pre-processing of the data significantly improves the signal to noise ratio, and consequently ensures the success of pole extraction.

V. Conclusion

Experiments of extracting the poles of various targets have been conducted; four of these targets are shown in Figure 2. The targets are buried in homogeneous dry sand at a depth of 3-10 cm; the receiving and transmitting antennas are separated by a distance of 40 cm, and the sampling interval is 0.078 ns. The measured and recovered waveforms are shown in Figure 3, where the pole locations are also indicated.

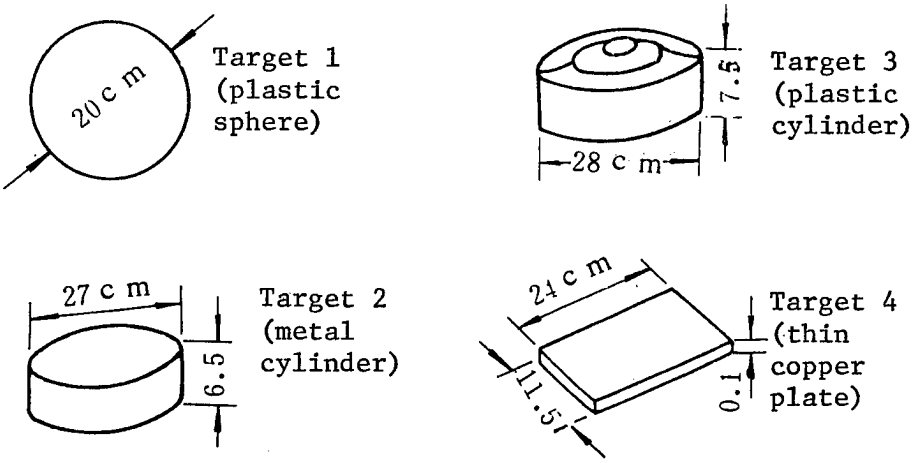


Figure 2. Physical Properties of Four Different Targets

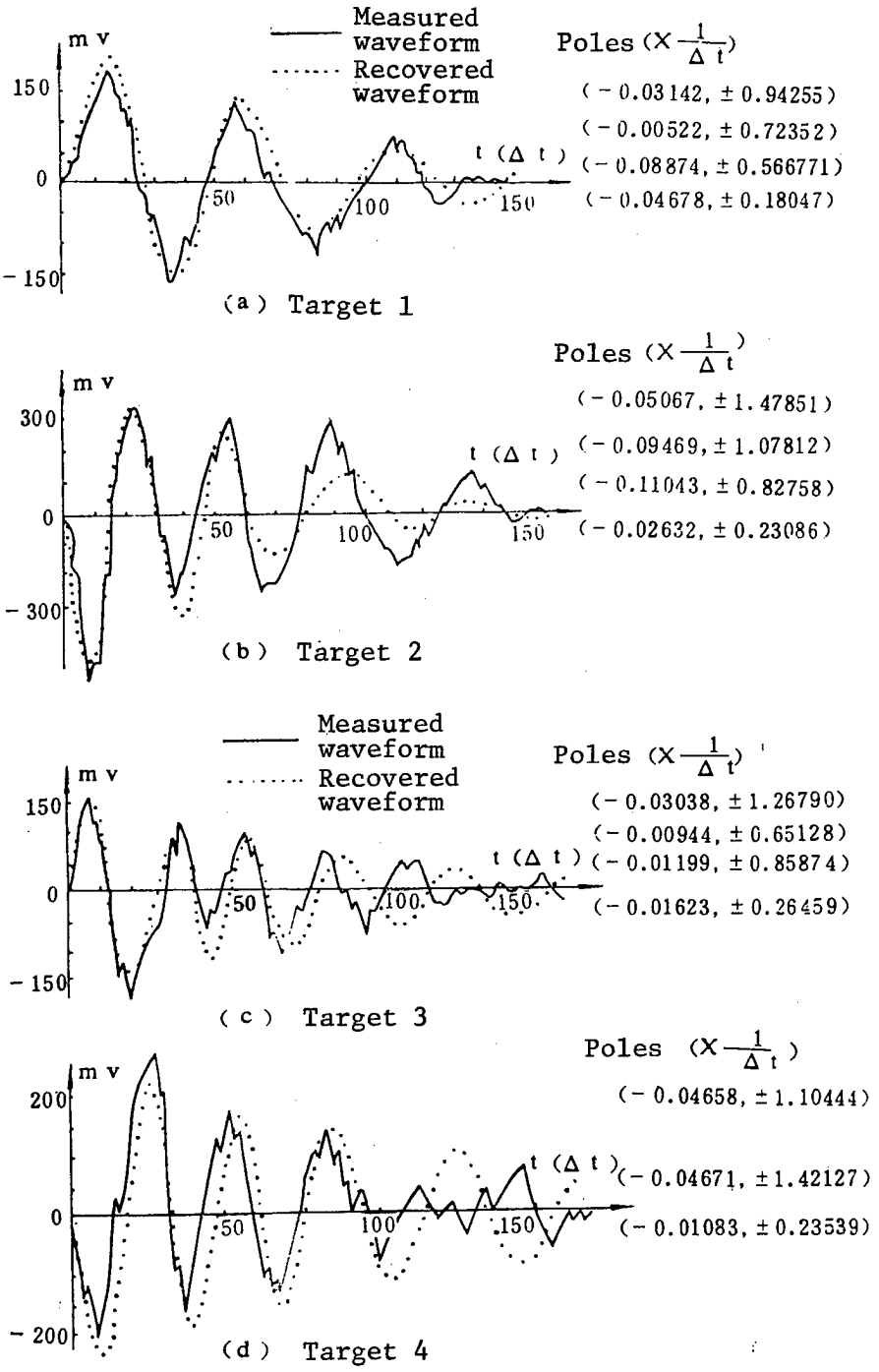


Figure 3. Poles of Buried Targets, Measured Waveforms and Recovered Waveforms

On the basis of analysis and experimental results, the following conclusions can be drawn:

1. The eigenvalue method always yields smaller error of approximation than classical methods.

2. In the presence of noise, the eigenvalue method can still be used to determine the number of poles and to extract the poles as long as two basic parameters of the noise statistics are known.

3. Changes in target depths (within 10 cm) has little effect on the poles; however, the measured results with respect to target depth are not presented here due to space limitation.

4. The initial waveform must be pre-processed to facilitate pole extraction.

5. The initial phases of conducting and non-conducting targets show clearly opposite sense (see Figure 3-a and Figure 3-b).

6. It can be seen from Figure 3 that the larger the scattering object, the smaller the real part of the pole and the larger the imaginary part; conversely, the smaller the scattering object, the larger the real part and the smaller the imaginary part.

The authors wish to thank Mr Chen Jinxiong and Mr Huang Peikang of the No 2 Institute of the Ministry of Aeronautics and Astronautics Industry for their guidance and support; they also wish to thank Comrades Cao Changjun, Zi Haiyen, Liang Tao, Li Xun, and Fang Yi of this institute for their assistance in conducting the experiment.

References

1. Baum, C. E., "The Singularity Expansion Method in Transient Electromagnetic Fields," L. B. Felsen, ed., Heidelberg Springerverleg, 1976, Chapter 3.

2. Chan Luen, C., "A Characterization of Subsurface Radar Targets," PROCEEDINGS OF THE IEEE, Vol 67, No 7, 1979.

3. Moffatt, D. L., "Characterization of Subsurface Electromagnetic Soundings," PB81-163495.

4. Blaricum, M. V., "Problem and Solution Associated With Prony's Method for Processing Transient Data," IEEE TRAN. ON AP, Vol AP-26, No 1, 1978.

5. Xu Hai, "Relationships Between Pole Distribution Based on the Singular Expansion Method and the Shape of the Scattering Object," XIAN JIAOTONG UNIVERSITY BULLETIN, Vol 21, No 3, 1987.

Pure Digital Control for the Outer Axis of a Three－Axis Flight Simulation Table

40100067B Harbin HARBIN GONGYE DAXUE XUEBAO[JOURNAL OF HARBIN INSTITUTE OF TECHNOLOGY] in English No 3, June 90, pp 121-127

Wang Mao *Zhao Keding* *Liu Qinghe*

(Department of Mechanical Engineering)

Abstract In this paper, a pure digital control method for the outer axis of a three－axis flight simulation table is worked out and a state observer is used to obtain the acceleration signal .The method is simple and easy to be emplayed, and especially suitable for the electrohydraulic servo－system with fast response. The simulation results show that the system given in this paper has optimal three－order system performance.

Key Words: Electrohydraulic position servo system; state feedback; digital control.

0 Introduction

A flight simulation table is a kind of experimental equipment used for simulating the flying attitude of a tactical missile. Therefore, it is required to have the performance of high accuracy, fast response, less following time and little overshoot. Now a days, the optimal linearization tested by an analog potentiometer can only reach ±0.02% (and the best position accuracy is ±4′32″). Besides, the deficiencies of parameters shifting and nonlinearity often occur in other analog components. Therefore, the analog control system today is difficult to satisfy the performance required. Furthermore, because the classical digital control is only restricted to single feedback loop system, digital design is very difficult, and especially there is no sufficient theory for a multi－loop feedback system, the hybrid electrohydraulic servo system is widely used in almost all the control system design of the table. No doubt, this scheme is unfavorable to improving its accuracy and reliability. Of caurse, optimal control and adaptive control are convenient for digital implementation, but they will take a long time to be carried out in software and this can not satisfy the expected bandwidth of a system. In addition, since the wordlength of a microprocessor is constant, a complex control algorithm will make large numbers of accumulated errors. In order to avoid the above difficulties, a simple algorithm based on a state observer and state feedback is given in this paper. The digital simulation results indicate that its performance is satisfactory.

1 Algorithm of the Controller

In order to obtain acceleration signals and implement the system conveniently by

a microprocessor , a state observer is used to gain all the states in this design. The system block diagram is shown in Fig. 1.

The state space representation of the system is given by

$$\dot{X}=AX+BU$$

$$Y=CX \tag{1}$$

where $\{A, B, C\}$ can be both controllable and observable.

Assume the sampling period to be T_0, then the discretization of Eq. (1) is as follows:

$$X(k+1)=FX(k)+GU(k)$$

$$Y(k+1)=CX(k+1) \tag{2}$$

where $F=e^{AT_0}$,$G=\int_0^{T_0} e^{At}dt\cdot B$ and Eq. (2) can be both controllable and observable.

Introducing state feedback, we obtain

$$U(k)=r(k)-KX(k) \tag{3}$$

Substituting Eq. (3) into Eq. (2), we have

$$X(k+1)=FX(k)+G[r(k)-KX(k)]=[F-GK]X(k)+Gr(k)$$

$$Y(k+1)=CX(k+1) \tag{4}$$

From the linear system theory, we know that the poles of the system can be assigned to any expected position by selecting different feedback matrix K, and the introduction of state feedback will not destroy the controllablity of the system. Since Eq. (2) is observable, we can introduce a closed loop state observer to obtain all the states . Generally , the observer has two forms:

(1)

$$\bar{X}(k+1)==F\hat{X}(k)+GU(k)$$

$$\hat{X}(k+1)=F\bar{X}(k+1)+L[Y(k+1)-\bar{Y}(k+1)]$$

$$\bar{Y}(k+1)=C\bar{X}(k+1) \tag{5}$$

where $\hat{X}(k)$ is the reconstructed state of $X(k)$, and L is the gain matrix of the observer.

(2)

$$\hat{X}(k+1)=F\hat{X}(k)+GU(k)+L[Y(k)-\hat{Y}(k)]$$

$$\hat{Y}(k+1)=C\hat{X}(k+1) \tag{6}$$

where L is the same as the above-mentioned.

Substituting $\hat{X}(k)$ into Eq. (3), we have

$$U(k) = r(k) - K\hat{X}(k) \tag{7}$$

Substituting Eq. (7) into Eq. (2), we obtain

$$\begin{aligned} X(k+1) &= FX(k) - GK\hat{X}(k) + Gr(k) \\ Y(k+1) &= CX(k+1) \end{aligned} \tag{8}$$

In the continuous-time system, the separation theorem shows that the dynamics of the plant is independent of the dynamics of the observer, and vice versa. For a discrete-time system, there is also a same theorem.

Let $\Delta X(k) = X(k) - \hat{X}(k)$, then

$$\begin{aligned} X(k+1) &= (F - GK)X(k) + GK\Delta X(k) + Gr(k) \\ \Delta X(k+1) &= (F - LC)\Delta X(k) \end{aligned} \tag{9}$$

Rewrite Eq. (9) in the vector form as follows

$$\begin{bmatrix} X(k+1) \\ \Delta X(k+1) \end{bmatrix} = \begin{bmatrix} F-GK & GK \\ 0 & F-LC \end{bmatrix} \begin{bmatrix} X(k) \\ \Delta X(k) \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} r(k) \tag{10}$$

From the matrix theory, the characteristic polynomail of the discrete transfer matrix of Eq. (10) is given by Eq. (11)

$$\det[(\lambda I - F + GK)(\lambda I - F + LC)] = \det[\lambda I - F + GK] \det[\lambda I - F + LC] = 0 \tag{11}$$

So the poles of the plant and the observer can be assigned to the expected position by separately selecting K and L. Consequently the controller and the observer can be designed separately.

The stability of the system is discussed as follows. From Eq. (2), so long as the whole eigenvalues of the transfer matrix are inside the unit circle, Eq. (10) is stable. When the system comes into a steady-state, Eq. (12) is obtained.

$$\lim_{k\to\infty} \Delta X(k) = \lim_{k\to\infty} (F - LC)^k \Delta X(0) = 0 \tag{12}$$

Eq. (12) shows when k is sufficiently large, the state error of the observer and the plant is zero, i.e., the states estimated by the observer are the real states of the plant.

Generally, the whole state initial values of the plant are impossible to be known, so, when $X(0) \neq 0$, at the beginning of the system transient procedure the observed states are not the actual states of the plant; and when $X(0) = 0$, the observed states are the actual ones of the plant at the beginning. Analysing Eq. (6), we find that if $X(0) = 0$, then the observer matrix L is useless and an unclosed-loop observer will satisfy the requirements.

2 Pure Digital Control for the Outer Axis of a Three-Axis Flight Simulation Table

The electrohydraulic servo system used in the outer axis of a three-axis table can be expressed approximately as follows.

$$G(s)=\frac{K\cdot\frac{1}{D_m}\cdot\omega_n^2}{s(s^2+2\xi\omega_n s+\omega)}$$

where

$D_m=800\text{ml/rad}$ — displacement of the motor

$\omega_n=207.35\text{rad/s}$ — characteristic angular frequency of the motor

$\xi=0.1$ — damping ratio of the motor

K— gain of amplifier (determined by speed errors).

The following problems will be considered in design.

2.1 Selecting Expected Poles

From Eq. (13), we know that the system is of three-order. Then, we can make the expected poles of the plant equal to the closed-loop poles of a three-order optimal system. The expected system is expressed as follows.

$$G_e(s)=\frac{\omega_n^3}{s^3+2\omega_n s^2+2\omega_n^2 s+\omega_n^3} \tag{14}$$

where $\omega_n=K_v$, and K_v is the open loop gain of the system.

Here, we take $K_v=50$ and substitute it into Eq. (14), and then we obtain the expected poles of the plant as follows.

$$\lambda_1=-106.0639,\quad \lambda_{2,3}=-47\pm j85$$

In order to obtain the actual states of the plant as fast as possible, the expected poles of the observer should be as far away from the Im axis in [s] plane as possible, but in practice it is limited by the oscillating frequency of the clock. Generally its real parts are taken as (2 ~ 3) times as the real parts of the expected poles of the plant. Here, we take $\lambda_1'=2\lambda_1=-212$, $\lambda_{2,3}'=2\text{Re}[\lambda_{2,3}]\pm j40$

2.2 Selecting the Sampling Period

The sampling period T_0 must satisfy the sampling theorem. When T_0 is too large, sometimes we can obtain an expected robust discretization system, but after state feedback being introduced the closed-loop system will be an unstable positive feedback system (the unclosed-loop system must be stable). When T_0 is too little, the accumulated error will make the system unstable and reduce its accuracy.

So, from the viewpoint of engineering, T_0 will be selected as follows.

$$\omega_s > \max(5\sim10)\times(\omega_n \text{ or the bandwidth of the system})$$

In this paper, the characteristic frequency of the observer is larger than that of the expected system, and it has decisive influence on the system. So ω_n should be the characteristic frequency of the observer. Then

$$T_0 < \frac{\pi}{5\omega_n}$$

Here, T_0 is selected as 0.0005s.

2.3 Construction of the Control System

The implementation block diagram is shown in Fig. 2. The controller and the observer inside the closed dotted line are implemented by a microprocessor. Because the introduction of state feedback damps the output of the system, we introduce a compensator into the input pass. Its parameters are determined as follows:

From the viewpoint of the continuous system and assuming that there is no observer, the system shown in Fig. 2 is equivalent to the system shown in Fig. 3. Because k_1 is the feedback gain of state x_1 and considering that $Y=c_1x_1$, the output feedback gain in Fig. 3 should be k_1/c_1. For the unit step input, the output of the system is

$$Y(s)=\frac{c_1}{(sG'(s)+\frac{k_1}{c_1}\cdot c_1)}\cdot\frac{1}{S}=\frac{c_1}{sG(s)+k_1}\cdot\frac{1}{s}$$

Then

$$y(\infty)=\lim_{s\to0} s\cdot Y(s)=c_1/k_1$$

So, in order to ensure $y(\infty)=1$, the block ? in Fig. 3 should be replaced by an amplifer with a gain k_1/c_1.

3 Simulation Results

Based on Eq. (13), digital simulation has been done. The step and sine response of the plant with zero and little and large initial states are discussed respectively. When the state initial values of the plant are zero, the system is equal to one without an observer. Its response is shown in Fig. 4. The unit step response of the plant with initial states is shown in Fig. 5. When the state initial values of the plant is not zero, a non-normal balance point occurs in the position following time. Comparing Fig. 4 with Fig. 5, we find that the position following time is about 0.05s and 0.1s respectively for the system without and with an observer. The sin (79t) response (for the gain frequency bandwidth test) and the sin (69t) response (for the phase frequency banduidth test) of the system are shown in Fig . 6a and Fig . 6b respectively.

4 Conclusions

The pure digital control system for the outer axis of a three-axis flight simulation table is

designed in this paper. It is simple and convenient for implementation. Simulation results show that although a non－normal balance point occurs in the system with an observer, the scheme given in this paper can ensure perfect dynamic performance for the three－axis table and suit for a most electrohydraulic position servo system.

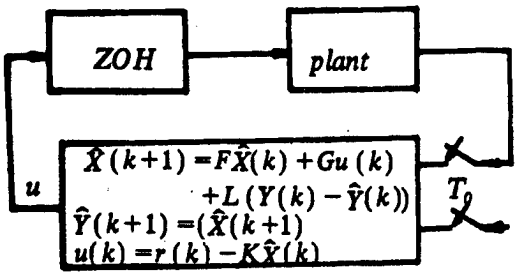


Fig . 1 A discrete－time control system including an observer

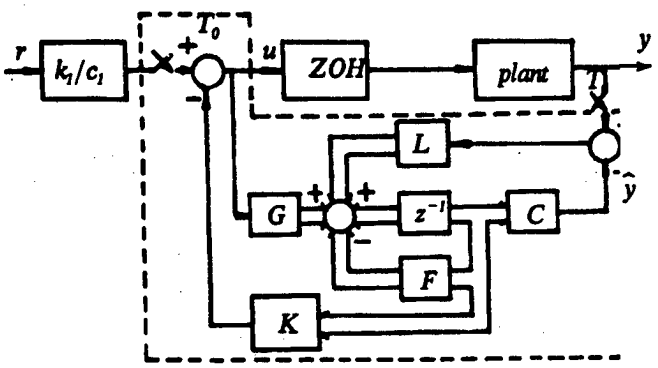


Fig. 2 A block diagram of the system in implementation

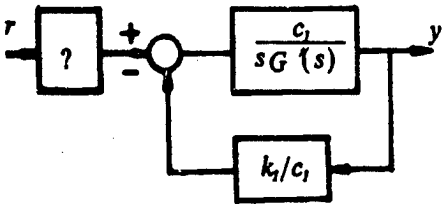


Fig . 3 The equivalent system

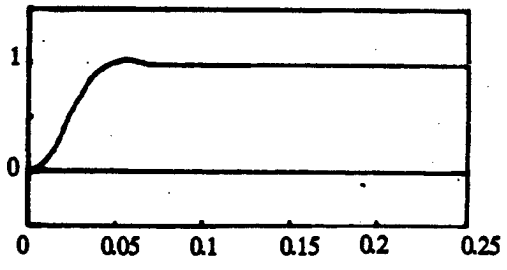


Fig . 4 The unit step response of the plant with a zero initial value

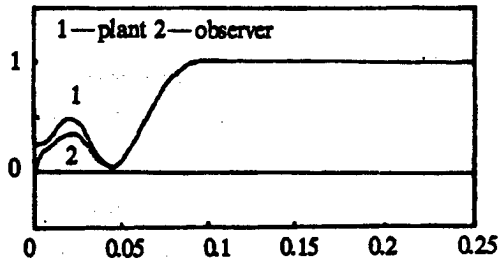
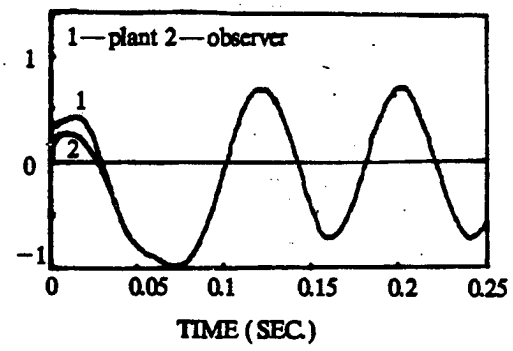
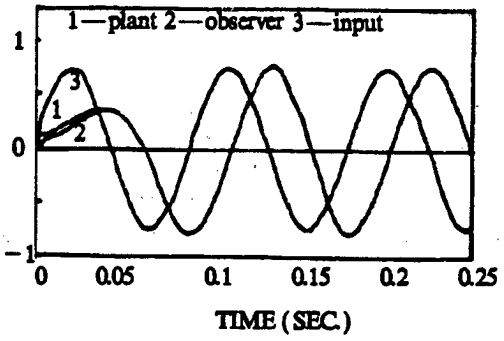


Fig . 5 The unit step response of the plant with a little initial value



(a)



(b)

Fig . 6 The sinewave response of the plant with an initial value

References

1 Chitsong Chen. Linear System, Theory and Design. Holt Rinehart and Winston. 1988

2 Paul Katz. Digital Control Using Microprocessor. Technion－Israel Institute of Technology. 1981

3 Astrom K J, Wittenmark B. Computer Controlled Sysetm — Theory and Design. Prentice － Hall, Inc. 1984

4 Li Hongrei. Hydraulic Control System. National Defence Industry Publishing House. Beijing. 1981 (In Chinese)

A Hybrid Learning Method for Multilayer Perceptrons

40100067A Harbin HARBIN GONGYE DAXUE XUEBAO[JOURNAL OF HARBIN INSTITUTE OF TECHNOLOGY] in English No 3, June 95, pp 52-61

Zhao Meide Huang Wenhu Hong Jiarong

(School of Astronautics)

Abstract A Newton learning method for a neural network of multilayer perceptrons is proposed in this paper. Furthermore, a hybrid learning method is legitimately developed in combination of the backpropagation method proposed by Rumelhart et al with the Newton learning method. Finally, the hybrid learning algorithm is compared with the backpropagation algorithm by some illustrations, and the results show that this hybrid learning algorithm has the characteristics of rapid convergence.

Key Words: Computer; Newton learning method; neural network; multilayer perceptron

0 Introduction

Multilayer perceptrons are feed－forward nets with one or more layers of nodes between the input and output nodes. These additional layers contain hidden units or nodes that are not directly connected to both the input and output nodes. Multilayer perceptrons overcome many of the limitations of single layer perceptrons [1], but were generally not used in the past because effective training algorithms were not available. This has changed with the development of new training algorithm, i. e., the backpropagation algorithm [2]. Since its emergence, they have been quite successfully used for a wide variety of tasks such as information processing, coding and pattern recognition [3-4]. Unfortunately, in many applications, the number of iteration required before convergence can be large and there exists local minima problem. Some modifications to the backpropagation algorithm were given by Vogl et al [5] in order to accelerate the convergence of the backpropagation algorithm. But we found that these modifications cannot accelerate the convergence by many numerical simulations. In Section 2 of the paper the backpropagation algorithm and some modifications to it are briefly described. In Section 3 the recursive formulas for calculating Hessian matrix of multilayer perceptrons are derived and a Newton learning algorithm is given. A hybrid learning algorithm for multilayer perceptrons is proposed in Section 4, and in Section 5 two examples are given to compare the two methods.

1 The Backpropagation Method and Its Modification

1.1 Model of Neural Nets

In general, a neural network consists of N nodes (processors), each of which is connected to all the other nodes. In the present work, however, only layered feedforward networks are used .Use of such a layered structure sacrifices a certain amount of parallelism since only processors in the same layer operate in parallel but adjacent layers compute serially. Although some parallelism is lost, these networks are popular because there exists a sophisticated learning scheme for training the network. A typical network using the layered structure is shown in. Fig. 1. It consists of an input layer, an output layer, and hidden layers in between. The nodes in the hidden layer are necessary to implement nonlinear mappings between the input and output patterns.

Associated with each node is a threshold value. The output of each node is determined by the level of the input signal in relation to the threshold value. Signals can propagate through the network only between adjacent layers and only in a single direction. These signals are modified by the connection strengths (.weights) between the nodes. Let

$\theta_{i,k}$ threshold of the ith node in the kth layer,

$W_{i,j,k}$ weight between jth node in the $(k-1)$th layer to ith node in kth layer,

$net_{i,k}$ input to the ith node in the kth layer,

$O_{i,k}$ output of the ith node in the kth layer.

The input to a node is given by

$$net_{i,k} = \sum_{j} [\, W_{i,j,k} O_{j,k-1} \,] + \theta_{i,k} \tag{1}$$

The summation index j extends over all nodes in the $(k-1)$th layer. The output of a node is given as

$$O_{i,k-1} = j(net_{i,k}) \tag{2}$$

where f is a sigmoid function given by

$$j(net_{i,k}) = \frac{1}{1+e^{-net_{i,k}}} \tag{3}$$

1.2 The Backpropagation Method and Its Modification

In order to describe conveniently, here we use vector notation . The basic iteration formula of the backpropagation algorithm is :

$$W^{k+1} = W^{k} - \eta_k \nabla E(W^k) + \alpha_k (W^k - W^{k-1}) \tag{4}$$

where W is a column vector which involves all the weights and the thresholds, W^k represents the column vector of W in kth iteration, $E(W)$ is the total error, i. e., the sum of the squares of the error of the output signal , $\nabla E(W)$ is the gradient column vector of $E(W)$ with respect to W, η_k is a learning constant , and α_k is a momentum factor.

In the backpropagation algorithm, η_k and α_k are preset constants . Let $E(W) = E_p(W)$, where $E_p(W)$ is the error function under the pth input patterns, i. e.,

$$E_p(W) = \frac{1}{2}\sum_j (d_{pj} - O_{pj,n})^2 \tag{4}$$

where the summation index j extends over all nodes in the output layer. d_{pj} is the desired output (provided by a teacher via the input−output samples) at the ith output node and $O_{p,j,n}$ is the actual output at the jth output node. The element of $\nabla E(W)$ is :

$$\frac{\partial E_p}{\partial W_{i,j,k}} = -\delta_{pi,k} O_{pj,k-1} \tag{5}$$

$\delta_{p,i,k}$ is the error signal for unit i of the kth layer under the pth input samples.

$$\delta_{pi,k} = \begin{cases} j'(net_{p,i,k})(d_{pi} - O_{pi,k}) & for\ k = n \quad (6a) \\ j'(net_{p,i,k}) \sum_j (\delta_{pj,k+1} W_{j,i,k+1}) & otherwise \quad (6b) \end{cases}$$

where the index s indicates summation over all nodes of the kth layer. The detailed description of the backpropagation algorithm is given in [2] and [4].

Some modifications to the backpropagation algorithm was given in [5]. These modifications are as follows: Instead of updating the network weights after each pattern is presented to the network, the network is updated only after the entire repertoire of patterns to be learned has been presented to the network; Instead of keeping η constant, it is varied dynamically so that the algorithm utilizes a near−optimum η ; the momentum factor α is set to zero when, an signified by a failure of a step to reduce the total error, the information inherent in prior steps is more likely to be misleading than beneficial. Only after the network takes a useful step, i. e. , one that reduces the total error, does α again assume a non−zero value.

Here, we propose a new scheme for selecting α_k and η_k, they are selected such that

$$\alpha_k, \eta_k = \min_{\alpha,\eta} E(W^k - \eta \nabla E(W^k) + \alpha(W^k - W^{k-1}))$$

Theoretically we have solved the problem of how to select α and η in the backpropagation algorithm. However, by many numerical simulations we find that both the modifications to the backpropagation algorithm given in [5] and above scheme cannot accelerate the convergence of the backpropagation method. In the next section we propose a new learning method for multilayer perceptrons.

2 The Newton Learning Method for Multilayer Perceptrons

In order to solve above problems, we propose a Newton learning method for multilayer perceptrons.

2.1 The Hessian Matrix of Multilayer Perceptrons

The Hessian matrix of the network under the pth input patterns is denoted as H_p Hence we have

$$H_p = \nabla^2 E_p$$

where $\nabla^2 E_p$ is a second order derivative matrix of E_p with respect to W. For convenience, we introduce the following notation:

$$H^{r,q}_{i,k,j,l} = \frac{\partial^2 E_p}{\partial W_{i,k,r} \partial W_{j,l,q}} \tag{7a}$$

$$\gamma_{i,j,r} = \frac{\partial \delta_{pi,r}}{\partial net_{pj,r}} \tag{7b}$$

$$\zeta_{i,j,r,q} = \frac{\partial \delta_{pi,r}}{\partial net_{pj,q}} \tag{7c}$$

$$\varphi_{i,j,r,q} = \frac{\partial O_{pi,r}}{\partial net_{pj,q}} \tag{7d}$$

Since H_p is a symmetric matrix, we can discuss it in two cases :

1) The diagonal elements of H_p , $H^{r,r}_{i,k,j,l}$

From (5), we have,

$$\frac{\partial E_p}{\partial W_{i,k,r}} = -\delta_{pi,r} O_{pk,r-1}$$

By (7a), we obtain

$$H^{r,r}_{i,k,j,l} = \frac{\partial^2 E_p}{\partial W_{i,k,r} \partial W_{j,l,r}} = O_{pk,r-1} \frac{\partial \delta_{pi,r}}{\partial W_{j,l,r}}$$

since

$$\frac{\partial \delta_{pi,r}}{\partial W_{j,l,r}} = \frac{\partial \delta_{pi,r}}{\partial net_{pj,r}} \frac{\partial net_{pj,r}}{\partial W_{j,l,r}}$$

According to (1) and (7b), we can obtain

$$\frac{\partial \delta_{pi,r}}{\partial W_{j,l,r}} = O_{pl,r-1} \gamma_{i,j,r}$$

Hence

$$H^{r,r}_{i,k,j,l} = -O_{pk,r-1} O_{pl,r-1} \gamma_{i,j,r} \tag{8a}$$

2) The non-diagonal elements of H_p , $H^{r,r}_{i,k,j,l}$ (r > q)

From (7a) and (5), we have

$$H^{r,r}_{i,k,j,l} = \frac{\partial (-\delta_{pi,r} O_{pk,r-1})}{\partial net_{pj,q}} \cdot \frac{\partial net_{pj,q}}{\partial W_{j,l,q}}$$

$$= -\frac{\partial net_{pj,q}}{\partial W_{j,l,q}} \left(O_{pk,r-1} \frac{\partial \delta_{pi,r}}{\partial net_{pj,q}} + \partial \delta_{pi,r} \frac{O_{pk,r-1}}{\partial net_{pj,q}} \right)$$

According to (1), (7c) and (7d), we can obtain

$$H^{r,q}_{i,k,j,l} = -O_{pl,q-1}(O_{pk,r-1}\zeta_{i,j,r,q} + \delta_{pi,r}\varphi_{k,j,r-1,q}) \quad (8b)$$

where the recursive formulas for calculating $\gamma_{i,j,r}, \zeta_{i,j,r,q}$ and $\varphi_{i,j,r,q}$ are given in Appendix.

2.2 The Newton Learning Method for Maltilayer Perceptrons

According to the Hessian matrix derived above, we can obtain the basic iteration formula of the Newton learning method for multilayer perceptrons as follows:

$$W^{k+1} = W^k - \psi_k H_k^{-1} G_k \quad (9)$$

where $H_k = \sum_p H_p = \sum_p \nabla^2 E_p(W^k)$, and

$$G_k = \sum_p \nabla E_p(W^k)$$

The index p represents summation over all input−output samples used for training. ψ_k is a constant which is called a Newton lea ning constant in the paper.

The Newton learning method has been implemented on a Sun3 Workstation in Clanguage. We find by many illustrations that this algorithm cannot converge to a solution. There are possibly several reasons for it through our analysis: 1) Although the Newton algorithm can quadratically converge near the exact solutions, it is very difficult to find a better initial guess; 2) After several iterations, the matrix H will become singular; 3) We cannot ensure that for each iteration the quadratic model can adequately approximate the objective function E.

To apply the Newton learning method proposed in the paper, we propose a hybrid learning method which combines the backpropagation method with the Newton learning method. The hybrid learning method has the advantages of the two methods. The illustrations given in Section 5 show that this new learning algorithm has a surprising speed of convergence.

3 The Hybrid Learning Method for Multilayer Perceptrons

Since the backpropagation method is a steepest decent method, it can globally converge to a solution. The Newton learning method can quadratically converge in a local region of the solution. Therefore, the hybrid learning method can rapidly globally converge. The basic idea of the hybrid learning method is: after each input pattern is presented to a network, the weights and thresholds are updated by the backpropagation algorithm; after entire input patterns are presented to the network, the weights and thresholds are updated by the Newton learning method. The detailed description of the algorithm is as follows.

Step1 Initialize weights and thresholds (set all weights and node thresholds to small random values).

Step2 According to (1) and (2), calculate the actual outputs of each node under the pth input pattern in forward direction.

Step3 Calculate the error signals of the output nodes in terms of (6a) and update the weights and thresholds of the output layer according to (3)

Step4 Calculate the $\gamma_{i,j,r}$ of the output layer in terms of (A. 1a) and calculate the corresponding elements of the matrix H_p by (8a).

Step5 Calculate $\zeta_{i,j,r,q}$ and $\varphi_{i,j,r,q}$ in terms of (A.3) and (A.4), then calculate the corresponding elements of the matrix H_p by (8b).

Step6 By(A.1)calculate the $\gamma_{i,j,r}$ of the penultimate nodes, and calculate the corresponding elements of the matrix H_p by (8a).

Step7 Calculate the error signals of the penultimate nodes in terms of (6b), and update the weights and thresholds of the penultimate layer according to (3).

Step8 Repeat Step 5, Step 6 and Step 7, until the entire elements of matrix H_p is calculated.

Step9 Repeat Step 2 – Step 8, until all input patterns are presented to the network.

Step10 Decide whether or not matrix H has an inverse matrix. If it has, then go to Step 11, otherwise go to Step 12.

Step11 Update all weights and thresholds of the network according to (9).

Step12 If the maximum of the differences between the actual output and the desired output of the output layer for all input patterns is less than preset value, then terminate, otherwise go to Step 2.

4 Illustrations and Comparison

The backpropagation algorithm and the hybrid algorithm have been implemented on Sun3 Workstation in C and FORTRAN languages. Two examples are given as follows.

4.1 The XOR Problem

The XOR problem is a difficult classical problem, which can't be solved by single layer perceptrons, Since it is a simplest nonlinear seperatable problem. The XOR problem, i.e., Exclusive OR problem is the problem that a neural net can generate the following outputs for the following input samples by training.

Input	Pattern	Output
0	0	0
1	1	0
0	1	1
1	0	1

Fig. 2 and Fig. 3 show the solution solved by the backpropagation algorithm and the hybrid algorithm respectively. The solution in Fig. 2 is reached after 532 sweeps through the four stimulus patterns with learning rate of $\eta = 0.5$. and momentum factor of $\alpha = 0.9$. The solution in Fig. 3 is reached after 39 sweeps with $\eta = 0.5$, $\alpha = 0.9$ and $\psi = 0.11$.

Through many numerical simulations we find the parameter η and α effects on the number of iterations before convergence for the two methods. Fig. 4 and Fig. 5 show these effects

respectively. The Newton learning constant ψ affects the number of iterations before convergence for the hybrid algorithm. Fig.6 shows this effect. ψ is now varied between 0.005 and 0.02 by experience.

In numerical simulation we also find that the Newton updation for the weights and thresholds only occur in the first several sweeps in the hybrid algorithm. Table 1 is obtained by assuming that after 10 sweeps the Newton updating stops.

Table 1. Set $\alpha=0.95$ and $\psi=0.011$

η	Backpropagation		Hybrid Learning	
	sweep	CPU time	sweep	CPU time
0.1	2736	246	127	6
0.2	1011	87	124	6
0.3	581	51	68	4
0.4	383	35	29	2
0.5	274	25	28	2

BP—Backpropagation HL—Hybrid Learning

4.2 The Parity Problem

The parity problem is a problem which can computer the parity of any input pattern, i.e., the output required is 1 if the input pattern contains an odd number of 1s and 0 otherwise. This is a very difficult problem because the most similar patterns require different answers. The XOR problem is a parity problem with input patterns of size two. Here we have done a parity problem with patterns of size 3. Fig. 7 and Fig. 8 show the actual solutions with three input lines and three hidden units obtained by the backpropagation algorithm and the hybrid algorithm respectively. Let $\eta=0.5$, $\alpha=0.9$, and $\psi=0.006$. The solutions in Fig. 7 and Fig. 8 are reached after 1441 and 52 sweeps through the eight stimulus patterns respectively.

5 Conclusions

In this paper, a Newton learning algorithm for multilayer perceptrons is derived and a hybrid learning algorithm is developed in combination of the backpropagation algorithm with the Newton learning algorithm. The results of numerical simulations given in the paper show that the hybrid algorithm has much faster convergence than the backpropagation algorithm. The future study of the project is to generalize this new algorithm to high order neural nets and apply it to solving practical problems.

Appendix

The Recursive Calculation For $\gamma_{i,j,r}$, $\zeta_{i,j,r,q}$ and $\varphi_{i,j,r,q}$

1. The Recursive Calculation For $\gamma_{i,j,r}$
We derive the formula of $\gamma_{i,j,r}$ in two cases as follows.

(i) Layer r is an output layer

From (6a), we have

$$\delta_{pi,r}=(d_{pi}-O_{pi,r})j\,'(net_{pi,r})$$

By (7b), we can obtain.

$$\gamma_{i,j,r}=\begin{cases}0 & for\ i\neq j\\ j\,''(net_{pi,r})(d_{pi}-O_{pi,r})-[(j\,'(net_{pi,r})]^2, & otherwise\end{cases} \qquad (A1a)$$

(ii) Layer r is a hidden layer

From (6b), we have

$$\delta_{pi,r}=j\,'(net_{pi,r})\sum_j \delta_{pj,r+1}W_{j,i,r+1}$$

(1) For $i\neq j$,

$$\gamma_{i,j,r}=j\,'(net_{pi,r})\sum_s W_{s,i,r+1}\cdot\frac{\partial\delta_{ps,r+1}}{\partial net_{pj,r}}$$

By (7c), we obtain,

$$\gamma_{i,j,r}=j\,'(net_{pi,r})\sum_s W_{s,i,r+1}\,\zeta_{s,j,r+1,r}) \qquad (A1b)$$

(2) For i=j,

$$\gamma_{i,j,r}=j\,''(net_{pi,r})\sum_s(\delta_{ps,r+1}W_{s,i,r+1})+j\,'(net_{pi,r})\sum_s(W_{s,i,r+1}\zeta_{s,i,r+1,r})$$

2. The Recursive Calculation For $\zeta_{i,j,r,q}$ (A1c)

According to (7c), we have.

$$\zeta_{i,j,r,q}=\frac{\partial\delta_{pi,r}}{\partial net_{pj,q}}=\sum_s\left(\frac{\partial\delta_{pi,r}}{\partial net_{ps,q+1}}\frac{\partial net_{ps,q+1}}{\partial net_{pj,q}}\right)$$

$$\because\ \frac{\partial net_{ps,q+1}}{\partial net_{pj,q}}=\frac{\partial net_{ps,q+1}}{\partial O_{pj,q}}\frac{\partial O_{pj,q}}{\partial net_{pj,q+1}}=W_{s,j,r+1}f\,'(net_{pj,q}) \qquad (A.2)$$

$$\text{and}\ \because\ \frac{\partial\delta_{pi,r}}{\partial net_{ps,q+1}}=\begin{cases}\gamma_{i,s,r} & ,\ for\ \ r=q+1\\ \zeta_{i,s,r,q+1} & ,\ for\ \ r>q+1\end{cases}$$

$$\therefore\ \zeta_{i,j,r,q} = \begin{cases} j'(net_{pj,r}) \sum_s (W_{s,j,q+1}\, \gamma_{i,s,r}), & for\ r=q+1 \\ j'(net_{pj,r}) \sum_s (W_{s,j,q+1}\, \zeta_{i,s,r,q+1}), & for\ r>q+1 \end{cases} \qquad (A.3)$$

3. The Recursive Calculation For $\varphi_{i,j,r,q}$

(i) For r=q,

By (7d), we obtain,

$$\varphi_{i,j,r,r} = \frac{\partial O_{pi,r}}{\partial net_{pj,r}} = \begin{cases} 0, & for\ i \neq j \\ j'(net_{pj,r}) & for\ i=j \end{cases} \qquad (A,4a)$$

(ii) For r > q

According to (7d) and by chain rule, we have

$$\varphi_{i,j,r,q} = \frac{\partial O_{pi,r}}{\partial net_{pj,r}} = \sum_s \frac{\partial O_{pi,r}}{\partial net_{ps,q+1}} \cdot \frac{\partial net_{ps,q+1}}{\partial net_{pj,r}}$$

By (A . 2), we obtain

$$\varphi_{i,j,r,q} = j'(net_{pj,q}) \cdot \sum_s (W_{s,j,q+1}\, \varphi_{i,s,r,q+1}) \qquad (A.4b)$$

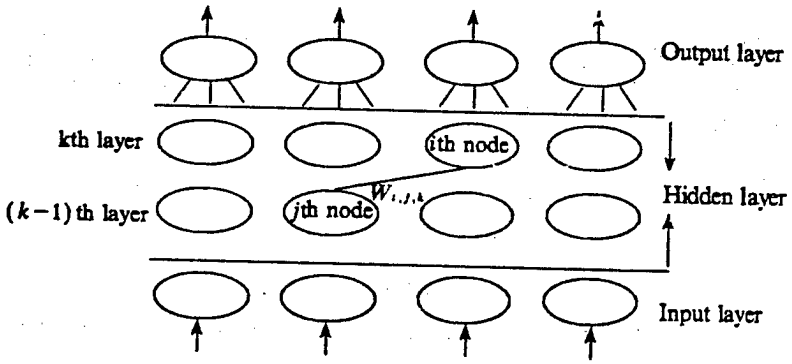


Fig . 1 The structure of a feed−forward net

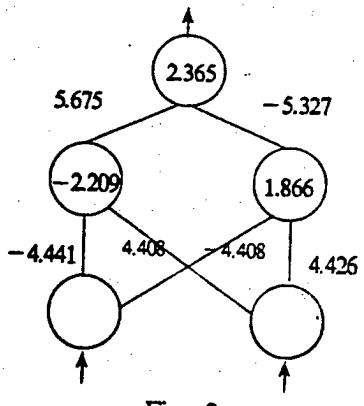


Fig . 2

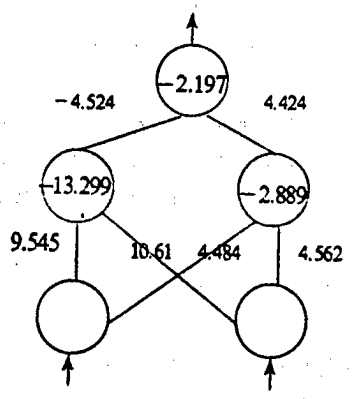


Fig . 3

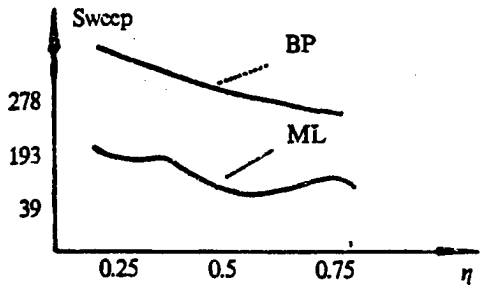


Fig . 4

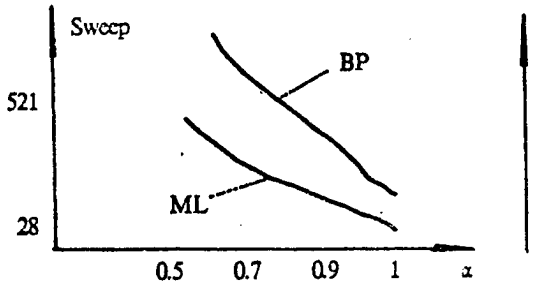


Fig . 5

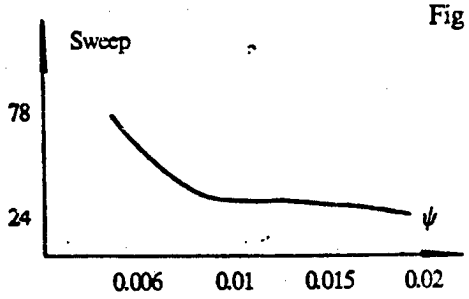


Fig . 6

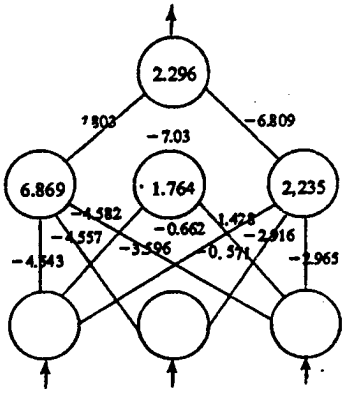


Fig . 7

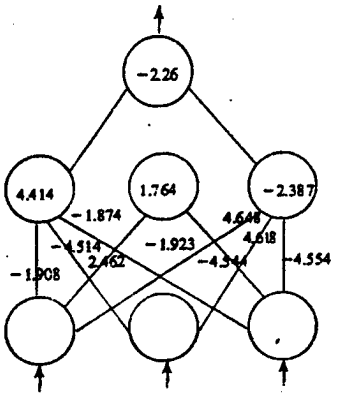


Fig . 8

References

1 Minsky M , Papert S. Perceptrons: An Introduction to Computational Geometry. MIT Press. 1969

2 Rumelhart D E, Hinton G E, Williams R J. Learning Internal Representations by Error Propagation. In: D. E, Rumelhart, and J. L. McClelland. (eds). Parallel Distributed Processing. Exploration in the Microstructure of Cognition. Vol. 1. Foundations, MIT Press. Cambridge. 1986

3 Rumelhart D E, McClelland J L, the PDP Research Group. Parallel, Distributed Processing, Exploration in the Microstructure of Cognition. Vol. 1－2. MIT Press Cambridge. 1986

4 Lippman R P. An Introduction to computing with Neural Nets. IEEE ASSP Magazine. 1987; 4(2): 4－22

5 Vogl T P et al. Accelerating the Convergence of the Backpropagation Method. Biol. Cybern. 1988; 59 . 257－263

6 Hu Zhongying. A modified Memory－Gradient Method, J. of Numerical Mathematics for Colleges. 1989; 2 (In Chinese)

- END -

22161
63

NTIS
ATTN: PROCESS 103
5285 PORT ROYAL RD
SPRINGFIELD, VA 22161

Foreign Broadcast Information Service (FBIS) and Joint Publications Research Service (JPRS) publications contain political, military, economic, environmental, and sociological news, commentary, and other information, as well as scientific and technical data and reports. All information has been obtained from foreign radio and television broadcasts, news agency transmissions, newspapers, books, and periodicals. Items generally are processed from the first or best available sources. It should not be inferred that they have been disseminated only in the medium, in the language, or to the area indicated. Items from foreign language sources are translated; those from English-language sources are transcribed. Except for excluding certain diacritics, FBIS renders personal names and place-names in accordance with the romanization systems approved for U.S. Government publications by the U.S. Board of Geographic Names.

Headlines, editorial reports, and material enclosed in brackets [ ] are supplied by FBIS/JPRS. Processing indicators such as [Text] or [Excerpts] in the first line of each item indicate how the information was processed from the original. Unfamiliar names rendered phonetically are enclosed in parentheses. Words or names preceded by a question mark and enclosed in parentheses were not clear from the original source but have been supplied as appropriate to the context. Other unattributed parenthetical notes within the body of an item originate with the source. Times within items are as given by the source. Passages in boldface or italics are as published.

SUBSCRIPTION/PROCUREMENT INFORMATION

The FBIS DAILY REPORT contains current news and information and is published Monday through Friday in eight volumes: China, East Europe, Soviet Union, East Asia, Near East & South Asia, Sub-Saharan Africa, Latin America, and West Europe. Supplements to the DAILY REPORTs may also be available periodically and will be distributed to regular DAILY REPORT subscribers. JPRS publications, which include approximately 50 regional, worldwide, and topical reports, generally contain less time-sensitive information and are published periodically.

Current DAILY REPORTs and JPRS publications are listed in *Government Reports Announcements* issued semimonthly by the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, Virginia 22161 and the *Monthly Catalog of U.S. Government Publications* issued by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

The public may subscribe to either hardcover or microfiche versions of the DAILY REPORTs and JPRS publications through NTIS at the above address or by calling (703) 487-4630. Subscription rates will be provided by NTIS upon request. Subscriptions are available outside the United States from NTIS or appointed foreign dealers. New subscribers should expect a 30-day delay in receipt of the first issue.

U.S. Government offices may obtain subscriptions to the DAILY REPORTs or JPRS publications (hardcover or microfiche) at no charge through their sponsoring organizations. For additional information or assistance, call FBIS, (202) 338-6735,or write to P.O. Box 2604, Washington, D.C. 20013. Department of Defense consumers are required to submit requests through appropriate command validation channels to DIA, RTS-2C, Washington, D.C. 20301. (Telephone: (202) 373-3771, Autovon: 243-3771.)

Back issues or single copies of the DAILY REPORTs and JPRS publications are not available. Both the DAILY REPORTs and the JPRS publications are on file for public reference at the Library of Congress and at many Federal Depository Libraries. Reference copies may also be seen at many public and university libraries throughout the United States.