Proc. 2002 Australasian Conference on Robotics and Automation
Auckland, 27-29 November 2002

Mechatronic Controller for Tele-Operated Camera Platform

N.S. Tlale, G. Bright, J. Potgieter
Massey University, Institute of Technology and Engineering
Auckland, New Zealand
Fax: +64 – 9 – 4439640
E-mail: g.bright@massey.ac.nz

Abstract

The New Zealand Film industry requires skilled operators to control the orientation of camera platforms. This is one of the crucial factors in producing movies with precision cinematography. Film footage of moving landscapes makes it difficult for the camera operator to efficiently operate the camera platform e.g. an inclined mountain terrain with surrounding trees and bushes. Such scenes require tele-operation of the camera platform.

Tether tele-operation increases the time required to set-up the camera platform for the scene shoots, and hence increases the cost of making a movie. Tether connections limit the manner in which the camera platform can be mounted in order to achieve the best scene shooting. This paper reports on the project undertaken by the Mechatronics and Robotics Research Group at Massey University, in conjunction with the New Zealand film industry, on the development of wireless control for tele-operated camera platforms.

1 Introduction

A typical camera platform used in the film industry can be described as a three degree of freedom, revolute manipulator. The three degrees of freedom correspond to the yaw, roll and pitch orientations. Figure 1 shows the camera platform at rest position, with the mentioned degrees of freedom. The base is normally fixed on a stationery or moving platform. The camera platform consists of three links; joints O, O' and C' join these links.

Using the Denavit-Hartenberg (D-H) representation let the axis of rotation of the yaw rotation (γ) to be Z_o axis of the reference co-ordinate system $X_oY_oZ_o$. The other three co-ordinate systems can be located as per D-H criteria [Fu et al, 1986]. The reference co-ordinate system does not rotate with the camera platform. The $X_1Y_1Z_1$ co-ordinate system can be rotated about the Z_1 axis at an angle θ (pitch). The origins of both the $X_0Y_0Z_0$ and the $X_1Y_1Z_1$ co-ordinate systems is O.

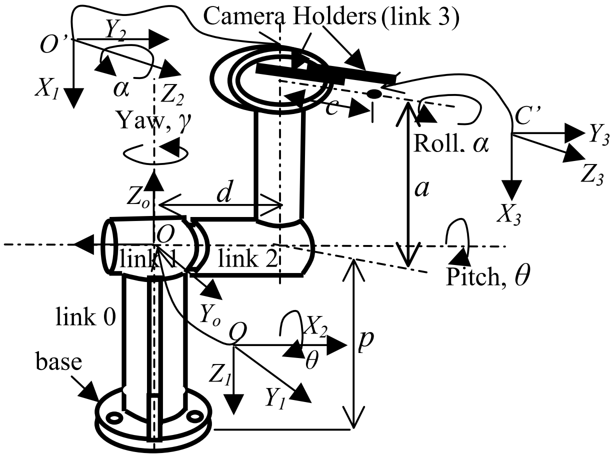


Figure 1. Schematic of the Camera Platform.

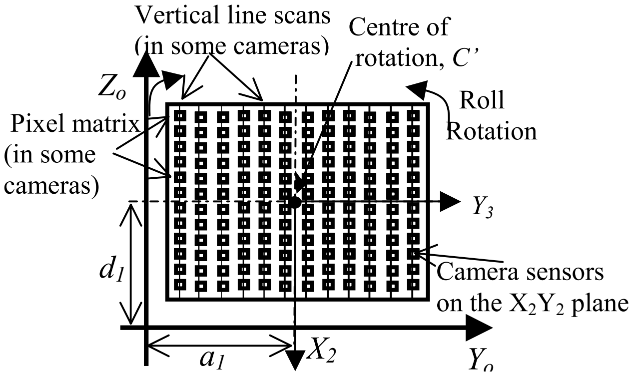


Figure 2. The Roll Rotation Rotates Other Camera Sensor Points About Point D'.

2 Kinematics of the Camera Platform

The Denavit-Hartenberg (D-H) representation simplifies the kinematics analysis of the camera platform. The kinematics analysis was used to gain insight of the camera platform, so that a robust, 'knowledge-based' control of the camera platform could be implemented. The kinematics model will be further used to develop a simulation model of the camera platform. This will enable the camera platform operator to simulate the shoots and analyse the best shoots strategy offline. This optimises the time taken on the shoot site.

Table 1 indicates the values for a D-H representation for the camera platform, where,

θ_i = angle of rotation from X_{i-1} axis to the X_i axis about the Z_{i-1} axis,
α_i = angle of rotation from Z_{i-1} axis to the Z_i axis about the X_i axis,
a_i = the distance from the intersection of Z_{i-1} axis and the X_i axis to the origin of the ith co-ordinate system along the X_i axis, and
d_i = the distance from the origin of the (i-1)th co-ordinate system to the intersection of Z_{i-1} axis and X_i axis along Z_{i-1} axis.

Link	θ_i	α_i	a_i	d_i
1	γ	-90^0	0	0
2	θ	90^0	a	d
3	α	0	c	0

Table1. D-H Parameters for the Developed Camera Platform

To get the camera to focus on a scene, the camera operator would adjust the yaw rotation followed by the pitch rotation, and finally the roll rotation. To find the composite homogenous transformation matrix, **T**, which will relate a point $\mathbf{p}_3$ in the $X_3Y_3Z_3$ frame with respect to the base co-odinate system $X_oY_oZ_o$, then:

$$\mathbf{p}_0 = \mathbf{T}\mathbf{p}_3$$

where $\mathbf{p}_3 = (x_3, y_3, z_3, 1)^T$ and $\mathbf{p}_0 = (x_0, y_0, z_0, 1)^T$. **T** can be obtained by the chain product of successive co-ordinate transformation matrices of ${}^{i-1}\mathbf{A}_i$. ${}^{i-1}\mathbf{A}_i$ is the basic homogenous rotaion-translation matrix for a D-H reperesentation, where i represents the number of the link or co-ordinate frame concerned. Using ${}^{i-1}\mathbf{A}_i$ matrix, one can relate a point $\mathbf{p}_i$ at rest in link i, as expressed in homogenoeus co-ordinates with respect to co-ordinate system i, to the co-ordinate system i-1 established at link i-1 by,

$$\mathbf{p}_{i-1} = {}^{i-1}\mathbf{A}_i\,\mathbf{p}_i$$

where $\mathbf{p}_{i-1} = (x_{i-1}, y_{i-1}, z_{i-1}, 1)^T$ and $\mathbf{p}_i = (x_i, y_i, z_i, 1)^T$. Thus, **T** can be determined as,

$$\mathbf{T} = {}^{0}\mathbf{A}_1\,{}^{1}\mathbf{A}_2\,{}^{2}\mathbf{A}_3$$

${}^{i-1}\mathbf{A}_i$ can be found by rotating and traslating the co-ordinate frame i-1 so that its axes are the same as the axes of the co-ordinate frame i. The D-H procedure for achieving this is:

i) rotate about the Z_{i-1} axis an anlge θ_i to to align X_{i-1} axis with the X_i axis
ii) translate along the Z_{i-1} axis a distance of d_i to bring the X_{i-1} and X_i axes into coincidence
iii) translate along the X_i axis a distance of a_i to bring the two origns as well as the X_{i-1} axis into coincidence
iv) rotate about the X_i axis an anlge α_i to bring the two co-ordinate systems into coincidence [Denavit et al, 1995].

Defining the matrix $\mathbf{B}_{\zeta,\xi}$ to represents the rotation/ranslation of ξ about/along the ζ axis. Then, each of above operations can be represented by a basic homogeneous rotation-translation matrix. The product of these four basic homogeneous transformation matrices results in a ${}^{i-1}\mathbf{A}_i$, i.e.

$${}^{i-1}\mathbf{A}_i = \mathbf{B}_{Z_{i-1},d_i}\,\mathbf{B}_{Z_{i-1},\theta_i}\,\mathbf{B}_{X_i,a}\,\mathbf{B}_{X_i,\alpha}$$

Using the general homogenous matrix of the form

$$\mathbf{B}_{\varsigma,\xi} = \left[\begin{array}{c|c} \text{rotation matrix} & \text{vector position} \\ \underline{(3\text{x}3)} & \underline{(3\text{x}1)} \\ \text{perspective} & \\ \text{transformation (1x3)} & \text{scaling (1x1)} \end{array}\right]$$

where the rotation matrix represents the rotation of the ith co-ordinate system with respect to i-1th co-ordinate system, vector position represent the starting postion of the ith co-ordinate in the i-1th co-ordinate frame, scaling is a scaling factor, and perspective transformation represents a tranformation martrix for a perspective projection. It can be shown that,

$${}^{i-1}A_i = \begin{bmatrix} \cos\theta_i & \cos\alpha_i \sin\theta_i & \sin\alpha_i \sin\theta_i & a_i \cos\theta_i \\ \sin\theta_i & \cos\alpha_i \cos\theta_i & -\sin\alpha_i \cos\theta_i & a_i \sin\theta_i \\ 0 & \sin\alpha_i & \cos\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Using Table 1, we can evaluate the matrices ${}^{i-1}\mathbf{A}_i$ as,

$${}^{0}\mathbf{A}_1 = \begin{bmatrix} \cos\gamma & 0 & -\sin\gamma & 0 \\ \sin\gamma & 0 & \cos\gamma & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$${}^{1}\mathbf{A}_2 = \begin{bmatrix} \cos\theta & 0 & \sin\theta & a\cos\theta \\ \sin\theta & 0 & -\cos\theta & a\sin\theta \\ 0 & 1 & 0 & d \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and $${}^{2}\mathbf{A}_3 = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 & c \\ \sin\alpha & \cos\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus **T** can be determined to be,

$$\mathbf{T} = {}^{0}\mathbf{A}_1\,{}^{1}\mathbf{A}_2\,{}^{2}\mathbf{A}_3$$

$$= \begin{bmatrix} C\gamma C\theta C\alpha - S\gamma S\alpha & C\gamma C\theta S\alpha - S\gamma C\alpha & C\gamma S\theta & C\gamma(cC\theta + aC\theta) - dS\gamma \\ S\gamma C\theta C\alpha + C\gamma S\alpha & S\gamma C\theta S\alpha - C\gamma C\alpha & S\gamma S\theta & S\gamma(cC\theta + aC\theta) - dC\gamma \\ -S\theta C\alpha & -S\theta S\alpha & C\theta & -cS\theta - aS\theta \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Knowing the transfomation matrix of the camera sensors, **H**, we can find the camera sensors' position in the reference cordinate frame by multiplying **T** by **H**. Thus the orientation of the camera sensors on the area of interest is known.

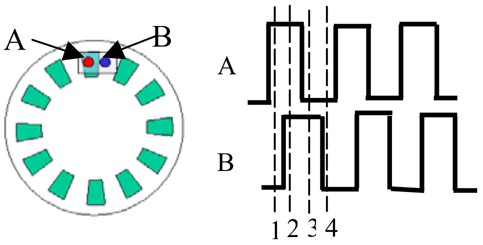
3 System Review

A camera platform that was controlled with a hard-wired control system was developed (ref Figure 3). It consisted of two hand wheels that control the roll and the yaw orientations. The joystick was used to control the pitch orientation. The two hand-wheels had adjustable viscous dampers to maintain the feel of a typical hand-wheel for conventional camera platforms. The two hand-wheels, the joystick and associated electronics were mounted on a separate box (operator side) from the camera platform (camera platform side, ref Figure 3).

The encoders implemented produced a quadrature signal, which could be used for position and direction sensing. Figure 4 illustrates the principle of quadrature signals. The phase of the quadrature signals was used to determine the direction of turn. Relative positional control, instead of absolute positional, of the camera platform was required. Incremental encoder with a counter was implemented to facilitated positional control and inference. Encoder signals were also used to determine the speed. Data acquisition of the encoder signals was triggered by the 'falling' and 'rising' edges of the two-encoder signals. Therefore, four signals were required to fully describe the motion of the encoder.

The servo-controller implementing a PID routine to control the position and velocity of the motors within the required response time, determinism and with stability. This closed-loop position control can be seen in Figure 4. The PID constant were tuned to achieve the required response times experimentally.

The hand-wheels had infinite rotation in both directions, while the rotations of the camera platform mechanisms were limited. To stop the camera platform mechanisms from rotating beyond the possible ranges of angular rotations when the hand-wheels were being turned infinitely, the limit switches were installed at the limits of possible rotation of the three rotating mechanism of the camera platform. The limit switches were used to input a zero control signal to the servo-controller when the ranges of the angular rotations of the camera platform mechanism were reached. The control system was designed so as not to respond to control signals above a certain threshold. For example, when the hand-wheels were being freely rotated at a high rotational speed.



State	Ch A	Ch B
S_1	High	Low
S_2	High	High
S_3	Low	High
S_4	Low	Low

Figure 4. Quadrature Signals from the Incremental Encoder.

The developed wireless communication system was required to send the position, speed and time information to the camera platform servomotor controller communication system.

3.1 Requirements on the Wireless Communication System Performance

The average sight reaction time of trained camera operators was used as a benchmark for the minimum response time of the integrated system [Washington University]. The average sight reaction time of a trained operator, t_o, was determined to be 0.2446 seconds. The developed hard-wired system was determined to have a lag/response time, t_{sh}, of 0.062 seconds. Since $t_{sh} >> t_o$, the developed hard-wired system had a satisfactory performance. The developed wireless system lag/response time had to be smaller than t_o = 0.2446 sec to achieve satisfactory control of the camera platform.

The lag/response time of the hard-wired system was measured by determining the time difference between the hand-wheel/joystick movement and the relevant anticipated response motion of the camera platform. This is illustrated in Figure 5. The encoder signals of the hand - wheels (or analogue signal of the joystick) and camera

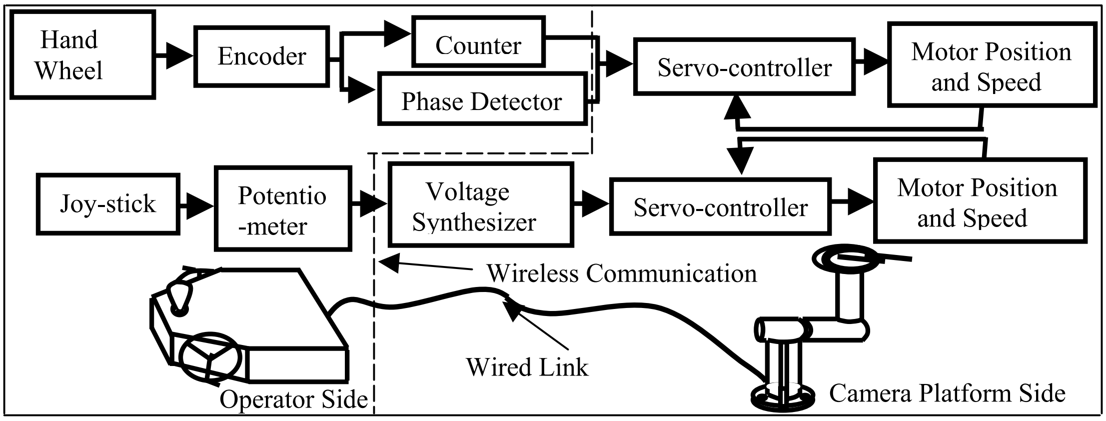


Figure 3. Control of Hard-wired Camera Platform.

platform DC servomotor were used to determined the response time t_{sh}, by determining the times the signals were initially observed. The lag times could be determined by implementing an oscilloscope or a simple measurement system. The response time of the hard-wired controlled camera platform represented the minimum response time that was achievable. The wireless control system's response time would be limited by the narrow bandwidth of the wireless communication system.

The incremental encoders implemented had 2000 counts per revolution. That was a total of 8000 counts per revolution for a quadrature signal. The maximum rotational speed of the hand-wheels that was set to produce a response on the DC servomotor was 10 revolutions per second. This was twice the maximum rotational speed that the hand-wheels could be rotated by the trained camera operators. Thefore, for both encoders, the wireless communication system had to be able to transmit:

$$\{2000\ (counts) \times 8\ (bits/byte/count) \times 4 (signals/count) + 1\ bit\ (direction)\} \times 2\ (encoders) \times 10\ (rev/second) = 1280.02\ Kbits/second$$

The signal from the joystick circuitry had to be sent across the wireless communication system. The position of the joystick could be represented as an eight bit word. The joystick directional information was processed by the voltage processing circuit on the camera platform side (refer Figure 3) Assuming the maximum frequency of 10 Hz, then the maximum number of bits per second that the wireless communication system needed to trasmit about the joystick information was:

$$8\ bits * 10\ Hz = 80\ bits/second$$

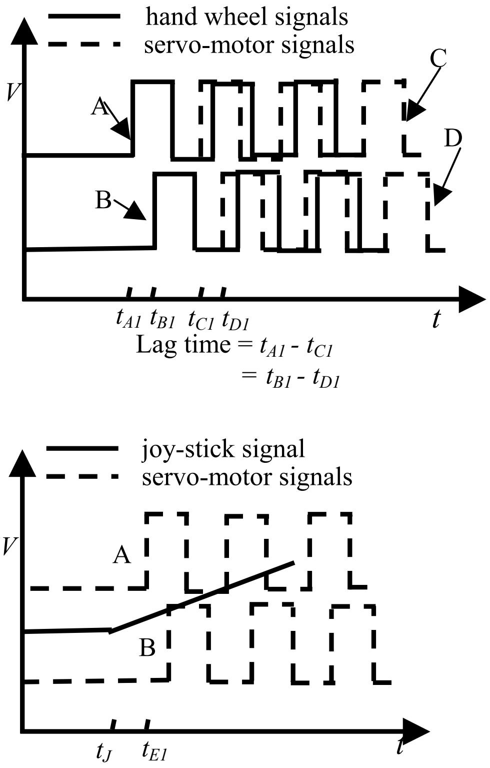


Figure 5. Quadrature Signals from the Incremental Encoder.

The total number of bits per second required to be handled by the wireless communication system, in order to effectively conrol the camera platform was:

$$1\ 280\ 020 + 80 = 1280.1\ Kbits/sec$$
$$= 160.0125\ Kbytes/sec$$

4 Wireless Communication System

Factors that were considered when choosing the technology for the wireless comminucation solution were: versatility, high perfomance and low-cost. Belkin wireless Ethernet implementing a carrier radio frequency of 2.4 GHz and IEEE 802.11b standard was used on two single board computers (the sender and the receiver). The physical layer throughtput of the system was given as 11 Mbits/sec, 5.5 Mbits/sec and 2 Mbits/sec depending on the distance between the the two data points. Figure 6 shows this relationship.

Comparing the data in Figure 6 and the calculated total number of bits per second (1280.1 Kbits/sec) required to achieve satisfactory control of the camera platform, it was expected that the camera platform response time would be satisfactory.

However, the *ping* tests, using data packets of different sizes, indicated that the performance of the wireless communication system was inferior to the one shown in Figure 5. Figure 7 shows the results of the *ping* tests. The *ping* tests gave an indication of the round-trip times for data packets of different sizes. Round trip-time is the time a data packet takes to be received by the transmitter from the time of transmission i.e. time from transmitter to receiver to transmitter [Comer, 2000]. The round-trip times were adequate to give an indication of the wireless communication system performance, as TCP/IP protocol was implemented. Issues relating to TCP/IP protocol are discussed in the following sections.

The results in Figure 6 were obtained under ideal conditions. There was a line of sight between the transmitter and the receiver. As soon as the line of sight was lost, the system became unreliable. A response time

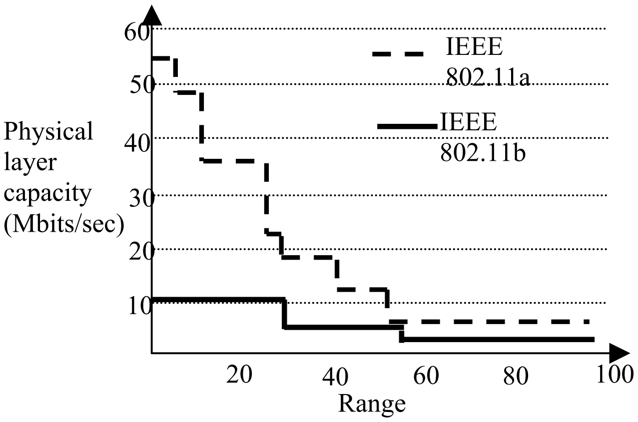


Figure 6. Physical Layer Capacity Wireless Network Implementing IEEE 802.11b Standard [Comer, 2000].

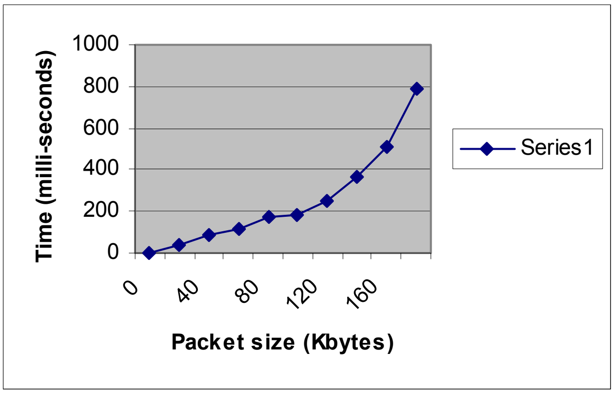


Figure 7. Round-times for data packets of different sizes implementing TCP/IP with line of sight.

of 0.72 seconds for a data packet of 120 Kbytes indicated that our system would not have had a satisfactory performance with data throughput of 160.0125 Kbytes/second as calculated in the previous section. In order to improve the system reliability and performance, techniques for reducing the minimum data transfer rate acceptable to optimally control the camera platform had to be used. A mechatronic design of the integrated camera platform was implemented in order to achieve this.

5 Mechatronic Design

Mechatronics can be defined as the synergy of mechanical, electrical and computer engineering to design systems that can have a superior performance with minimal resources [Bolton, 1995].

Several mechatronics techniques were implemented in order to design the wireless camera platform for optimum control. One of the strategies implemented was to make the platform end intelligent, and to process the raw data on the operator end so as to send as minimum data as possible across the wireless network, to efficiently control the camera platform.

In order to reduce the rate of data that the wireless communication system was to handle, firstly, a step-down gear mechanism (with fixed step-down ratio of ¼) was implemented on the operator side. The step-down gear mechanisms were inserted between the hand-wheels and their relative encoders. This reduced the amount of data that was needed from the encoders in order to optimally control the camera platform. The servo-controller algorithms were changed accordingly to accommodate this change.

Another strategy implemented was to design the software architecture of the wireless camera platform so as to be able to control the data throughput of the system. Figure 9 shows a schematic of the implemented architecture. Both the operator and camera platform ends were fitted with a single board computer. TCP/IP protocol was implemented as a network protocol under ‘peer-to-peer’ comminication mode between the two single board computers.

Data from circuits of the encoders of the hand-wheels and the joystick was supplied to dynamic RAM. Dynamic RAM supplied a sending socket buffer with its contents using the *first-in-firs-out* (FIFO) principle. Programming with threads was implemented to achieve this. Sending socket buffer supplied the circuitry of the wireless ethernet (physical layer of the protocol) of TCP/IP with its data contents as required by the TCP/IP. Wireless ethernet was implemented in the TCP/IP as a data-link layer. TCP/IP received data on the platform end and supplied the receiving socket buffer with its data. The receiving socket buffer supplied the dynamic RAM of the platform side with data, which was in turn supplied to the circuitry of the servo-controller. The application layer of TCP/IP protocol, File Transfer Protocol (FTP), handled this last communication action. FTP was used to transfer ASCII character sets. FTP implements transparent bit streams that permit exchanging any sort of data or text file [Roughan et al, 2001]. FTP could be also used to compress data in order to reduce the amount data handled by the wireless communication system. Word/identifier mechanisms for controlling user access would then be used.

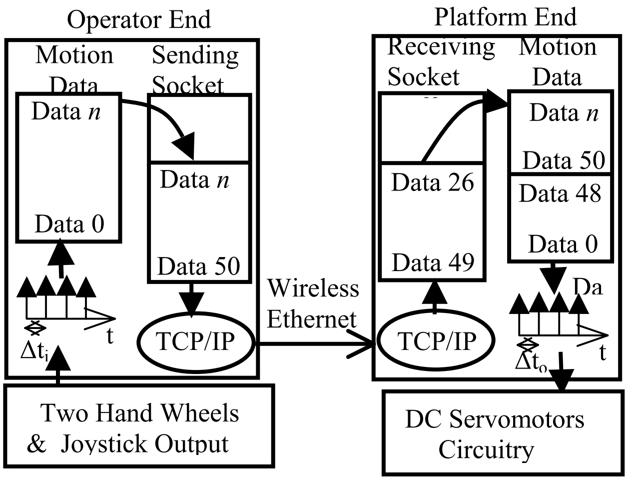


Figure 9. Flow of Control Data from the Encoders to the DC Motor Servo- controller.

The transport layer of TCP/IP protocol, TCP, transports an unstructured string of 8-bit bytes. On the transmit side, TCP accepts the data-stream from the originating host computer, segments it into packets, numbers each, and attaches an error control mechanism to each before sending it on its way. On the receiving side, TCP accepts each packet, checks each for errors, and re-sequences them in order to ensure that the file is reconstituted in its original form [Naugle, 1999]. Instances of missing or errored packets are resolved through requests for re-transmission. TCP assumes responsibility for the integrity of the entire data-stream, from end-to-end, which clearly adds value, but can decrease the wireless system throughput. This is the reason why the performance of wireless system implementing TCP/IP is not as shown in Figure 6.

Mechatronic design of the camera platform architecture was implemented by using software timers to trigger data acquisition/data transfer. On the operator’s side, a timer was used in order to trigger data acquisition from encoders’ and joystick’s circuitries to the dynamic RAM and receiver memory buffer. On the camera platform side, a timer was implemented in order to output data to the servo-controller from the receiver's buffer memory and dynamic RAM. Thus, the wireless system throughput was determined by the frequency of the timer. The frequency of the timers was variable (software) and could be adjusted to give optimum performance. The timers were programmed in such a way that the memory usage was a maximum in the buffers and the RAM.

Other software techniques could be used and are discussed in the latter sections

6 Performance of the Control System

The total response time of the PID servo-controller and DC servomotor was determined by exciting the system with a ramp input of about 3000 Kbytes/sec. This was twice the anticipated input signal as calculated previously. Figure 9 shows the response times of the system and error. The error increased with increasing input signal. The system exhibited a degree of second order response to a ramp input [Ogata, 1990]. It can be seen that the response times were in milli-seconds. The maximum response in lag occurred at 140 Kbytes/second and was equal to 25 milli-seconds. This was satisfactory for our application

By implementing the gear mechanism with a ¼ step down ration, the data that the wireless communication system had to handle was reduce by a factor of 4 i.e.

160.0125/4= 40.004 Kbytes/sec

From Figure 8, the round-trip time for a packet of 40 Kbytes was 0.087 seconds, which was satisfactory for our application.

The factors of affecting the total response time of the integrated system were:

(Response time of servomechanism and its controller) + (response time of wireless communication system) + (total response time of the different processes implemented) < t_o (0.2446 sec)

The determine response times were equal to

0.025 + 0.087 = 0.112 milli-seconds

Therefore the total response time of the different processes implemented in the system had to be less than 0.2446 - 0.112 = 0.1326 seconds. This was acceptable for the system's application.

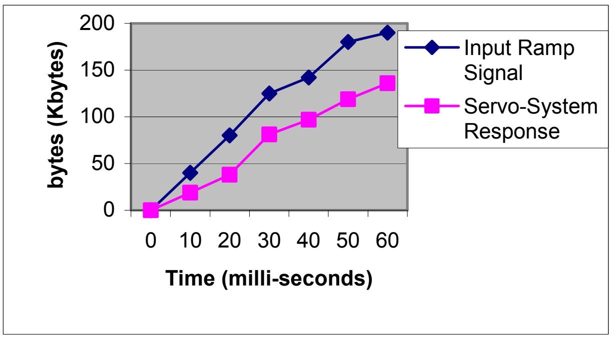


Figure 10. Response of DC servomotor and its PID controller to ramp input.

7 Conclusion

The mechatronic controller for tele-operated camera platform was developed implementing wireless network technologies. Individual sub-systems, consisting of developed integrated systems, were optimised. The benchmark used in determining the system's performance was the sight reaction time of the trained camera operators.

Research is continuing with tuning TCP/IP wireless networks. The UDP protocol, which does not have packet acknowledgment mechanisms, could be used. The next stage will entail the development of more computer software, using the kinematics equations of the system, so that the camera operators can simulate/program the camera platform offline.

References

[Fu et al, 1986] K.S.Fu, R. C. Gonzalez, C.S.G. Lee, *Robotics: Control, Sensing, Vision, and Intelligence*, McGraw-Hill Int. Ed. NY, 1986

[Denavit et al, 1995] Denavit, J. and Hartenberg, R. S.*, Kinematic Notation for Lower-Pair Mechanisms Based on Matrices*, Journal of Applied Mechanics, June 1995, pp. 215-221

[Washington University] http://faculty.washington.edu/chudler/java/reacttime.html

[Roughan et al, 2001] M. Roughan, A. Erramilli and D. Veitch, *Ne work performance for TCP networks. Part 1: Persistent sources*, Proceedings of Seventeenth International Tele-traffic Congress, Salvador da Bahia, Brazil, September 24-28, 2001.

[Comer, 2000] D. E. Comer, *Internetworking with TCP/IP :Principles, Protocols, and Architectures*, Vol.1, 4th Ed. ,Prentice-Hall, New Jersey, 2000.

[Bolton, 1995] W. Bolton, *Mechatronics: Electronic Control Systems in Mechanical Engineering*, Addison-Wesley Longman Ltd, Essex, England, 1995

[Naugle, 1999] M. Naugle, *Illustrated TCP/IP: A graphic Guide to the Protocol Suit*', John Wiley & Sons Inc., New York, 1999.

[Ogata, 1990] Katsuhiko Ogata, '*Morden Control Engineering*', 2nd Ed., Prentice-Hall Int Eddtions, New Jersey, 1990.