Memorias del XVI Congreso Latinoamericano de Control Automático, CLCA 2014
Octubre 14-17, 2014. Cancún, Quintana Roo, México

Modelado y control de una camara de video activa en modo seguimiento: Parte 1

Antonio Marin Hernández * **Elvia Palacios Hernández** ** **Daniel Murrieta Alvarez** *

* *Depto. Inteligencia Aritificial, UV Sebastian Camacho No. 5, C.P. 91000, Xalapa, Ver., México. (e-mail: anmarin@uv.mx,).*
** *Facultad de Ciencias, UASLP Av. Salvador Nava s/n, Zona Universitaria, C.P. 78290, S.L.P., México (e-mail: epalacios@fciencias.uaslp.mx)*

Abstract: Este trabajo en dos partes presenta el modelado y control de una cámara de video activa en modo de seguimiento. La parte 1 propone un modelo matemático del funcionamiento de la cámara de video en PanTilt de un robot móvil. Primero se propone un modelo en simulink de la cámara en PanTilt, para obtener señales de entrada y salida. Con esta información se estudian dos estructuras (ARX y OE) para la identificación de los coeficientes de modelos polinomiales correspondientes al eje horizontal y vertical. Finalmente se obtienen dos funciones de transferencias para los sistemas Pan y Tilt. La parte 2 mostrara que estos modelos pueden ser utilizados para el control del la cámara de video.

Keywords: .

1. INTRODUCCIÓN

La interacción humano -robot (IHR) es el estudio de las interacciones entre humanos y robots. La IHR es un campo multidisciplinar con la colaboración de la interacción persona-computadora, la inteligencia artificial, la robótica, la comprensión del lenguaje natural, diseño y ciencias sociales. Entonces el objetivo fundamental de la IHR es generar prinicipios y algoritmos para que los robots sean capaces de interactuar directa, segura y efectiva con los humanos. El problema principal en IHR es que la forma de comunicarse con el robot y la información que obtiene el robot del mundo que lo rodea difiere de como la realiza el humano. La comunicación entre el humano y robot tiene que ser lo más parecido posible a la comunicación que se produce entre humanos. Por que el robot tiene que ser capaz de reconocer a un humano y obtener información de su entorno. La comuniación visual en un robot es la que le proporciona la mayor cantidad de información. El sentido de la vista de un robot es proporcionado por cámaras, las cuales permiten reconocer personas, objetos, gestos, etc. En este trabajo se utiliza un robot móvil Pionner P3DX, el cual cuenta con diferentes sensores y una cámara de video con movimiento horizontal y vertical independiente de los movimientos del robot. Esta cámara es utilizada para explorar el entorno del robot. Una de las aplicaciones de esta cámara de video es la detección de rostros humanos para propositos de interacción. Entonces el objetivo de este trabajo es disear dos controladores, uno para cada eje (horizontal y vertical) para ubicar el rostro de un humano centrado en el espacio de trabajo del robot. En la Fig.1 se presenta en forma gráfica el esquema de control que se debe realizar en la plataforma robótica. Para realizar esto primero se propone determinar un modelo matemático de la cámara de video, el cual servira para ajustar un par de controladores PID. Además estos controladores son comparados con un controlador difuso.

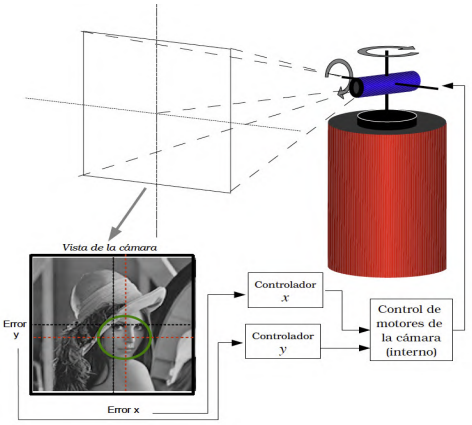

Fig. 1. Esquema general del sistema de control de seguimiento de rostro.

2. CÁMARA

La cámara empleada en esta plataforma es el modelo SNC-RZ25 de la marca Sony la cual se presenta en la Fig. 2.

Esta equipada con una interfaz de red 100Base-Tx\ 10 Base-T y un servidor web interno, permitiendo configurar la unidad en cualquier red TCP\IP. Para obtener las imágenes video será necesario una computadora contectada a la red y un navegador de internet. En la Tab.1 se presentan algunas de las características de la dinámica de los motores de los ejes que proveen el movimiento de la cámara. También las características del video de la cámara

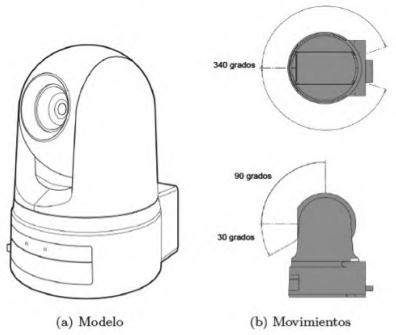

Fig. 2. Cámara SNC-RZ25

Table 1. Características dinámicas de los motores de video SNC-RZ25N

Rango de mov. horizontal total	340°
Posición inicial horizontal	170°
de mov. hacia la derecha	$+170^o$
de mov. hacia la izquierda	-170°
Velocidad máxima horizontal	$100^o/s$
Pasos de vel. horizontal	18
Rango de mov. vertical total	120°
Posición inicial vertical	0°
de mov. hacia arriba	$+90^o$
de mov. hacia abajo	-30°
Vel. máxima vertical	$90^o/s$
Pasos de vel. vertical	14

Table 2. Características del video SNC-RZ25N

Acercamiento Optico	18x
Acercamiento Digital	216x
Disp. de imagen	CCD
N de pixeles efectivos	380000
Dist. focal	f=4.1mm a 73.8mm
Iris	Auto/Manual
Dist. mín. de objetivo	300mm(wide)-800mm(tele)
Interfaz serial	RS-232C
Puertos I/0	$entradasdesensor$ $2salidasdealarma$
Salida de video analógica	BNC composite (75Ω)
Sistema de la señal	NTSC
Resolución	470 lineas horizontales
Relación S/N	Ms de 50db
Masa	1.3 kg
Dimensiones	140×200×148 mm
Voltaje de entrada	12VDC o 24 VAC
Consumo	18 W

se presentan en la Tab.2. La comunicación entre la cámara SNC-RZ25 y la computadora del robot P3DX se realiza a través de dos conexiones TCP/IP para el video y mediante una conexión serial para los comandos de movimiento de la cámara. También se incluye un protocolo de control y comunicaciones transparente, manejable a través de un puerto serial. Este protocolo es llamado ViSCA por sus siglas en ingles (Video System Control Architecture). En este caso la comunicación RS232 tiene las siguientes características:

- Velocidad de comunicación: 9600 bps.
- Bits de inicio: 1.
- Bits de paro: 1.
- Paridad: Sin paridad.
- MSB primero

Las comunicaciones desde el controlador comienzan con un encabezado el cual comprende la dirección del remitente y la dirección del destinatario, seguido del mensaje y finaliza con un terminador. La parte del mensaje consta de dos bytes que indican el modo de las comunicaciones, la categoría del código y algunos parámetros opcionales. La longitud máxima del mensaje es de 14 bytes. El terminador esta predeterminado a FFh y el controlador debe verificar siempre su presencia en las comunicaciones.

3. MODELADO DE LA CÁMARA SNC-RZ25

La cámara SNC-RZ25 se encuentra colocada en un robot móvil Pionner P3DX, todo el sistema se utiliza para reconocer los rostros de las personas que el robot se encuentra en su camino. Donde le procedimiento para realizar esta actividad se puede describir con los siguientes pasos:

(1) Leer imágenes de la cámara.
(2) Detectar si hay rostros dentro de la imagen.
(3) De ser afirmativa la orden anterior, ubica la posición del rostro dentro de la imagen, donde el punto de referencia es el centro de la imagen.
(4) Si el rostro de encuentra en el centro, la diferencia entre la posición actual de este y el centro de la imagen es enviado a un controlador.
(5) El controlador realiza las estimaciones de la velocidad y dirección a la que hay que mover los motores de la cámara con la finalidad de que el rostro quede centrado en la imagen y una vez obtenidos estos, envía comandos a la unidad de manejo de los motores.
(6) Se repite el ciclo.

Para realizar estas tareas es necesario contar con un modelo matemático de la cámara para poder diseñar un controlador para cada uno de los motores de la cámara. Este modelo se concebirá como un sistema en función de entradas y salidas. Primero se propone un modelo del funcionamiento de la cámara en simulink y después se realiza una identificacion de parámetros para obtener el modelo matemático. De acuerdo a las propiedades de la cámara descritas anteriormente es posible considerar los dos ejes de la cámara como subsistemas independientes (ver Fig.3) donde la entrada del sistema horizontal es un número de 0 a 18 el cual corresponde a una velocidad entre $0^o/s$ a $100^o/s$ mientras que para el sistema vertical la entrada varia de 0 a 14 correspondiendo a una velocidad de $0^o/s$ a $90^o/s$. Las salidas de los dos sistemas son la posición horiozontal x y la posición vertical y, respectivamente.

Para determinar las posiciones actuales $x(t)$ y $y(t)$ de la cámara se utiliza la velocidad promedio en una dimensión $v = \frac{\Delta x}{\Delta t}$ quedando de la forma

$$x(t) = 5.50 v_x(\Delta t) + x(t-1) \quad (1)$$

$$y(t) = 6.42 v_y(\Delta t) + y(t-1) \quad (2)$$

donde v_x y v_y son las velocidades, Δt es el incremento del tiempo y $x(t-1)$ y $y(t-1)$ son las posiciones en el incremento del tiempo anterior y las constantes $5.55 = \frac{100^o/s}{18}$ y $6.42 = \frac{90^o/s}{14}$ son valores de escalamiento

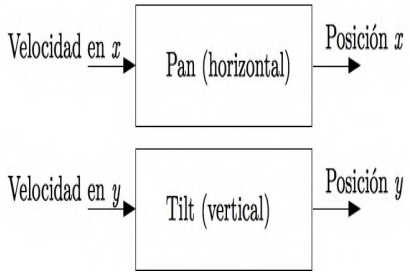

Fig. 3. Cada eje de la cámara se considerará como un sistema independiente, cada subsistema representa el movimiento de un eje (horizontal y vertical)

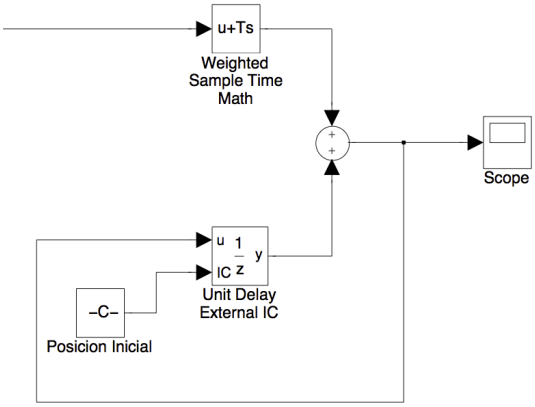

Fig. 4. Bloques de simulink que realizan las ecuaciones 1

de la velocidad. Para implementar las ecuaciones (1) en simulink es necesario un bloque de aritmética de tiempo de muestreo que realiza el producto $u * \Delta t$, un bloque sumea y un elemento de memoria que determine la posición en el instante anterior conectandose de la forma siguiente. Para representar los valores máximos (± 170 y $+90 \rightarrow -30$) de los movimientos horizontal y vertical se le agregan bloques de saturación. También se tiene que limitar la velocidad entre $-100/s \rightarrow +100/s$ y $-90/s \rightarrow +90/s$ para los sistemas PAN y TILT, respectivamente. Otro efecto que se produce en el sistema y tiene que considerarse es la cuantificación de los valores de entrada. El comando de movimiento enviado solo puede modificar la velocidad a través de 18 valores discretos para el sistema PAN y 14 pasos para el sistema TILT. Estos valores se presentan en la Tab.3.

Este efecto es representado por un bloque de cuantificación. Otro efecto que se debe considerar es el retardo en el transporte el cual se estima considerando el retardo generado por la comunicación serial entre la computadora y la cámara, i.e. el tiempo que toma entregar los datos el controlador a la cámara lo cual depende del tamaño del paquete de datos. El mensaje de comando consta de 18 dígitos hexadecimales y el comando ACK de 6 dígitos hexadecimales, entonces el mensaje total será de 24 dígitos hexadecimales o de 12 palabras digitales si por cada palabra se envían 10 bits, la cuenta total será de 120 bits enviados. Además la velocidad es de 9600 kbps por lo tanto el retardo será de 0.013 segundos, el cual se representa

Table 3. Valores de velocidad correspondientes a cada valor discreto.

# de paso	Vel. horizontal	Vel. Vertical
00	$0/s$	$0/s$
01	$5.5/s$	$6.4/s$
02	$11.1/s$	$12.8/s$
03	$16.6/s$	$19.3/s$
04	$22.2/s$	$25.7/s$
05	$27.7/s$	$32.1/s$
06	$33.3/s$	$38.5/s$
07	$38.8/s$	$49.9/s$
08	$44.4/s$	$51.36/s$
09	$50/s$	$57.8/s$
10	$55.5/s$	$64.2/s$
11	$61.1/s$	$70.6/s$
12	$66.6/s$	$77.1/s$
13	$72.2/s$	$83.5/s$
14	$77.7/s$	$90/s$
15	$83.3/s$	$90/s$
16	$88.8/s$	$90/s$
17	$94.4/s$	$90/s$
18	$100/s$	$90/s$

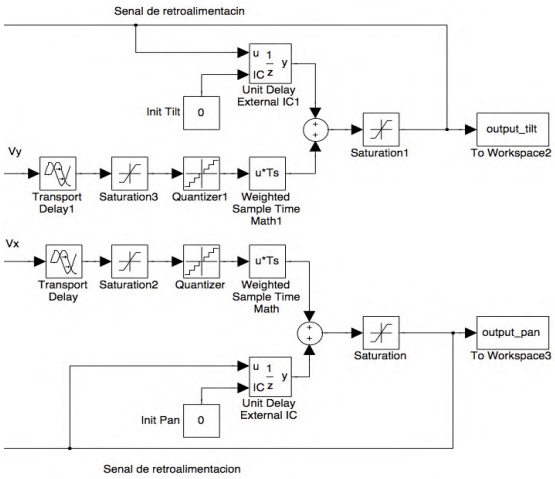

Fig. 5. Diagrama de bloques de la cámara en Simulink

en el modelo con un bloque de retardo de transporte. Para complementar el modelo se le agrega ruido de muy baja potencia el cual es determinado por un bloque de generación de ruido blanco de banda limitada. El modelo en simulink propuesto para la cámara se presenta en la Fig. 5.

Las respuestas del modelo de la Fig.5 a una entrada arbitraria se muestran en la Fig.6. En la Fig.6 la señal de color azul corresponde a la velocidad, se observa que para los valores positivos de esta, la pendiente de las respuestas (señal verde) es positiva y para valores negativos de velocidad, la pendiente es negativa. También se observa que el ruido no afecta la respuesta de los sistemas. Estas señales se utilizaran para la identificación del modelo, como se presenta a continuación.

4. IDENTIFICACIÓN DEL MODELO

En esta sección se describe el procedimiento realizado para determinar el número de coeficientes n_x de la estructura del modelo elegida. Primero se someten los dos modelos a simulaciones con datos de entrada especialmente preparados para obtener un conjunto de datos entrada-salida,

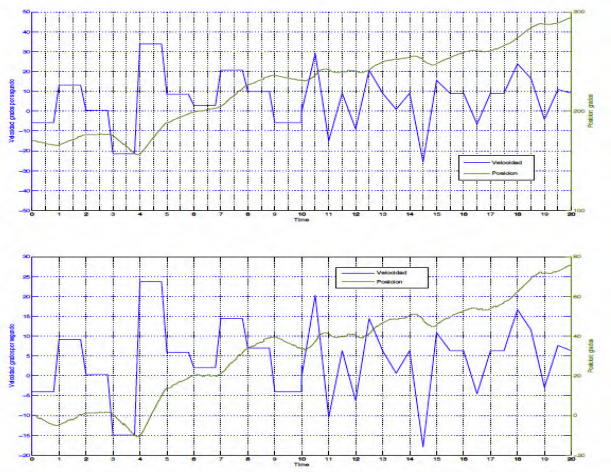

Fig. 6. Respuesta del modelo, figura superior eje horizontal, figura inferior eje vertical.

los cuales se dividaran para las etapas de identificación y validación.

4.1 Preparación de datos de entrada

Para esta etapa es importante seleccionar los datos de entrada (velocidades) que nos permitan identificar el modelo correctamente. En este caso se utilizarán secuencias pseudo-aleatorias binarias (PRBS). Una señal PRBS es periódica, deterministica con propiedades similares al ruido blanco. Esta señal consiste en una sucesión de impulsos rectangulares modulados en su ancho, de forma que tiene un contenido abundante de frecuencias. El análisis de densidad de potencia espectral es un método para observar la potencia de las distintas frecuencias que posee alguna señal. Realizando este análisis al ruido blanco es posible encontrar que debido a sus características aleatorias contiene un rango constante de potencia sobre toda su banda de frecuencias, en este sentido, una señal PRBS tiene una densidad espectral aproximada a la del ruido blanco, con lo cual satisface las necesidades de la teoria de indentificación de sistemas. Entonces las entradas de los modelos en simulink (5) consistirán en una señal PRBS y rampas cuyos valores se encuentran en torno a las velocidades permitidas por la cámara.

Con estas nuevas señales se realizó un análisis espectral donde se compara con una señal PRBS pura y con la densidad espectral del ruido blanco para poder observar los fenomenos descritos previamente. En la Fig.7 se muestran estas señales, la señal preparada (azul) tiene un contenido frecuencial constante a comparación de la señal PRBS pura (roja). Considerando que la frecuencia de muestreo de las simulaciones de los modelos no sera mayor de 20Hz, de acuerdo con el teorema de muestreo, el máximo componente frecuencial estará alrededor de los 10Hz. En la Fig.7 la densidad de potencia espectral en ese rango de frecuencias es casi constante como la del ruido blanco (verde).

4.2 Simulaciones

Una vez determinados los generadores de datos de entrada para excitar los modelos, se obtienen los datos de salida, para realizar esto se implementó el siguiente algoritmo:

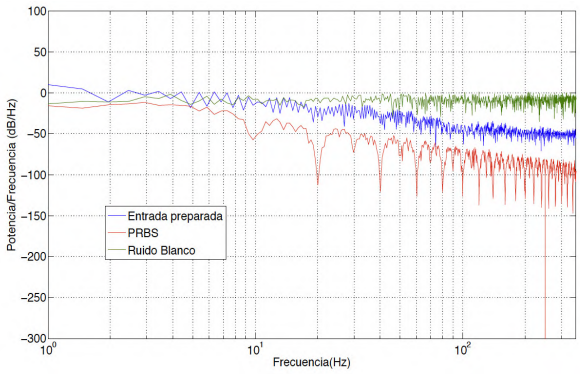

Fig. 7. Densidades espectrales de los datos que se prepararon para la cámara

(1) Contruir los datos de entrada ejecutando el modelo generador de entradas y guarda los datos en el espacio de trabajo.
(2) Simular el modelo PAN y TILT con los datos de entrada del espacio de trabajo, los datos de salida también son colocados en el espacio de trabajo.
(3) Procesar los datos. Aquí se extrae los vectores de datos de las estructuras, se re-muestrea los datos y convierte a columnas.
(4) Guadar los datos en un archivo *.mat.

4.3 Herramienta de identificación de Sistemas

En esta etapa se utiliza la interfaz grfica para importar los datos en el dominio del tiempo desde el espacio de trabajo y construir datos para identificación dentro de un objeto `iddata` necesario para Matlab. Después se eliminó la tendencia lineal de los datos con el comando `detrend`. Se necesitaran 4 estructuras de datos `iddata` para llevar a cabo la identificación. Como el tiempo total de simulación es de 20 segundos con un tiempo de muestreo de 10 ms, se tienen 2001 muestras. De este total de muestras se utilizarán el 75% para la identificación y el 25 % para validación.

4.4 Estructura del modelo y número de coeficientes

El modelo a identificar es un modelo polinómico a partir de los datos. Las estructuras elegidas para este problema son tipo ARX (AutoRegressive eXogenous) y OE (Output Error). La estructura tipo ARX no representa el flujo más natural de señales debido al momento en el que se inserta el ruido, este modelo es la primera opcin de identificación ya que el calculo de los parámetros se realiza por medio de regresión lineal, concretamente por el método de mínimos cuadrados. En el tipo OE se usa el método de minimizacioón de la predicción del error. Ahora el problema es encontrar el mejor número de parámetros de los polinomios en los denominadores y numeradores de los modelos, para ello se realizan los siguientes pasos:

(1) A partir de los objetos de identificación, definir un vector que contiene el numero de parámetros máximos por cada polinomio.

Table 4. Ajuste del modelo ARX para PAN.

n_a	n_b	n_k	% de ajuste
100	100	5	97.89
50	46	5	95.76
50	64	5	95.04
30	25	5	72.92
20	18	5	38.77
15	15	5	21.11
10	7	5	13.96
5	4	5	5.242
3	3	5	3.825
2	2	5	3.263

Table 5. Ajuste del modelo OE para PAN.

n_a	n_b	n_k	% de ajuste
43	7	5	95.18
13	24	5	90.01
41	22	5	89.54
20	10	5	69.62
66	60	5	59.63
1	6	5	50.58
11	6	5	48.68
35	1	5	-39.34
1	1	5	-50.46

(2) Estimar el retardo nk mediante la orden `delayest`.
(3) Verificar todas las combinaciones posibles para el número de parámetros máximos, seleccionando la mejor combinación y almacenando estos datos en una matriz.
(4) Guardar el espacio de trabajo e imprimir las matrices de resultados

En los arreglos resultantes ARX cada columna representa el número máximo de coeficientes que se probaron n_a, n_b y el n_k es fijo. Para los arreglos OE los coeficientes son: n_b, n_f y n_k es fijo. Con estos valores se realiza la estimación de los modelos.

4.5 Estimación de los modelos

Esta etapa se realiza a través de la interfaz gráfica donde se verificaron todas las combinaciones de los coeficientes anteriores obteniendo los resultados para los modelos ARN y OE mostrados en la Tab.4 y Tab.5. En la Tab.4 se observa una tendencia en la estimación de los modelos en conseguir un mejor porcentaje de ajuste. En la Tab.5 se observa una diferencia entre los ajustes de la Tab.4, aqui no se observa una tendencia, el mejor porcentaje de ajuste se consigue con $n_f = 43$ y $n_b = 7$.

En la Fig.8 la referencia es la linea discontinua de color negro la cual muestra la respuesta obtenida con los datos de validación. Se observa que los modelos **arx1001005**, **arx50455** y **arx59645** tienen un ajuste mayor al 90%, los demás son descartados. En la Fig.9 se muestran los errores de los modelos ARX para el sistema PAN. De igual forma se muestran las respuestas y errores de los modelos OE para PAN en las Fig.10 y Fig.11.

En la Fig.10 se observan las respuestas de los modelos con estructura OE a diferente número de coeficientes, los datos de validación se muestran con una linea discontinua, el mejor ajuste lo consigue el modelo **oe4375**, otros modelos con aceptable nivel de ajuste son **oe41225** y **oe13245**. Esto se confirma en la Fig.11 donde se muestran los errores

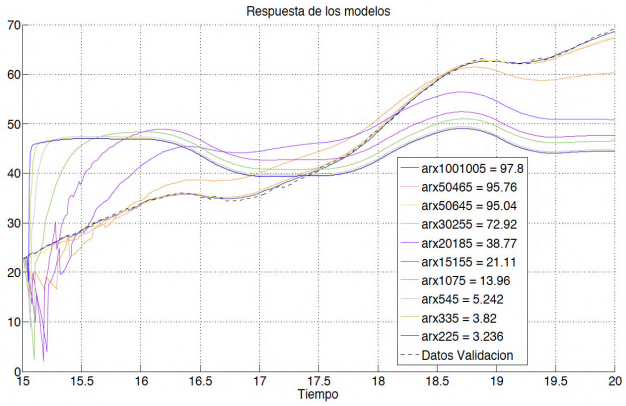

Fig. 8. Respuestas de los modelos ARX del sistema PAN.

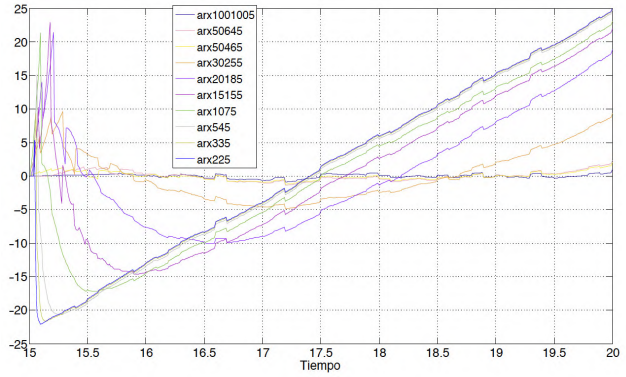

Fig. 9. Errores de los modelos ARX del sistema PAN.

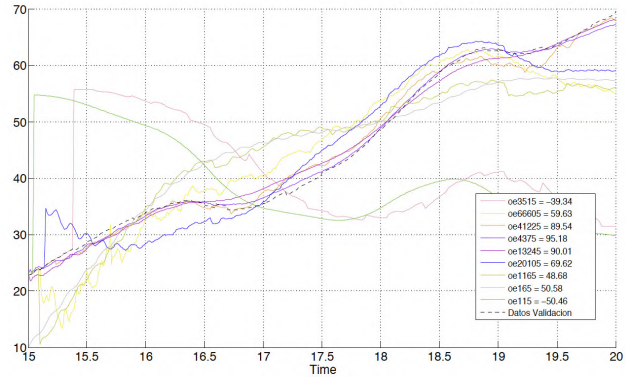

Fig. 10. Respuestas de los modelos OE del sistema PAN.

de los modelos. Este mismo procedimiento se realizo para el sistema TILT. El modelo ARX para el sistema TILT que presentó un mejor ajuste tiene parámetros $a = 100$ y $b = 97$ con un porcentaje del 92.01%. Con respecto al modelo OE es mejor ajuste fue del 93.8% con $k = 75$ y $b = 5$.

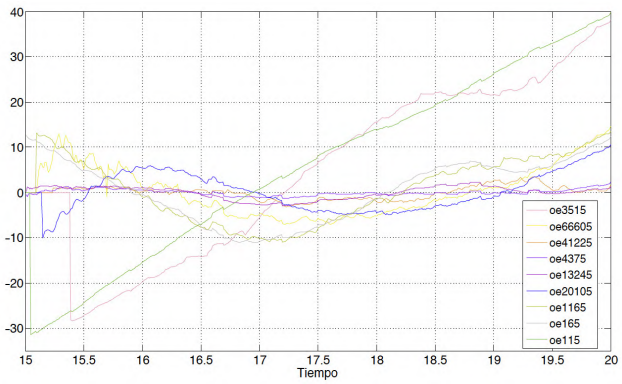

Fig. 11. Errores de los modelos OE del sistema PAN.

Table 6. Modelos Seleccionados.

Sistema	ARX n_a,n_b,n_k	%	OE n_a,n_b,n_k	%
PAN	50 46 5	95.76	4 37 5	95.15
TILT	100 97 9	92.01	75 5 9	93.8

4.6 Seleccin de los modelos finales

Los resultados presentados anteriormente muestran el comportamiento para diferentes parámetros, para elegir los mejores modelos se establecen varios criterios para esta selección:

(1) Seleccionar modelos con un ajuste del 95%. Si no lo hay, elegir el más alto.
(2) Cuando más de dos modelos cumplen con el criterio 1, se selecciona el modelo con el menor número de parámetro n.

Teniendo en cuenta estos criterios se seleccionan los modelos presentados en la Tab.6.

Estos modelos se exportan al espacio de trabajo.

4.7 Cálculo de los modelos finales

Finalmente, utilizando toda la información guardada en el paso anterior se convierten los modelos parámetricos polinomiales en funciones de transferencia utilizando `th2tf`. Como los modelos parámetricos polinomiales son discretos las funciones de transferencia resultantes son discretas (funciones en el dominio z). Las funciones de transferencia obtenidas tienen orden muy alto, posiblemente provocado por una redundancia de parámetros durante la identificación. El orden de estas funciones de transferencia se reduce a través de `balred` , y después de varias pruebas el orden mínimo encontrado fue de 2. Por último, las funciones de transferencia discretas se convierten al dominio continuo, ya que el verdadero muestreo se realiza en la entrada y salida del controlador.

La validación de los modelos reducidos en tiempo continuo se lleva a cabo por medio de una comparación con los modelos originales, sometiendo ambos modelos a una prueba de respuesta al escalón, esto se muestra en la Fig.12.

Como se observa en la Fig.12 la estructura que presenta mejor ajuste para los sistemas PAN y TILT es la ARX, la

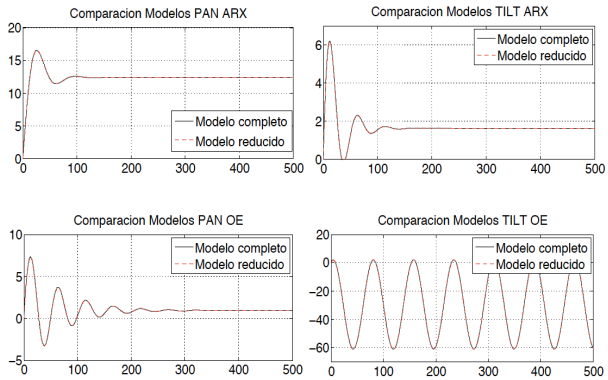

Fig. 12. Comparación de los modelos obtenidos a una respuesta escalón

cual es la estructura elegida. Los modelos finales obtenidos son

$$G(s)_{PAN} = \frac{-0.03398s^2 + 0.9906s + 0.1158}{s^2 + 0.08295s + 0.0009373} \quad (3)$$

$$G(s)_{TILT} = \frac{-0.03499s^2 + 0.9654s + 0.02665}{s^2 + 0.0746s + 0.01649} \quad (4)$$

5. CONCLUSIONES

En este artículo se presentó el modelado de una cámara de video activa en modo de seguimiento. Las principales contribuciones en esta parte son: primero se propone un modelo en simulink simulando el funcionamiento de la cámara en panTilt. Debido a los movimientos de la cámara se dividió el sistema en dos subsistemas: uno para el movimiento en vertical (pan) y otro para el horizontal (Tilt). Además se propusieron dos estructuras de identificación a partir de datos (ARX y OE) del modelo en simulink para estimar modelos polinomiales, obtuviendo mejores resultados con la estructura ARX en ambos subsistemas. Como las funciones de transferencia resultantes fueron de orden muy alto se realizó una reducción a funciones de transferencia de segundo orden. Se validaron estas funciones de transferencia con el modelo del funcionamiento en simulink de la cámara con la respuesta escalón, comprobando que estas funciones de transferencia representan la dinámica de la camara de video en panTilt.

REFERENCES

A. Marín Hernández. Visión Dinámica para la navegación de robots móviles. Tesis Doctoral. *Institut National Polytechnique de Toulouse*, 2004.

L. Ljung. System Identification: Theory for the User. *Prentice Hall, 1998.*

R. Ferreyra. Estudio comparativo entre Control PID y Difuso. *SOMI XV, Congreso Nacional de Instrumentación*, 2000.

R.H. Rahmat. Performance comparision between PID and fuzzy logic controller in position control system od DC servomotor. *Journal Teknologi*, 45, 2006.

J. Godjevac. Comparison between PID and fuzzy control. *Internal Report:Ecole Polytechnique Federal de Lausanne*, R93.36i, 1993.