DISEÑO DE MÓDULO DE VISION PARA CELDA DE MANUFACTURA FLEXIBLE.

Luis Alberto López Estrada [1], Guillermo Ramírez Prado [2]

[1] Alumno de Posgrado, [2] Profesor Investigador, Universidad Autónoma de Aguascalientes

Contacto: imt.luis.lopez@gmail.com

RESUMEN

La visión artificial es un elemento fundamental para el control de calidad en los sistemas de manufactura actuales, en el presente artículo se presenta los fundamentos teóricos, la aplicación y el diseño de un módulo de visión para la universidad tecnológica de Aguascalientes el cual se tiene planeado implementar en la actualización de sus equipos e instalaciones. Por medio del uso de la plataforma de Matlab [1] se diseña una interfaz de usuario que sea capaz de identificar formas (clasificación de elementos) y se contrasta la velocidad del módulo con lo que se espera en sistemas industriales de similar aplicación.

Palabras clave: Diseño, Visión Artificial, Manufactura Flexible,

1.-INTRODUCCIÓN

Para poder mantener el paso con las tendencias tecnológicas la Universidad Tecnológica de Aguascalientes se ha dado a la tarea de actualizar los distintos laboratorios con los que cuenta para de esta manera capacitar de una mejor manera a los alumnos. Se ha planeado la ampliación de la celda de manufactura actual, dentro de esta ampliación se plantea la necesidad de contar con un módulo de visión. El objetivo es desarrollar un sistema capaz de realizar inspecciones visuales cuyo funcionamiento sea continuo, usando la plataforma de Matlab [1] como sistema de procesamiento de imágenes. El cual debe de contar con una interfaz gráfica de usuario y comunicación serial por medio del estándar USB.

2.-FUNDAMENTOS

Para poder definir el desempeño que se busca en el módulo de visión es necesario como primer paso definir donde será aplicado, en este caso será en una celda de manufactura flexible, con esto en mente el diseño se enfoca en funciones de control de calidad de media a baja velocidad [2].

El objetivo de un sistema de inspección por visión artificial en la industria suele ser el control de calidad de una pieza, lo que puede abarcar sus dimensiones, números de serie (clasificación), la presencia de componentes defectuosos, etc. En una celda de manufactura flexible su función es la de realizar inspecciones en procesos donde existe

diversidad de piezas con cambios frecuentes de producción. Para poder integrar un sistema de visión es necesario cumplir con una serie de prerrequisitos básicos, los cuales incluyen una combinación de hardware y software, que en su conjunto realizaran una función según la meta que se plantee.

La visión artificial según Ballard y Brown [3] es un pre-requisito para poder realizar una interpretación, manipulación y apreciación de los objetos físicos que se están analizando. Los elementos de un sistema de visión artificial en manufactura incluyen la adquisición de la imagen, el procesamiento de las imágenes, el despliegue de la imagen para el usuario, y la comunicación con un actuador.

González, R. y Woods [4] modelaron una serie de pasos fundamentales para el procesamiento de imágenes digitales. (Fig.2)

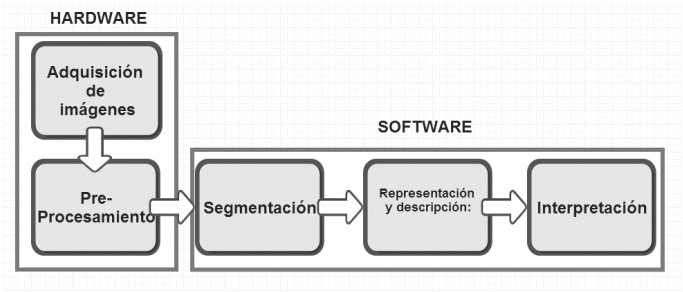


Figura 2. Diagrama de los pasos para el procesamiento de imágenes

Este modelo sirve como referente de guía de los pasos que debe de tomar el módulo de visión para poder lograr su función.

La adquisición y el pre-procesamiento de las imágenes se logra por medio de Hardware, donde la electrónica integrada de la cámara ayuda a corregir posibles problemas de iluminación o de geometría, M. Sonka, V. y R. Hlavac Boyle [5] recomiendan tener algún tipo de pre-procesamiento de imágenes para garantizar la obtención de características clave de los objetos analizados.

Chazelle et al [6] destaca tres características básicas en los pixeles que forman la imagen para su segmentación, los cuales son: La similitud de pixeles, la conectividad de los pixeles y la discontinuidad de contornos.
Con estas características en mente, se puede decir que los principios fundamentales para decidir cuales segmentos de una imagen son importantes o no, es por discontinuidad o similitud de características. Esto es de gran ayuda para la selección del mejor algoritmo para nuestro módulo de visión.

K.S. González et al [7], en su estudio definen que la descripción de un sistema de visión consiste en representar las regiones en términos de sus características externas (frontera) o bien, representarla en términos de sus características internas (los píxeles que componen a la región). Una representación externa se elige cuando el enfoque principal está en las características de la pieza. Una representación interna es seleccionada cuando interesan las propiedades de los píxeles. Para nuestra aplicación nos interesa una representación de características externas de las imágenes.
Para poder lograr el objetivo del módulo de visión es necesario implementar los pasos anteriores en una plataforma computacional (SOFTWARE), por lo que es necesario usar un algoritmo que encapsule todos los pasos del procesamiento de imágenes.

Actualmente los sistemas computacionales se han hecho más robustos y accesibles, la tecnología de teléfonos celulares y computadoras personales de última generación han puesto en nuestras manos este poder que usamos de manera cotidiana, inclusive los equipos de telefonía celular son capaces de procesar con una velocidad muy aceptable el reconocimiento de patrones, y aplicaciones básicas de reconocimiento facial.

El reconocimiento facial es un área de aplicación del procesamiento de imágenes que usa algoritmos que encapsulan la segmentación, presentación e interpretación de imágenes por medio de rutinas computacionales, y que en los últimos años han avanzado de manera excepcional. Estos algoritmos no se limitan al uso exclusivo de reconocimiento facial, pues en sí, son algoritmos de reconocimiento de formas, de los cuales sus principios son de utilidad para cualquier tipo de sistema de visión artificial.

El modelo de apariencia activa (Active Appearance Model, AAM) [8], permite reproducir de forma sintética imágenes de superficies que incluyen deformaciones no rígidas y cambios de apariencia. Están basados en la obtención, mediante una fase de entrenamiento, de un modelo estadístico de la forma y la apariencia del objeto de interés.

Paul Viola y Michael Jones [9] diseñaron un algoritmo de detección de formas en el cual se han basado la mayoría de las diferentes técnicas de reconocimiento facial. Consta primordialmente de dos elementos claves, el primero es que basa su análisis en Imágenes integrales, lo que hace posible una rápida evaluación de características clave, pues no toman en cuenta la intensidad de la imagen (ayuda a corregir problemas en iluminación), y el segundo elemento es un clasificador en cascada. La detección del objeto es la combinación en cascada de grupos de clasificadores cada vez más complejos (Fig. 4). Esta

estrategia incrementa de forma considerable la velocidad de ejecución del algoritmo, ya que permite detectar rápidamente las zonas de la imagen que no contienen la forma a identificar y dedicar mayor esfuerzo a aquéllas en que hay más probabilidad de existencia de imágenes positivas en las primeras etapas.

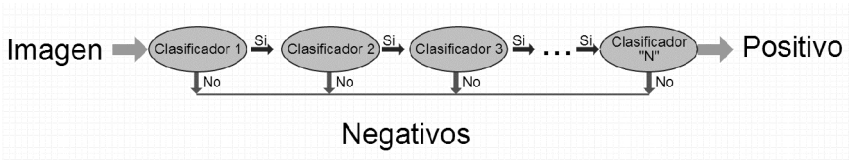


Figura 4. Representación esquemática del algoritmo Viola-Jones.

Por último se agrupan las características extraídas que nos devuelve la etapa anterior según la zona de la imagen y se van clasificando en cascada para poder seleccionar aquellas zonas de la imagen más prometedoras e ir eliminando aquellas que no tienen características relevantes. Para poder extraer las características que servirán de patrón en el clasificador, se debe realizar un entrenamiento con un conjunto de imágenes adecuado, tanto con imágenes de los objetos a detectar como con varias imágenes que contengan o no los objetos a identificar, esto asegura el correcto funcionamiento del algoritmo.

Nuestra plataforma de análisis de imágenes cuenta con un toolbox que nos ayuda a integrar la parte del clasificador con el algoritmo de Viola-Jones.

3.- PROCEDIMIENTO.

Para comenzar con el diseño del módulo de visión debemos de partir de la función a la que será sometido, en este caso será la celda de manufactura flexible de la universidad , por lo que es prioridad la flexibilidad y adaptabilidad del módulo a los equipos existentes, además que estará sometido a velocidades bajas de producción. Tomando en cuenta estos elementos, se ha creado un diagrama esquemático del diseño del sistema.

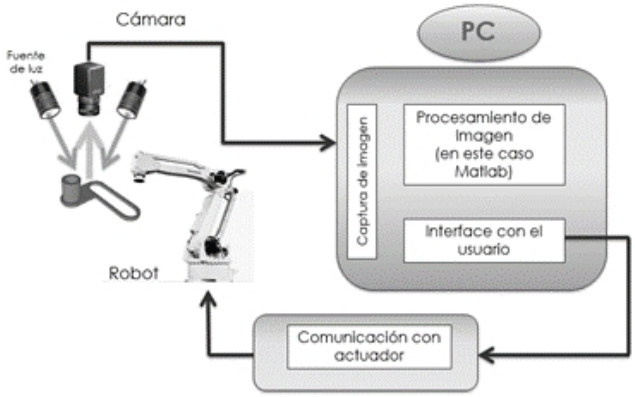


Figura 5. Esquema de los elementos del sistema de visión.

Para la selección de la cámara que hará la adquisición de imágenes para el módulo de visión se busca obtener imágenes en los formatos estándares de la industria, los cuales incluyen: Camera Link, GigE Vision, DirectShow compliant WDM (básico para sistemas basados en Windows) y UVC standard interface. Debido a que se usara una plataforma de computadora personal, la transmisión de los datos de la cámara debe de ser por medio de los estándares de USB (2.0 o 3.0).

Se han seleccionado dos cámaras, ambas tienen una resolución comparable, una con USB 2.0 y otra con USB 3.0, con la finalidad de comparar la velocidad de captura de ambas. Con la captura de imágenes como fuente de información, se programa por medio de la librería de herramientas de procesamiento de imágenes de Matlab un programa para la identificación de objetos por medio del algoritmo de Viola-Jones, el diagrama de flujo (Fig.6) muestra la secuencia lógica del programa.

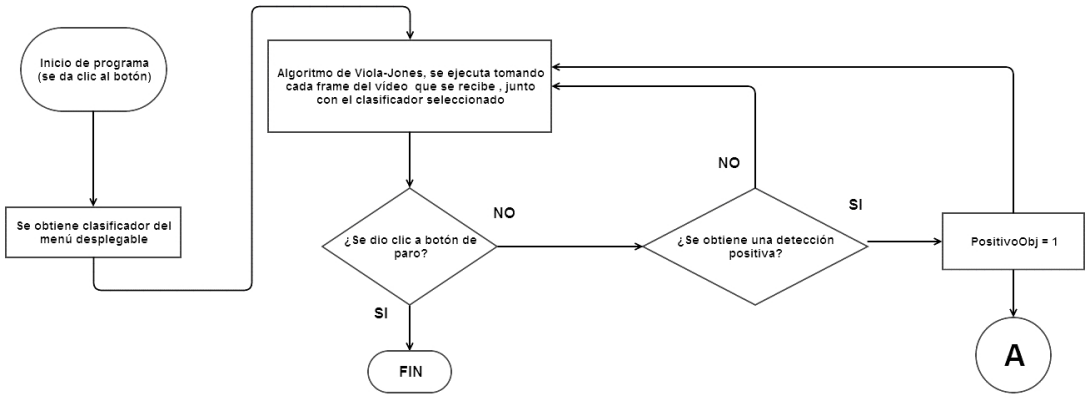


Figura 6. Diagrama de flujo del programa del módulo de visión.

El programa espera a que se oprima el botón de inicio, por default siempre está cargado un clasificador (entrenado previamente), con estos dos requisitos empieza a correr el algoritmo de reconocimiento, leyendo cada cuadro de la cámara y pasándola por los clasificadores en cascada. Al terminar cada cuadro revisa si no se ha dado clic en el botón de paro, si no se registra el cambio el programa sigue, de lo contrario lo detiene. La siguiente condicional revisa si el algoritmo arrojo alguna detección positiva, si la comparación es negativa se regresa al algoritmo para procesar el siguiente cuadro. Si el objeto es detectado se manda un valor booleano "1" a una variable la cual tiene una conexión interna con un subprograma para la comunicación USB, al terminar de hacer el cambio de valor se regresa al algoritmo para procesar el siguiente cuadro. La siguiente figura muestra el diagrama de flujo de la comunicación del actuador por medio del puerto USB. (Fig.7).

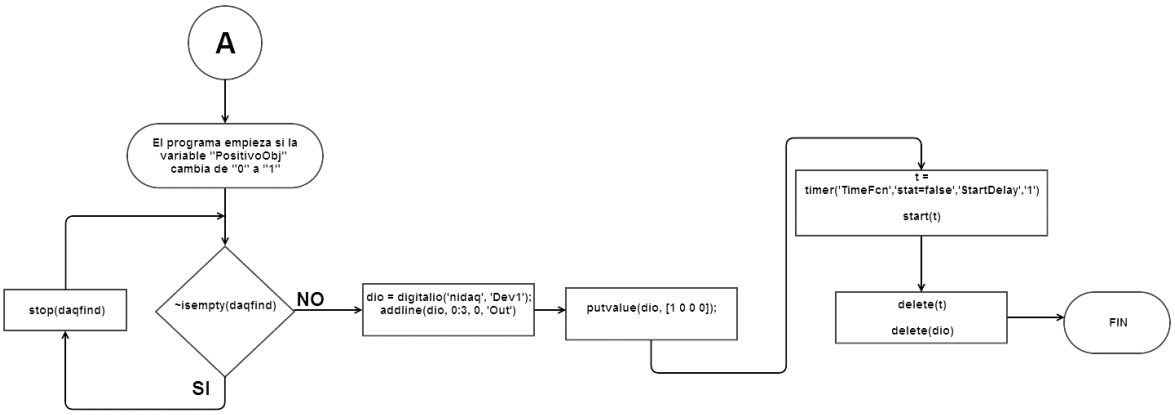
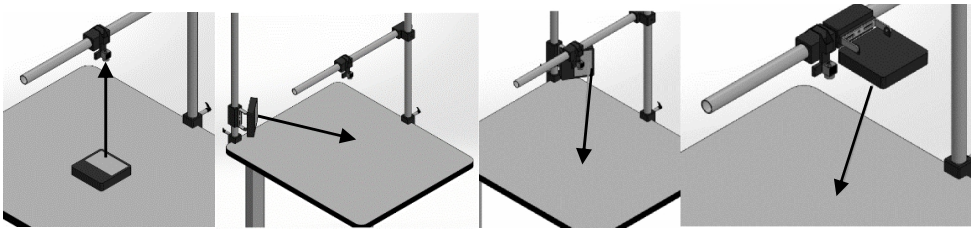


Figura 7. Diagrama de flujo del programa de comunicación USB

Este subprograma recibe un valor booleano de "A" (el cual es el programa principal) lo cual indica que se identificó con éxito un objeto , esto inicializa el programa , primero se checa que no esté funcionando el dispositivo de adquisición de datos, la condición es que "sea distinto a vacío" (si no está funcionando esta "vacío"), después se crea la variable que habilita la comunicación con la tarjeta de adquisición de datos y paso seguido se manda un bit a dicha variable, esta señal será interpretada por el actuador para accionarse. El siguiente comando es un temporizador, el cual mantiene la señal lógica por un segundo, esta manera nos aseguramos que el actuador registre la identificación del objeto, al final se cierra la comunicación y se le da un valor de "0" a la variable interna del programa principal ("A").

Para poder implementar el sistema de visión en la celda de manufactura se diseña un sistema mecánico tubular que alojaran todos los elementos que lo conforman, se diseña

tomando en cuenta la modularidad de la celda. Debe de ser capaz de ser configurable para las distintas posiciones de la fuente de iluminación. (Fig. 8)



A) **B)** **C)**

Figura 8. Configuraciones de elementos según el tipo de iluminación, Iluminación tipo back (A), Iluminación tipo direccional (B), Iluminación tipo axial (C).

3.-RESULTADOS

Para poder validar el la velocidad de procesamiento del módulo de visión es necesario tomar el tiempo que tarda en reconocer los objetos deseados y contrastarlos con equipos del mercado en situaciones similares. Encontrar un parámetro de tiempo de reconocimiento de objetos no es fácil pues en la industria cada equipo se diseña y configura para una tarea en específico, lo que engloba muchas variables y situaciones. J. Enciso et al [10] en su investigación de visión activa de fisuras obtuvo procesamiento final a una resolución de 256 pixeles de 634 milisegundos. En otra investigación elaborada por J. Giraldo et al [11] por medio también de un chip dedicado a procesamiento de imágenes aplicado a un sistema de visión para robótica, con una resolución de 1024x1280 pixeles obtuvo una detección de caracteres deseados de entre 100 a 600 milisegundos.

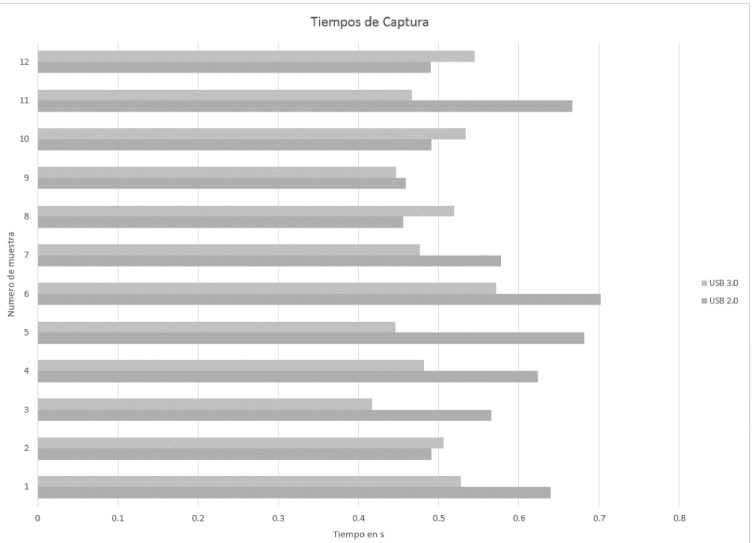
Considerando la función del módulo de visión en la celda de manufactura flexible y la información antes presentada, se considera que una velocidad de reconocimiento de objetos de entre 300 a 600 milisegundos es adecuada y comparable con los estudios antes presentados.

Para recopilar la información de cuánto tiempo le toma al sistema detectar un objeto se toma un video en el cual se quitan y ponen engranes en el módulo cada 5 segundos, de esta manera tenemos una fuente de video calibrada y controlada de sucesos que se deben detectar. Para empezar a analizar la información que se obtiene, se registran los tiempos en la tabla 1, los cuales arrojaron lo siguiente.

Tabla 1.- Resultados de tiempos de captura

USB 2.0	
Tiempo de Colocación de Engranes (s)	*Tiempo de Captura (s)*
5	0.639366968
10	0.49037119
15	0.565642087
20	0.623735011
25	0.681192143
30	0.701691148
35	0.577714308
40	0.455878479
45	0.458653949
50	0.490967872
55	0.666192277
60	0.490241228

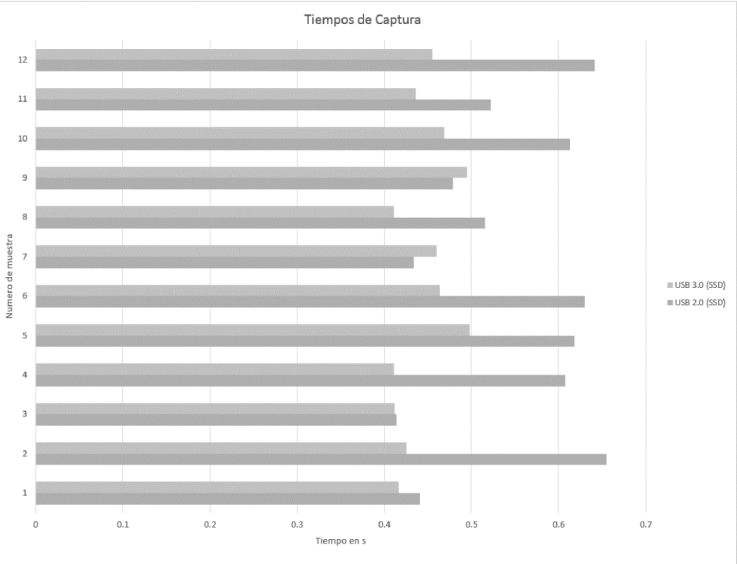
USB 3.0	
Tiempo de Colocación de Engranes (s)	*Tiempo de Captura (s)*
5	0.527931697
10	0.505561461
15	0.416876306
20	0.481384695
25	0.446482324
30	0.571475708
35	0.476186754
40	0.519555954
45	0.44719716
50	0.533949116
55	0.466468728
60	0.544594845



Con Ambas cámaras se obtienes resultados similares, la cámara de USB 2.0 está en un rango de 500 a 700 milisegundos aproximadamente, mientras que la de USB 3.0 tiene un rango de 400 a 500 milisegundos aproximadamente, pero, si se aprecia una ligera ventaja con la cámara USB 3.0, esto se puede interpretar que la resolución en este caso, no es factor del porque la cámara antes mencionada tiene una ventaja, pues ambas tienen una resolución muy cercana. Entonces el cuello de botella de información no lo dan ni la resolución de la cámara ni los cuadros por segundos que manejan, sino que tan rápido el programa puede tener acceso a esa información, y el standard USB 3.0 es más rápido para transmitir información que el USB 2.0. Al tener una arquitectura basada en PC el disco duro es un elemento clave para poder tener acceso a la información de captura de imágenes y a su procesamiento, los datos anteriores se tomaron con un disco duro convencional. Para contrastar resultados y mejorar el desempeño del módulo se vuelven a tomar lecturas de tiempo de detección , pero esta vez usando un disco duro de estado sólido , el cual elimina el cuello de botella de transmisión de datos entre la cámara y la escritura de datos en el programa. En la tabla siguiente se muestran los datos recopilados con el disco duro de estado sólido.

Tabla 2.- Resultados de tiempos de captura con un SSD

USB 2.0 (SSD)	
Tiempo de Colocación de Engranes (s)	*Tiempo de Captura (s)*
5	0.440717778
10	0.65523417
15	0.414342844
20	0.607453788
25	0.618003228
30	0.62968215
35	0.434188567
40	0.515162526
45	0.478718916
50	0.613295314
55	0.52198912
60	0.640904172
USB 3.0 (SSD)	
Tiempo de Colocación de Engranes (s)	*Tiempo de Captura (s)*
5	0.416239024
10	0.425425113
15	0.411713233
20	0.410688135
25	0.49813309
30	0.46379258
35	0.459744803
40	0.41097604
45	0.495073361
50	0.468869514
55	0.435877917
60	0.455608097



La ganancia en el tiempo de reconocimiento es poca, pero si la hay, la cámara de USB 2.0 tiene un rango de 400 a 600 milisegundos, mientras que la de USB 3.0 está entre 400 y 450 milisegundos aproximadamente. Con el disco duro de estado sólido se ha logrado la mayor velocidad de captura y procesamiento de imágenes con los elementos que se han seleccionado para la construcción del módulo de visión.

3.-CONCLUSIONES.

El módulo es capaz de procesar de manera eficaz y es de uso amigable. No solo es capaz de reconocer las formas, sino compensar problemas en las imágenes capturadas, además de que el equipo de captura no necesita ser especializado, siempre y cuando cumpla con los formatos y estándares de la industria, lo que le da una flexibilidad muy amplia al módulo. Además con un proyecto de esta naturaleza el beneficio financiero es evidente pues la inversión es mínima y se obtienen resultados que pueden competir con equipos industriales que son diseñados bajo pedido a un costo elevado. Pero la importancia de este proyecto no solo cae en cumplir con las metas que aquí se plantearon, van más allá, pues el verdadero objetivo de este proyecto es proveer a la comunidad estudiantil con una herramienta más para su formación.

3.-REFERENCIAS BIBLIOGRÁFICAS.

[1] Matlab R2010b. Ver.7.11.0.584. Licencia 648755. Universidad Tecnológica de Aguascalientes. Agosto 16, 2010.

[2] Serope, Kalpakjian. Steven, R. Schmid. Manufacturing Engineering and Technology. United States of America: Prentice Hall, 2000. 1086-1087p.

[3]D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, N.J., 1982. 1-12p.

[4] González, R. y Woods, R., (2002) Digital image processing. Prentice Hall. 25-26p.

[5] M. Sonka, V. y R. Hlavac Boyle, Image Processing, Analysis, and Machine Vision, Cole Publishing Company, 1999. 57-58p.

[6] Chazelle et al. Application challenges to computational geometry: CG impact task force report. Technical Report, Princeton University, 1994. 13p.

[7] Fu, K.S. González, R.C. y Lee, S.G., Robótica, control, detección visión e inteligencia. Mc Graw - Hill. 1998. 165-159p.

[8] Cootes, T. Edwards, G. y Taylor, C. Active appearance models, Computer Vision-ECCV'98, 1998. 484 p.

[9] Viola, P. y Jones, M. Rapid object detection using a boosted cascade of simple features. CVPR, 2001. 2-6p.

[10] Enciso, J., García-Alegre, Mª.C., Guinea, D. Visión Activa en Detección de Fisuras Superficiales con Arquitecturas DSP Especializadas. AIDO'96. Foro Iberoamericano de Tratamiento Digital de Imágenes: Visión Industrial e Imagen Médica. Valencia. Octubre. 1996. p.9-10.

[11] Giraldo, Juan J., Estrada, Eliud, Pineda, Diego F., López, Alexander. Diseño e implementación de un sistema en un solo chip para la navegación y reconocimiento de señales de tránsito en un sistema robótico móvil Ingeniería y Competitividad. Ingeniería y Competitividad, 2012, Volumen 14, No. 2,19 – 21p.