

**CENTRO DE INVESTIGACION Y DE
ESTUDIOS AVANZADOS DEL IPN**

UNIDAD ZACATENCO

DOCTORADO EN COMPUTACIÓN

**“SEGUIMIENTO DE OBJETOS EN SUPERFICIES
IRREGULARES SUAVES MODELADAS CON
INFORMACION DE SU APARIENCIA”**

Protocolo para sustentar solicitud de admisión

Presenta: M. en C. Farid García Lamont

Asesor

Dr. José Matías Alvarado Mentado

Julio de 2006

INDICE

RESUMEN	2
INTRODUCCION	3
RETOS DE LA TESIS	12
OBJETIVOS DE LA TESIS	14
RELEVANCIA	15
PROPUESTA	16
METODOLOGIA	17
POSIBLES RESULTADOS	18
PLAN DE TRABAJO	19
REFERENCIAS	20

RESUMEN

Los robots han adquirido más libertad y mayor capacidad de procesar información lo que en consecuencia les ha permitido ser más autónomos e "inteligentes". Los robots jugadores de fútbol del Dr. Raúl Rojas (R3), por ejemplo, lo demuestran. Sin embargo estos mismos robots tienen limitaciones, quizás la más importante sea su incapacidad de desplazamiento en superficies que no son lisas, es decir, están diseñados para funcionar en superficies sin desniveles e irregularidades, comunes en un ambiente real de juego de fútbol y, en general el común de superficies de desplazamiento. Se le debe enseñar al robot, por medio de visión artificial y redes neuronales (RN), a distinguir y librar (no caer debido a) estas irregularidades en el terreno. Es aquí donde es aplicable un "novedoso" e ingenioso método de reconocimiento de imágenes basado en apariencias el cual puede ser de mucha utilidad para ayudar al robot a reconocer en sus detalles el entorno en donde se ubica y de tal manera que pueda desplazarse óptimamente y sin sufrir accidentes. Combinar las ventajas que ofrece cada uno de los enfoques comentados y lograr robots con capacidad de desplazarse fluidamente en superficies irregulares es la problemática en que se propone circunscribir mi investigación doctoral.

INTRODUCCION

En la actualidad los avances tecnológicos son rápidos y constantes, i.e., lo que el día de hoy es un producto de vanguardia, en seis meses muy probablemente será desplazado por uno nuevo y mejor. La tendencia de la tecnología es la creación de productos cada vez más baratos, eficientes y también que sean autónomos.

Al mencionar autónomos nos referimos a que tales productos funcionen con un mínimo o nula intervención .del ser humano. Por ejemplo, los robots en la industria tienen un papel muy importante ya que permiten reducir tiempo y costos en la producción además de realizar actividades que pueden ser peligrosas para una persona.

En las últimas décadas se ha trabajado en este aspecto estudiando la Inteligencia Artificial donde se han desarrollado líneas de investigación como visión, control difuso, redes neuronales, seguimiento, etc.

Pero ¿en dónde se ha aplicado la Inteligencia Artificial?, ¿cuáles han sido los avances logrados? Cada año, desde 1997, se lleva a cabo el torneo mundial de jugadores robóticos de fútbol mejor conocido como RoboCup. En dicho torneo diferentes equipos muestran sus robots donde todos tienen en común la misma herramienta para el control de los robots, tal herramienta es la Inteligencia Artificial.

Redes neuronales, visión, seguimiento y programación en tiempo real son algunas de las ramas de la Inteligencia Artificial que se han utilizado para dar vida a los equipos robotizados de fútbol. Cabe mencionar que en el caso específico de los equipos robotizados de fútbol se involucra la teoría del control automático. El objetivo del torneo es desarrollar nuevas tecnologías y conocimientos que sean útiles no solo en el área de entretenimiento sino también, por ejemplo, en el campo de la industria o la medicina.

La visión artificial ha avanzado considerablemente, actualmente ha ganado terreno el reconocimiento de objetos utilizando modelos basado en apariencias. Anteriormente el

enfoque geométrico era el mas común pero tal enfoque podía omitir información necesaria ya que el modelado extrae solo los contornos y los detalles como letras escritas sobre el objeto no los podía distinguir, mientras que el enfoque basado en apariencias puede obtener esa información.

Desplazamiento y seguimiento en robots móviles

Diversos trabajos han tratado sobre Inteligencia Artificial aplicada a visión y desplazamiento de robots [3], [10], [11] pero en especial existe uno que llama la atención debido a los buenos resultados obtenidos, es el caso de los robots jugadores de fútbol del Dr. Raúl Rojas y su equipo de investigación de la Universidad Libre de Berlín que han obtenido el primer lugar en los años 2004 y 2005 (categoría small-size), segundo lugar en los años 1999, 2000, 2002 y 2005 (en 2005 en la categoría middle-size) y tercer lugar en 2003 [11].

Descrito a grades rasgos, para la visión de los robots pequeños (small-size) se tiene una cámara de video digital colocada sobre la cancha de fútbol, a unos 4 - 5 m. de altura, mediante la cual se especifican las dimensiones del campo y la ubicación de los robots. Toda esta información es enviada a una computadora central que procesa esa información. Dicha computadora tiene programados los algoritmos que le permiten tomar decisiones y armar estrategias que deben seguir los robots. Los robots tienen comunicación con la computadora de manera inalámbrica. En la figura 1 muestra una perspectiva del campo de juego.

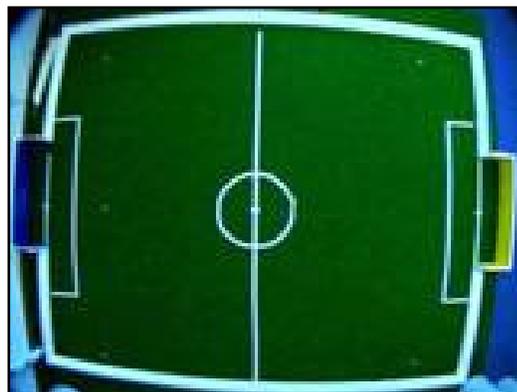


Figura 1

Los robots funcionan con un programa que han desarrollado llamado “Neural Coach”. Este programa le enseña al robot a jugar igual que un entrenador le enseña a jugar a un humano, utilizando diagramas estáticos de lo que constituye una buena y una mala jugada (movimientos).

Primeramente para entrenar al robot, se capturan diversas situaciones posibles de juego. La información relevante de cada ejemplo es guardada en un archivo XML. En las situaciones de juego se consideran tanto la velocidad del robot como del balón igual a cero. Sin embargo la misma aproximación general puede ser aplicada a situaciones dinámicas. A cada situación almacenada se le asigna información que indica si es conveniente dar un pase o no. A estas situaciones le es asignado un numero real $-1 \leq x \leq 1$, donde se toma la convención que cuando $x \leq 0$ es mas conveniente esquivar o hacer un “dribbling”, y cuando $x > 0$ es mejor pasar el balón. La figura 2 muestra (utilizando el Neural Coach) la asignación de un valor a una situación de juego en específico.

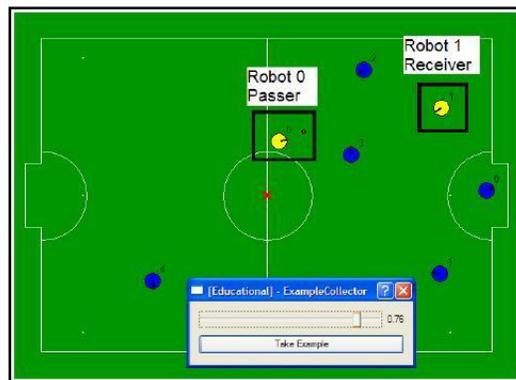


Figura 2

El campo es codificado utilizando un vector de parámetros, por ejemplo, un parámetro es el espacio libre alrededor de un robot receptor. Este parámetro es muy relevante porque influye en cuanto a si es correcto dar un pase que puede ser interceptado. Al codificar estas situaciones como un vector, se pierde información debido a que no se puede recuperar todas las posiciones de los robots pero se obtiene una perspectiva más abstracta del campo de juego. Al jugador que tiene el balón se le llama “pasador” y al que recibe “receptor”.

Los parámetros utilizados se dividen en dos, aquellos que intervienen con el dribbling, con el balón y aquellos con la recepción del balón. Los parámetros asociados con el dribbling son:

1. ***Libertad de dribbling:*** Espacio disponible para dribbling con el balón. El parámetro refleja el tiempo que tomaría al oponente más cercano para interferir el dribbling del robot.
2. ***Angulo de dribbling:*** Angulo definido por la posición del pasador, la mitad de la línea de gol del oponente y la esquina más cercana del campo. Por ejemplo un robot en una esquina del oponente tiene un ángulo de dribbling de 0° .
3. ***Distancia del dribbling a la meta:*** Distancia desde el pasador a la línea de gol.

Los siguientes incisos corresponden a los parámetros asociados a la realización de pases:

4. ***Espacio disponible:*** Espacio entre el receptor y el oponente mas cercano.
5. ***Angulo de pase:*** Es la misma característica que en el ángulo de dribbling pero ahora para el receptor.
6. ***Distancia mínima tangencial:*** Distancia más corta que un oponente tiene que recorrer para bloquear un pase.
7. ***Tiempo de espera durante el pase:***
8. ***Libertad de dribbling para el receptor después de un pase:*** Es el mismo parámetro de “libertad de dribbling”, pero ahora es para el robot receptor.

El receptor es desplazado sobre una malla de 17×21 puntos como se ve en la figura 3.

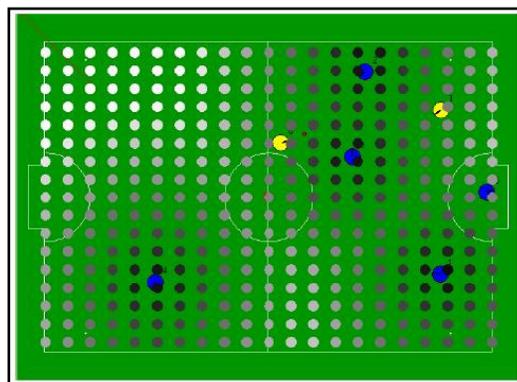


Figura 3

En la figura anterior, por medio de la malla, muestra el espacio disponible de la cancha, los puntos mas claros indican el espacio libre de oponentes mientras que los puntos oscuros indican lo contrario.

Los ejemplos son agrupados en varias categorías para tener una visión general de ellos y activar o desactivar grupos particulares de ejemplos. Al activar y desactivar estos grupos, pueden entrenarse diversas redes neuronales para codificar diferentes comportamientos. Por ejemplo, si un oponente tiene la habilidad de bloquear pases largos, se puede desactivar los ejemplos donde los pases largos son considerados como jugadas buenas.

La red neuronal tiene ocho nodos de entrada (parámetros enlistados anteriormente) y una capa “escondida” cuya dimensión puede ser arbitraria, para este caso utiliza catorce nodos “escondidos” (ver figura 4). El nodo de salida emite un valor que indica la acción que debe tomarse. Un valor en el intervalo $(-\infty, 0]$ es interpretado como “no dar pase”, mientras que un valor en el intervalo $(0, \infty)$ significa “realizar el pase”.

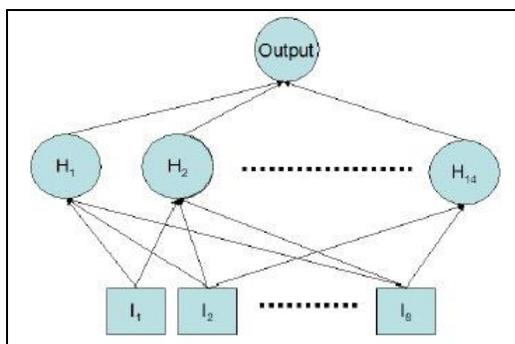


Figura 4

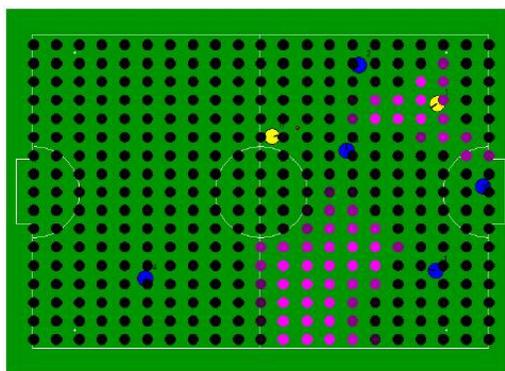


Figura 5

En la figura 5 muestra el resultado o salida de la red neuronal (neural coach). Si un compañero se coloca en alguna de las dos áreas de color rosa entonces el jugador de la izquierda que lleva el balón debe dar el pase.

La posición del balón es información que no solo los robots deben conocer sino también que deben predecir. Para predecir la posición del balón utiliza el filtro de Kalman el cual permite calcular estados futuros. La principal aplicación del filtro de Kalman en la visión del robot es el seguimiento de un objeto, en este caso es el balón. Es necesario conocer la posición y velocidad del objeto en cada instante. Una secuencia de imágenes es capturada por una cámara, después un método de procesamiento de imágenes segmenta el objeto y luego calcula su posición en la imagen, esta información se toma como el estado del sistema en cierto instante. Considerando tal información se puede utilizar en el filtro de Kalman.

Modelado de objetos basado en apariencias

Por otra parte han existido avances significativos en visión artificial no solo con el Dr. Raúl Rojas y su equipo. Actualmente ha ido ganando terreno el modelado de objetos basados en apariencias. En su tesis doctoral el Dr. Luis Altamirano Robles [2] propone una metodología que consiste en obtener modelos de objetos con un mínimo de imágenes el cual es muy útil para nuestros propósitos ya que nos posibilita obtener modelos robustos de objetos y superficies, en sus detalles, y de manera eficiente. Pero, ¿exactamente donde radica la fuerza de este nuevo enfoque de modelado? Esta radica en que reduce los cálculos para elaborar el modelo de objetos y superficies, y por consecuencia el costo computacional.

La forma que describen en [2] es, se coloca un objeto sobre una mesa giratoria, luego una cámara es colocada enfrente del objeto, el objeto se hace girar por medio de la mesa con desplazamientos angulares constantes (figura 6).

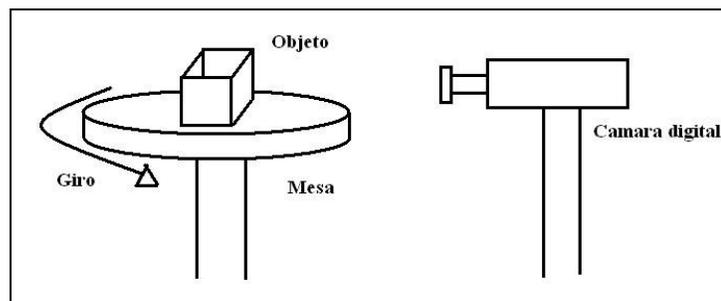


Figura 6

Se hace una toma digital X del objeto en el ángulo inicial L_1 , después una segunda toma Y en el ángulo final L_2 , seguido se hace una toma Z del objeto a la mitad del intervalo angular $L_m = (L_1 + L_2) / 2$, posteriormente utilizando las imágenes X y Y se construye una imagen Z' por medio de una interpolación lineal que corresponde al ángulo L_m . Una vez realizado lo anterior se calcula la diferencia entre la imagen real Z con la imagen interpolada Z' correspondiente, se compara el resultado con un valor ϵ de precisión, si la diferencia de las imágenes es menor o igual que el valor de precisión entonces se guardan las imágenes de los ángulos L_1 y L_m como imágenes necesarias para obtener el modelo, en el caso contrario, se recalcula el intervalo angular y se realizan los mismo pasos. A continuación se muestra el algoritmo que resume lo anteriormente explicado.

1. Asignar la precisión deseada ϵ ,
2. $L_i=0$,
3. $L_s=360$,
4. Adquirir imágenes para las posiciones L_i y L_s ,
5. Si $(L_s-L_i) < 2\theta$ hacer $Error=1$ e ir a 16,
6. $L_m=(L_s+L_i)/2$,
7. Adquirir la imagen correspondiente a la posición L_m ,
8. Interpolar linealmente entre las imágenes L_i y L_s usando la expresión $g(\lambda)=\lambda X+(1-\lambda)Y$, $0 \leq \lambda \leq 1$, λ es un número real,
9. Si es posible interpolar la imagen L_m con las imágenes L_i y L_s dentro del margen de error ϵ usando $g(0.5)$ y usando SSD como criterio para comparar L_m y $g(0.5)$ esto es, $\|L_m-g(0.5)\|^2 \leq \epsilon$, entonces ir al paso 13,
10. Almacenar en una pila el intervalo (L_m, L_s) ,
11. $L_s=L_m$,
12. Ir al paso 5,
13. Guardar las imágenes L_i y L_s como imágenes necesarias para construir el modelo del objeto.
14. Si la pila no esta vacía, extraer de ella una pareja de valores y asignarla a (L_i, L_m) , e ir al paso 5,
15. $Error=0$,
16. Fin.

Donde ϵ es la precisión, L_i y L_s son las cotas inferior y superior respectivamente del intervalo de desplazamiento angular de la mesa, θ es el valor de los incrementos angulares, L_m es la imagen con cierto ángulo de giro de la mesa, X y Y son los vectores que representan las imágenes tomadas en ángulos diferentes, SSD es la suma de diferencias cuadráticas que se define como $\|X - Y\|^2 = \sum_{i=1}^n (x_i - y_i)^2$. Los resultados mostraron que se reduce el número de imágenes necesarias para modelar los objetos.

En la tabla 1 se muestra un ejemplo de los resultados de este algoritmo, donde se hacen tomas de los objetos en 0, 10 y 20 grados, finalmente se calcula, por medio de interpolación lineal, la imagen a 10 grados.

 <p data-bbox="313 762 508 800">Imagen 1 (0°)</p>	 <p data-bbox="820 772 1031 810">Imagen 2 (20°)</p>
 <p data-bbox="313 1129 667 1167">Imagen interpolada (10°)</p>	 <p data-bbox="820 1136 1068 1173">Imagen real (10°)</p>

Tabla 1

RETOS DE LA TESIS

Los R3, como se ha mencionado, son eficientes y han dado buenos resultados. Pero es importante señalar que el funcionamiento óptimo de estos robots pueden darse solo si se cumplen ciertos requisitos, uno de ellos es el tipo de terreno sobre la cual se desplazan los robots: debe ser una superficie totalmente plana y sólida.

Pero ¿qué puede pasar si estos robots son colocados en superficies irregulares?, por ejemplo, terrenos donde existan baches y/o desniveles. Lo mas seguro es que los robots no puedan identificar dichas variantes ya que no han sido diseñados para reconocer esos ambientes y en consecuencia no pueden desplazarse correctamente.

Actualmente, los robots pueden identificar a un compañero, un adversario o el balón, porterías y el campo de juego. La información que obtiene de estos es limitada. No tiene capacidad de reconocer baches o desniveles, situación inusual de una cancha de fútbol.

Para lograr desplazamiento en tales condiciones, el sistema de reconocimiento de objetos y superficies debe hacerse más robusto. Esto implica que los algoritmos de las redes neuronales que controlan el desplazamiento de los robots tengan que ser modificados. Rediseñando o expandiendo los algoritmos de R3 para que evalúen el movimiento del robot tomando en cuenta las irregularidades del terreno. Modificar la estructura lógica de las redes neuronales que en R3 ha dado buenos resultados. La figura 7 muestra una idea general del problema.

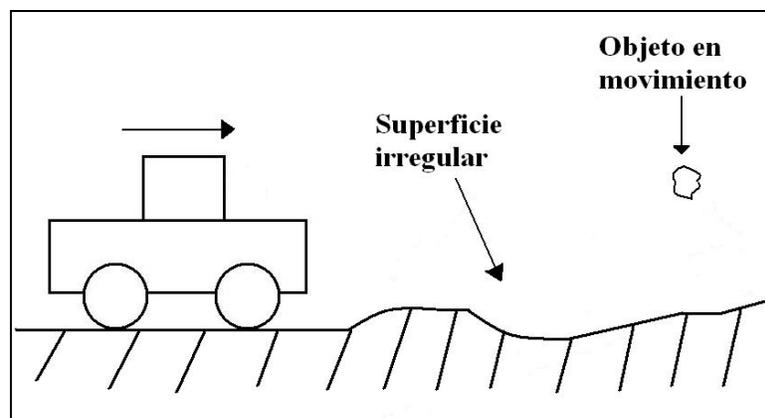


Figura 7

El dibujo muestra un robot móvil que se desplaza sobre una superficie no plana para seguir un objeto en particular que se está moviendo, en este ejemplo es la figura que se encuentra a la extrema derecha. Nótese que la superficie si bien no es total, las irregularidades no son tan abruptas. Dichas irregularidades pueden ser prominencias o depresiones. Otro aspecto importante a considerar es la textura del terreno, es decir, que tan áspera es la superficie.

OBJETIVOS DE LA TESIS

General:

Desarrollar un algoritmo con un sistema de visión en superficies irregulares con seguimiento de objetos.

- I. Modelar superficies utilizando la descripción de su apariencia.
- II. Realizar el algoritmo para determinar la ruta óptima en superficies irregulares, utilizando visión artificial y redes neuronales.
- III. Simular el algoritmo del seguimiento de objetos en superficies irregulares.

Específicos:

- I. Modelar superficies utilizando la descripción de su apariencia.
 - Definir el concepto de obstáculos (baches, desniveles y/o irregularidades).
 - Modelar superficies con el método basado en apariencias con especificaciones de iluminación, escala de grises y colores.
- II. Realizar el algoritmo para determinar la ruta óptima en superficies irregulares, utilizando visión artificial y redes neuronales.
 - Determinar el seguimiento de objetos con Redes Neuronales, utilizando el modelo obtenido en I.
 - Predecir estados utilizando el filtro de Kalman.
- III. Simular el algoritmo del seguimiento de objetos en superficies irregulares.
 - Con la información obtenida en I y II realizar simulación de algoritmo.

RELEVANCIA

Es importante aportar nuevos conocimientos que sean útiles en el campo científico y tecnológico así como dar un uso práctico al conocimiento ya existente, por ejemplo, en este caso la visión artificial y RN cuentan con muchos años de investigación donde se han dado implementaciones físicas, [2], [3], [11]. En este trabajo se pretende utilizar lo mejor de visión (utilizando el modelado basado en apariencias), y RN para lograr un algoritmo que aplicado a un robot le permita ser lo mas autónomo posible.

Científica

Mejorar el diseño de las redes neuronales y el modelado de objetos basado en apariencias, dotar al robot de más "inteligencia" tal que pueda desplazarse de manera robusta en superficies irregulares:

- Extender algoritmos con redes neuronales para aprendizaje de robots.
- Extender algoritmos para el modelado de objetos con información de la apariencia.

Tecnológica

El algoritmo aplicado a un robot (bien diseñado) puede utilizarse en actividades que sean peligrosas al ser desarrolladas por personas o para inspeccionar lugares de difícil acceso a los humanos, pero también pueden ser utilizados en actividades que no sean puramente industriales. Más adelante se enuncia un ejemplo.

PROPUESTA

Como ya se señalo los R3 funcionan bajo condiciones específicas, se desplazan sobre terrenos sin irregularidades. Se propone para investigación doctoral, hacer más robusto el sistema de visión del robot aprovechando las metodologías de reconocimiento de objetos basado en apariencias. De esta forma es más fácil identificar los objetos estáticos (obstáculos) porque es posible obtener información, como colores, matices, textura, etc., que permite distinguir los obstáculos de los objetos en movimiento.

➤ **Desarrollar un algoritmo para determinar la ruta óptima del seguimiento de objetos con un *Sistema de visión* para terrenos irregulares *suaves*, utilizando redes neuronales.**

- Modelar superficies suaves basado en su apariencia,
 - Superficies cóncavas (prominencias),
 - Superficies convexas (depresiones).
- Pendientes no mayores de 15°.
- Considerar la textura de la superficie (con y sin asperezas).
- Modelar las irregularidades más comunes de las superficies.
- Entrenar las RN para generalizar las irregularidades utilizando los casos más comunes anteriormente mencionados.
- Modelar objetos en movimiento con el enfoque basado en su apariencia.
- Entrenar las RN para predecir el movimiento de objetos.
- Calcular la ruta optima de seguimiento del objeto con RN.

METODOLOGÍA

Lo primero es determinar la cantidad de luz que se necesita para el correcto modelado de objetos porque esto repercute directamente en la información a obtener, es decir, no es igual hacer una toma de un objeto en la mañana a una que se hace en la tarde donde el sol se encuentra en diferente posición respecto a su ubicación inicial.

Después se debe hacer un modelo de la superficie desde una perspectiva superior y después a nivel del robot. La razón de realizar primero tomas aéreas es debido a que estas proporcionan más información del ambiente lo que facilita al robot realizar los cálculos para el desplazamiento. Para el modelado de las superficies se utilizará el reconocimiento de imágenes basado en apariencias [1] y [2].

Una vez determinada la mejor ruta se debe identificar los objetos que están en movimiento para determinar cual se seguirá [10] y así estimar los estados futuros del objeto [11] con el fin de calcular la trayectoria que debe tomar el robot para perseguir el objeto [8] y sin sufrir alguna colisión con su entorno [3].

A continuación se enumeran de manera resumida los pasos de la metodología:

1. Determinar cantidad de luz y tipos de iluminación.
2. Modelar superficies desde una perspectiva superior o aérea.
3. Reconocer obstáculos del modelo (situaciones más comunes).
4. Identificar objetos en movimiento.
5. Calcular estados futuros de objetos que estén en movimiento.
6. Determinar la ruta óptima de seguimiento con redes neuronales basado en trabajos previos.
7. Simulación.

POSIBLES RESULTADOS

Los posibles resultados se pueden dividir en dos partes, resultados en el modelado basado en apariencias y en el desarrollo de redes neuronales, a continuación se enuncian.

- Modelado basado en apariencias
 - Modelar irregularidades más comunes.
 - Modelar superficies completas con irregularidades.
 - Modelar objetos en movimiento.
 - Modelar el tipo de rugosidad de la superficie.

- Desarrollo de redes neuronales
 - Generalizar obstáculos.
 - Identificar objetos en movimiento.
 - Predecir estados futuros del objeto en movimiento.
 - Calcular trayectoria necesaria para el seguimiento de objetos (velocidad y posición) considerando **desniveles no mayores de 15° de inclinación**.

REFERENCIAS

1. Altamirano, L.C., Altamirano, L., Alvarado, M., *Non-Uniform Sampling For Improved Appearance-Based Models*, Pattern Recognition Letters, 521-535 pp., 2002.
2. Altamirano, L., *Metodologías para la reducción del número de imágenes requeridas para el reconocimiento de objetos articulados*, Centro de Investigación en Computación - IPN, Tesis doctoral, 2002.
3. Yang, S.X., Meng, Q.H., *Real – Time Collision – Free Motion Planning Of A Mobile Robot Using A Neural Dynamics – Based Approach*, IEEE Transactions On Neural Networks, Nov. 2003, 1541-1552 pp.
4. Mendoza, J.R., *Diseño del control de un robot de dos grados de libertad para aplicaciones de seguimiento de objetos*, Instituto Nacional de Astrofísica, Óptica y Electrónica, Tesis de maestría, 2003.
5. Tannenbaum, A., Angement, S., *Dynamic active contours for visual tracking*, IEEE Transactions on automatic control, Apr. 2006, 562-579 pp.
6. Rife, J.H., Rock S.M., *Design and validation of a robotic control law for observation of deep—ocean Jellyfish*, IEEE Transactions on robotics, Apr. 2006, 282-291 pp.
7. Salas, J., Avalos, W., Castaneda, R., Maya, M., *A machine vision system to measure the parameters describing the performance of a Foucault pendulum*, Machine vision and applications, May 2006, 133-138 pp.
8. Han, Y., Hahn, H., *Visual tracking of a moving target using active contour based SSD algorithm*, Robotics and autonomous systems, Dec. 2005, 265-281 pp.
9. Atev, S., Arumugan, H., Masoud, O., Janardan, R., Papanikolopoulos, N., *A vision-based approach to collision prediction at traffic intersections*, IEEE Transactions on Intelligent Transportation Systems, Dec. 2005, 416-423 pp.
10. Chang, W. Ch., Lee, S.A., *Autonomous vision-based pose control of mobile robots with tele-supervision*, Proceedings of the 2004 IEEE, International Conference on Control Applications, Taipei, Taiwan, Sept. 2004, 1049-1054.
11. www.fu-fighters.de