

Sistema Híbrido para la identificación y seguimiento de robots móviles basado en técnicas de procesamiento de imágenes y redes neuronales recurrentes

Percy Wilianson Lovon Ramos
School of System Engineering
San Agustín University
Arequipa, Perú
Email: percylovon@gmail.com

Dennis Barrios Aranibar
School of Computer Science
San Pablo University
Arequipa, Perú
Email: dennisbarrios@gmail.com

Abstract—el paper debe ser aceptado :D

I. INTRODUCCIÓN

A medida que la ciencia avanza, el desarrollo tecnológico también lo hace; sobretodo en lo que implica facilitar y mejorar la calidad de vida del ser humano, es así que los robots han tomado un papel importante en el cumplimiento de este propósito. Los robots tienen diversas aplicaciones, siendo una de ellas la industria, en la cual se pretende ayudar en el proceso de fabricación de productos y/o servicios utilizando robots para tal cometido [1]. Otra aplicación importante es la milicia, donde los robots tienen mayor repetitividad y precisión [2]. También se tiene una creciente influencia en la agricultura, debido a la falta de mano de obra en tareas que hacen peligrar la integridad del ser humano [3].

Para que los robots no necesiten intervención de la mano humana en su interacción con el mundo exterior (autonomía) es necesario que tengan sensores que sean eficientes. Una de las formas de sensoramiento más importantes es la de visión. Baltes y Anderson mencionan que hay dos tipos de visión, una que es la visión local y la otra global. En la Visión Local el robot tiene una perspectiva de primera persona; en esta el robot tiene una cámara puesta sobre su propia estructura, este tipo trae ciertas dificultades como que el robot sólo puede ver lo que su campo de visión le permite, además si el ambiente multirobot se requeriría cámara para cada uno de los robots. [4]

Por otro lado, la visión global, es la que contiene una o más cámaras (multicámara) las cuales cubren todo el espacio de trabajo, sería mejor llamada espacio inteligente. Un espacio inteligente contiene *sensores* los cuales tienen el rol de identificar los objetos y recibir información del mundo exterior, *procesadores* los cuales son el núcleo del proceso de visión ya que procesan la información, *actuadores* y *dispositivos de comunicación* [5].

El principal objetivo de la visión global es la de detectar e identificar el objeto y a partir de allí hacerle el seguimiento (Moving Object Tracking-MOT). Según Acharya y Ray se tienen dos enfoques para realizar el MOT: el enfoque basado en el reconocimiento y el basado en el movimiento. En el

primero se estudia bajo las características del objeto en el otro se usa las características del movimiento del objeto [6]. El presente trabajo tiene como fin presentar un nuevo abordaje para el problema de seguimiento de objetos en ambientes multicámaras, para lo cual primero se deberá revisar el estado del arte del problema y después se explicará las técnicas que se tomará: para la detección del objeto se utilizará un sistema de visión global basado en técnicas de procesamiento de imagen, con las características obtenidas por el sistema de visión global se realizará, mediante las redes neuronales recurrentes, el seguimiento de los robots.

El artículo está organizado de la siguiente manera: En la sección 1 se hará una introducción al tema en general. En la sección 2 se hará una revisión a los antecedentes y estado del arte de los sistemas de visión global. En la sección 3 se hará una revisión de el tema del seguimiento de objetos móviles en ambientes multicámaras, sus ventajas dificultades. En la sección 5 se hará una revisión de lo que se propone en el presente artículo, se explicará las técnicas utilizadas. En la Sección 6 se verán los resultados de la presente investigación, además se explicará el ambiente de trabajo utilizado. En Sección 7 se darán a conocer las conclusiones del presente trabajo además se mencionará lo futuros abordajes en el tema de los autores.

II. SISTEMAS DE VISION GLOBAL

Como ya mencionamos en la sección anterior se prefiere el uso de sistemas globales para tener un panorama de todo el ambiente, este tipo de visión nos evita la oclusión, pero sin embargo acarrea problemas como la sincronización, si es que fuera en una sola computadora el problema de conectar varias cámaras es solucionado a veces cámaras de determinado estándar de dispositivos. Generalmente los sistemas de visión global para los robots, se apoyan en marcas sobre los robots de forma de círculos los cuales les sirven en el reconocimiento. El problema de visión global, tiene como uno de sus focos principales el problema de la calibración de cámaras que permite establecer parámetros intrínsecos y extrínsecos (internos y externos) para que el funcionamiento de la

camara sea correcto y no variable, dicho problema puede ser abordado con clustering por ejemplo el K-Means, en el que cada clase del algoritmo K-means es representado por un valor RGB(red,green,blue) llamado centro y se escoge aleatoriamente al inicio. Despues de tener los centros se podria agrupar todos los que tienen centros mas cercanos, despues se saca la media y ese es el nuevo centro, el algoritmo terminaria cuando los centros ya no varian [7]. Este mismo enfoque nos puede servir para la segmentacion, esto dividiendo el conjunto de pixeles y agruparlos segun su semejanza, ademas usar de las redes neuronales para clasificar el color, en este caso tendríamos que darle a la red como valores de entrada los valores RGB de cada pixel para que nos pueda retornar la identificacion del color[8].

El uso del color para obtener el reconocimiento del objeto deseado. Por ejemplo el sistema RoboRoos el cual es dividido en un total de 9 capas donde cada una tiene su tarea asignada. En este sistema se utiliza el color para diferenciar los diferentes objetos (robots, pelota), ademas se apoyan de camaras que cumplen con el estandar IEEE Firewire, salvandose asi del problema de bus del USB cuando se trata de conectar varias camaras[9]. Ademas se puede aplicar el color usando mapas de color por ejemplo se puede realizar mapas de color para encontrar la crominancia (el componente que contiene informacion del color en cualquiera video), despues de aplicar los mapas de color se verifica a traves de los contornos si es una pelota, o un jugador[10].

III. SEGUIMIENTO DE OBJETOS EN AMBIENTES MULTICÁMARA

Si bien los sistemas de vision globales nos proporcionan la ubicación, identificación de los objetos, es necesario hacerles un seguimiento, esto se hace por distintos métodos que se tratarán en esta sección. Esta tarea se puede realizar con una sola cámara pero no se puede cubrir todo el espacio bajo este enfoque, es por ello que se recomienda necesario implementar sistemas multicámara.

Uno de los enfoques que realizan los investigadores en el tema es el de la lógica fuzzy el cual nos otorga un grado más cercado a la percepcion humana, esto debido a las funciones de membresía que nos indican grados de pertenencia a los cuales puede pertenecer un color. Se pueden usar la lógica fuzzy de diferentes formas por ejemplo utilizar autómatas fuzzy donde cada estado de la automática S_i representa en la cámara i la aptitud para que sea esta cámara candidata a que en ella este el objeto que se sigue[12], sin embargo este enfoque no es lo suficientemente robusto contra las oclusiones, ademas de solo realizar con una cámara. Tambien apoyados por un grafo de región de adyacencia para representar las imagenes del foreground, se pueden utilizar histogramas de color fuzificados asociados a cada region de adyacencia, los cuales combinados tienen bajo costo computacional y son aplicables en tiempo real[17].

Sin embargo existen otros enfoques relacionados también con el color como los que se basan en histogramas de color, los cuales son fuertes contra el problema de oclusión (los objetos tienden a confundirse mientras más alejados estan de la cámara). Una forma es utilizando el coeficiente de Bhattacharya el cual apoya a la decisión sobre cual cámara

es la que tiene en su enfoque el objeto a seguir mediante un histograma [11]. Otra forma en la se puede hacer el seguimiento basado en histogramas es apoyarse del COBMAT (Color-Based Multiple Agent Tracking), este algoritmo nos permite distribuir el procesamiento del seguimiento de objetos, evitando asi la centralización. Además de apoyarnos en camaras inalámbricas[18].

Algunos autores abordar el problema desde el punto de vista estocástico, en el cual es necesario a veces tener las probabilidades a priori, esto se puede hacer un entrenamiento previo mediante una simulacion de objeto para que el metodo reconozca todo el espacio de trabajo utilizan una persona cargando una pelota y se reconoce la pelota, entonces despues de esto se puede utilizar una cadena de markov para dividir el espacio cubierto en una matriz de estados y poder asi realizar el seguimiento[13], sin embargo este enfoque tiene debilidades de cuando hay dos personas caminando en sentido contrario, se le puede confundir. Otra manera de hacerlo es con planos homográficos, los cuales no permiten reducir el costo computacional, esto apoyado con la programación dinamica para realizar el seguimiento [19]

Además existen abordajes que se basan en las características de determinados dispositivos y estandares para realiza el seguimiento. Por ejemplo las camaras Firewire, basadas Estandar IEEE 1394, el cual da facilidades para conectar varias cámaras, un problema vital para el enfoque de multicámaras[15]. Tambien se pueden usar sistemas de visión ya realizados por ejemplo el sistema de Visión Anafocus Eye-RIS con un algoritmo que combina la deteccion de objetos en movimiento con la extracción de características, el Eye-Ris utiliza el sensor de imagen inteligente (SIS)Anafocus [16]

Un problema parecido al que se aborda en este paper, seria un enfoque distribuido en el que la comunicacion se realiza mediante el protocolo de comunicacion UDP(User Datagram Protocol), esto con el objetivo de seguir varios robots en ambiente multicámara, para facilitar el proceso de identificación del robot les ponen unas marcas con formas geométricas(circulos). Utilizan dos tipos de enfoques para realizar la identificación uno es basado en el color y otro es basado en formas geometricas . Tienen dos programas diferentes, en uno es para controlar las cámaras y la adquisición de imágenes y el otro programa es para centralizar la información obtenida de las cámaras [14].

IV. PROPUESTA

Se propone un abordaje basado en la tecnicas de procesamiento de imagenes que sirvan para implementar el sistema de vision global, el cual tendra como salida la posicion y orientacion actual en cada frame del robot. Luego de este proceso se utilizara una red neuronal recurrente que aprendera de los estados y las posiciones y predecira los siguientes estados para hacerle el seguimiento.

A. Sistema de Vision

En el sistema de vision proponemos utilizar los metodos de el filtro de Gauss y la transformada de Gauss. Para realizar esto utilizamos la libreria de OpenCv, la cual nos da soporte para primero aplicar un filtro de gauss para eliminar el ruido en

nuestra imagen, luego convertimos nuestra imagen a la escala de grises, y finalmente aplicamos la transformada de Gauss para obtener los círculos en la imagen, cuyos centros también son hallados.

Una vez hallados los centros y radios de cada círculo procedemos a realizar el cálculo de la ubicación de nuestro robot, el cual utilizará unos círculos marcados como se muestra en la figura. Entonces como tenemos dos círculos de ubicación conocida procedemos a aplicar el método que nos dice que la posición basado en esos dos círculos estará en el punto medio de la recta que une los dos centros c_i y c_j de los dos círculos [7]. Entonces aplicamos la siguiente igualdad:

$$\cos(\alpha) = \frac{y_1 - y_2}{a} \quad ; \cos(\alpha) = \frac{x}{\frac{a}{2}} \quad (1)$$

de la misma manera se puede hallar la coordenada y :

$$\sin(\alpha) = \frac{x_1 - x_2}{a} \quad ; \sin(\alpha) = \frac{y}{\frac{a}{2}} \quad (2)$$

Entonces al igual las ecuaciones anteriores, se puede hallar la posición actual del robot la cual es:

$$x = 2 \cdot (y_1 - y_2) \quad ; y = 2 \cdot (x_1 - x_2) \quad (3)$$

Una vez hallada la posición actual del robot procedemos a hallar la orientación con respecto al eje inicial de la imagen, de la misma forma utilizamos un método ya comprobado el cual consiste en comparar los puntos centrales y con su tangente hallar la orientación (ángulo) [7]. El ángulo θ de orientación sería dado por:

$$\theta = \arctan\left(\frac{y_1 - y_2}{x_1 - x_2}\right) \quad (4)$$

Así ya tenemos la posición actual en cada frame y la orientación que sigue cada robot, y estamos preparados para darle estos parámetros a la red neuronal recurrente y se pueda predecir su siguiente posición.

V. IMPLEMENTACION

VI. CONCLUSION

The conclusion goes here.

ACKNOWLEDGMENT

The authors would like to thank...

REFERENCES

- [1] E.F.Morales and L.E. Sucar, *Robots y su importancia para México*, 1st ed. Mexico, Instituto Nacional de Astrofisica, Optica y Electronica.
- [2] Pablo Garcia-Robledo and Jesus Torrijos *Robots de Seguridad y Defensa* Espaa, Universidad Politecnica de Madrid
- [3] J.A. Garcia and L.A. Vazquez *Robots de Seguridad y Defensa* Espaa, Departamento De automatica, Ingenieria Electronica e Informatica Industrial, Universidad Politecnica de Madrid
- [4] Jacky Baltes and John Anderson *Intelligent Global Vision for Teams of Mobile Robots, Mobile Robots: Perception & Navigation*, Sascha Kolski (Ed.), ISBN: 3-86611-283-1, InTech, Available from: <http://www.intechopen.com/books/mobile-robots-perception-navigation/intelligent-global-vision-for-teams-of-mobile-robots>, 1st ed. Pro Literatur Verlag, Germany/ARS Austria 2007.
- [5] Misel Brezac, Ivan Petrovic and Edouard Ivanjko, *Robust and accurate global vision system for real time tracking of multiple mobile robots*, *Robotics and Autonomous Systems* 56 (2008) 213-230. Elsevier, 2008
- [6] Tinku Acharya and Ajoy K. Ray, *Image Processing Principles and Applications*, pag. 6, 1st ed. B. Michaelis and G. Krell (Eds.): DAGM 2003, LNCS 2781, pp. 591599, Springer-Verlag Berlin Heidelberg 2003.
- [7] Kelson Romulo Teixeira Aires, Pablo Javier Alsina, Adelardo Adelino Dantas de Medeiros, *A GLOBAL VISION SYSTEM FOR MOBILE MINI-ROBOTS* ed. SIMPSIO BRASILEIRO DE AUTOMAO INTELIGENTE, 5, Canela, 2001
- [8] Manu Chhabra, Anusheel Nahar, Nishant Agrawal, Tamhant Jain, Amitabha Mukerjee, Apurva Mathad and Siddhartha Chaudhuri, *Novel Approaches to Vision and Motion Control for Robot Soccer* ed. Proceedings of the National Conference on Advance Manufacturing and Robotics, India
- [9] Ball, David and Wyeth, Gordon and Nuske, Stephen, *A global vision system for a robot soccer team* ed. SAustralasian Conference on Robotics and Automation, 6-8 December 2004, Canberra
- [10] Gnnr, Claudia and Rous, Martin and Kraiss, Karl-Friedrich, *Real-Time Adaptive Colour Segmentation for the RoboCup Middle Size League* ed. RoboCup 2004: Robot Soccer World Cup VIII, Springer Berlin Heidelberg
- [11] Kajta Nummiaro, Esther Koller-Meier, Tomas Svoboda, Daniel Roth, and Lucas Van Gool and Ajoy K. Ray, *Color-Based Object Tracking in Multi-camera Environments*, 1st ed. B. Michaelis and G. Krell (Eds.): DAGM 2003, LNCS 2781, pp. 591599, Springer-Verlag Berlin Heidelberg 2003
- [12] Kazuki Morioka, Silvester Kovacs, Peter Korondp, Joo-Hoo Lee and Hideki Hashimoto, *Adaptive camera selection based on fuzzy automation for object tracking in a Multicamera System* ed. Journal of Engineering Annals of Faculty of Engineering Hunedoara, 2008
- [13] Anthony R. Dick and Michael J. Brooks, *A Stochastic Approach to Tracking Objects Across Multiple Cameras* ed. Springer -Verlag Berlin Heidelberg, 2004
- [14] Renato F. Garcia, Pedro M. Shiroma, Luiz Chaimowicz, Mario F.M. Campos, *Um Arcabouco para Localizacao de Enxames de Robos* ed. Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
- [15] Piyush Kumar Rai, Kamal Tiwari, Prithwiji Guha and Amitabha Mukerjee, *A cost-Effective Multiple Camera Vision System using Firewire Cameras and Software Synchronization* ed. IIT Kanpur, UP, India.
- [16] Fethullah Karabiber, Paolo Arena, Luigi Fortuna, Sabestiano De Fiore, Guido Vagliasindi and Sabri Arik, *Implementation of a Moving Target Tracking Algorithm Using Eye-RIS Vision System on a Mobile Robot* ed. Springer Science+Business Media, LLC 2010
- [17] Amir Hossein Khalili and Shohreh Kasaei, *Object Modeling for Multicamera Correspondence Using Fuzzy Region Color Adjacency Graphs* ed. Springer Berlin Heidelberg, Advances in Computer Science and Engineering 2009
- [18] Enre Oto, Frances Lau, and Hamid Aghajan, *Color-Based Multiple Agent Tracking for Wireless Image Sensor Networks* ed. Proceedings of the 8th international conference on Advanced Concepts For Intelligent Vision Systems, Springer-Verlag, 2006
- [19] Yildiz, Alparslan and Akgul, YusufSinan, *A Fast Method for Tracking People with Multiple Cameras* ed. Trends and Topics in Computer Vision, Springer Berlin Heidelberg, 2012