



“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

21 y 22 de abril de 2016, Cortazar, Guanajuato, México

Sistema de seguimiento múltiple con Camshift en ambientes controlados para la detección de personas en salones de clases

Juan Paulo Sánchez Hernández, Gerardo Escamilla Rojas, Miguel Ángel Ruiz Jaimes, Sandra Elizabeth León Sosa

Universidad Politécnica del Estado de Morelos

{[juan.paulosh,ergo113949,mruiz,lsandra](mailto:juan.paulosh,ergo113949,mruiz,lsandra}@upemor.edu.mx)}@upemor.edu.mx

Resumen: En este artículo se presenta, un sistema de seguimiento múltiple en salones de clases con entornos controlados y no controlados. El algoritmo que se utiliza para realizar seguimiento múltiple es CAMSHIFT, uno de los algoritmos más eficientes. El histograma de color fue definido a partir de un conjunto de 30 muestras de piel. Con el fin de mejorar el proceso de seguimiento se utilizan algunas técnicas de procesamiento de imágenes, tales como: filtro gaussiano, operaciones morfológicas, histogram back projection y detección de contornos con códigos de cadena. Para comprobar la eficiencia del algoritmo se establecieron test de pruebas, en donde el algoritmo funciona muy bien en entornos controlados.

Palabras claves: Algoritmo CAMSHIFT, histogramas de color, seguimiento múltiple

Abstract: In this paper a multiple tracking system is presented. This approach is applied into classroom in controlled and uncontrolled environments. The algorithm for multiple tracking was CAMSHIFT, an efficient algorithm which use a color histogram to obtain a set search windows. The histogram of color was built using a set of 30 skin samples. This process is improved through image processing technics such as Gaussian filter, histogram back projection method, morphological operations and edge detection with chain code. To evaluate the efficiency of the algorithm a set test were applied where this process shows better results in controlled environments.

Key words: Multiple Tracking, CAMSHIFT, color histogram,

I. Introducción

La visión computacional, se refiere a la capacidad de dotar a un dispositivo electrónico de interpretar los objetos de su entorno por medio de un sensor (una cámara por ejemplo). Es decir, los humanos tenemos la capacidad de poder reconocer objetos de nuestro espacio visual de una manera muy natural, por lo que la visión computacional, intenta realizar las mismas tareas que hace un ser humano, por ejemplo, reconocer objetos, seguirlos, detectar movimiento, reconstruirlos, localizarlos en su entorno, reconocer intenciones y reconocimiento de acciones humanas (Poppe, 2010) .



“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

21 y 22 de abril de 2016, Cortazar, Guanajuato, México

Además, es una de las áreas con gran proyección, debido a que se han realizado una gran cantidad de aplicaciones interesantes (de Campos, 2014), como ejemplo, se encuentra el reconocimiento de objetos (Arévalo-Vázquez, Zúñiga-López, Villegas-Cortez, & Avilés-Cruz, 2015), clasificación de objeto (Sánchez-Hernández & Hernández Rabadán, 2014) o el seguimiento de objetos (tracking) (Breitenstein, Reichlin, Leibe, Koller-Meier, & Van Gool, 2009), este último es el objeto de interés de esta investigación.

Cuando se menciona el Tracking de objetos, se refiere a detectar y estimar la localización de un objeto particular (persona u objeto) dentro de una escena, la cual puede ser a partir de una secuencia de imágenes digitales capturadas, de tal manera, que dicho objeto no se pierda de vista. Realizar Tracking es un proceso muy complicado dentro de la visión computacional debido a que se tienen que resolver varios problemas de procesamiento de imágenes, el cual es un paso fundamental en todo sistema de visión computacional. Es decir, en el procesamiento de imágenes se tienen que realizar tareas de limpieza de imagen, mejoramiento y transformación de espacios de color. Todo lo anterior, con el fin de obtener una imagen lo mejor posible que nos facilite la visión computacional.

El Tracking, se puede realizar con uno o varios objetos, en el caso de varios objetos existen trabajos muy interesantes, como el desarrollado por (Jog & Halbe, 2012) en donde se realiza un seguimiento múltiple de objetos en movimiento utilizando el algoritmo CAMSHIFT (Bradski, 1998), así también otro trabajo relacionado con el seguimiento múltiple utilizando CAMSHIFT y color (Hidayatullah & Konik, 2011).

En este trabajo se pretende realizar tracking múltiple en personas con el fin de detectarlas en entornos controlados y no controlados dentro de salones de clases.



“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

21 y 22 de abril de 2016, Cortazar, Guanajuato, México

Lo anterior, debido a que se podría utilizar dicho sistema para detectar a las personas en contingencias naturales como incendios o temblores, esto permitiría al personal de seguridad poder localizar a personas y poder rescatarlas.

Este artículo se encuentra organizado de la siguiente manera: en la sección I, Introducción se presenta la introducción en donde se revisan los conceptos importantes de la visión computacional, así como el tracking múltiple y sus aplicaciones; en la sección II, se presenta la metodología de visión computacional que se utiliza para realizar el seguimiento múltiple, así como la descripción de los algoritmos y modelos matemáticos utilizados; en la sección III, se presentan los resultados obtenidos en los ambientes de prueba elegidos, al igual que la eficiencia obtenida; en la sección IV, se presentan las conclusiones del trabajo realizado, así como los trabajos futuros generados de esta investigación.

II. Metodología

En esta sección se presenta la metodología que se sigue para la realización de esta investigación (Figura 1), así como se describen las etapas de la que están conformadas, de igual manera se describen los algoritmos y modelos matemáticos que se utilizaron.

1. Adquisición de Imagen: En esta etapa se realiza la extracción de la muestra inicial de tipos de pieles con el objetivo de tener representadas dichas muestras en un histograma. Esta muestra será de vital importancia para las etapas posteriores.

a) **Muestra inicial:** Se eligió un conjunto de 30 personas de las que se extrajo una muestra de piel, con el fin de crear una imagen que representara todas las muestras de piel (Figura 2a).

“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

21 y 22 de abril de 2016, Cortazar, Guanajuato, México

b) Cámara: Para realizar la adquisición de imágenes de esta investigación se utilizó una cámara Logitech HD Pro Webcam c920.

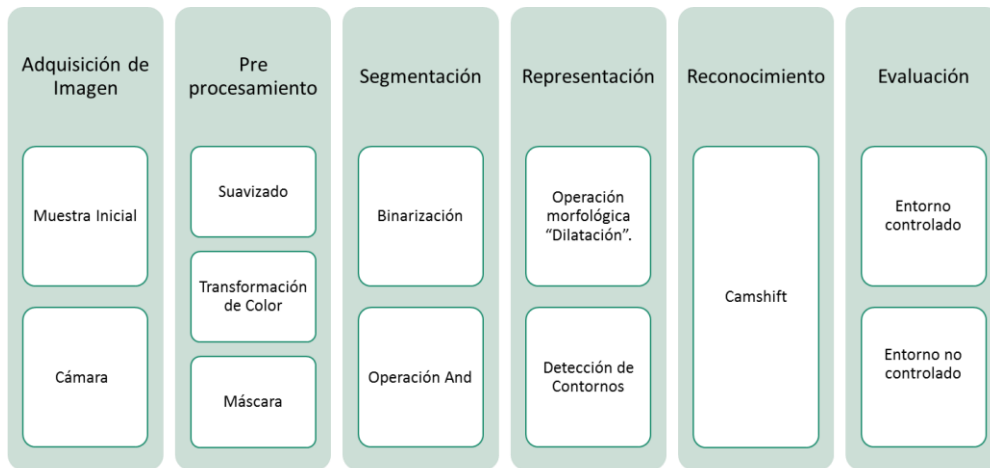


Figura 1. Metodología de visión computacional

2. Pre procesamiento: En esta etapa se utilizan técnicas de visión computacional para mejorar la calidad de las imágenes capturadas y poder realizar las etapas posteriores de mejor manera.

a) Suavizado: Se utiliza un suavizado gaussiano con el fin de resaltar las características de la imagen (Figura 2b)

b) Transformación de Color: Se realiza la transformación de color del espacio de colores de RGB a HSV con la finalidad de que la imagen resalte (Figura 2c) mucho mejor las características de la piel.



Figura 2. Muestra de piel de 30 personas



“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

21 y 22 de abril de 2016, Cortazar, Guanajuato, México

- c) **Máscara:** Se obtiene una máscara a partir de la retroproyección.
3. **Segmentación:** Una vez obtenida la muestra de 30 personas de piel, se realiza una segmentación, es decir, en la captura de la imagen, en donde se detecte piel, se segmentará la imagen, con el fin de identificar la piel.
 - a) **Binarización:** Obtenido el histograma de piel, se establecen los umbrales que nos permitirán binarizar la imagen.
 - b) **And:** Con el fin de obtener una imagen mucho mejor segmentada se realiza una operación **And** entre la máscara y la imagen binarizada.
4. **Representación:** En esta etapa se realizan operaciones de dilatación y detección de contornos con el fin de detectar los contornos de piel de las personas.
 - a) **Dilatación:** Este operador morfológico nos permite dilatar un poco los pixeles con el fin de realizar una mejor detección de contornos.
 - b) **Detección de contornos:** Se realiza la detección de contornos con el fin de establecer las ventanas de búsqueda para el algoritmo de seguimiento. Para extraer los contornos se utilizaron los códigos de cadena.
5. **Reconocimiento:** Una vez realizadas las etapas previas, se procede a realizar el tracking, para esto, se utiliza el algoritmo de seguimiento denominado CAMSHIFT.
 - a) **Algoritmo CAMSHIFT.** Para que el algoritmo funcione es necesario alimentarlo con una muestra inicial de color, por lo que en esta etapa se le envían varias muestras segmentadas de color con el fin de que pueda realizar el tracking de las personas detectadas en la escena.
6. **Evaluación:** Para verificar la validez del tracking múltiple, se tuvieron que definir dos entornos, uno controlado y el otro no controlado.



“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

21 y 22 de abril de 2016, Cortazar, Guanajuato, México

a) **Entorno controlado:** Se eligió un salón pequeño, con iluminación homogénea y un fondo en color “azul”. En la Tabla 1, se presentan las características del entorno controlado.

Salón	2x2 metros
Fondo	Azul oscuro
Iluminación	Lámpara de luz blanca de neón
Distancia iluminación al objeto	2 metros
Distancia cámara-objeto	1.5 metros

Tabla 1. Características del entorno controlado

b) **Entorno no controlado:** El entorno no controlado, se realizó en salones de clases en donde no se controló el espacio, el fondo, y la iluminación.

Salones	1) Salón 1: 5x6 metros 2) Salón 2: 8X8 metros
Fondo	1) Salón 1: No controlado 2) Salón 2: No controlado
Iluminación	Salón 1: Lámpara de luz blanca Salón 2: Luz natural
Distancia iluminación al objeto	Salón 1: indeterminada Salón 2: 2 metros
Distancia cámara objeto	Salón 1: Variada Salón 2: Variada

Tabla 2. Características del entorno no controlado

III. Resultados

En esta sección se presentan los resultados obtenidos en dos entornos y con un conjunto definido de personas, se realizaron 30 pruebas en entornos controlados y 30 en entornos no controlados.

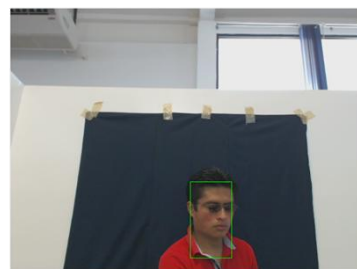
“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

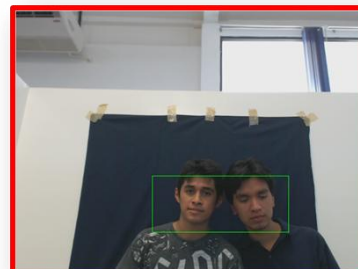
21 y 22 de abril de 2016, Cortazar, Guanajuato, México

En la Figura 4, se puede observar la ejecución del seguimiento múltiple en un entorno controlado. Se definió un fondo de color azul, un salón de dimensiones pequeñas y se controló la iluminación mediante lámparas fluorescentes cálidas. Se puede observar que el sistema funciona muy bien debido a que puede segmentar la piel de las personas en el entorno, por lo que con dichas muestras se alimenta el algoritmo CAMSHIFT y puede realizarse el tracking múltiple correctamente.

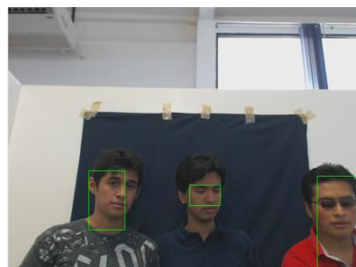
En la Figura 5, se presentan ejemplos de pruebas en entornos no controlados, como se aprecia en la Figura 5a, se logra detectar a una persona en un ambiente en donde no existe un control ni de iluminación, ni de fondo. En la Figura 5b, se realiza el Tracking múltiple en el salón 1 y se observa que detecta a las personas en el entorno, sin embargo, también detecta otros objetos, como la mesa y las



a) Tracking con una persona



b) Tracking con dos personas y tracking erróneo



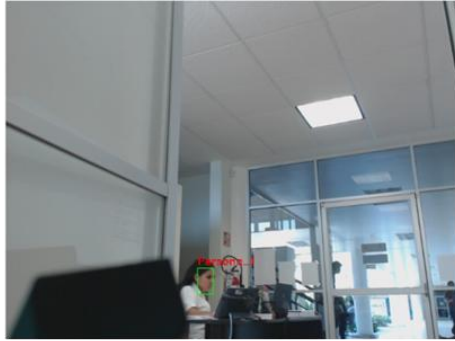
c) Tracking con tres personas



d) Tracking con cinco personas

Figura 4. Pruebas realizadas en entornos controlados

“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”
Multidisciplinario
21 y 22 de abril de 2016, Cortazar, Guanajuato, México



a) Tracking con una persona



b) Tracking múltiple en salón uno



c) Tracking múltiple en salón dos



d) Tracking múltiple en salón dos

Figura 5. Pruebas en ambientes “no controlados”

conexiones de corriente. En la Figura 5c y 5d, se presenta una imagen de prueba en el salón dos, con dos y cuatro personas. En donde se observa que realiza el seguimiento de personas, sin embargo se detectan muchos falsos positivos, es decir, el sistema confunde partes de la mesa como persona.

De los resultados de eficiencia obtenidos, se obtuvo una eficiencia del 76.9%. Para obtener dicha eficiencia se realizaron 30 pruebas en entornos controlados y se probaron con 1 a 5 personas y se tomaba como positivas las pruebas en donde el sistema lograba detectar correctamente a la persona y como negativas si se confundía y no realizaba correctamente el tracking. Para el caso de entornos no



“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

21 y 22 de abril de 2016, Cortazar, Guanajuato, México

controlados, fue muy difícil obtener una eficiencia debido a que cuando se realizaba el tracking lograba seguir a la persona, sin embargo se detectaban muchísimos falsos positivos, esto debido a los cambios de iluminación tan bruscos en entornos no controlados. Por lo que es necesario mejorar el proceso de seguimiento de tal manera que se minimicen los falsos positivos en entornos no controlados.

IV. Conclusiones

En este trabajo de investigación se presenta una aplicación de Tracking múltiple utilizando el algoritmo CAMSHIFT, así como algunas de las técnicas de procesamiento de imágenes para mejorar el proceso de Tracking. Se realizaron las pruebas en dos entornos, uno controlado y el otro no controlado, los dos se realizaron en salones de clases, naturalmente, el algoritmo funciona mucho mejor en entornos controlados, lo que dio muy buenos resultados. Por el contrario, en las pruebas en entornos no controlados, se realizó un buen seguimiento, sin embargo, cuando las personas se encontraban muy lejos el sistema no lograba identificarlas, de igual manera, el cambio de iluminación fue un factor importante para que el sistema no realizara el Tracking de manera efectiva.

Dado los resultados mostrados, lo siguiente es trabajar en mejorar el seguimiento de personas dentro de salones de clases en donde no se pueda controlar el entorno, así mismo, aplicarlo en diferentes áreas de la universidad.

V. Bibliografía

Arévalo-Vázquez, E. E., Zúñiga-López, A., Villegas-Cortez, J., & Avilés-Cruz, C. (2015). Implementación de reconocimiento de objetos por color y forma en un robot móvil. *Research in Computing Science*, 21-31.

Bradski, G. (1998). Computer Vision Face Tracking For Use in a Perceptual User interface. *Intel Technology Journal Q2 98*, 1-15.



“CONGRESO INTERNACIONAL DE INVESTIGACIÓN E INNOVACIÓN 2016”

Multidisciplinario

21 y 22 de abril de 2016, Cortazar, Guanajuato, México

Breitenstein, M. D., Reichlin, F., Leibe, B., Koller-Meier, E., & Van Gool, L. (2009). Robust Tracking-by-Detection using a Detector Confidence Particle Filter. *IEEE 12th International Conference on Computer Vision (ICC)*, 1515-1522.

de Campos, T. (2014). A survey on computer vision tools for action recognition, crowd surveillance and suspect retrieval. *XXXIV Congresso da Sociedade Brasileira de Computação – CSBC 2014*, 1123-1132.

Hidayatullah, P., & Konik, H. (2011). CAMSHIFT Improvement on Multi-Hue and Multi-Object Tracking. *International Conference on*, 1-6.

Jog, A., & Halbe, S. (2012). Multiple Objects Tracking using CAMshift Algorithm in OpenCV. *IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)*, 1, 41-46.

Poppe, R. (2010). A survey on vision-based human action recognition. *Image and Vision Computing* (28), 976-990.

Sánchez-Hernández, J. P., & Hernández Rabadán, D. L. (2014). Comparison of classification methods applied to the diagnosis of malignant melanoma using asymmetry. *Programación Matemática y Software*, 51-56.