

DetECCIÓN DE ACTIVIDADES CON CADENAS DE SÍMBOLOS Y MODELOS OCULTOS DE MARKOV

Activity detection using strings of symbols and Hidden Markov Models

Herlindo Hernández-Ramírez^{a*}, Luis A. Díaz-Jiménez^a, Jorge L. Pérez-Ramos^b

^aCIDESI, Av. Pie de la Cuesta No. 702, Desarrollo San Pablo, 76125 Santiago de Querétaro, Querétaro, México

^bFac. Informática UAQ, Av. de las Ciencias S/N, 76230 Juriquilla, Querétaro, México

*hhhermandez@posgrado.cidesi.edu.mx

Resumen

En este trabajo se presenta un enfoque integral para la generación de trayectorias en visión por computadora, aprovechando la representación del movimiento en cadenas de símbolos compactas. La metodología propuesta incorpora un proceso para preservar la información del movimiento a lo largo del tiempo e identificar áreas de alta probabilidad de movimiento. Mediante segmentación morfológica, se construyen trayectorias simplificadas para una representación más compacta. Se emplean Modelos Ocultos de Markov (HMM) para el modelado y reconocimiento de actividades, beneficiándose de la reducción de estados visibles gracias a la representación compacta del movimiento. Este trabajo demuestra cómo este enfoque permite una detección precisa y eficiente de actividades normales y anormales, proporcionando un marco robusto para sistemas de visión por computadora.

Abstract

This paper presents a comprehensive approach to trajectory generation in computer vision, leveraging the representation of motion as compact chains. The proposed methodology incorporates a process to preserve motion information over time and identify areas of high probability of movement. Through morphological segmentation, simplified trajectories are constructed for a more compact representation. Hidden Markov Models (HMMs) are employed for activity modeling and recognition, benefiting from the reduction of visible states due to the compact representation of motion. This work demonstrates how this approach allows for precise and efficient detection of normal and abnormal activities, providing a robust framework for computer vision systems.

Palabras clave:

Generación de Trayectorias, Modelos Ocultos de Markov, Reconocimiento de Actividades, Visión Artificial

Keywords:

Trajectory Generation, Hidden Markov Models, Activity Recognition, Computer Vision

Introducción

El Internet de las Cosas (IoT) y la reducción de costos en las tecnologías de monitoreo han popularizado el uso de cámaras en una variedad de entornos, como centros comerciales, estacionamientos, sistemas de vigilancia doméstica, autopistas e industrias [1], [2].

Tradicionalmente, el análisis de los videos se realiza manualmente por personal capacitado, que visualiza e interpreta la información para discernir entre lo importante y lo trivial. Sin embargo, este método presenta varias desventajas, incluyendo el error humano inherente a las tareas repetitivas y los costos asociados al mantenimiento y capacitación del personal [3], [4].

Estas limitaciones han impulsado la búsqueda de soluciones automatizadas, aprovechando la capacidad de procesamiento y los avances tecnológicos de las computadoras [1], [5], [6]. La metodología presentada en este trabajo se centra en la extracción de patrones de movimiento $\mathbf{P}_{i,j}$ y su conversión en trayectorias \mathbf{t}_i , que almacenan las posiciones de los objetos en movimiento en instantes de tiempo específicos, manteniendo así la información temporal del movimiento [7].

Una de las claves de este trabajo es la segmentación de las regiones \mathbf{r} , con alta pertenencia al movimiento, esto es posible con la utilización de una función de decaimiento temporal \mathbf{M}_T que preserva la información de los objetos en movimiento a lo largo del tiempo. Esta función permite identificar las áreas con movimiento en la escena, donde al crear el acumulado de evidencia de movimiento se crea una superficie \mathbf{S}_M que acumula la información de las áreas con mayor pertenencia al movimiento.

Para definir las regiones \mathbf{r}_i que forman las trayectorias \mathbf{s}_i , se aplica un operador de segmentación morfológica. Este proceso resulta en una máscara de estados \mathbf{S}_w , la cual delimita las áreas de mayor actividad en la escena. Estas regiones $\mathbf{r}_i \in \Sigma$, las cuales están directamente relacionadas con el conjunto de símbolos Σ , son utilizados para formar nuevas trayectorias como cadenas de símbolos \mathbf{s}_i . Se toma ventaja de la propiedad de definir las regiones pertenecientes al movimiento como \mathbf{r}_i , donde cada $\mathbf{P}_{i,j} \in \mathbf{r}_i$, da como resultado la región \mathbf{r}_i a la que pertenece $\mathbf{P}_{i,j}$. Entonces al aplicar un operador γ a las cadenas de símbolos \mathbf{s}_i , estas representan una versión más compacta de las trayectorias originales \mathbf{t}_i [8].

El uso de Modelos Ocultos de Markov (HMM) permite modelar y reconocer actividades basadas en las trayectorias generadas. Un HMM se define por sus parámetros $\lambda = \{A, B, \pi\}$, donde cada uno de estos parámetros se establece a partir de la información de los estados ocultos y visibles del sistema. La cardinalidad de las matrices A, B, π , está directamente relacionada con el número de estados del sistema. Al utilizar trayectorias s_i como cadenas de símbolos compactas, se reduce el número de estados del sistema y, en consecuencia también se reduce la complejidad del modelo [9].

En las siguientes secciones, se detallará el proceso de generación de trayectorias, la validación de la metodología propuesta y los resultados obtenidos. Se demostrará cómo la representación compacta del movimiento aplicada a los HMM permite una definición de los estados del modelo más precisa y eficiente, proporcionando un marco robusto para el análisis en sistemas de visión artificial.

Marco teórico

Se abordan las ideas principales relacionadas con la elaboración este trabajo.

Detección de movimiento y generación de trayectorias

La detección de movimiento en visión artificial es una tarea fundamental para el proceso de generación de trayectorias, existen diferentes técnicas para realizar esta tarea, algunas simples otras complejas, aunque con un mismo objetivo: obtener la información de los objetos que se mueven dentro de una escena, A partir de esta información, se forman trayectorias que representan la dinámica de estos objetos [10], [11].

El proceso para la generación de trayectorias propuesto implica tomar la información de secuencias de video y realizar la extracción de los patrones de movimiento, es decir, las posiciones de los objetos en la escena. Con esta información se aplica una técnica para obtener la evidencia temporal del movimiento y así crear un acumulado de toda esta evidencia para resaltar las regiones que han experimentado mayor pertenencia al movimiento. Esta información, a su vez facilita la aplicación de un proceso de segmentación para definir regiones más acotadas, al intersectarse con los patrones de movimiento forman una nueva representación de trayectorias. El proceso para la generación de estas nuevas trayectorias puede observarse en la Fig. 1.

En este trabajo la extracción de los patrones de movimiento $M_O : I_{m \times n} \times I_{m \times n} \rightarrow I_{m \times n}$ se realiza aplicando la siguiente ecuación [12], [13]:

$$M_O(I^t_{P_{i,j}}, I^{t-1}_{P_{i,j}}) = I^t_{P_{i,j}} - I^{t-1}_{P_{i,j}} \quad (1)$$

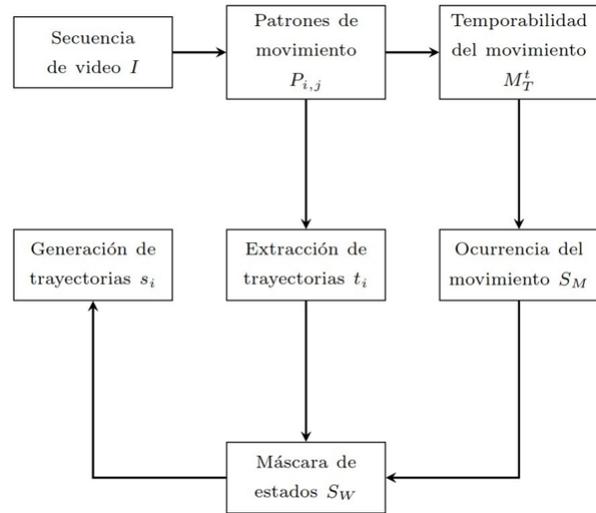


Fig. 1 Diagrama de flujo para el proceso de generación de trayectorias S_i

De la ec. (1), se obtiene la información de los objetos en movimiento M_O , identificando el pixel $P_{i,j}$ correspondiente al objeto en movimiento de la escena I^t en el instante de tiempo t , esta información se obtiene al comparar los pixeles $I^t_{P_{i,j}}$ y $I^{t-1}_{P_{i,j}}$ en los instantes de tiempo t y $t-1$. La información de movimiento obtenida se utiliza para formar los patrones de movimiento t_i , estos patrones de movimiento forman el conjunto de trayectorias $T = \{t_1, t_2, t_3, \dots, t_n\}$ donde cada trayectoria $t_i = P^t_{i,j}$, almacena las posiciones i, j del objeto en movimiento en el instante de tiempo t [14].

Una vez obtenida la información de los objetos en movimiento $P_{i,j}$, es necesario definir como se mueven estos objetos en la escena a lo largo del tiempo, para esto se defina una función $M_T : I_{m \times n} \times I_{m \times n} \rightarrow I_{m \times n}$ que preserva la información temporal del movimiento de los objetos. Para un pixel $P_{i,j}$ con el método de plantillas temporales el movimiento se define como [15]:

$$M_{TP_{i,j}}^t = \alpha M_{TP_{i,j}}^{t-1} + (1 - \alpha) M_O(I^t_{P_{i,j}}, I^{t-1}_{P_{i,j}}) \quad (2)$$

Donde la constante de decaimiento α esta acotada a los valores $[0,1]$. La función M_T preserva la información de los objetos en movimiento capturada por M_O a medida que los objetos se desplazan por la escena. La función de movimiento M_O almacena la diferencia absoluta que existe entre la imagen I^t en los instantes de tiempo t y $t-1$.

Los pixeles $P_{i,j}$ correspondientes a las zonas donde no se detectó movimiento tienen el valor $P_{i,j} = 0$, mientras que las zonas donde se detectaron objetos en movimiento tienen valor $P_{i,j} > 0$, lo que significa que para los instantes de tiempo t y $t-1$, existió una variación en los valores de $P_{i,j}$ [16].

Con la información de la temporabilidad de los objetos en movimiento M_T , es necesario conocer la distribución de probabilidades (pdf) de la ocurrencia del movimiento de cada pixel, esto es posible al crear una superficie S_M que contenga la información de movimiento acumulado, correspondiente

a cada pixel de la escena en el instante de tiempo t , cómo se define a continuación.

$$S_M = \sum_{i=1}^n M_T^i \quad (3)$$

Las zonas con alta probabilidad de pertenencia al movimiento r_i , son definidas al aplicar un operador de segmentación morfológica a S_M , este proceso consiste en aplicar *Watershed*(S_M) [17] a un conjunto de información de movimiento, correspondiente a un periodo de tiempo n . Con la aplicación de este proceso de segmentación a S_M se obtiene la máscara de estados S_W .

De la máscara de estados $S_W = \text{Watershed}(S_M)$ se obtiene $\Sigma = \{r_1, r_2, r_3, \dots, r_n\}$ que es el conjunto de regiones r_i con las que se formaran las trayectorias s_i [8], la visualización de cómo están distribuidas las zonas pertenecientes al movimiento en la escena es posible con $M_W = S_W \cup I_b$. Para obtener las trayectorias $\hat{S} = \{s_1, s_2, s_3, \dots, s_n\}$, se aplica lo siguiente.

$$s_i = M_W \cap t_i \quad (4)$$

Donde s_i es la trayectoria generada por $s_i = \{r_i^t\}$. Esta nueva representación de trayectorias tiene la propiedad $s_p = \{r_1, r_1, r_1, r_2, r_2, r_2, r_3, r_3\}$, lo que significa que algunos patrones de movimiento P_{ij} , comparten una misma región r_i , por lo tanto, para una trayectoria t_p , se tiene $|s_p| = |t_p|$, lo que implica que las trayectorias s_i tienen la misma cardinalidad que t_i .

Al aplicar un operador ∇ a s_p , es posible eliminar los elementos r_i repetidos en cada trayectoria s_p . Esto puede observarse al aplicarse este operador a la trayectoria $\nabla(s_p) = \{r_1, r_2, r_3\}$ [15], con esta propiedad se obtiene $|\nabla(s_p)| < |t_p|$ lo que significa que al ser una representación más compacta, se necesita menos información para representar las trayectorias de los objetos en la escena.

Detección de actividades

Con la ventaja obtenida en la representación del movimiento, ahora es posible hacer el modelado de los objetos utilizando Modelos Ocultos de Markov (HMM).

Un HMM es un modelo probabilístico definido como $\lambda = \{A, B, \pi\}$, dado que los HMM obedecen el comportamiento de un sistema doblemente estocástico, para poder aplicarlos primero se definen los estados ocultos S_i y los estados visibles v_i , estas definiciones permiten desarrollar las reglas de transición a_{ij} y emisión $b_i(k)$.

En la Fig. 2 se muestra una representación de un HMM con dos estados ocultos, las transiciones entre los estados ocultos a_{ij} y las relaciones que existen entre los estados visibles v_i con los estados ocultos S_i [18], [19].

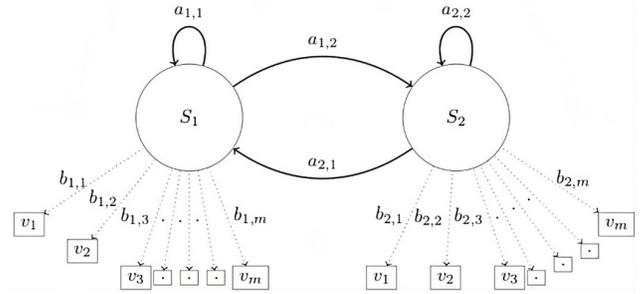


Fig. 2 HMM con dos estados ocultos (S_1, S_2) y v_m estados visibles

La definición de los parámetros de un HMM incluye $A = \{a_{ij}\}$, donde A es la matriz de transición de un estado oculto S_i en el tiempo t a un estado oculto S_j en el tiempo posterior $t+1$. La denominación de un estado en un instante de tiempo t , donde $t=1, 2, 3$ es q_t . La transición entre estados ocultos se expresa como $Q = \{q_1, q_2, q_3, \dots, q_T\}$, donde T representa el número total de instantes de tiempo que forma Q .

La selección de un estado oculto S_i en un instante de tiempo t , también puede expresarse como $q_t = S_i$ donde $1 \leq i \leq n$, siendo n el número de estados ocultos del modelo. Con la información de la transición de los estados ocultos Q en los instantes de tiempo $q_t = S_i$, ahora es posible calcular la matriz de transición entre estados ocultos de la siguiente manera [20]:

$$A_{i,j} = \frac{\chi(S_{i,j})}{\sum_{k=1}^n \chi(S_{i,k})} \quad (5)$$

En la ec. (5), $\chi(S_{i,j})$ cuenta el número de transiciones del estado oculto S_i al estado oculto S_j , mientras que $\chi(S_{i,k})$ cuenta el total de transiciones del estado oculto S_i hacia cualquier estado oculto del modelo, incluyendo el estado oculto S_j .

La dimensión de la matriz de transición A es igual a $n \times n$, donde n es el número de estados ocultos definidos en el modelo.

El siguiente parámetro del modelo es la matriz de emisión $B = \{b_i(k)\}$. Los posibles estados visibles del sistema se definen como $V = \{v_1, v_2, v_3, \dots, v_m\}$, donde m es el número de símbolos que conforman los estados visibles del modelo. Estos estados visibles están relacionados con cada estado oculto S_i del modelo.

La secuencia de observaciones en un instante de tiempo t se denota como O_t , para una secuencia de longitud T esta secuencia se representa como $O = \{O_1, O_2, O_3, \dots, O_T\}$, donde cada elemento $O_t = v_k$. Esto significa que dentro del conjunto de observaciones O se pueden repetir elementos v_k , la probabilidad de una observación v_k asociada a un estado S_i , se representa como $b_i(k)$ que es la probabilidad de pertenencia de una observación a un estado oculto, estas probabilidades se calculan con la información de las observaciones O como sigue [20]:

$$B_{i,k} = \frac{\chi(b_i(k))}{\sum_{m=1}^M \chi(b_i(m))} \quad (6)$$

En la ec. (6), $\chi(b_i(k))$ cuenta el número de veces que el i -ésimo estado oculto emitió la observación k , y $\chi(b_i(m))$ cuenta el total de observaciones emitidas desde el i -ésimo estado oculto hacia cualquiera de las observaciones v_m , incluyendo la observación v_k .

La dimensión de la matriz de emisión B es igual a $n \times m$, donde n es el número de estados ocultos y m es el número de estados visibles definidos en el modelo.

Por último, la definición de la matriz de probabilidades iniciales $\pi_i = P(q_1 = S_i)$, representa la distribución de probabilidades de iniciar en el estado oculto S_i donde $1 \leq i \leq n$.

La dimensión de la matriz π es igual a $1 \times n$ y esta acotada por el número de estados ocultos del sistema [20].

Con la definición de los parámetros del modelo y sus características, para la aplicación que se plantea en este trabajo los estados ocultos del HMM, corresponden a la detección de actividades normales y actividades anormales, los estados visibles del modelo corresponden a la evidencia del movimiento con el que se forman las trayectorias y están directamente relacionados con las regiones r_i definidas para representar la información del movimiento como cadenas de símbolos s_i .

La relación, entre la evidencia del movimiento de los objetos en la escena en su representación más compacta como regiones r_i y los estados visibles del modelo v_k , puede expresarse como $r_i = v_k$. Esta relación, aplicada a los HMM, ofrece una ventaja significativa al momento de calcular los parámetros intrínsecos del modelo.

En la siguiente sección se realiza la validación de la propuesta y se resaltan las ventajas de usar la representación de movimiento como cadenas de símbolos compactas s_i .

Experimentación y resultados

Para la sección de experimentación y resultados se sigue el proceso descrito en el diagrama de flujo de la Fig. 1. La primera acción consiste en tomar una secuencia de imágenes con el objetivo de analizar la información de $I = \{I_1, I_2, I_3, \dots, I_n\}$. Con esta información ahora es posible extraer los patrones de movimiento P_{ij} al aplicar la ec. (1) a cada una de las imágenes de I , para extraer la información de movimiento M_O , como se observa en la Fig. 3.

Con la información de P_{ij} , se hace el seguimiento del objeto a lo largo de la escena para formar la trayectoria t_i .

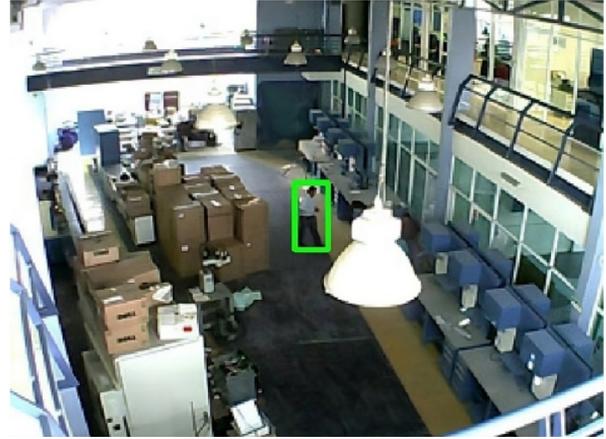


Fig. 3 Extracción de la información de los objetos en movimiento M_O

Por ejemplo, para el objeto de la Fig. 3, la trayectoria formada es $t_1 = \{[85, 88], [84, 89], [81, 90], [79, 91], [82, 94], [79, 93], [78, 94], [80, 96], [83, 98], [84, 100], [84, 102], [84, 103], [84, 106], [84, 110], [84, 113], [85, 115], [88, 118], [90, 121], [93, 124], [95, 127], [97, 131], [102, 135], [106, 140], [108, 143], [110, 146], [111, 151], [113, 155], [115, 159], [118, 166], [119, 171], [118, 176], [118, 181], [115, 185]\}$.



Fig. 4 Trayectoria formada con la información de movimiento para t_1

En la Fig. 4 se observa la recta que forma el objeto de la Fig. 3 al desplazarse por la escena. La recta que forma la trayectoria t_1 como resultado del seguimiento de un objeto en la escena está formada por las posiciones i, j , que se obtuvieron con los patrones de movimiento $P_{i,j}$. Esta información es posible al aplicar la ec. (1) sobre la secuencia del objeto a lo largo del tiempo.

Con la información de los objetos en movimiento, es posible conocer la temporabilidad del movimiento de los objetos al aplicar la ec. (2) a t_1 . Con esto se obtienen los píxeles asociados al movimiento para cada objeto, como se observa en la Fig. 5.

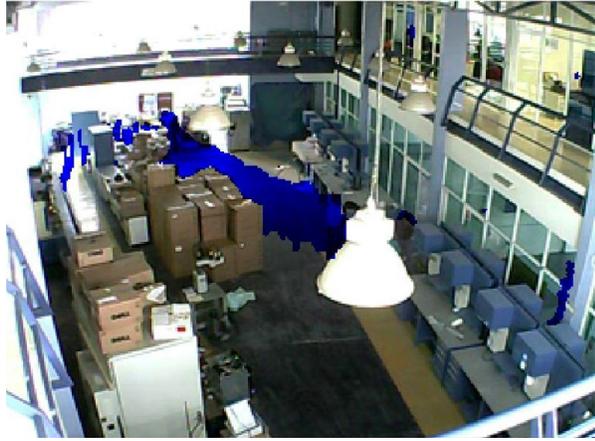


Fig. 5 Temporabilidad del movimiento M_T de la trayectoria t_i

Para conocer cómo se distribuye la pertenencia al movimiento en la escena en un instante de tiempo t , se analizan 5 minutos de información para obtener la evidencia de movimiento que puede observarse en la Fig. 6.

En las zonas resaltadas en azul de la Fig. 6 se observan las regiones con mayor ocurrencia de movimiento.

Una característica de S_M , es que proporciona evidencia de cómo está distribuido el movimiento en la escena. Esta distribución, para el proceso de generación de trayectorias, debe abarcar la mayor área de movimiento permisible en la escena, es decir todos los pasillos, escaleras y zonas sin obstrucciones son candidatas para iluminarse de azul y así representar la totalidad de zonas validas con pertenencia al movimiento en la escena.

Para el proceso de generación de regiones r , es necesario identificar cuáles fueron los píxeles $P_{i,j}$ con mayor pertenencia al movimiento, ya que estos nos darán información de las regiones que son más visitadas por los objetos. Esta información puede apreciarse al visualizar a S_M , como una superficie de valles y montañas, esta visualización de la información de la distribución del movimiento es necesaria para la visualización de las regiones r con mayor pertenencia al movimiento.



Fig. 6 Distribución del movimiento S_M para 5 minutos de video

La visualización de la superficie obtenida al aplicar la ec. (3), se muestra en la Fig. 7(a). Para generar las regiones r_i , se utiliza $Watershed(S_M)$, Este proceso implica aplicar una técnica de inundación a la superficie invertida de S_M , como se ilustra en la Fig. 7(b). A medida que se realiza esta operación, las regiones adyacentes se van segmentando, como se observa en la Fig. 7(c), lo que da como resultado una máscara S_M que contiene las regiones r_i las cuales están asociadas con áreas de alta actividad de movimiento.

Con las regiones r_i definidas, se forma $M_W = S_W \cup I_b$, lo que permite visualizar la relación que existe entre la escena y las nuevas regiones definidas, y así, formar las trayectorias s_i . El proceso completo se muestra en la Fig. 7.

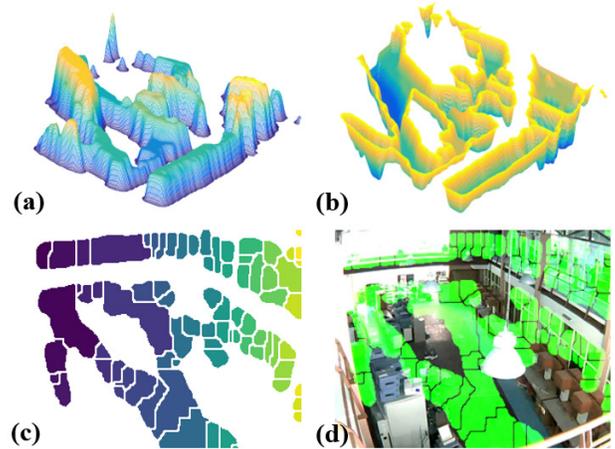


Fig. 7 Generación de las regiones r_i ; (a) información de las regiones r en S_M ; (b) S_M invertida; (c) Máscara de estados S_W ; (d) distribución de las zonas pertenecientes al movimiento M_W

Estas regiones tienen un símbolo definido, lo cual es una característica que permite la creación de nuevas trayectorias s_i con una representación acotada por el total de símbolos $r_i \in \Sigma$ asociados a las regiones r que fueron definidas por el proceso de segmentación, descrito en la sección anterior, esta definición de símbolos Σ correspondiente a cada región r_i puede observarse claramente en la Fig. 8.

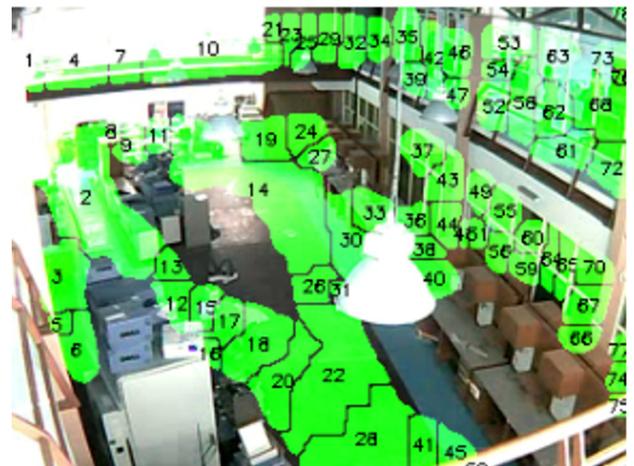


Fig. 8 Asignación de símbolos a cada r_i en M_W

$$B = \begin{bmatrix} 0.015 & \dots & 0 \\ 0.02 & \dots & 0.0001 \end{bmatrix}$$

Por último, se calcula la matriz de probabilidades iniciales, la cual nos da la probabilidad de iniciar en un estado S_i , esta matriz queda como sigue:

$$\pi = [1 \quad 0]$$

Ahora que se tiene la información de los parámetros del HMM, es posible realizar la detección de actividades. Se analizaron 610 trayectorias aplicando HMM, con estas trayectorias analizadas, se comprobó una efectividad del 83.6 % es decir 510 de las trayectorias fueron reconocidas satisfactoriamente.

Esta información demuestra la efectividad del modelo para el reconocimiento de actividades normales y anormales.

Conclusión

En este trabajo, se ha validado un método efectivo para la detección de actividades mediante la generación de trayectorias representadas como cadenas de símbolos. Este enfoque demuestra cómo se puede utilizar la información de regiones específicas r_i para formar trayectorias s_i , reduciendo significativamente los estados visibles en el Modelo Oculto de Markov (HMM).

Se ha comprobado que la propiedad $v_k = r_i$, descrita en secciones anteriores, permite una representación más eficiente del movimiento, disminuyendo el número de estados visibles del HMM. Para esta aplicación, se ha observado que la información necesaria para representar el movimiento varía considerablemente entre t_i y s_i . En particular para este trabajo, la representación del movimiento mediante s_i utiliza solo el 0.1% de la información disponible, comparado con la representación basada en las posiciones $P_{i,j}$.

El uso de una resolución de cámara de 240×320 píxeles implica que hay 76,800 posibles posiciones para el movimiento en una imagen. Sin embargo, al emplear la información de las regiones r_i , que se reducen a 78 elementos, se demuestra una notable optimización en la representación de los patrones de movimiento. Este enfoque no solo simplifica el modelo, sino que también mejora la eficiencia computacional y la capacidad de generalización del HMM en la detección de actividades.

En resumen, la metodología presentada aprovecha una representación más compacta y eficiente del movimiento, logrando una detección de actividades más precisa y con menor costo computacional. Este avance es significativo para aplicaciones en visión por computadora y reconocimiento de patrones, ofreciendo un marco más robusto y eficiente para el análisis de trayectorias y detección de actividades.

Referencias

- [1] I. Lee and K. Lee, "The Internet of Things (IoT): Applications, investments, and challenges for enterprises," *Bus. Horiz.*, vol. 58, no. 4, pp. 431–440, 2015.
- [2] J. Sengupta, S. Ruj, and S. Das Bit, "A Comprehensive Survey on Attacks, Security Issues and Blockchain Solutions for IoT and IIoT," *J. Netw. Comput. Appl.*, vol. 149, p. 102481, 2020.
- [3] L. Wu, Q. Wang, M. Jian, and Y. Q. B. Zhao, "A Comprehensive Review of Group Activity Recognition in Videos," SPRINGER, p. 17, 2021.
- [4] B. T. Morris and M. M. Trivedi, "A survey of vision-based trajectory learning and analysis for surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 8, pp. 1114–1127, 2008.
- [5] H. Zhao, C. Tan, E. J. O'Brien, N. Uddin, and B. Zhang, "Wavelet-based optimum identification of vehicle axles using bridge measurements," *Appl. Sci.*, vol. 10, no. 21, pp. 1–19, 2020.
- [6] M. L. Hoang, "Smart Drone Surveillance System Based on AI and on IoT Communication in Case of Intrusion and Fire Accident," *Drones*, vol. 7, no. 12, 2023.
- [7] S. Rao and P. S. Sastry, "Abnormal activity detection in video sequences using learnt probability densities," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 1, pp. 369–372, 2003.
- [8] J.-M. García-Huerta, H. Jiménez-Hernández, A.-M. Herrera-Navarro, T. Hernández-Díaz, and I. Terol-Villalobos, "Modelling dynamics with context-free grammars," *Video Surveill. Transp. Imaging Appl.* 2014, vol. 9026, no. 702, p. 902611, 2014.
- [9] J. Pohle, R. Langrock, F. M. Van Beest, and N. M. Schmidt, "Selecting the Number of States in Hidden Markov Models" arXiv: 1701.08673v2 [stat.ME] 14 Apr 2017."
- [10] A. W. M. Smeulders, D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, and M. Shah, "Visual tracking: An experimental survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 7, pp. 1442–1468, 2014.
- [11] T. Bouwmans, R. Advanced, S. Background, and T. Bouwmans, "Recent Advanced Statistical Background Modeling for Foreground Detection " 2015.
- [12] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1337–1342, 2003.

- [13] A. Yilmaz, O. Javed, and M. Shah, “*Object tracking*,” ACM Comput. Surv., vol. 38, no. 4, pp. 13-es, 2006.
- [14] M. Daldoss, N. Piotto, N. Conci, and F. G. B. De Natale, “*Learning and matching human activities using regular expressions*,” Proc. - Int. Conf. Image Process. ICIP, no. OCTOBER, pp. 4681–4684, 2010.
- [15] D. Hucke, M. Lohrey, and C. P. Reh, “*The smallest grammar problem revisited*,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9954 LNCS, no. 7, pp. 35–49, 2016.
- [16] and S. M. Weiming Hu, Tieniu Tan, Fellow, IEEE, Liang Wang, “*A Survey on Visual Surveillance of Object Motion and Behaviors*,” IEEE Trans. Syst. Man. Cybern., vol. 34, no. 4, pp. 4068–4079, 2004.
- [17] F. M. S. Beucher, *The Morphological Approach to Segmentation: The Watershed Transformation*, vol. 18, no. 4. 2001.
- [18] E. Fosler-lussier, “*Markov Models and Hidden Markov Models: A Brief Tutorial*,” vol. 1198, no. 510, 1998.
- [19] B. H. Juang, “*An Introduction to Hidden Markov Models*,” no. January, 1986.
- [20] LAWRENCE R. RABINER, “*A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*,” 1989.