VOLUMEN 37
XXX Verano De la Ciencia
ISSN 2395-9797
www.jovenesenlaciencia.ugto.mx

Reconocimiento de Acciones Mediante Integrales Múltiples

Recognition of Actions Using Multiple Integrals

Pedro Fernando Lara Rodríguez[1], Arturo Jaramillo Gil[2], Claudia E. Esteves Jaramillo[1]
[1] Departamento de Matemáticas, División de Ciencias Naturales y Exactas, Universidad de Guanajuato.
[2] Centro de Investigación en Matemáticas.
pf.lararodriguez@ugto.mx, jagil@cimat.mx, ce.esteves@ugto.mx

Resumen

La clasificación automática de acciones humanas es un problema importante en áreas tan diversas como salud, seguridad, entretenimiento, robótica, etc. El objetivo del proyecto es proponer una adaptación de la metodología de integrales múltiples con la finalidad de clasificar poses humanas mediante la función obtenida con la signatura de esta acción en el espacio tridimensional. Se busca generar una representación funcional de poses mediante un análisis espectral de la matriz de conectividad entre puntos característicos del cuerpo humano. Posteriormente se busca resumir el resultado de este análisis utilizando una adaptación adecuada de la metodología de integrales múltiples para obtener una función que represente cada acción, y con esto hacer una clasificación con un método conveniente.

Palabras clave: reconocimiento automático de acciones; métodos de clasificación.

Introducción

Poder clasificar acciones humanas de manera precisa es un problema de vital importancia en nuestro mundo moderno; las aplicaciones en el sector de la salud

El análisis de trayectorias o caminos en espacios multidimensionales es una herramienta fundamental en el estudio y clasificación de dichas acciones, así como una interesante aplicación de la Teoría de signatura de caminos (del inglés path signature). Esta ofrece una representación matemática robusta y compacta de trayectorias, permitiendo capturar la estructura y dinámica de un movimiento sin depender directamente de la parametrización temporal o de traslaciones en el espacio.

Estas propiedades hacen a esta herramienta especialmente útil en contextos donde la velocidad de ejecución o la frecuencia de muestreo pueden variar, como en la detección de acciones humanas con distintos métodos de captura, como lo es el MOCAP. Aunque existe bibliografía sobre el clasificamiento de acciones humanas con herramientas computacionales como el aprendizaje de máquina[1], la metodología descrita en este trabajo propone una alternativa a los métodos populares que se usan para clasificar.

Metodología

Basamos nuestra investigación en la base de datos de la Universidad Carnegie Mellon; esta tiene una base de datos de diversas acciones humanas como correr, trotar, caminar, saltar, interactuar con objetos, etc., que han capturado con equipo MOCAP.

[1] Tevet et al., 2022.

En este conjunto de datos, cada persona está representada por 31 puntos clave (joints) en tres dimensiones. Por lo tanto, el esqueleto en un cuadro individual puede considerarse como un punto en $\mathbb{R}^{31} \times \mathbb{R}^{3} \simeq \mathbb{R}^{93}$. Así, podemos ver una secuencia de esqueletos $x = (x_0, \dots, x_N) \subset \mathbb{R}^{93}$ proveniente de un videoclip con $N+1$ cuadros como una discretización de un camino continuo subyacente $P: [0,1] \to \mathbb{R}^{93}$, donde el esqueleto de cada cuadro x_i es una muestra del movimiento humano en un tiempo. Clasificar la acción ahora significa entender este camino.

Signatura de un camino

La signatura del camino es uno de los conceptos centrales en la teoría matemática de caminos rugosos (rough path theory). Es una herramienta poderosa que describe de manera eficiente y concisa un camino como el continuo $P: [0,1] \to \mathbb{R}^{93}$ mencionado anteriormente. La signatura del camino describe el camino independientemente de su parametrización temporal, pero conserva el orden en el que ocurren las cosas. Por lo tanto, es invariante, por ejemplo, a la tasa de muestreo, una propiedad deseable ya que una acción humana es invariante respecto a la velocidad de grabación del video o la rapidez con la que se realizó la acción. Además, la sifnatura es robusta ante cuadros faltantes.

La signatura se define como integrales repetidas del camino P. Si denotamos los d componentes del camino P en el tiempo t como $X = (X_t^1, \dots, X_t^d)$ (en nuestro caso $d = 93$), entonces las integrales de primer orden del camino P están dadas por

$$S(P)^i \coloneqq \int_{0<s<1} d\,X_s^i = X_1^i - X_0^i$$

para cada $i \in \{1, \dots, d\}$. De manera similar, las integrales de segundo orden están dadas por

$$S(P)^{i,j} \coloneqq \iint_{0<t_1<t_2<1} d\,X_{t_1}^i dX_{t_2}^j$$

para cada $j, i \in \{1, \dots, d\}$.. Repitiendo este proceso, para cualquier entero $k \geq 1$, la integral repetida de orden k está dada por

$$\mathrm{S}(P)^{i_1,\dots,i_k} \coloneqq \int \cdots \int_{0<t_1<\cdots<t_k<1} \mathrm{d}X_{t_1}^{i_1} \dots dX_{t_k}^{i_k}$$

para cada colección de índices $i_1, \dots, i_k \in \{1, \dots, d\}$. La signatura del camino P se define ahora como la sucesión de todas estas integrales repetidas:

$$Sig(P) \coloneqq (1, S(P)^1, \dots, S(P)^d, S(P)^{1,1}, S(P)^{1,2}, \dots).$$

El número de integraciones k sobre el camino P se conoce comúnmente como el nivel de la signatura. Cada una de las k integraciones se realiza sobre un eje de coordenadas. El conjunto de índices $i_1, \dots, i_k$ describe el orden y los ejes sobre los cuales se realizan estas integraciones. Para un camino en d dimensiones, existen d^k posibles combinaciones de índices para estas integraciones. Esto implica que el número de términos en la sucesión $Sig(x)$ crece exponencialmente con el nivel k.

Dado que la signatura completa es una sucesión infinita de integrales repetidas, en la práctica se elige una truncación finita, calculando las integrales solo hasta un nivel finito N; es decir, $0 \leq k \leq N$ en la ecuación anterior. A esto se le llama la signatura de nivel N, y se denota como $Sig^N(x)$. Dado que el número de integrales crece exponencialmente con cada nivel adicional, usualmente solo se calculan pocos niveles (menos de 10).

En la práctica, para calcular la signatura a partir de una secuencia de esqueletos $x = (x_0, \dots, x_N) \subset \mathbb{R}^{93}$, se supone que el camino P es la interpolación lineal por tramos de dicha secuencia, es decir, que $P\left(\frac{i}{N}\right) = x_i$ para cada $i \in \{0, \dots, N\}$, y lineal entre los puntos de muestreo.

Más detalles sobre la signatura de caminos se pueden encontrar en (Lyons & McLeod, 2025).

Visualización de Signaturas de Caminos Unidimensionales

Habiendo definido formalmente la signatura de un camino, para ayudarnos a comprender mejor sus propiedades exploraremos algunos ejemplos de caminos y sus respectivas signaturas. Consideremos primero los siguientes caminos unidimensionales, cada uno definido en el intervalo $[a,b] = [0,1]$:

- $W_t \coloneqq t.$
- $X_t \coloneqq 2t$
- $Y_t \coloneqq \sin\left(\frac{\pi}{2}t\right) + 1$
- $Z_t \coloneqq \begin{cases} -t, & \text{si } t < 0.5, \\ 3t - 2, & \text{en otro caso.} \end{cases}$

La figura 1 muestra la signatura hasta nivel $K = 3$; excluyendo el término de nivel cero, la signatura truncada consiste en $1^1 + 1^2 + 1^3 = 3$ términos, cuyos índices son respectivamente $(1), (1,1), (1,1,1)$.

Para facilitar la comprensión, en lugar de simplemente calcular $S(\cdot)_{a,b}$ consideramos la signatura $S(\cdot)_{a,t}$ como una función de $t \in [a,b]$. Además, graficamos la primera derivada del camino con respecto a t para complementar la figura.

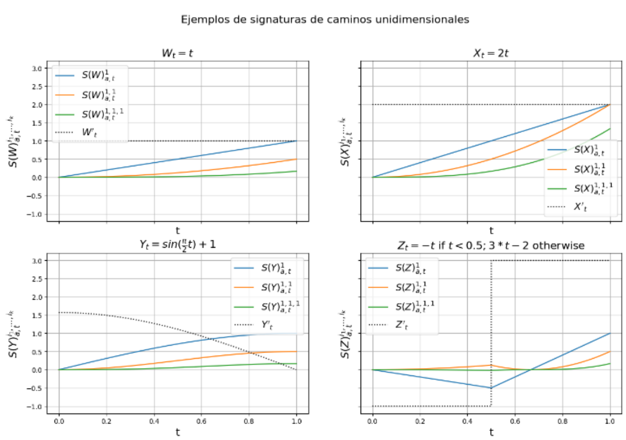

Figura 1. Visualización del cálculo de signaturas para distintos caminos.

Transformaciones de los Caminos

El término k-ésimo de la signatura es del orden de $\mathcal{O}\left(\frac{1}{k!}\right)$, por lo que muchas veces esto no representa un problema, ya que los primeros N niveles de la signatura contienen los términos más grandes. Sin embargo, si la función que intentamos aprender depende fuertemente de términos de orden superior, entonces se ha perdido información importante.

Para abordar este problema, se han introducido diversas transformaciones de caminos que ayudan a capturar propiedades importantes del camino en términos de menor orden dentro de la signatura. Una descripción más amplia de las distintas transformaciones de caminos y sus efectos se puede encontrar en (Terry Lyons, Andrew D. McLeod, 2022). Aquí describimos la selección de transformaciones de caminos que consideramos para nuestro método.

Recorte y Normalización

Estas no son transformaciones específicas de la signatura del camino. Como es común en aprendizaje automático, recortamos y normalizamos las coordenadas de los puntos clave para que estén dentro del intervalo $[-1,1]$, recortando al cuadro delimitador de los puntos clave, extendiendo el lado más corto de forma simétrica hasta formar un cuadro, y luego escalando dicho cuadro al intervalo objetivo.

Camino Espacio-Temporal

Existen dos tipos de transformaciones de caminos: espaciales y temporales, que asumen una estructura distinta del camino.

Las transformaciones como el recorte y la normalización anteriores son transformaciones espaciales, ya que operan sobre las coordenadas espaciales en cada instante de tiempo, no sobre una serie temporal. De forma similar, es posible ver un conjunto de n puntos clave como n muestras de un camino en $\mathbb{R}^2$, y luego tomar la signatura de cada uno de esos caminos espaciales de longitud n. Por ejemplo, se pueden tomar todos los pares de joints posibles de puntos clave dentro de un solo cuadro y considerar cada par como un camino con 2 puntos de muestra. Así, las transformaciones espaciales esperan una secuencia con estructura `[cuadro][punto clave][coordenadas]`, para recorrer los cuadros y procesar las coordenadas espaciales.

Por otro lado, las transformaciones temporales operan sobre todas o algunas de las coordenadas espaciales a lo largo del tiempo. Por esta razón, estas transformaciones esperan una secuencia con la estructura `[elemento][cuadro][coordenadas]`, lo que se conoce como un camino espacio-temporal. Por lo tanto, después de aplicar todas las transformaciones espaciales que se deseen usar, se debe transformar de un camino espacial a un camino espacio-temporal antes de continuar con las transformaciones temporales.

Transformación con Múltiples Retardos

Para codificar mejor las dependencias temporales, se puede usar la transformación de múltiples retardos del camino. Para esto, el camino se amplía con varias versiones retrasadas de sí mismo, de manera que en cualquier punto del camino esté disponible la información de varios pasos de tiempo. Las copias individuales del camino se rellenan con ceros según sea necesario.

Desintegración Temporal en Signaturas de Caminos Diádicos

Para caminos largos, a menudo puede ser más eficiente (con un rendimiento comparable pero usando un conjunto de características más pequeño) considerar términos de bajo nivel de las signaturas en intervalos de tiempo intermedios, en lugar de términos de orden alto de la signatura del camino completo. Una forma eficiente de hacer esto es tomar signaturas dentro de una estructura jerárquica diádica.

En cada nivel diádico, dividimos los intervalos del nivel anterior a la mitad. Es decir, el “nivel 0” contiene todo el camino, el “nivel 1” contiene sus dos mitades, el “nivel 2” sus cuatro cuartos, etcétera.

Para evitar cortar eventos que ocurren cerca de las divisiones en subintervalos, se pueden considerar intervalos diádicos superpuestos, por ejemplo:

- Tres mitades en el nivel 1: $[0,T/2]$, $[T/4,3T/4]$ y $[T/2,T]$,
- Siete cuartos en el siguiente nivel: $[0,T/4]$, $[T/8,3T/8]$, ..., $[3T/4,T]$, etcétera.

Abajo ilustramos el uso de la función LieIncrementStream de la librería RoughPy con el propósito de calcular una signatura truncada de una secuencia dada de incrementos. Supongamos que $N,d \in \mathbb{Z}_{\geq 1}$ y que tenemos una secuencia de incrementos $x \in \mathcal{S}(\mathbb{R}^d)$ dada por $x = (I_1, \ldots, I_N)$ con $I_1, \ldots, I_N \in \mathbb{R}^d$. El código siguiente calcula la signatura de x truncada a profundidad $k \in \mathbb{Z}_{\geq 1}$.

Prototipo de Clasificación

Para la tarea de clasificación se necesita manipular correctamente la base de datos considerada (https://mocap.cs.cmu.edu/ y https://sites.google.com/a/cgspeed.com/cgspeed/motion-capture?authuser=0). La idea es generar una base de datos grande con suficientes animaciones de acciones humanas específicas (como sería caminar, correr y saltar).

Una vez que se pueda acceder a dicha base de datos de manera eficiente, se calculan las signaturas de cada dato con el método descrito anteriormente.

Después, con las signaturas obtenidas, se emplea un método de clasificación conveniente para obtener el resultado. El artículo de referencia usa métodos lineales y de aprendizaje de máquina para esta tarea; en dicha bibliografía se explica por qué es conveniente usar estos métodos.

Conclusiones

Mediante la aplicación de transformaciones espaciales y temporales, es posible capturar información relevante del movimiento incluso en niveles bajos de la signatura, lo que permite una representación eficiente y discriminativa.

Las visualizaciones realizadas sobre caminos unidimensionales ayudaron a interpretar de forma más intuitiva las propiedades de las signaturas truncadas, mostrando cómo distintas trayectorias generan diferentes patrones en sus componentes. Además, las transformaciones como la multi retardada y la desintegración temporal diádica demostraron ser estrategias efectivas para enriquecer la representación sin recurrir a términos de orden elevado, lo que favorece su uso en tareas de aprendizaje automático.

Las signaturas de caminos destacan por su capacidad para abstraer patrones de forma independiente del tiempo, lo que las convierte en una herramienta prometedora en aplicaciones donde la temporalidad es variable o irrelevante. Esta metodología tiene un gran potencial en áreas como el reconocimiento de acciones, análisis biométrico, interfaces interactivas y diagnóstico médico, entre otras.

Finalmente, no se realizaron experimentos y comparaciones numéricas debido a que nuestro tiempo de investigación sólo nos permitió desarrollar un prototipo de esta metodología. No obstante, se espera que tenga un desempeño bueno en comparación a otros métodos de clasificación pues la naturaleza de los cálculos de signaturas de caminos es consistente y determinista.

Este enfoque abre la puerta a futuros desarrollos, como el diseño de modelos de aprendizaje profundo adaptados a signaturas, o la integración con técnicas de atención y clasificación basadas en características geométricas y temporales aprendidas.

Referencias

Lyons, T., & McLeod, A. D. (2025). Signature Methods in Machine Learning. *arXiv,* 7. https://doi.org/10.48550/arXiv.2206.14674

Tevet, G., Raab, S., Gordon, B., Shafir, Y., Cohen-Or, D. & Bermano, A. H. (2022). Human Motion Diffusion Model. *arXiv,* 2. https://doi.org/10.48550/arXiv.2209.14916

CMU Graphics Lab. (2003). CMU Graphics Lab Motion Capture Database. https://mocap.cs.cmu.edu/info.php