Prior Espacial-Temporal Difuso (PETD) para la corrección del Seguimiento Articular en Natación

David Muñoz Valero[1][0000−0002−2509−9911], Juan Moreno Garcia[1][0000−0003−2430−145X], Eusebio Angulo[2][0000−0001−6731−3229], and Alejandro Lopez Hernandez[3][0009−0006−2890−5292]

[1] Escuela de Ingeniería Industrial y Aeroespacial de Toledo (UCLM) david.munoz@uclm.es,juan.moreno@uclm.es
[2] Escuela Superior de Informática de Ciudad Real (UCLM) eusebio.angulo@uclm.es
[3] Facultad de Ciencias del Deporte de Toledo (UCLM) alejandro.lopez@uclm.es

Resumen La estimación de pose en entornos acuáticos sufre de inestabilidad temporal debida a oclusiones, turbulencia y fluctuaciones lumínicas. Para tratar esto, en este trabajo en progreso se propone un prior espacio-temporal difuso interpretable, que genera para cada punto y cada instante información sobre dónde es probable que se ubiquen las marcas en el siguiente *frame* mediante un conjunto difuso. Esto se consigue mediante un método difuso que utiliza la trayectoria reciente y la estimación actual de la marca.

Keywords: estimación de pose · representación espacio-temporal difusa · natación.

1. Introducción

El análisis de pose en natación ha ganado notable impulso gracias a los avances en estimación de pose humana. No obstante, el entorno acuático introduce perturbaciones específicas, como salpicaduras y oclusiones, que degradan la estabilidad de las marcas (*keypoints*) detectados fotograma a fotograma que compromete el cálculo posterior, generando una incertidumbre sobre las posiciones de las marcas. Para atenuar el ruido se han propuesto múltiples filtros temporales: el filtro de Kalman [1], el adaptativo One-Euro [2], el suavizador de Whittaker-Eilers [3] y la Regresión por Procesos Gaussianos (GPR) [4]. Aunque todos reducen fluctuaciones de alta frecuencia, presentan una limitación compartida: ninguno construye de manera explícita un *prior espacial*, entendido como dónde es probable que aparezca el *keypoint* (marcas o puntos articulares del nadador) en el siguiente fotograma. En este trabajo en progreso se propone un prior espacio-temporal difuso interpretable, que genera para cada punto y cada instante información sobre dónde es probable que se ubiquen las marcas en el siguiente *frame* mediante un conjunto difuso. Esto se consigue mediante un método difuso que utiliza la trayectoria reciente y la estimación actual de la marca.

2. Propuesta inicial

La propuesta se fundamenta en tres pilares: una representación difusa de la trayectoria en una ventana temporal de k fotogramas, operadores espaciales difusos basados en *Region Connection Calculus* difusos (Fuzzy RCC-8) para evaluar la coherencia entre *keypoints*, y operadores temporales difusos inspirados en lógica temporal para detectar eventos destacados en la secuencia. El procesamiento de cada keypoint sigue un ciclo de tres pasos:

Paso 1. Construcción de una representación difusa de las k posiciones vecinas de la marca generando el conjunto de trayectoria $\mathcal{T}_{i,t}$: En lugar de considerar únicamente el fotograma inmediatamente anterior, se define una ventana deslizante de tamaño k centrada en el instante t. Para cada keypoint i, las posiciones de los k fotogramas anteriores y posteriores se representan como un conjunto difuso de trayectoria $\mathcal{T}_{i,t}$. Este conjunto captura la noción de dónde se ha estado y hacia dónde se dirige de forma gradual, sin asumir una dinámica lineal rígida. La ventana permite que el sistema tolere breves oclusiones o saltos de detección, ya que la información de los fotogramas vecinos actúa como contexto. Esta ventana puede ayudar a predecir dónde deberían encontrarse *keypoints* ocluidos.

Paso 2. Uso de operadores espaciales difusos entre los *keypoints* relacionados anatómicamente, produciendo restricciones en el espacio (Sección **??**):

Region Connection Calculus (RCC-8) proporciona un conjunto de relaciones topológicas entre regiones [5]. En su versión difusa, cada *keypoint* se modela como una región difusa en el espacio de la imagen. Inicialmente se consideran tres relaciones fundamentales:

- **DC** (*Disconnected*): grado en que dos regiones difusas están separadas. Permite detectar si dos *keypoints* anatómicamente próximos (por ejemplo, codo y muñeca) han divergido anormalmente.
- **EC** (*Externally Connected*): grado en que dos regiones se tocan en el borde sin solaparse. Es útil para evaluar la continuidad de la cadena cinemática sin forzar superposiciones.
- **PO** (*Partially Overlapping*): grado de solapamiento entre dos regiones difusas. Detecta situaciones de oclusión parcial donde dos *keypoints* se confunden en la imagen.

Para cada par de *keypoints* anatómicamente conectados (i, j) —por ejemplo, codo y muñeca—, se evalúan las relaciones RCC-8 difusas sobre sus regiones espaciales asociadas. Los grados μ_{DC}, μ_{EC} y μ_{PO} capturan el grado de coherencia topológica entre ambas marcas, actuando como restricciones esqueléticas que complementan el prior temporal. A continuación se mostrará un ejemplo considerando que la distancia entre dos puntos es la métrica euclídea. Sean las posiciones de referencia para el codo y muñeca izquierda en el *frame* t:

- Codo: $z_{\text{codo}} = (100, 200)$ píxeles (se denotará por px.)
- Muñeca: $z_{\text{muñeca}} = (130, 200)$ px.
- Distancia anatómica: $d_t = 30$ px (se asume que es la distancia correcta).

Y sean las siguientes medidas del detector en el *frame* $t+1$:

- Codo detectado: $z'_{\text{codo}} = (102, 205)$ (ligero movimiento, coherente)
- Muñeca detectada: $z'_{\text{muñeca}} = (180, 210)$ (salto brusco de 50 px)
- Distancia anatómica: $d_{t+1} \approx 78$ px.

Basándose en lo anterior, el problema se puede modelar de forma difusa de la siguiente manera. Se persigue que las regiones estén algo conectadas (DC muy bajo), la conexión externa de las fronteras sea alta (EC alto) y el solapamiento sea bajo (PO muy bajo). Para ello se modela cada marca como una región difusa de radio $\sigma = 16$ px, lo que provoca que se toquen levemente considerando la distancia anatómica correcta entre puntos.

En el *frame* $t+1$ se tiene que $78 \gg 32$, por lo que las regiones no se solapan; por tanto se tiene:

- $\mu_{\text{DC}} = 1$ (grado de desconexión total porque no hay solapamiento)
- $\mu_{\text{EC}} = 0$ (no se tocan las regiones difusas)
- $\mu_{\text{PO}} = 0$ (sin solapamiento)

Paso 3. Se aplican los operadores temporales difusos (F, G, U) sobre la secuencia de conjuntos difusos, detectando eventos destacados que modulan la confianza de la estimación final. La salida del ciclo es una posición corregida que integra la coherencia esquelética y la evidencia temporal: Inspirados en los operadores de lógica temporal [6], se definen tres operadores difusos sobre la ventana de k fotogramas:

- **Finally / Eventually (F)**: mide el grado con el que una propiedad (por ejemplo, "el keypoint está en reposo") se cumple en algún momento de la ventana. Permite detectar eventos transitorios como el inicio de una brazada.
- **Globally (G)**: mide el grado con el que una propiedad se mantiene durante toda la ventana. Es útil para identificar segmentos de trayectoria estables que merecen mayor confianza.
- **Until (U)**: mide el grado con el que una propiedad se sostiene hasta que otra ocurre. Permite modelar transiciones como "el keypoint se mantiene cerca de la línea de agua hasta que emerge".

Estos operadores se evalúan sobre la secuencia de conjuntos difusos $\mathcal{T}_{i,t}$ produciendo un conjunto de eventos destacados $\mathcal{E}_{i,t}$ que caracterizan la evolución temporal de cada keypoint.

A modo de ejemplo, considérese la detección del evento de inicio de tracción en la brazada de crol. El operador U (Until) evalúa sobre la ventana de k fotogramas el grado con el que la muñeca se ha mantenido en la zona de entrada hasta iniciar la tracción. Si la marca permanece en la zona de entrada durante varios *frames* consecutivos y posteriormente desciende con velocidad creciente, el valor de activación será alto, señalando con firmeza la transición al evento. Este evento detectado modula la confianza del prior espacial para los fotogramas posteriores, confirmando que la trayectoria observada sigue el patrón biomecánico esperado de la brazada. Este tipo de funcionamiento de descubrimiento de eventos puede generalizarse para detectar otros instantes destacados que los expertos en natación identifiquen como relevantes.

3. Conclusiones

En este trabajo en progreso se ha articulado una propuesta de prior espacio-temporal difuso para la corrección de marcas articulares en el seguimiento articular de la natación. Se están utilizado tres conceptos claves para conseguir la corrección, una representación difusa de la trayectoria mediante ventanas de k fotogramas, operadores espaciales RCC-8 difusos para evaluar la coherencia esquelética (entre otras cosas), y operadores temporales difusos inspirados en lógica temporal para detectar eventos destacados en la secuencia. La combinación de estos elementos puede permitir generar un prior espacial interpretable que anticipa dónde esperar una marca en el siguiente fotograma, fusionando esa expectativa con la nueva medición de forma gradual.

Acknowledgments

Este trabajo ha sido financiado por la ayuda SBPLY/24/180225/00003, concedida por la Junta de Comunidades de Castilla-La Mancha. Los autores quieren agradecer al departamento de Tecnologías y Sistemas de Información (TSI) de la Universidad de Castilla la Mancha (UCLM) y al proyecto ATENEA: Aprendizaje Profundo, Metaheurísticas y Conocimiento Experto para Sistemas Multiagente en Entornos Ferroviarios Liberalizados Parcialmente Observables (Referencia 2025-GRIN-38352), cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER), la financiación necesaria para la difusión de este trabajo.

Referencias

1. Welch, G., Bishop, G.: An introduction to the Kalman filter. Technical Report TR 95-041, Univ. of North Carolina at Chapel Hill (1995)
2. Casiez, G., Roussel, N., Vogel, D.: The One-Euro filter: A simple speed-based low-pass filter for noisy input in interactive systems. In: Proc. CHI, pp. 2527–2530. ACM (2012)
3. Eilers, P.H.C.: A perfect smoother. Anal. Chem. **75**(14), 3631–3636 (2003)
4. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. MIT Press (2006)
5. Schockaert, S., De Cock, M., Cornelis, C, Kerre, E.E.: Fuzzy region connection calculus: Representing vague topological information. International Journal of Approximate Reasoning, Vol. 48, 1, pag. 314-331, 2008
6. Frigeri, A., Pasquale, L., Spoletini, P.: Fuzzy Time in Linear Temporal Logic. ACM Transactions on Computational Logic, Vol. 15 (2014)