

Tracciamento 3D della palla da punti di vista multipli nella pallavolo

Sergio Povoli^{1,†}, Luigi Riz^{1,†}, Davide Boscaini¹, Stefano Messelodi¹ and Fabio Poiesi¹

¹*Technologies of Vision (TeV), Fondazione Bruno Kessler (FBK), Italy*

Abstract

Il progetto 5VREAL si propone di realizzare un sistema per l'estrazione automatica di statistiche nel contesto della pallavolo, al fine di fornire strumenti analitici sia agli allenatori, per una valutazione dettagliata delle dinamiche di gioco, sia agli spettatori, per arricchire la loro esperienza di fruizione. Per raggiungere questo obiettivo, adottiamo un approccio innovativo basato sull'utilizzo di algoritmi di Intelligenza Artificiale e tecnologie di Visione 3D, integrati all'interno di una pipeline che sfrutta le potenzialità del 5G e del cloud computing. In questo lavoro, esaminiamo le ragioni che motivano la nostra ricerca e presentiamo un'analisi dettagliata dell'architettura per il tracciamento in 3D della palla sviluppata fino a questo punto.

Keywords

sport analytics, object detection, object tracking, real world digitalization

1. Introduzione

L'intelligenza artificiale (IA) sta rivoluzionando numerosi settori, e lo sport non fa eccezione. L'ambiente sportivo, con le sue regole precise, la puntuale misurazione dei miglioramenti delle performance e la costante ricerca di miglioramenti, offre un terreno fertile per l'applicazione dell'IA. Dal miglioramento delle prestazioni dei giocatori all'ottimizzazione delle strategie di gioco, dall'arbitraggio alle previsioni delle partite, l'IA sta assumendo un ruolo sempre più centrale nello sviluppo e nell'innovazione nel mondo dello sport. L'introduzione dell'intelligenza artificiale nello sport non solo permette di migliorare le prestazioni e l'esperienza degli attori direttamente coinvolti o degli spettatori, ma offre anche nuove opportunità in settori connessi, come i giochi per computer e la robotica. Un esempio emblematico di questo fenomeno è rappresentato dalla partnership tra DeepMind e il Liverpool Football Club, che ha portato alla creazione di TacticAI [1], un sistema che fornisce insights tattici sui calci d'angolo mediante l'utilizzo di tecnologie predittive e generative basate sull'IA.

Il progetto 5VREAL si inserisce in questo contesto, mirando a introdurre l'intelligenza artificiale nel mondo della pallavolo, un ambiente che tradizionalmente si è evoluto sulla base dei dati, ma che spesso ha incontrato difficoltà nell'acquisizione di informazioni dettagliate e significative. L'utilizzo dell'IA offre l'opportunità di superare queste sfide, consentendo un'analisi più approfondita e una comprensione più completa e quantificata

Ital-IA 2024: 4th National Conference on Artificial Intelligence, organized by CINI, May 29-30, 2024, Naples, Italy

[†]Questi autori hanno contribuito in egual misura.

✉ spovoli@fbk.eu (S. Povoli); luriz@fbk.eu (L. Riz);

dboscaini@fbk.eu (D. Boscaini); messelod@fbk.eu (S. Messelodi);

poiesi@fbk.eu (F. Poiesi)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

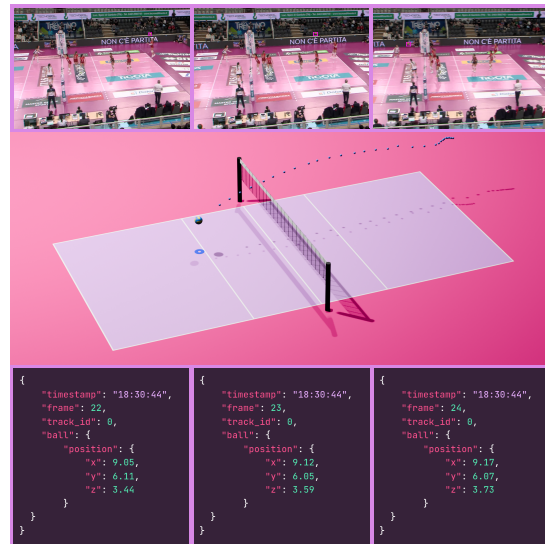


Figure 1: Ricostruzione 3D della posizione della palla e della sua traiettoria percorsa durante una battuta. Questa ricostruzione è stata generata in maniera automatica sfruttando i flussi video raccolti da 3 smartphone, calibrati e sincronizzati, durante una gara ufficiale della Lega Pallavolo Serie A Femminile.

delle dinamiche di gioco.

1.1. Obiettivi progetto

Scopo principale del progetto 5VREAL è l'estrazione automatica in real-time di statistiche e analytics riguardanti la pallavolo, tramite algoritmi di intelligenza artificiale. Il tutto deve essere possibile utilizzando un'infrastruttura agile e facilmente installabile, che possa essere

trasportata in luoghi diversi ed adattata a diverse situazioni senza difficoltà. Per questo, 5VREAL prevede l'utilizzo di due o più smartphone abilitati alla connettività 5G per l'acquisizione di video da punti di vista diversi. I dati raccolti da tali dispositivi vengono trasmessi in cloud dove, tramite algoritmi di intelligenza artificiale, vengono estratte le diverse statistiche previste dal progetto. La centralizzazione di tutti i dati in cloud irrobustisce il concetto di infrastruttura portabile e permette una distribuzione più efficiente degli analytics ai diversi attori che li presenteranno ad allenatori e spettatori.

Utilizzando la pipeline proposta in 5VREAL, sarà possibile ottenere il tracciamento del pallone in 3D, assieme ad ulteriori parametri di interesse per spettatori ed allenatori, come ad esempio la velocità della palla in battuta, l'elevazione del giocatore in schiacciata o il ritmo di gioco. È prevista anche l'estrazione di statistiche più complesse, come ad esempio il riconoscimento automatico del tipo di azione messa in atto da una squadra o l'estrazione di pattern frequenti nel susseguirsi dei vari fondamentali. Per raggiungere tutto ciò, 5VREAL prevede use case diversi, che possano soddisfare le necessità degli allenatori, più interessati ad aspetti tecnico-tattici, e quelle degli spettatori, più incentrate sul coinvolgimento tramite statistiche di facile fruizione. Lo use case "coach" prevede lo sviluppo di un app mobile, utilizzabile in fase di allenamento e in partita, che fornisca allo staff tecnico un'ampia gamma di informazioni e statistiche che descrivano in real time e in modo misurabile le prestazioni degli atleti. Lo use case "fan engagement" si concentra invece sulla presentazione agli spettatori di statistiche di facile comprensione, tramite l'utilizzo di grafiche sovrimpresse alle immagini in diretta di una partita.

L'estrazione automatica dei KPI identificati nel contesto del progetto richiede l'integrazione di diverse componenti hardware e software in un'unica pipeline, come rappresentato nella Figura 2. La pipeline consiste in due o più telecamere, calibrate e sincronizzate, opportunamente posizionate nei pressi del campo da gioco, che acquisiscono e trasmettono i dati attraverso la rete 5G verso un'unità di elaborazione remota. All'interno di questa struttura, i dati ricevuti vengono decodificati e la loro qualità è aumentata attraverso l'uso di algoritmi di intelligenza artificiale. Il prodotto migliorato viene quindi dato in input sia al modulo di IA dedicato al tracciamento in tempo reale della palla in tre dimensioni, che alla componente destinata al riconoscimento delle pose, delle azioni e dei giocatori. Quest'ultimo modulo riceve come input addizionale le informazioni estratte dal modulo di tracciamento della posizione della palla in 3D. Successivamente al processo di elaborazione, una piattaforma di raccolta e analisi dei dati è al servizio di un'applicazione che soddisfa i requisiti dello use case "coach", garantendo una efficace visualizzazione dei KPI. Contemporaneamente, una piattaforma di produzione

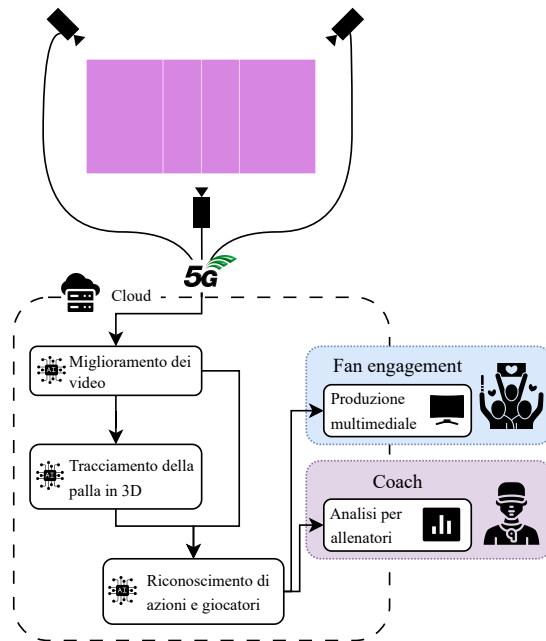


Figure 2: Pipeline ad alto livello del progetto 5VREAL per analisi real time delle pallavolo. Il sistema si basa su K camere calibrate e sincronizzate che acquisiscono la scena inviando i flussi video tramite la rete 5G ad un'unità computazionale esterna. Questi stream video vengono migliorati grazie a tecniche IA prima di essere processati da algoritmi intelligenti per l'estrazione del tracciamento della palla nel mondo 3D e il riconoscimento di azioni, pose e giocatori. Il prodotto dei moduli IA è poi utilizzato per soddisfare entrambi gli use cases tramite un modulo di raccolta e analisi dati per gli allenatori ed un sistema di produzione multimediale per la creazione di video aumentati per spettatori e fan.

multimediale, alimentata dalle immagini televisive, permette l'arricchimento dei video in diretta con grafiche e dati, al fine di migliorare il coinvolgimento dei fan.

1.2. 3D ball tracking

Il modulo di tracciamento automatico della palla in tre dimensioni e di estrazione dei parametri che ne descrivono il moto in modo robusto ricopre un ruolo fondamentale all'interno della pipeline di 5VREAL. A tale scopo, questa componente richiede l'utilizzo di flussi video provenienti da diverse camere, per le quali deve essere garantita la calibrazione geometrica e la sincronizzazione a livello temporale. Per ogni flusso, il modulo di visione sviluppato nel contesto di 5VREAL estrae una o più rilevazioni del pallone, tramite l'utilizzo di tecniche di object detection. Queste rilevazioni vengono utilizzate per effettuare triangolazioni e produrre ipotesi di palloni localizzati nel mondo tridimensionale, chiamate *ball proposals*. Analiz-

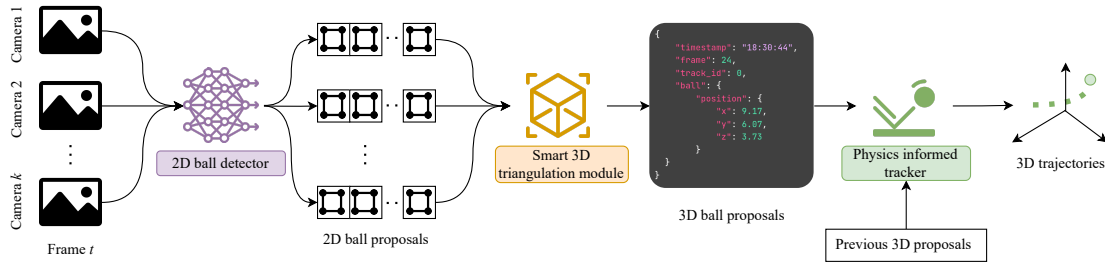


Figure 3: Il modulo di tracciamento della palla in 3D è composto principalmente da tre componenti. Il primo modulo produce delle proposte di palloni nel mondo 2D per le immagini in analisi. Queste proposte vengono elaborate dal secondo modulo, lo “Smart 3D triangulation module”, il cui obiettivo è triangolare le rilevazioni 2D nello spazio 3D, sfruttando la calibrazione delle telecamere. Si ottengono così proposte di posizionamento della palla in 3D. Correlando temporalmente varie proposte della posizione del pallone in 3D, il terzo modulo sviluppato si occupa di creare associazioni e ricostruire una traiettoria, facendo uso anche di un modello fisico del moto della palla. L’output finale della pipeline consiste, ad ogni frame, nella posizione di uno o più palloni appartenenti a una o più traiettorie.

zando le rilevazioni 3D estratte ad ogni istante temporale, è poi possibile filtrare le rilevazioni errate e correlare rilevazioni dello stesso pallone in timestamp consecutivi, ottenendo quale output finale il tracciamento in tre dimensioni del pallone. Per ogni traccia sono anche fornite le statistiche richieste, come velocità di picco o altezza massima raggiunta.

2. Metodo di tracciamento del pallone in 3D

La pipeline adottata in 5VREAL per il tracciamento in tre dimensioni del pallone è riportata in Fig. 3. Ad ogni istante di tempo t , la componente software riceve K diverse immagini, una per ogni dispositivo di acquisizione. L’allineamento temporale può essere assicurato tramite tecniche di sincronizzazione, come il *genlock*, oppure l’acquisizione può avvenire in modo asincrono per ogni camera. In quest’ultimo caso è necessario lo sviluppo di un gestore di camere che, ad ogni istante t , fornisca per ogni dispositivo di acquisizione l’immagine il cui timestamp di acquisizione minimizza la distanza con t . Le K immagini sono quindi processate con algoritmi per la rilevazione del pallone, che produce in output un insieme di *2D ball proposals* nella forma di bounding boxes. Uno dei possibili metodi applicabili per la rilevazione della palla in 2D si basa sull’utilizzo di algoritmi classici di computer vision. Ad esempio, dopo aver localizzato gli oggetti in movimento tramite *frame-by-frame difference* o *background difference*, si possono utilizzare tecniche di morfologia e applicare filtri basati su forma e dimensione degli oggetti oppure applicare opportune trasformate (e.g. Hough transform [2]) per selezionare ipotesi di rilevazione della palla. Un’alternativa all’utilizzo di tali metodi è l’applicazione di reti neurali dedicate ad *object detection* [3, 4], opportunamente adattate tramite



Figure 4: Esempio di *ball proposal* estratta con il *2D ball detector*, su immagini pubblicamente disponibili online. In questo esempio è riportato solamente il *proposal* con la confidenza maggiore. I frame riportati in figura sono disponibili su Youtube (Canale Youtube Volleyball Watchdog: <https://www.youtube.com/@VolleyballWatchdog>).

fine tuning per il rilevamento del pallone. All’interno del progetto 5VREAL, la rilevazione del pallone è effettuata utilizzando YOLOv8 [5]. Tale modello è allenato sul VolleyVision dataset [6], in modo da ottenere una specializzazione sul rilevamento della palla. Fig. 4 riporta alcuni risultati ottenuti applicando il modello proposto su video disponibili online.

L’output di tale elaborazione è una lista di *ball proposals* per ognuna delle K immagini in input. Tali proposte

sono filtrate tramite sogliatura sulla confidenza e tenendo conto di eventuali vincoli geometrici fra viste diverse. Le proposte sopravvissute a tale filtraggio vengono date in input ad un modulo di triangolazione, che combina i dati provenienti dalle K differenti viste, tenendo in considerazione i parametri di calibrazione geometrica di ognuna delle camere. Tale processo di triangolazione avviene tenendo conto delle relazioni fra ogni coppia delle K viste, risultando decisamente robusto anche nei casi in cui il numero di viste K è particolarmente ridotto (nel setting di 5VREAL sono previste solamente 3 dispositivi di acquisizione). La componente di triangolazione produce in output *3D ball proposals*, che raccolgono semplicemente le posizioni in tre dimensioni di diverse rilevazioni considerate plausibili. Le rilevazioni ottenute fino a questo punto sono riferite solamente al timestamp t attualmente considerato e non tengono in considerazione informazioni provenienti da istanti di tempo precedenti. Il progetto 5VREAL richiede invece statistiche che possano essere riferite ad ogni traiettoria, intesa come insieme di rilevazioni in tre dimensioni correlate sull'asse temporale. Per questo, è necessaria l'applicazione di tecniche di tracciamento temporale, il cui sviluppo deve essere ancora portato a termine nel progetto 5VREAL. Il tracker, nella sua versione finale, non correrà semplicemente *3D ball proposals* fra istanti temporali consecutivi, ma terrà anche traccia di parametri fisici che descrivano il moto di ogni pallone in modo analitico. Così facendo, sarà possibile sopperire ad eventuali errori di triangolazione o di rilevazione in 2D, eliminare *ball proposals* errate ed aumentare la robustezza ad eventuali occlusioni. Sarà inoltre possibile estrarre da ogni traiettoria le statistiche di interesse per il progetto e prevedere in anticipo le coordinate di ogni pallone nei successivi istanti di tempo $t + 1, \dots, t + n$. Fig. 5 e 6 riportano esempio di traiettorie estratte con la versione attuale del tracker.

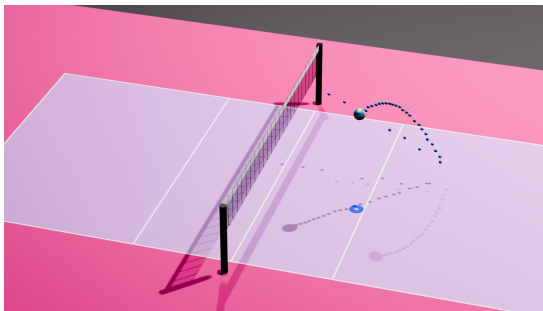


Figure 5: Esempio di ricostruzione 3D della traiettoria e della posizione della palla di una ricezione avvenuta a seguito di un attacco.

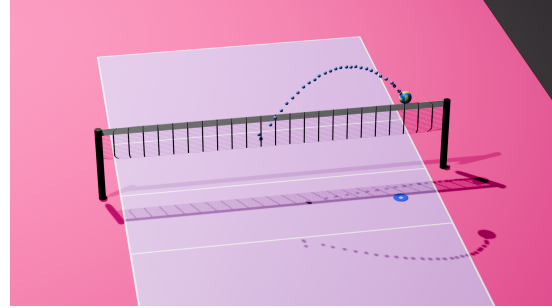


Figure 6: Esempio di ricostruzione 3D della traiettoria di un'alzata.

3. Conclusioni

Abbiamo presentato i progressi raggiunti nel contesto del progetto 5VREAL, concentrando l'attenzione sullo sviluppo della parte relativa alla visione per il tracciamento della palla in 3D. Questi primi risultati rappresentano un avanzamento significativo verso l'obiettivo finale del progetto, che mira a creare un sistema integrato in grado di estrarre automaticamente le informazioni fondamentali per generare nuove e interessanti statistiche per gli allenatori di pallavolo, nonché nuove grafiche coinvolgenti per gli spettatori. Il progetto sarà terminato entro la fine di ottobre 2024, integrando tutte le componenti sviluppate e verificando la corretta funzionalità del sistema. Durante questa fase, saranno condotte valutazioni quantitative e qualitative del prototipo finale, raccogliendo dati sul suo utilizzo e conducendo test per garantire il soddisfacimento completo delle specifiche richieste per promuovere l'avanzamento dell'innovazione nel campo della pallavolo attraverso l'applicazione dell'intelligenza artificiale.

Acknowledgments

Il progetto "5VREAL: 5G Volley Reality Experience & Analytics Live" è stato finanziato dal Ministero delle Imprese e del Made in Italy (CUP I53C23001340005) e vede la collaborazione della Libera Università di Bolzano (capofila di progetto), Fondazione Bruno Kessler, Vodafone Italia s.p.a., EMG Italy s.p.a. e Small Pixels s.r.l..

References

- [1] Z. Wang, P. Veličković, D. Hennes, N. Tomašev, L. Prince, M. Kaisers, Y. Bachrach, R. Elie, L. K. Wenliang, F. Piccinini, et al., Tacticalai: an ai assistant for football tactics, *Nature Communications* 15 (2024) 1–13.

- [2] R. O. Duda, P. E. Hart, Use of the hough transformation to detect lines and curves in pictures, *Communications of the ACM* 15 (1972) 11–15.
- [3] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, *Advances in neural information processing systems* 28 (2015).
- [4] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
- [5] G. Jocher, A. Chaurasia, J. Qiu, Ultralytics YOLO, 2023. URL: <https://github.com/ultralytics/ultralytics>.
- [6] S. S. Sabzalê, VolleyVision, 2023. URL: <https://github.com/shukkkur/VolleyVision>.