Re-enacting Football Matches in VR using Virtual Agents' Realistic Behaviours

Weilai Xu*, Ismail Alarab*, Charlie Lloyd-Buckingham*, Steve Bowden†, Benjamin Noer†,
Fred Charles*, Simant Prakoonwit*, Andrew Callaway*, Shelly Ellis*, Chris Jones†
**Bournemouth University*
Poole, UK
Email: {wxu, fcharles}@bournemouth.ac.uk
†*In Reality Ltd.*
Yeovil, UK
Email: {chris.jones}@inrealityltd.com

***Abstract*—Football (or Soccer)[1] matches are often analysed from 2D video footage that lacks any spatial details of the players on the pitch. On the other hand, the real-time positions of the players derived from recorded tracking data are not easily understandable without an immersive and effective visualisation of the players located on a pitch with the ability of the system to review actions over a timeline. To address some of these drawbacks, we propose a novel "Believable Agent Behaviour" (BAB) engine, based on virtual reality equipped with artificial intelligence for the generation of realistic virtual agent animations to enact the data-driven events from football matches. This system implements an immersive, interactive, believable, and automated end-to-end simulation to analyse and replay a football match using immersive 3D graphics.**

***Index Terms*—Virtual Reality — Artificial Intelligence**

I. Introduction

Analysts in football teams currently use 2D video footage to review and analyse matches. Such videos are often captured from a single fixed point located off-field and as such are far from providing a flexible view of the football matches' actions from all players on the pitch. The 2D video technology lacks any spatial awareness, which plays an essential role whilst analysing the moves of the players in a football match. Increased spatial awareness results in better tactical awareness and improved decision making which ultimately results in better team performance [1]. Better performance is fundamental for elite football clubs, as poor results affect revenue and brand. It includes being able to review past performances as well as establish possible new strategies based on individual players' performance. Football teams are increasingly using technology and data science to improve team performance which will give them a competitive advantage. The proposed engine provides an end-to-end simulation of a football match which utilises virtual reality (VR) and artificial intelligence (AI). The aim of the VR interface outlined here is to focus on replaying football matches within an immersive and interactive visualisation for users to explore, combined together with a virtual agents' behaviour engine to increase the believability of the players on the virtual pitch [2]. For the animation of the virtual players, an AI model is adopted, which is built upon the work in [3], to perform real movements of the virtual players. The overall framework (see Figure 1) aims to transform positional tracking data captured during sporting events (here, a live football match) into a VR simulation which can be replayed for the purpose of the review of the re-enactments [4]. The simulation allows users (e.g. football team analysts, coaches and players) to review an entire match in fully immersive 3D and through the lens of the "virtual camera" offering the users with the view of the actions from any position on the pitch over the duration of the match interacting through a timeline interface [5].

[1] referred to as *football* thereafter.

II. Related Works

VR training has been a long-standing domain of applications for the purpose of using immersive visualisation to enhance skills learning [6], sports' performance [7], and their effect on overall training performance [8]. Overall, VR is a technology which has been well accepted in a variety of sports exercise activities, and review of those activities post-training [9].

One of the requirements [10] to allow for an efficient means of reviewing football matches within VR requires the ability to recreate believable virtual agent behaviours due to the characteristics linked to the rich animations required for the visualisation of virtual football players. To this effect, we used AI-based motion retargetting algorithms from [11] and [12] as the basis for our extension of the football actions available to be replayed through our own animation controller within the context of football matches [13].

III. System Overview

The overall architecture of the VR replay interface (see Figure 1) provides an efficient mean to combine both motion capture data from actual football players and football match tracking and events data provided into a single flexible VR interface which allows for the review of existing matches. The VR interface requires the need to generate a mechanism to incorporate specialist animations for the purpose of visualising believable agent behaviours in the context of team sports within a real-time, interactive and immersive simulation of

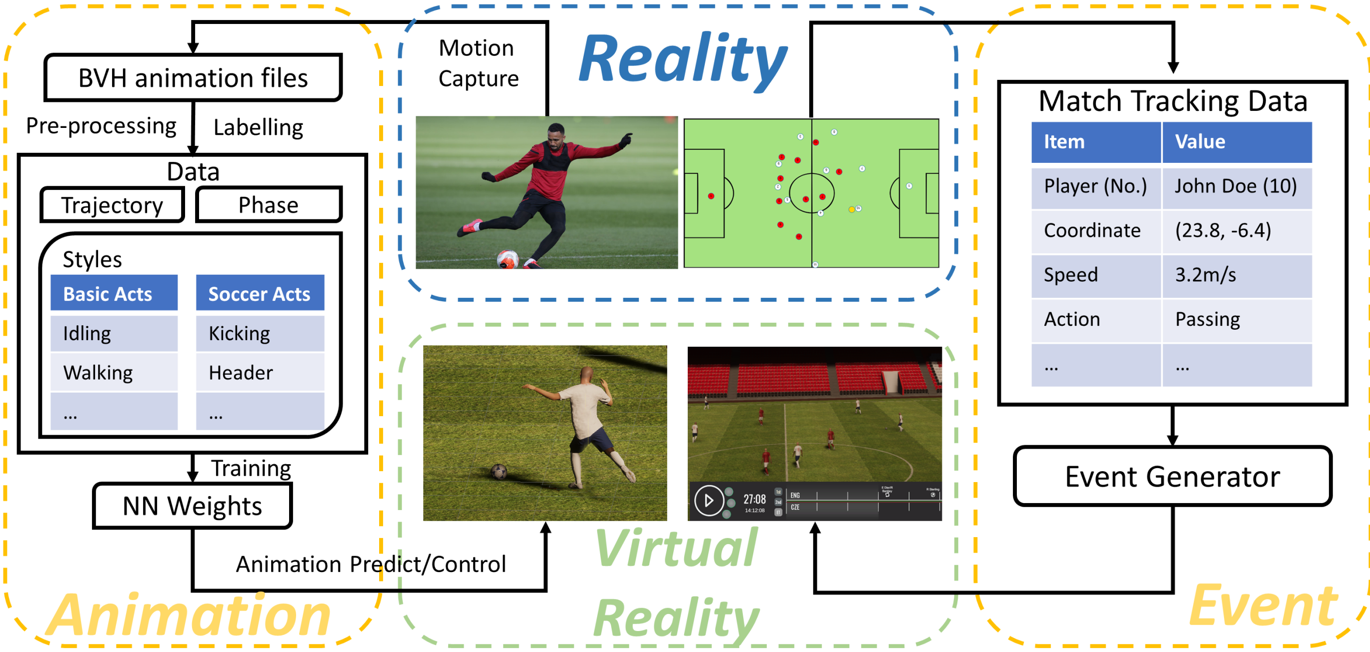

Fig. 1. System Overview: **Animation** of virtual agents (i.e. football players) is generated from a limited set of specific football player actions captured through motion capture; **Event data** is extracted from recorded football match data available and used as the generation of the replay events in the **VR simulation**.

TABLE I
OVERVIEW OF OUR MIXED MOTION CAPTURE DATASET.

Movement Styles	Clips	Frames	Time Duration
Forward Walking	1	7,244	2 min
Backward Walking	1	7,856	2 min 11 sec
Forward Running	1	3,993	1 min 7 sec
Backward Running	1	5,144	1 min 26 sec
Idling	1	1,198	20 sec
Side Walking	1	5,377	1 min 30 sec
Side Running	1	2,727	45 sec
Sprinting	4	1,704	28 sec
• Kick	29	4,781	1 min 20 sec
⌞ Normal Kick	5	1,196	18 sec
⌞ Corner Kick	4	750	13 sec
⌞ Moving Kick	10	1.263	21 sec
⌞ Pass	4	522	9 sec
⌞ Stand Pass	6	974	16 sec
• Header	13	2,520	42 sec
⌞ Jump Header	4	901	15 sec
⌞ Stand Header	9	1619	27 sec
Chested Control	4	949	16 sec
Defending	10	1,259	21 sec
Dribbling	8	2,786	46 sec
Total	76	47,478	13 min 11 sec

the players for the purpose of the exploration of match events by users.

IV. Animation Implementation

To achieve the animation of a single virtual agent, our implementation is based on the standard approach using local motion phases proposed by [3], [11].

We use our own motion captured data from actual football players where we obtained the variety of standard and complex movements, such as locomotion (e.g. walk, run, sprint) and the more specialised football movements (e.g. kick, jump header). To further increase the realism of the more basic movements, we included the open-source locomotion data from 100STYLE [12] to our training dataset. The data generated from the long continuous motion capture sessions is labelled and automatically processed frame by frame, and finally saved as tabular data to train the AI model. Table I shows an overview of our datasets. All motion capture data segments are recorded with a framerate of 60fps. Thereafter, we describe the labelling process and the neural network architecture used in our system.

A. Labeling

There are three parts of the labelling process according to [3].

1) Trajectory Detection: The trajectory of a virtual agent contains the information of the root locations, directions, and heights within a sliding window of 1.9s. This process can be completed automatically.

2) Phase Labelling: This parameter is originally used to describe the cycle of feet when the virtual agent is doing movements with recurrently footwork, e.g. walking. This process is completed automatically to the most extent. For more complicated football-style movements, the locomotion phase [11] would be applied in the future.

3) Style Labelling: We manually labelled our motion capture clips extracted from the motion capture library. Particularly, 4 keyframes are labelled to describe the *starting*, *ongoing*, and *ending* stages of each specialist football movement. Before the starting and after the ending stages, the value of each style is assigned the value of 0, while the value of 1 is

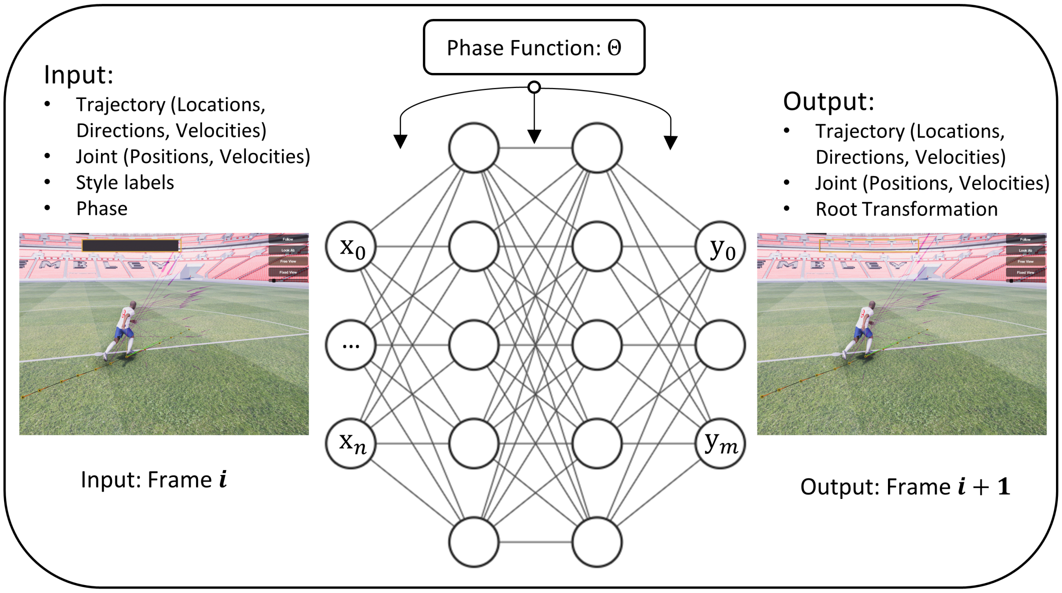

Fig. 2. The overview of the PFNN network and the parameterization of the input and output.

assigned for the on-going stage. During starting and ending stages, the value is linearly interpolated between 0 and 1.

B. Phase-Function Neural Network: PFNN

We use PFNN as stated in [3] which is a regression neural network harnessed by a phase function for learning an appropriate manifold of human motion as schematised in Figure 2. This network maps an input of 504 features, which are embedded in 2 consecutive hidden layers of 512 neurons each, into 376-dimensional space.

1) Input: The PFNN model is formed of input features from the previous frames and the phase function.

The input features of PFNN represent the spatial information of the joints positions, velocities and rotations of the virtual agent in 3D besides the style labelling where each style is accompanied by 12 features under the past, current and future trajectories (6 past; 1 current; 5 future).

For the phase function, we use the cubic Catmull-Rom spline as defined in [14] with four control points (corresponding to left and right toe site joints) to obtain the appropriate regression in the system. Mathematically, the phase function can be expressed as:

$$\Theta(p;\beta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2}w & 0 & \frac{1}{2}w & 0 \\ w^2 & \frac{-5}{2}w^2 & 2w^2 & -\frac{1}{2}w \\ -\frac{1}{2}w^3 & \frac{3}{2}w^3 & -\frac{3}{2}w^3 & \frac{1}{2}w^3 \end{bmatrix} \begin{bmatrix} \alpha_{k_0} \\ \alpha_{k_1} \\ \alpha_{k_2} \\ \alpha_{k_3} \end{bmatrix} \quad (1)$$

where $\beta = [\alpha_0\, \alpha_1\, \alpha_2\, \alpha_3]$, p is the phase, w $= \frac{4p}{2\pi}(mod\, 1)$ and $k_n = \lfloor w \rfloor + n - 1 (mod\, 4)$.

In our implementation, we feed our processed data into PFNN neural network [3] and train the model for 50 epochs which takes around 1.5 hours on a NVidia RTX A4000 GPU. We opt for the same hyperparameters from [14] as follows: learning rate of $1\text{x}10^{-4}$, weight decay $2.5\text{x}10^{-3}$, batch size of 32, and dropout of 0.7.

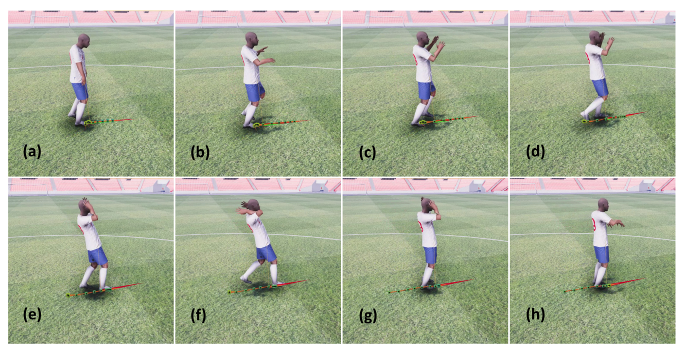

Fig. 3. Sequence of virtual agent's animation showing the inclusion of the specialised football move of a **throwing** into the flexible animation retargeting framework.

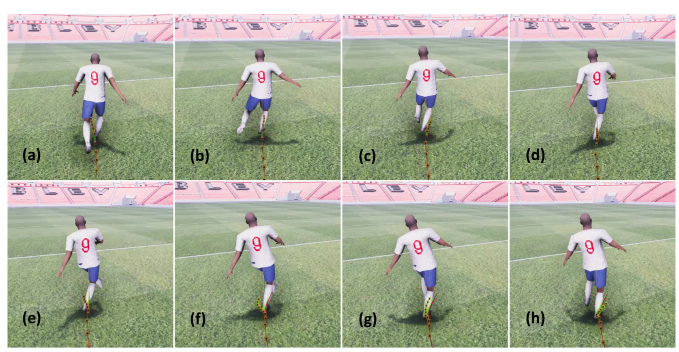

Fig. 4. Sequence of virtual agent's animation showing the inclusion of the specialised football move of a **normal kick** into the flexible animation retargeting framework.

2) Output: The output of the regression network corresponds to the joints' positions, velocities and rotations of the virtual agent. In addition, the output provides the root translational and rotational velocities which update the root transformations of the position and rotation via the neural network predictions. We present two examples of the animation for football movements (throwing: Figure 3 and kicking: Figure 4) predicted by the trained PFNN model from our dataset.

V. Replay VR Interface

A. Positional Tracking Data Processing

The final stage of our overall end-to-end pipeline requires the need to process the positional tracking data which is generated from actual football matches - this data is generated for the football teams and available for them to use as a review tool. Unfortunately, there doesn't exist any user-friendly interface for the purpose of reviewing this data. The data includes timestamped positions of all players on the pitch, as well as timestamped events occurring throughout the football match.

This stage requires the mapping of the identified players' actions on the pitch and their dynamic movements from one event to the next, through the definition of the trajectories of the agents driven by our neural network-based animation system.

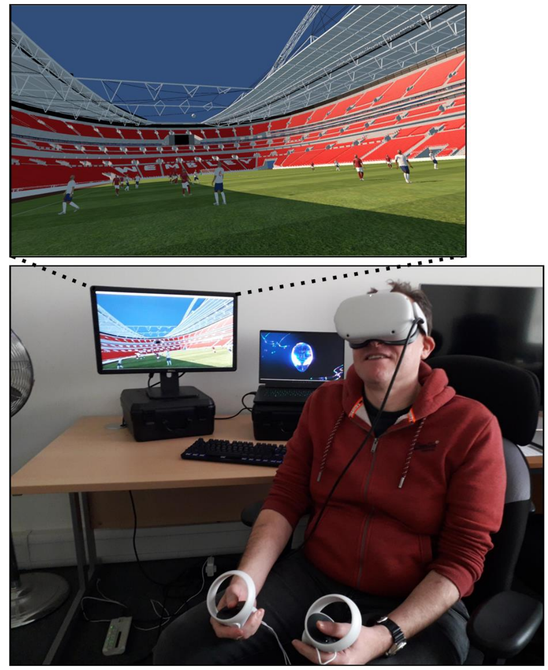

Fig. 5. Overview of the VR interface with a football coach reviewing the match from any position within the virtual environment, whilst being able to play and replay the players' actions.

B. User Interface

Figure 5 provides a view of the actual interface as seen by the users who review the match data using inexpensive VR technologies (here, an Oculus Quest 2 headset). The VR interface affords the users with the ability to review any part of the football games from any position on the pitch or from any players' point of view.

This overall automated process allows for the visualisation of any data generated from recorded positional tracking and event data provided to football clubs. This data-driven approach offers the flexibility for the performance analysts to review and analyse any match data to better devise strategies for their players.

VI. Conclusions

Our VR interface intertwined with AI allows to transform positional tracking data captured during sporting events (e.g. a live football match), combined with extracted match events information, into a VR simulation where football team analysts, coaches and players can review any entire match in fully immersive 3D environments. This flexible VR interface affords them with the ability to review actions' on the pitch at any time and anywhere within the match on the pitch. The VR interface already provides a high-level of realism of the virtual agents' behaviours generated from our collected motion capture datasets of players' football movements. An early user evaluation of the overall system provided positive response from team analysts. A full evaluation of the VR interface offering the ability for coaches and players to review the same match data through a multi-user VR interface will provide an evaluation of the performance improvements, over a longer period of training. The flexibility of the framework created demonstrates its potential to reproduce VR-based simulations for team sports for a larger number of events.

Acknowledgments

The project was supported in part by Innovate UK grant 10004907. We would like to thank Areej Hassan, Sophie Wain and Liam Cotmore for their help and support.

References

[1] K. Fujii, "Data-driven analysis for understanding team sports behaviors," *Journal of Robotics and Mechatronics*, vol. 33, no. 3, pp. 505–514, 2021.

[2] M. Cavazza, F. Charles, and S. J. Mead, "AI-based animation for interactive storytelling," in *Proceedings Computer Animation 2001. Fourteenth Conference on Computer Animation (Cat. No. 01TH8596)*. IEEE, 2001, pp. 113–120.

[3] D. Holden, T. Komura, and J. Saito, "Phase-functioned neural networks for character control," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, pp. 1–13, 2017.

[4] F. Vidal-Codina, N. Evans, B. E. Fakir, and J. Billingham, "Automatic event detection in football using tracking data," *arXiv preprint arXiv:2202.00804*, 2022.

[5] J. Porteous, J. Teutenberg, F. Charles, and M. Cavazza, "Controlling narrative time in interactive storytelling," in *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, 2011, pp. 449–456.

[6] D. Checa and A. Bustillo, "A review of immersive virtual reality serious games to enhance learning and training," *Multimedia Tools and Applications*, vol. 79, no. 9, pp. 5501–5527, 2020.

[7] N. Mascret, G. Montagne, A. Devrièse-Sence, A. Vu, and R. Kulpa, "Acceptance by athletes of a virtual reality head-mounted display intended to enhance sport performance," *Psychology of Sport and Exercise*, vol. 61, p. 102201, 2022.

[8] A. Shimi, A. Papantoniou, K. Neokleous, and M. N. Avraamides, “Athletic performance in immersive virtual reality: The effect of training intensity.” *European Journal of Psychology Open*, vol. 81, no. 1, p. 24, 2022.

[9] L. Capasa, K. Zulauf, and R. Wagner, “Virtual reality experience of mega sports events: A technology acceptance study,” *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 17, no. 2, pp. 686–703, 2022.

[10] K. Tuyls, S. Omidshafiei, P. Muller, Z. Wang, J. Connor, D. Hennes, I. Graham, W. Spearman, T. Waskett, D. Steel *et al.*, “Game Plan: What AI can do for Football, and What Football can do for AI,” *Journal of Artificial Intelligence Research*, vol. 71, pp. 41–88, 2021.

[11] S. Starke, Y. Zhao, T. Komura, and K. Zaman, “Local motion phases for learning multi-contact character movements,” *ACM Transactions on Graphics (TOG)*, vol. 39, no. 4, pp. 54–1, 2020.

[12] I. Mason, S. Starke, and T. Komura, “Real-time style modelling of human locomotion via feature-wise transformations and local motion phases,” *arXiv preprint arXiv:2201.04439*, 2022.

[13] Z.-M. Ye, J.-L. Chen, M. Wang, and Y.-L. Yang, “Paval: Position-aware virtual agent locomotion for assisted virtual reality navigation,” in *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE Computer Society, 2021, pp. 239–247.

[14] D. Holden, “Reducing animator keyframes,” Ph.D. dissertation, 2017.