USING METAHUMANS IN LARGE-SCALE SPORTS PRODUCTION

Doris Retfalvi

disguise, United Kingdom

ABSTRACT

What if you could accurately simulate a major football match in a virtual environment and view your favourite moments of the game at a level of detail that would have been impossible through a traditional broadcast? With today's advances in game engine graphics, with realistic 3D humans, or MetaHumans, that future is closer than you'd think.

In this technical paper, we will investigate how MetaHumans and digital doubles are changing sports broadcasts today, and explore the possibilities that they'll bring to future match day coverage.

First, the paper will give an overview of broadcasters who are already using real-time virtual production for everything from match commentary to remote interviews in sports productions all around the world. As a case study, we'll then describe how MetaHumans were implemented during the Qatar World Cup coverage by TelevisaUnivision — the world's leading Spanish-language media and content company.

Together, the Televisa team worked with disguise Creative Services to use MetaHumans as part of their analysis show, TUDN. That meant journalists and former footballers could more accurately analyse gameplay, thanks to real-time data that was fed into 3D digital doubles of the players. This presents interesting possibilities for the future, including the potential to scan players into 3D digital replicas, then mirror their movements using markerless camera tracking technology enhanced by artificial intelligence.

INTRODUCTION

Mainstream sports networks had not yet widely adopted the use of MetaHumans and digital characters to recreate 3D gameplay for football matches. However, some sports networks have started to explore the benefits of 3D engines like Unreal Engine to enhance their coverage through virtual backdrops, augmented reality graphics and motion graphics.

For the 2022 FIFA World Cup coverage, one network has taken the possibilities of Unreal Engine one step further, deploying MetaHumans to create highly realistic interactive virtual representations of the game to provide viewers with a more engaging experience than ever before.

TelevisaUnivision is Mexico's largest broadcast network and the world's leading Spanish-language media and content company. For the 2022 FIFA World Cup, the broadcaster worked with the disguise Creative Services team to use MetaHumans as part of their La Jugada Qatar 2022 prime time analysis show on TUDN. This innovative collaboration enabled journalists and former footballers to more accurately analyse gameplay, thanks to real-time data that was fed into 3D digital doubles of the players, while offering viewers a deeper look at the biggest matches of the championship and an overall more entertaining viewing experience.

The TUDN programme ran during the entire duration of the World Cup - every weekday for two hours from 9pm in the evening, and for three hours on Thursday, Friday and Saturday.

TUDN was broadcast both on Televisa's 24h network and the open air network throughout Mexico, with a total of 9.4 million viewers tuning in. (1)

Following the World Cup, Televisa reported a 22% jump in pro-forma revenue for the fourth quarter in 2022, boosted by its 2022 FIFA World Cup coverage. (2)

CONSIDERATIONS

Innovative technologies in sports broadcasts: the landscape so far

Historically, several advanced technologies and innovations have been employed by sports networks to enhance their coverage and provide more engaging experiences for viewers. Some of these technologies include:

1. **Virtual studios and augmented reality**: Sports networks like ESPN, Fox Sports and DAZN have used virtual studios, augmented reality graphics and motion graphics to create immersive, data-driven presentations. These technologies, enabled by disguise's extended reality workflow and graphics control system could visually represent player statistics, game strategies, and real-time data in a compelling manner, helping viewers understand complex information more easily.
2. **Player tracking and performance analysis**: Sports networks have adopted player tracking technologies like SportVU, which uses cameras and computer vision to track player movements and analyse their performance. This data could be combined with advanced analytics to provide in-depth insights into player strengths, weaknesses, and contributions to the team.
3. **Virtual and remote camera systems**: Some networks, like Sky Sports, have implemented virtual camera systems that enable them to create unique camera angles and viewpoints that would be impossible to achieve with traditional camera set-ups. These systems could provide viewers with novel perspectives on the game, like bird's-eye views or close-ups of individual players.
4. **Teleportation**: Extended reality has also enabled Eurosport to experiment with "teleportation", where athletes or coaches were beamed into the virtual studio from remote locations to take part in an interview. This technology allowed for more engaging and interactive interviews, as the interviewee could respond to questions and interact with the environment in real-time.

As broadcasters around the world are constantly looking to boost viewer engagement and retain eyeballs on the screen, the use of game engines, motion capture, AI and real-time rendering is becoming increasingly mainstream, creating highly realistic and engaging virtual representations of broadcast events.

The brief

Televisa has always sought to cover global sports events in the most innovative way possible, enabled by virtual production technology. For the 2022 FIFA World Cup - an event with the highest viewership ratings in Mexico - the broadcaster wanted to take real-time 3D content and audience engagement to a new level.

As this was an evening game show where viewers would come together with family to review the matches that they saw during the day, the focus for TUDN was a lot less on the hard data analysis of the gameplay and more on the entertainment, giving a different view of the game that the wider public could understand and engage with.

With a multitude of sports analytics tools already on the market, Televisa wanted to go beyond the existing solutions. The mood of the show was a lot more festive, targeting a varied demographic, so the broadcaster wanted to ensure a programme that was accessible and engaging for a broader range of viewers, replacing a heavily technical analysis with a more light-hearted commentary that appealed to more and drove in a higher viewership.

Coming up with the idea

The aim was to gamify the matches, drawing on the power of the groundbreaking MetaHuman Creator in Unreal Engine to build photorealistic assets and animations that would almost turn the programme into a video game.

There was a lot of discussion around the most suitable Unreal Engine version for the project. With both UE5 and UE5.1 out, the former version had more stability while the latter had more innovative features to help meet the brief. Given the ambitious goals of the project and the unrivalled scale of the MetaHuman application, the team opted for UE5.1 to support the large computing power. The issue there was that other vendors that were integrated into the workflow hadn’t yet supported 5.1 so the team had to run multiple in-house beta testing on unreleased plugins to ensure a smooth performance for such a high volume of MetaHumans on the field at once.

The disguise Creative Services team created a custom workflow within UE5.1 to generate a 3D representation of the games.

As Televisa owns the largest football stadium, Estadio Azteca, in Mexico City, the broadcaster built a two-storey structure within the stadium to run the everyday production of the show. Meanwhile, the creative team 3D scanned the stadium to create a model of it that they would then place into a virtual Qatar Bay in Unreal where anchors and commentators could be transported to via their green screen studio, so they would cover the analysis in a fully computer generated environment. This virtual environment would also serve as the location where each day’s game highlights would be recreated.

The benefits of working with a real-time 3D engine

As the most advanced 3D graphics engine currently on the market, Unreal Engine offered the team numerous benefits for delivering the brief to the client's demands, including:

1. **Real-time rendering:** Unreal Engine excels at real-time rendering, allowing for high-quality visuals and smooth gameplay. The engine's advanced rendering capabilities enable realistic lighting, shadows, and reflections, which contribute to an immersive and believable football match experience.
2. **Animation tools:** Unreal offers a suite of native animation tools that the animators and developers collaborated on to create automated interpolations based on real world movements of the ball, making it fly through the air realistically, but also save time on set.
3. **Visual scripting:** Unreal's visual scripting system, Blueprints, enables non-programmers to create complex game mechanics and interactions without writing a single line of code. This makes it easier for designers and artists to collaborate on the project and iterate on gameplay features.

By leveraging these benefits, the team was able to create a custom workflow within Unreal that would accurately simulate the World Cup match highlights, while unlocking flexible animation editing for the producers, ultimately delivering a more engaging, realistic, and visually impressive experience for users.

METHOD

Typically, when a broadcaster runs an analysis programme for a football match, it would do so by overlaying computer-generated graphics on top of a video replay, using markers to track players and ball movements. Networks like Fox Sports have used common sports stats API data to create static AR graphics showcasing statistics like number of goals, penalties, possession and pass percentages.

But Televisa wanted to take this concept to a deeper level - achieving an even more comprehensive look at the gameplay in a way that has previously been impossible without continuously stopping and resuming the game. To address this particularly untapped need, the team developed a complex and highly customised solution that would generate a 3D sequence based on a play-by-play script of each match, translating every single movement of the players and the ball into a high-fidelity digital replica.

Technology overview

The aim was to make the solution as automated as possible.

TUDN's hybrid production, powered by UE5.1, integrated with sports data aggregator tools OptaStats (which is one of the official data providers of the FIFA World Cup) and GolStats. These data tools then applied AI to analyse the gameplay data with the help of a standard real-time sports statistics API. The data analysis, supplemented with the video clip of the replay, would enable the team to track every player on the field and then render a JSON

file that would describe every action in the timestamped replay (which player is running, their speed, the direction of the ball etc).

The team then built a custom toolset inside of Unreal that would make a data representation of the gameplay. Once this was done, they would end up with a very similar copy of what happened in the video clip, but in a completely virtual game engine environment. The creative team would then sit down with the producers to write a script about why that specific part of the game was relevant, creating a tool that would automatically insert different emojis and virtual characters throughout the 3D render of the game, adding a cinematic layer to the analysis.

The MetaHuman Creator tool within Unreal Engine 5.1 sat at the heart of the development of the digital players used to recreate the gameplay.

Data used

In the front-end, the team built a bespoke toolset that allowed them to load the data from the JSON file and generate the whole play at the click of a button. Based on the JSON data, this tool would automatically select the relevant teams with the right uniforms, put each player into their respective positions and run the gameplay exactly as it happened in reality.

In the back-end, a lot of work was done to crunch the data and to make it palatable for Unreal Engine. The team had to alter the original data from the OptaStats aggregator tool (which included information about the players, the teams, the tournament) and the player tracking data from GolStats, merge the two data sets and normalise them for Unreal with the help of Polygon Ipsum - disguise's Unreal Engine-native data feed aggregation tool that offers easy management of third-party data streams.

The team

The disguise Creative Services team working on delivering the project was split into two groups :
- the creatives responsible for creating custom MetaHuman rigs for each team and key player, as well as transforming the stadium scans into usable 3D assets and developing the entire virtual environment with lighting, secondary emoji characters, foliage, etc.
- the developers, working on translating the JSON files into Unreal sequences and building custom tools within Unreal to allow them the flexibility to edit the animations at the producers' request.

Both teams were based remotely with one dedicated support specialist on site in Mexico City assisting Televisa during the filming.

Planning and pre-production

The initial planning discussion started in January 2022, running through several creative rounds before finalising the concept of 3D simulation in May. The creatives then went to Mexico to scan the stadium, creating a 3D model of the stadium, including the pitch, goalposts, seating areas, and any other relevant elements, which they imported into Unreal Engine where they applied realistic textures, lighting, and shadows to create a believable environment, all while simultaneously building the virtual environment of Qatar Bay. They then merged the two environments together to have a final render of the complete setting.

By June, once the qualifiers concluded, three developers were already working to develop the Unreal Engine workflow, collaborating with GolStats to understand the data sets they would need to integrate into Unreal, and creating the MetaHuman versions for each team with three different sets of uniforms per team.

The challenge faced in pre-production was that much of the information required to start building the 3D assets was confirmed very close to the event kick-off, for instance the participants for each team were only announced a month before the start date. So the team had to rely on predictions around which players were most likely to be selected in the month leading up to that point. Once all the player details were confirmed, the team spent the remaining time prior to the event kick-off to fine-tune the digital characters.

Working with gameplay data to recreate the match highlights

Once the event started, the project turned into a very high-paced production running up to 20 hours each day throughout the duration of the championship. Every day the team had to produce at least three to four different plays to analyse in the evening show. The producers at Televisa would look at the biggest matches played on the day, select the most relevant sections of each match that they wanted to analyse and the creative team would work throughout the day with the data aggregators at GolStats to gather the relevant assets and interpret animated data to recreate those highlights in 3D.

To achieve this, Televisa would first send the developers an mp4 clip of the gameplay they wanted to reproduce as a reference. The team would then combine the mp4 file with the GolStats raw data from that specific match to render a large JSON file with all the existing information available for the selected gameplay (which player is running, their speed, the direction of the ball) and upload it into Unreal Engine.

Within Unreal, they would then create a single sequence for the entire segment and parse the JSON file for the following key data:
- what teams and players were on the field,
- what was their exact location,
- how fast were they running,
- who kicked the ball,
- what was the ball’s location during the entire time.

For any missing data from JSON, the team would refer back to the mp4 file to interpolate information such as where players were facing based on the direction they were running, or the speed of the ball based on its location on the field from one frame of the video to the next. The challenge here was to recreate moments not captured on film, at points when

the camera was focusing elsewhere. For this, the team had to rely on what the captured

players were doing and match the behaviour of the remaining players to theirs. Another unforeseen challenge here was the unusual frame rate of 20 FPS that the data played back once imported into the Unreal sequence, so the team had to implement a timecode translator to convert it to the PAL standard of 25 FPS.

Generating smooth playback of the ball was also challenging as the data generated in the JSON only specified certain moments of interest, leaving out details in the middle. For sports statistics APIs, the position of the ball between the passes is not relevant but for 3D animation, it's imperative. This introduced a nested challenge of making sure the ball's movement was realistic; a player couldn't pass the ball and have it appear at the other side of the field in an instant. However, positional interpolation was not always the solution either. Simply smoothing out the location of the ball from point A to B often left the ball appear as if it was floating through mid-air, instead of propelled by a fierce kick. The solution fell to the diversity of the team, where a seasoned animator would step in and break the interpolation, study the provided reference video from TUDN, and create the animation for the ball to achieve a realistic pass.

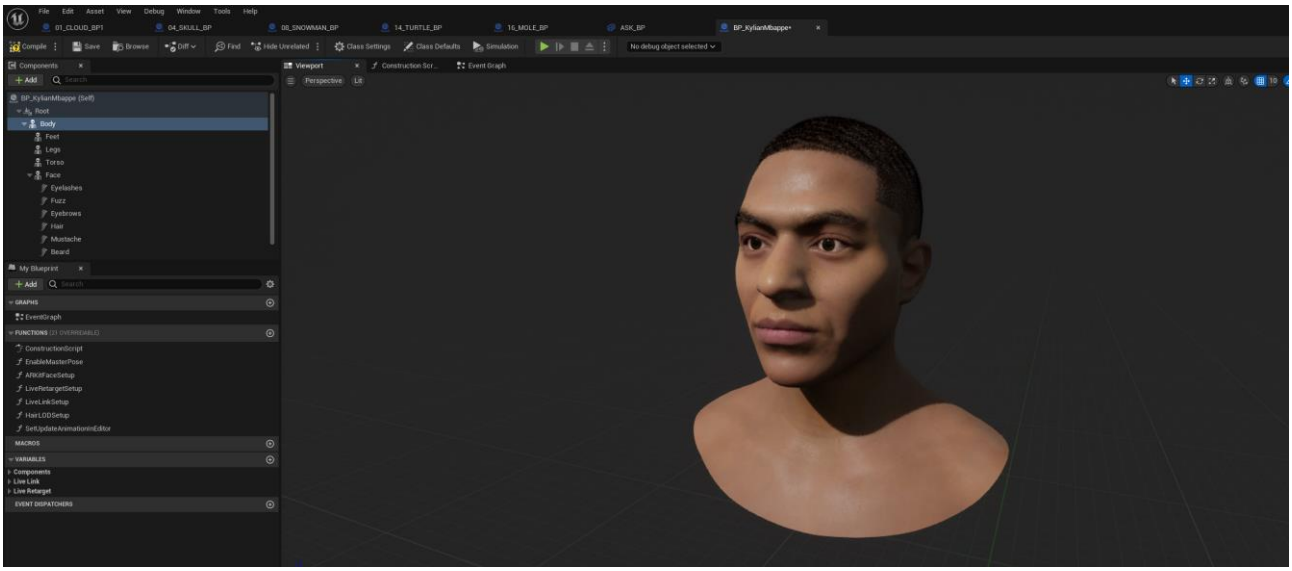
Bringing in the MetaHumans

disguise chose Unreal Engine 5.1’s MetaHuman Creator to recreate the players in 3D due to its state-of-the-art photorealism which was essential for player close-ups. The tool is also built for complex production and real-time rendering which was core to the development team. Although the inherent Level of Detail system (LOD) for MetaHumans within Unreal was powerful, the creatives had to modify it to fit its custom needs. Unreal's focus was video game playability but for the team, it was a mix of high-fidelity simulation as well as unhindered workflow in the hot seat. They opted to only use the highest level of detail during close-ups and keep a more reasonable level of detail for working.

The team initially considered lidar scanning the players using an iPhone and then importing the scan data into Unreal to morph into a MetaHuman. However due to the time constraints and availability of players to physically invite into the studio to scan, this option was abandoned. Instead, the producers opted for a more manual approach where the

team would research as much player footage as they could find to capture all the players' facial and physical features and then send the assets to a designated Unreal artist to model the MetaHumans.

Motion capture was also an option that was initially considered to mimic the players' movements, however finding a mocap workflow that seamlessly integrated with the MetaHuman Creator, by way of recording and editing the data into JSON and then importing it into Unreal, would have also required too much resource that was otherwise crucial in other parts of the workflow build. In the end, they proceeded with pre-made assets. The challenge there was that some of the animations required a lot of clean-up and editing but the new IK Rig Retargeting feature in UE5.1 allowed the creatives to map



animations from different sources to the same skeleton for a truly unified output.
For each national team that was analysed, three to five of their most relevant players were picked and a MetaHuman twin was created for each. The characters were then rigged with a skeletal structure that enabled them to move and interact within the 3D environment. These MetaHumans would then be placed automatically in the digital replays within Unreal (as part of the JSON file upload) to recreate the most pivotal moments of a match,

followed by brief moments of celebrations showing them dancing on the field after scoring a goal.

Due to time constraints, for the remaining players in each team, the team created a library of templated characters with several common characteristics linking them to the original team.

This project marked the largest public deployment of MetaHumans (22 per match analysed) in a broadcast programme running in a completely real-time Unreal Engine environment.

The team had to do a lot of development to get the 22 MetaHumans running, retargeting and animating all the players to the data to create an identical simulation of the real-life match.

MetaHumans require a very strict project structure, and can end up as very large files due to each character rig requiring its own unique set of assets, in addition to a number of shared assets between each character. This meant the team needed to set aside enough time from early in the project to start downloading the assets as soon as possible.

Secondary effects and characters

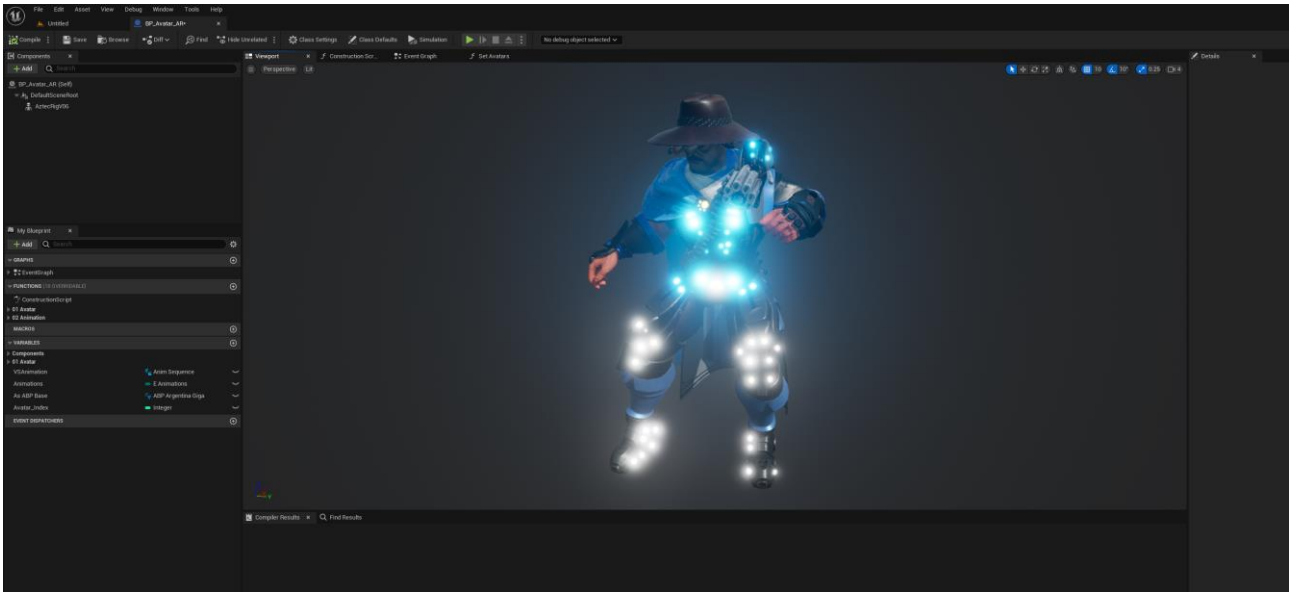
In addition to the players, the art director also created various effects and secondary characters such as emojis and team mascots that would react and animate at key moments in the game. To build these, the team drew references from assets that TUDN would usually include in their match replays and transcribed them to the 3D world to remain on brand. This also made it easy for the producers to select them at their preferred time and easy for the TUDN audience to recognise.

The team mascots were designed as AR avatars promoting each participating country. The creative team drew inspiration from the history and culture of each country when conceptualising the avatars, then sent the design brief to an external design agency to build the character rigs, 3D models and animations for the avatars. Once designed, the avatar files were sent back to the developers to rig using the Stype camera tracking system to turn them into AR mascots in Unreal Engine.

The goal here was to have a large foreground shot of the stadium in Qatar with the AR avatars composited into the scene symbolising when the matches were about square off.

As the team made progress with ingesting source animations, making state machines for these characters so the base rig could switch from idle, to standing, running, kicking or diving, the next step was to merge them into the stadium with the MetaHumans during gameplay. Matching them with the team’s uniform was a simple task as these could be

generated dynamically and on the fly, creating a static sequence. The team could scrub



through it and clean it up as necessary.
Once the characters and effects merged into the 3D environment, the producers working with the support team on site were able to easily select the portion of the gameplay sequence where these could be triggered (eg: when a player scored a goal, they could trigger an explosion of clover leaves and celebration emojis to display around the field). This created enhanced ease of use for the workflow, as the producers were already able to scrub through the sequence, pause it, make comments and leave notes around the

screen.

Building a customised Unreal Engine workflow

Typically when analysing match replays, broadcasters would pause the play and move the camera around. In a 3D match simulation, they can create dynamic camera angles and

replays to enhance the user experience. To achieve this, the following considerations apply:

1. Developing a virtual camera system capable of switching between multiple cameras placed around the stadium. This can include cameras following the ball, player-focused cameras, and wide-angle shots to provide a comprehensive view of the match. The system should smoothly transition between these cameras to maintain a cinematic and engaging experience.

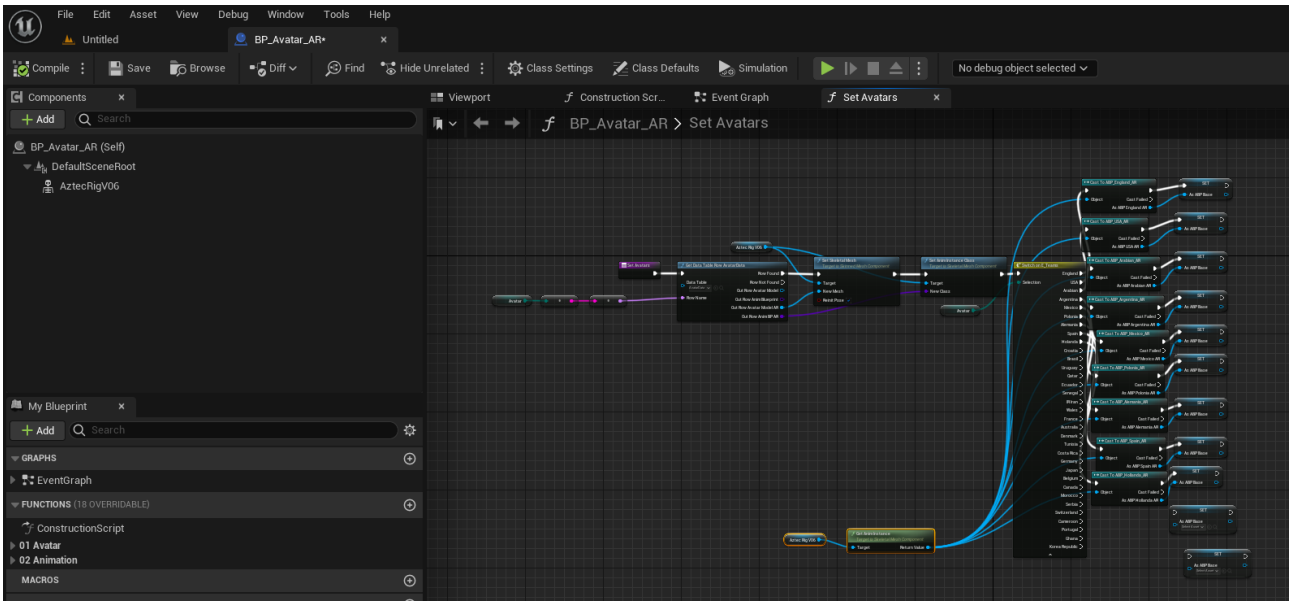
2. Designing camera movements and angles that mimic real-life football broadcasts. This includes slow pans, tracking shots, and aerial views. Using depth of field, camera shake, and other effects to create a sense of realism and excitement.

3. Implementing a replay system to capture and store key moments during the match, such as goals, fouls, and close calls. These moments can be replayed from different angles, with slow-motion and other visual effects to emphasise their significance. The replay system should also allow users to manually control camera angles and playback speed for further analysis.

For the TUDN project, as there was only one timeline in the Unreal Engine interface, the team had to not only pause the gameplay but to also create empty space on the timeline where they could move the virtual camera and animate the players' movements. They displayed the animation of every player on the field to create spaces where they could add in their own content animation on top (made up of hundreds of thousands of keyframes that they moved and shifted around). They also built a MetaHuman replay system to ingest data and then make a whole sequence in Unreal for the exact duration and portion of the match that the producers requested.

When developing the animation sequence, the team wanted to build a custom software that would allow operators from TUDN to easily edit the footage on a timeline but this

ultimately proved to be too challenging for the customer so, instead, disguise's art director



traveled ons site to support with editing.

Real-time rendering and optimisation

Real-time rendering was crucial for ensuring a smooth, high-quality 3D football match experience. The team had to optimise the scene to ensure smooth gameplay and high-quality visuals, adjust factors like level of detail, texture resolution, and lighting complexity, to achieve the desired balance between performance and visual fidelity.

They also implemented a Level of Detail (LOD) system to adjust the complexity of 3D models based on their distance from the camera. This reduced the rendering workload for objects further away, improving performance without sacrificing visual quality.

They also optimised lighting calculations by using techniques such as baked lighting, virtual shadow maps, and ambient occlusion. These methods reduced the computational demands of real-time lighting while still producing realistic results.

Finally, the frame rate was constantly monitored and adjusted to ensure a smooth and consistent experience. This involved dynamic adjusting of graphical settings and adaptive resolution scaling based on system performance with the help of Temporal Super Resolution (TSR) within Unreal.

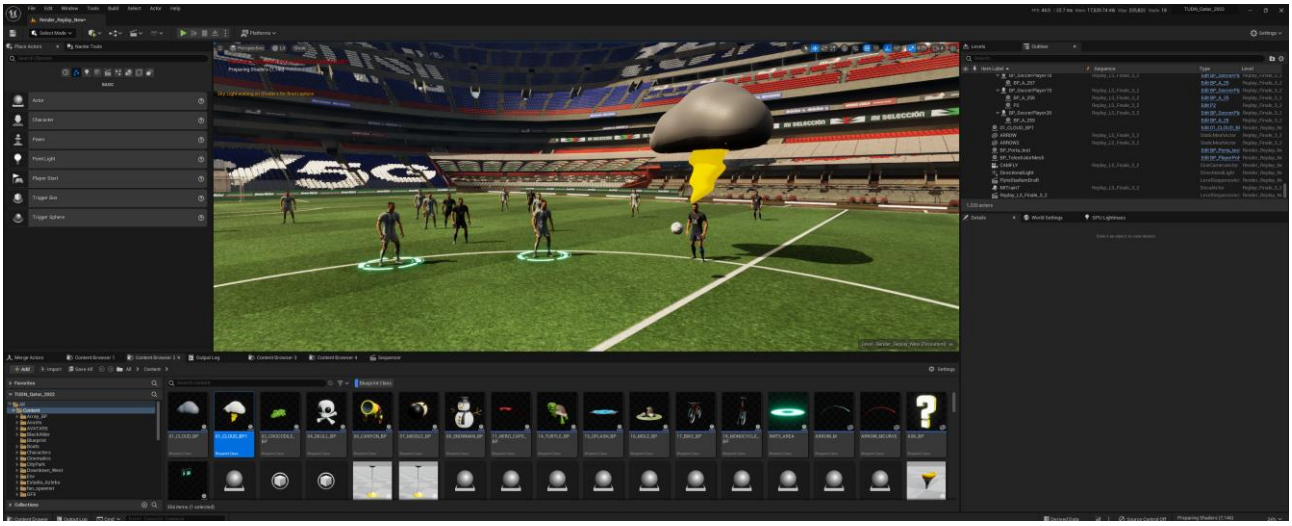
User input and interactivity

On top of that, the team created a tool to edit the animations on the fly, according to the producers' requests.

For this, they had to manipulate the data and move the Unreal timeline around the player to achieve the desired outcome. The main benefit here was the flexibility it unlocked for the

producer to choose what he wanted the final sequence to look like, removing or adding in players, characters or effects.

The tool also allowed producers to control the virtual camera movements. They could create a camera sequence that could place the camera in locations in and around the stadium where a real camera wouldn't be able to be placed without disrupting the players. This gave a range of unique viewpoints for the match simulation that the real-life match



could not offer.
When it came to implementing the AR avatars into the scene for a match square-off, the developers built a system applying Socket.io, multiple JSON packets and open SSH servers to integrate Vizrt controls into the workflow. This system would allow the operators to manually select the two teams about to play and their respective mascots (the avatars) would be triggered to start various actions like taunts, attacks and crouches ahead of the match, mimicking a video game scenario. The disguise team, made up of experienced broadcast operators, were well familiarised with traditional broadcast control tools like Vizrt, so they were able to quickly create simple templates for the operators to use.

Overall, the combination of photorealistic MetaHumans, 3D avatar animations, dynamic environments and effects, and the high level of customisation within Unreal Engine 5.1, resulted in a highly accurate simulation of football gameplay, creating an immersive and engaging viewing experience for TUDN audiences.

Watch the a short highlights reel of the programme here.

CONCLUSION

The groundbreaking solution developed for the TUDN 2022 FIFA World Cup broadcast, deploying Unreal Engine 5.1, gameplay statistics data and MetaHumans to recreate football gameplay, offered unparalleled experiences for both the audiences and the producers, including:

1. **Enhanced visualisation:** High-quality graphics, realistic animations, and dynamic camera angles helped bring the action to life and make audiences feel closer to the game.
2. **In-depth analysis**: By recreating football gameplay in a 3D environment, sports analysts could break down plays, strategies, and player performances in greater detail. This allowed for more thorough and visually compelling explanations of complex tactics, helping audiences better understand the game.
3. **Increased flexibility for producers:** 3D simulation allowed producers to control camera angles, playback speed, animations and other elements. This level of interactivity enhanced audience engagement and provided new editorial opportunities for producers.
4. **Accessibility:** The level of gamification added to the broadcast made the sport more accessible and entertaining for younger audiences or those with a more limited understanding of the sport. By offering a realistic and engaging virtual representation of the game, these simulations helped grow TUDN's viewership ratings.
5. **Real-time data integration**: By integrating real-time match data into 3D simulations, TUDN could provide up-to-date and accurate representations of the live matches.

TelevisaUnivison's fourth quarter revenue grew 22% to $1.5 billion, driven by a strong World Cup performance, while its full year revenue grew 13% to $4.7 billion, marking the second consecutive year of double-digit revenue growth (2).

This project proved that audiences, especially younger generations that traditional broadcasters are striving to attract the most, are open to new, gamified experiences offered by televised broadcast programmes. The opportunities that advanced 3D graphics engines and MetaHumans have offered TUDN can provide a blueprint for broadcasters around the world to follow in order to create a more immersive, engaging and visually compelling programme.

RESOURCES

1 - https://www.fifa.com/tournaments/mens/worldcup/qatar2022/news/fifa-world-cup-delivering-record-breaking-tv-audience-numbers
2 - https://s29.q4cdn.com/983326523/files/doc_financials/2022/q4/TelevisaUnivision-Announces-Fourth-Quarter-and-Full-Year-2022-Results.pdf
3 - https://www.sportsvideo.org/2022/11/17/fifa-world-cup-2022-televisaunivision-supplies-comprehensive-game-around-the-game-coverage/