From Recording to Reality: A Quarterback Simulator

George Kudrayvtsev
george.k@gatech.edu

Abstract — Modern virtual reality enables an unprecedented level of player immersion. Users often behave as they would in the real world; there are vast fields of untapped potential in using VR for learning and training. In particular, it can be applied towards athletics in training both mechanical and mental skills. Within the world of American football, such simulations already exist, but they require extensive manual effort in creating realistic gameplay scenarios. This work proposes and implements an automated pipeline that allows professional and casual players alike to tap into the existing monolithic library of in-game footage to effortlessly generate interactive VR simulations and rapidly hone their skills based on real gameplay.

1 PRESENTATION

The presentation is available here, or at this URL:

`https://1drv.ms/v/s!AvfgQ3Qt8rgWj6w2eCKy5lDoPMF6Pw`

Feel free to read the entire paper; it follows in full below.

2 INTRODUCTION

Consumer-ready virtual reality hardware that features full head and hand tracking in 6 degrees of freedom—such as the Oculus Quest or Rift S, the HTC Vive Pro, or the Valve Index—offers a safe, scalable way to create immersive simulations. The very nature of an immersive technology opens up the opportunity for innovation in any sector that lends itself to simulation. Though the entertainment industry is the darling of VR enthusiasts, there are plenty of horizons waiting to be broached. This work briefly explores virtual reality's application to education, learning, and training across many disciplines to justify investment in the field. Then, it presents an improvement over existing technologies specific to training quarterbacks on American football teams by eliminating the necessity for teams of artists, coaching staff, and other experts to manually craft simulated plays for practicing in virtual reality. It describes and demonstrates the algorithms that go into the transformation from 2D to VR, presenting pain points and suggesting future improvements.

2.1 VR as a Learning Medium

If a virtual reality environment fulfills a sufficient baseline of realism (McMahan et al. 2016; Mori et al. 2012), users often feel truly "present" in the simulation, allowing them to behave as they would in the real world (Schuemie et al. 2001). Because of that, VR is definitely far more than just a toy: it has offered revolutionary **methods for psychological research** (Blascovich et al. 2002), helped people **overcome anxieties** (in public speaking, competition, and other social phobias) (Klinger et al. 2005; M. North et al. 2015; Parsons and Rizzo 2008; Stinson and Bowman 2014), and trained people in "soft skills" like **empathy** and **leadership** (Gavarkovs 2019; Myers III 2019).

Virtual reality's potential to improve skills is not limited to the mind; it has been successfully applied to tasks that are physically demanding, as well. VR has helped the elderly improve their **balance** (Singh et al. 2012); allowed handball players to **practice goaltending** skills (Bideau et al. 2004); basketball players **practice freethrows** (Covaci et al. 2012) and **tactics** (Tsai et al. 2017); and helped quarterbacks practice **navigating** and **visualizing plays** and even **improve their passing** (Huang et al. 2015; Sports VTS 2019).

If it's not already abundantly clear, virtual reality offers a unique, effective way to teach a myriad of diverse ideas and skills. The work presented here focuses on the latter applications: virtual reality simulations for American football.

2.2 Related Work: VR Sports Simulators

Though there have been many forays into VR simulations for training athletes (as briefly covered above), two enterprise solutions catered towards American football stand out. They allow users to put on a VR headset and take part in a football play; however, they take remarkably different approaches.

2.2.1 *QBSIM*

The simulator created by Sports VTS (2019) stands at the cutting edge. With a sensor-laden ball, high-quality visuals, and full movement tracking, it hits all of the requirements for recreating an in-game scenario in virtual reality. The product

Figure 1. A screencap from QBSIM's trailer video showcasing their scenario-creation tool (*left*), and a screencap of the "in-headset experience" of a football play as seen by users of Strivr's platform (*right*).

has high-quality models and animations and is supported by real professional coaching staff and players as advisors. Their technology allows quarterbacks to increase their "mental reps" of in-game situations.

Its shortcoming, of course, which this work begins to addressed, is that plays still have to be defined by hand.

2.2.2 *STRIVR*

This company's simulator hits a different niche: rather than create a virtual world using 3D models, they actually capture 360° video of a *real* football play, then let users relive the actual footage in a headset (STRIVR 2019). Its more oriented towards tactics and "vision training" which the literature suggests is what truly separates amateurs from professionals (Appelbaum and Erickson 2018).

This has natural limitations at scale: every play needs to be drawn out, organized, shot, edited, etc. to recreate it as a simulation. Furthermore, users immediately encounter limitations of a single point-of-view; they can't see (or be shown) what else is going on that they may be able to infer. Further still, the world cannot react to user inputs; everything is pre-recorded. This tool does a great job of teaching "football vision," with team formations and movements being easier to highlight and replay, but falls short in teaching a quarterback how to physically react to a situation.

2.2.3 *3D Recreation*

Work by Roberts et al. (2007) has actually almost gone all the way with the work presented here: it presented a streamlined method to generate a 3D replay of a football play recording. Unfortunately, that progress has been lost into the ether; furthermore, it did not translate or apply the replays to virtual reality. However, the curt insights into the process of recreation that could be gleaned from abandoned demo videos were leveraged and built upon in this work.

2.3 Contribution

With an understanding of the limitations of existing technologies, the motivation behind this work should be abundantly clear. Training scenarios can be created with minimal user involvement by leveraging a potpourri of computer vision algorithms. Using professional game footage guarantees strategic value in the analyzed plays, and the massive library of footage available ensures that trainees can rapidly iterate through scenarios without a content-creation bottleneck.

3 COMPUTATIONAL PIPELINE

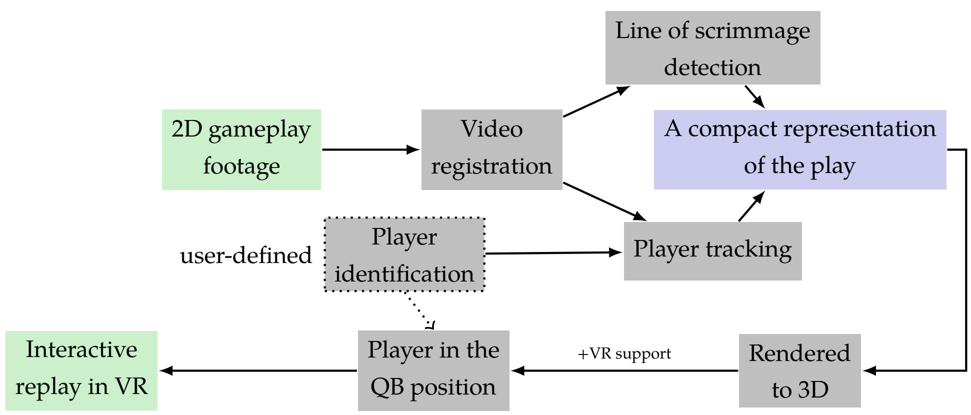

Figure 2. A high-level architecture of the automation pipeline, excluding the specifics of interactivity (path interpolation, throw simulation, etc.).

3.1 Video Registration

In order to ensure that all of the subsequent metrics are accurate, the perspective introduced by the camera needs to be removed. Recreating a "top-down" view of the play that preserves proportionality across the entire field. This is commonly referred to as *image rectification*, *unwarping*, or *registration* in the context of video.

The yard lines on the field are known to be parallel (see Figure 3); thus, they can be used as constraints to unwarp the perspective (Matas et al. 2000).

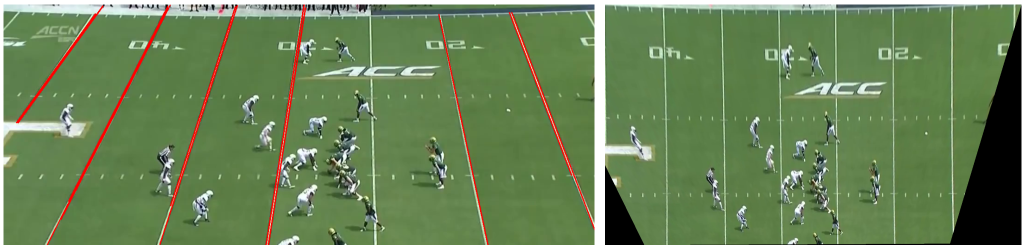

Figure 3. Identification of the yard lines on the football field for warping the image onto a rectilinear plane, and the use of that model to warp the image to preserve parallel lines.

Though a reliable rectification can be determined on the first few frames due to camera stability, the introduction of rapid pans and zooms quickly leads to degenerate homographies. This is due both to a lack of consistencies in line detection as well as insufficient keypoint matches across frames. The pipeline presented here is not yet robust enough to smoothly stitch the video into a global top-down view, which heavily affects the Player Tracking quality under significant camera motion.

3.2 Play Classification

Tracking high-level initialization data like the player starting positions, their teams, the initial size of the field, etc. is important for an accurate simulation. The field width is a hard-coded, per-play value, but future work could leverage in-footage data like yard line digits and tick marks to approximate the value.

3.2.1 *Line of Scrimmage Detection*

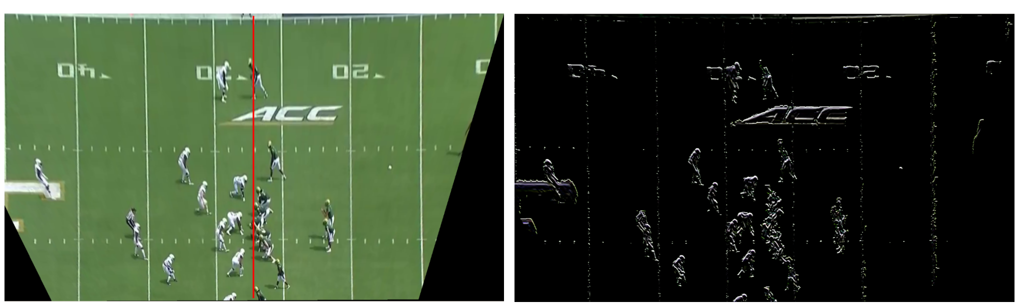

Figure 4. Identification of the line of scrimmage (left) via the densest vertical slice of ∇I_y (right).

To separate teams and orient the play properly in the field, it's important to detect the line of scrimmage. Because players often cluster at the line of scrimmage, the window with the highest total vertical gradient magnitude consistently offers a reliable result regardless of the team formations (Atmosukarto et al. 2013, pp. 995).

3.2.2 *Player Identification*

It's next-to-impossible to automatically identify the starting locations of the players without training a neural network. Even then, the training set would need to be robust to the variations in jersey colors, poses, etc. Simple convolution-based template matching (that is, identifying image patches that match a template) can be used to "suggest" possible locations of players. Unfortunately, due to high amounts of overlap between players at the line of scrimmage and the mathematical naïvety of the method, the suggestions are still imperfect (see Figure 5).

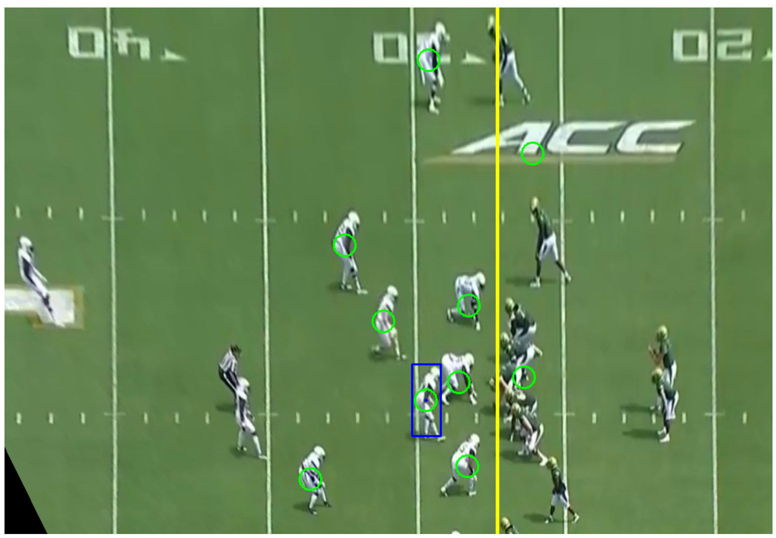

Figure 5. With a good initial choice (in blue here), it's possible to have a majority of good suggestions about player positions (in green) via rudimentary template matching. Obviously, it does not work well enough to be a reliable auto-initialization method, as evidenced by the handful of erroneous suggestions.

Because of that, the interface requires the users to specify starting positions for all players (including a special identifier for the quarterback). This is the only part of the processing pipeline that requires user interaction, and the line of scrimmage can still be used to determine a player's team.

3.3 Player Tracking

A variety of methods were experimented with for tracking players: neither template nor chromaticity-based particle filters could track players through rapid camera movements despite their claims (Dearden et al. 2006). However, tracking player centers-of-mass using sparse Lucas-Kanade optic flow proved to be far more effective (Bougue 2001; Lucas and Kanade 1981; Roberts et al. 2007; Shi and Tomasi 1994).

Specifically, at each frame, all of the current tracks (that is, points correlated across $\geqslant 2$ frames) are considered. If a track is within a player bounding box, it's temporarily associated with that player; the track is ultimately associated with the closer bounding box. Having multiple "candidate players" allows resolution of overlapping bounding boxes when players are close together. The center of mass of the player's bounding box is then adjusted to be at the mean position of its tracks.

After the entire video is tracked, the tracking paths are adjusted based on the unwarped perspective transformation determined during Video Registration.

3.3.1 *Smoothing*

Due to the fact that not every track gets a positional update every frame, the path may end up jittery do to non-linear fluctuations in the bounding box's center of mass. During simulation, such jitter is jarring, confusing, and impossible to interact with realistically. Hence, a post-process of the path to smooth the results is required.

An iterative method that minimizes a linear combination between each pair of neighboring points works well (Kudrayvtsev 2019; Thrun 2014). Given a point on the raw path, $\mathbf{p}_i \in \mathcal{P}$, the resulting point on the smoothed path, $\mathbf{s}_i$ is balanced by its distance from $\mathbf{p}_i$ (i.e. the amount of drift) and its distance from its smoothed neighbor(s), $\mathbf{s}_{i\pm1}$:

$$\min \begin{cases} \alpha \left\|\mathbf{p}_i - \mathbf{s}_i\right\| \\ \beta \left\|\mathbf{s}_i - \mathbf{s}_{i+1}\right\| \end{cases}$$

This can be solved iteratively via gradient descent until reaching some small

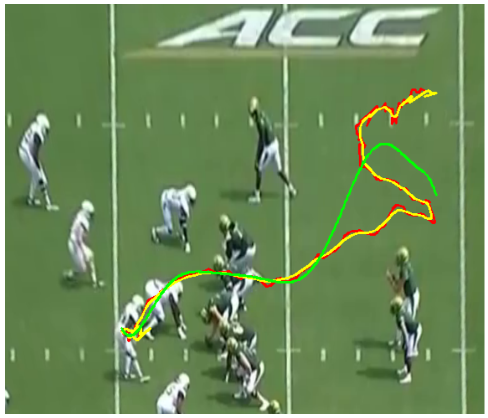

Figure 6. A single player track is rendered in three forms: the raw tracking data (in red), a 10^{th}-degree polynomial approximation (in green), and the iteratively-smoothed path (in yellow).

threshold (that is, until $\sum_i \Delta_i \leqslant \varepsilon$):

$$\mathbf{p_i} \in \mathcal{P}: \qquad \Delta_i = \alpha(\mathbf{p}_i - \mathbf{s}_i) + \beta(\mathbf{s}_{i-1} + \mathbf{s}_{i+1} + 2\mathbf{s}_i)$$
$$\mathbf{s}_i = \mathbf{s}_i + \Delta_i$$

Empirically, $\alpha = 0.1$, $\beta = 0.5$, and $\varepsilon = 0.1$ work well. As evidenced in Figure 6, this results in a path with little deviation from the raw data, yet a far smoother trajectory and overall jitter-free experience from the user perspective. The green trajectory was the result of an unsuccessful attempt to smooth paths by fitting them to n^{th} degree polynomials. Since polynomials are functions, they require that there is *only* one y-value for every x; this is only true for paths that don't backtrack. The smoothing method falls apart in the general case since there's no guarantee of that property.

4 RESULTS

All of the above information that is gleaned from the scene is serialized and fed into a game engine for the virtual reality simulation. Art quality aside, the reproduction of the footage is remarkably convincing, even with a prototype implementation. Putting on the headset provides a whole new perspective on the decisions the quarterback made: "objects in VR are closer than they appear." The scale of the game is completely different than when viewed from a camera hundreds of yards away.

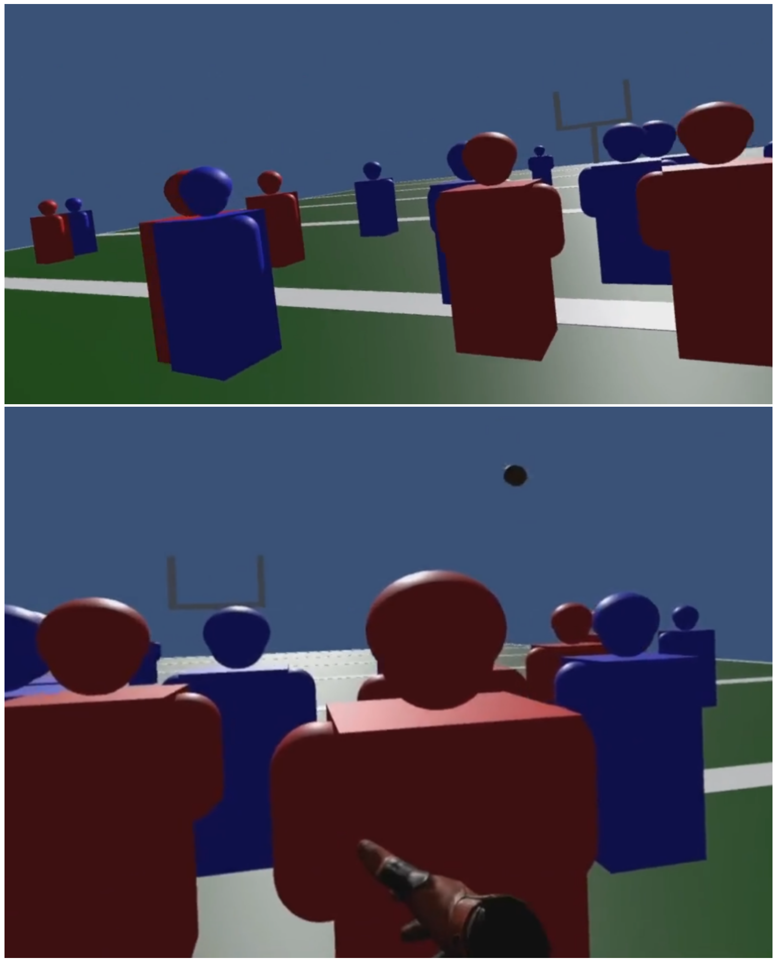

Figure 7. Views from the within the simulation. **Top**: the defender (in blue) rapidly approaches for a sack, moving much quicker than the video would suggest. **Bottom**: the player makes a successful pass to an open player.

5 CONCLUSIONS & FUTURE WORK

Volumes could be written on the potential for polish, expansion, and improvement of the basic framework presented here. The main areas of focus, however, can be categorized oriented around polishing **tracking** and **art** quality, as well as around further leveraging the inherent potential of a simulation by adding **analytics** and **replay control**.

5.1 Improvements: Polish

Tracking— As discussed at length in both the Video Registration and Player Tracking sections, the quality of player tracking when the footage undergoes

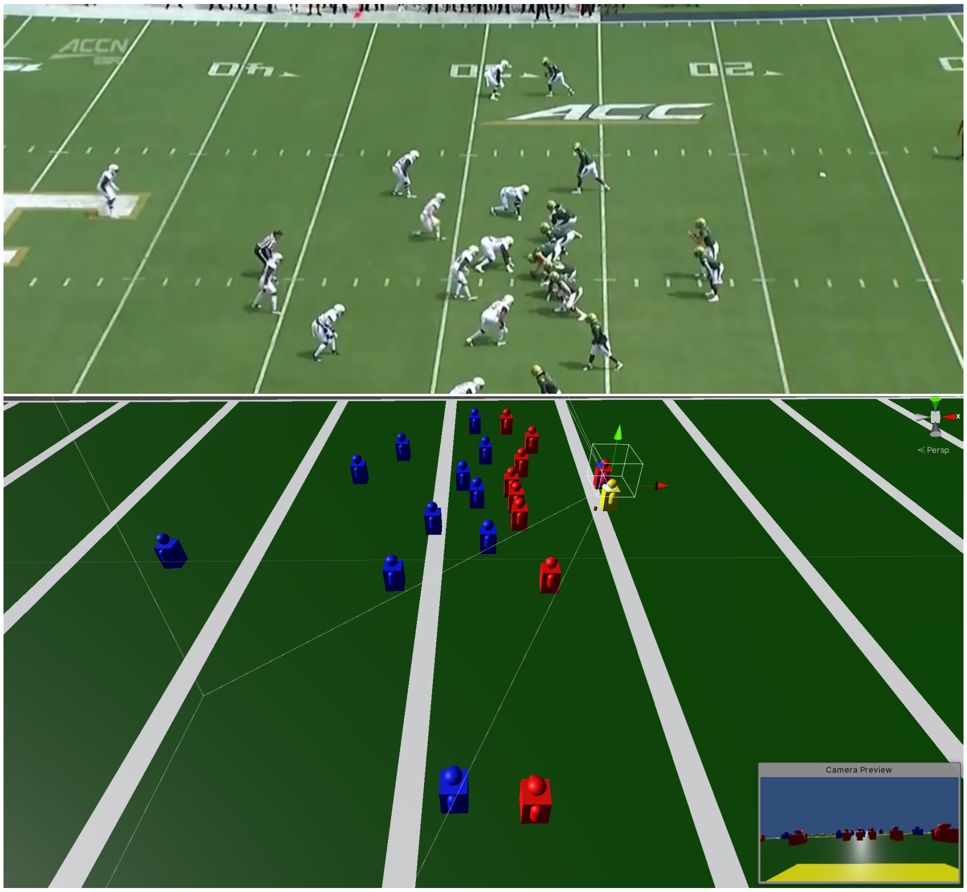

Figure 8. The first frame from a recording of a college football play (Georgia Tech vs. South Florida) video and the relative first frame recreated in a 3D environment.

rapid perspective change (specifically, zooming) can be improved considerably. The entire pipeline depends on finding accurate paths for players, hence any investment in this area pays off exponentially.

Art — The simulation's believability rests heavily on its ability to replicate the real world. Without even considering the "uncanny valley," in which a simulation can be criticized for minutae when "trying too hard," a baseline of "trying" that is better than geometric primitives is a requirement for a player's sense of "presence" in the simulation (Mori et al. 2012; Schuemie et al. 2001). The fidelity of interactions like throwing, ball collision, etc. is obviously directly correlated with a simulation's believability (McMahan et al. 2016; Warren 2018). Discrimination between players will also go a long way, since a professional will react differently depending on the class of player threatening his position

(Hudson and Hurter 2016).

5.2 Improvements: Features

Rather than having coaching staff spend time and energy crafting the play, they can now instead actively participate in the training process during the simulation. The ability to pause and rewind player movements, highlight specific players or actions, and view the play from any angle at any moment is an invaluable tool for professionals.

Analytics — By bringing the play from world of turf to world of pixels, everything can be enhanced with detailed analytics. From basic metrics like steps taken and passes completed, to complex analyses like identifying problematic players or strategies that lead to unsuccessful plays, the potential of digital analysis cannot be understated. The world of sports is already dominated by analytics to help maximize player potential and make micro-adjustments to their game; this adds an entire dimension to that.

Controls — By mixing principles of virtual and augmented reality, players can enhance their practice sessions with fine-grained control over the simulation. Pausing to observe a play from another perspective grants untapped potential for a player; much like how basketball players seeing their freethrow form can lead to improvements (Covaci et al. 2012), seeing decisions or mistakes from an outside view can lead to insights that are otherwise impossible to communicate. Whether quickly cycling through scenarios or trying the same one over and over, the immersive simulation grants the player unparalleled training efficiency by letting them experiment without waiting on anyone. There are dozens more potential improvements, but in summary, the fine-grain control over practice that an in-hand controller provides simply cannot be matched outside of the headset.

Generally-speaking, the work presented here is not intended to be an end-to-end training pipeline. Rather, it should be integrated or used alongside existing tools so that coaching staff and athletes have the ability to create their own plays **or** recreate existing ones; through integration, they can continue to leverage the platform support, feature sets, and art assets of existing, well-developed frameworks.

5.3 Conclusion

There is no doubt that virtual reality is a viable platform for education and training of both mental and physical skills; the literature supports the fact that humans behave remarkably similarly when present in a simulation compared to how they behave in real life. This, along with the scalability, flexibility, and power innately offered by software, can be leveraged by athletes and amateurs to practice more effectively.

The work presented here enables users of all technical and athletic skill levels to go from watching a football play on TV to experiencing it in virtual reality within minutes. The power and control that a simulation provides over a tradition training setting, the lower barrier-to-entry from idea to execution, and the potential to tap into a library of thousands of famous, unique, or challenging plays with little effort can revolutionize the way a professional quarterback approaches training.

6 REFERENCES

Appelbaum, L. G. and Erickson, G. (2018). Sports vision training: A review of the state-of-the-art in digital training techniques. *International Review of Sport and Exercise Psychology*, 11(1):160–189.

Atmosukarto, I., Ghanem, B., Ahuja, S., Muthuswamy, K., and Ahuja, N. (2013). Automatic recognition of offensive team formation in American football plays. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 991–998.

Bideau, B., Multon, F., Kulpa, R., Fradet, L., and Arnaldi, B. (2004). Virtual reality applied to sports: Do handball goalkeepers react realistically to simulated synthetic opponents? In *Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry*, VRCAI '04, pages 210–216, New York, NY, USA. ACM.

Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., and Bailenson, J. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2):103–124.

Bougue, J.-Y. (2001). Pyramidal implementation of the lucas kanade feature tracker.

Covaci, A., Postelnicu, C., Panfir, A. N., and Talaba, D. (2012). A virtual reality simulator for basketball free-throw skills development. In *Technological Innovation for Value Creation*, pages 105–112, Berlin, Heidelberg. Springer Berlin Heidelberg.

Dearden, A., Demiris, Y., and Grau, O. (2006). Tracking football player movement from a single moving camera using particle filters. In *Proceedings of the Conference on Visual Media Production*, pages 29–37. IET Press.

Gavarkovs, A. G. (2019). Behavioral counseling training for primary care providers: Immersive virtual simulation as a training tool. *Frontiers in Public Health*, 7:116.

Huang, Y., Churches, L., and Reilly, B. (2015). A case study on virtual reality American football training. In *Proceedings of the 2015 Virtual Reality International Conference*, pages 6:1–6:5, New York, NY, USA. ACM.

Hudson, I. and Hurter, J. (2016). Avatar types matter: Review of avatar literature for performance purposes. In Lackey, S. and Shumaker, R., editors, *Virtual, Augmented and Mixed Reality*, pages 14–21, Cham. Springer International

Publishing. Part of LNCS, Vol. 9740.

Klinger, E., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., and Nugues, P. (2005). Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. *CyberPsychology & Behavior*, 8(1):76–88.

Kudrayvtsev, G. (2019). Section 4.1: Smoothing. In *Algorithms for Robotics*, chapter 4.

Lucas, B. D. and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th Int'l Joint Conference on AI*, volume 2, pages 674–679, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

M. North, M., M. North, S., and R. Coble, J. (2015). Virtual Reality Therapy: An effective treatment for the fear of public speaking. *International Journal of Virtual Reality*, 03(3):1–6.

Matas, J., Galambos, C., and Kittler, J. (2000). Robust detection of lines using the progressive probabilistic Hough transform. *Computer Vision and Image Understanding*, 1(78):119–137.

McMahan, R. P., Lai, C., and Pal, S. K. (2016). Interaction fidelity: The uncanny valley of virtual reality interactions. In Lackey, S. and Shumaker, R., editors, *Virtual, Augmented and Mixed Reality*, pages 59–70, Cham. Springer International Publishing. Part of LNCS, Vol. 9740.

Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics Automation Magazine*, 19(2):98–100.

Myers III, G. L. (2019). *The V.A.L.O.R. leadership shift: The impact of identity verification through virtual platform interactions*. PhD thesis, Kent State University.

Parsons, T. D. and Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, 39(3):250–261.

Roberts, D., St. Clair, K., Wojtan, C., and Essa, I. (2007). Interactive 3D visualization of archived college football video. *DVFX 2007 Video Productions*.

Schuemie, M. J., van Der Straaten, P., Krijn, M., and van Der Mast, C. A. (2001). Research on presence in virtual reality: a survey. *Cyberpsychology & Behavior*, 4(2):183–201.

Shi, J. and Tomasi, C. (1994). Good features to track. In *Proceedings of the IEEE*

Conference on Computer Vision and Pattern Recognition, Seattle, WA.

Singh, D. K., Rajaratnam, B. S., Palaniswamy, V., Pearson, H., Raman, V. P., and Bong, P. S. (2012). Participating in a virtual reality balance exercise program can reduce risk and fear of falls. *Maturitas*, 73(3):239–243.

Sports VTS (2019). QBSIM: Real Training, Real Results.

Stinson, C. and Bowman, D. A. (2014). Feasibility of training athletes for high-pressure situations using virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 20(4):606–615.

STRIVR (2019). A whole new world of sports training.

Thrun, S. (2014). Lesson 5: Pid control. In *Artificial Intelligence for Robotics*, chapter 5. Udacity.

Tsai, W., Chung, M., Pan, T., and Hu, M. (2017). Train in virtual court: Basketball tactic training via virtual reality. In *Proceedings of the 2017 ACM Workshop on Multimedia-based Educational and Knowledge Technologies for Personalized and Social Online Training*, MultiEdTech '17, pages 3–10, New York, NY, USA. ACM.

Warren, L. E. (2018). The effect of interaction fidelity on user experience in virtual reality locomotion. Master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.