Computer Vision for Sports: Basketball

Ivan Martini
MSc in Computer Science, DISI
Trento, Italy
ivan.martini@studenti.unitn.it

***Abstract*—I usually like doing sports, but I never have been a huge fan of watching them live or in TV, as i find them mostly boring. There are some rare exception (mainly e-sports and some basketball), so i decided to seize the opportunity and work on the latter, elaborating also the shallow knowledge i've got of the sport itself. The literature has a lot of paper on the topic, but most of them are just experimental or not implemented, therefore did not have an impact on the sport, which is the target of this essay. Three interesting products will be analyzed, as they had in different ways quite some impact on the game itself.**

I. Introduction

This essay has the goal of providing a summary of the changes that the computer vision brought into sport fields, with the personal choice of analyzing basketball. The current literature is quite dense of sperimental papers or innovative researchs, but most of them are academic result and did not really change the sport itself. For this reason I decided to narrow down the focus mainly on the *National Basketball Association*, in order to produce a deeper and detailed report. The NBA is indeed the most important organization ruling over this sport and should be quite representative.

The structure is as follow: In the 2nd section there will a brief introduction on the general technologies that enhanced the sport and how, while in the 3rd there will be a more detailed and beefy discussion on the actual computer vision technologies. Finally the conclusion, with some personal reflection. In each section, where possibile, there will be a little debrief on interesting experimental papers.

II. The evolution of basketball through video technology

The power of computer vision, expecially if applied to machine/deep learning, is pretty obvious, however the whole referees' system takes only a little advantage from it. Despite being well trained and professional, the Officials that rule the NBA matches are still human, and operate with an accuracy of 95% [1] [2]. Following this leads, I found that the sports world (which is mostly unknown to me) of both athletes and fans has an heated debate on the use of Artificial Intelligence to rule the matches. The league is not new to change in the Officials setup, (until 1988 there were only 2 referees on the field, now they are 3) but it always took carefully the question, delaying the introduction of technologies for helping the referees [3] [4].

The only technology allowed, which is not computer vision related is the use of the *Replay Center* [5], a "room" with 94 screens and 20 operators that replay and review the matches while they happen. If the 3 officials are not sure of what decision take, the replay center guides them and provide a safe answer.

III. Computer Vision technologies in basketball

After an introduction at some shallow video applications, it is time to get deeper inside the computer vision products that really got an impact on the sport. In the previous section it has already been clarified why computers aid only a little the officials during the game, so most of the efforts are either in the offgame/training domain or in the broadcast one. Among the possibilities, I decided to pick three different technologies, the ones i found more interesting (and well documented), in order to produce a more detailed and focused report on all of them, avoiding a simple list of technologies briefly explained.

A. Stats Perform - SportVU

The first interesting product that I am going to analyze is the *SportVU* [6] system, developed by the data-driven company *Stats Perform* [7]. It has been introduced in the NBA in 2009 and widely adopted by the all 30 teams in late 2013. Being the first analysis tool adopted by the whole league it placed an important milestone in the story of the sport, providing for the first time an unbiased equal evaluation of all the players, which were deeper than just raw numbers (points, blocks, steals and so on) [8].

SportVU operates with 6 high definition camera (4K) divided in 2 clusters (3 camera each, all in the same position), placed above the hoops at 25ft ($\sim$ 7.60mt) from the ground and providing a perpendicular view. The camera are not particularly sofisticated and operates at 25FPS. Using this configuration, the pipeline is able to produce a consistent disposition of the players, analyzing also ball possession, speed, exerted pressure on the opponent and court control, which are hardly feasible without this technology. The introduction of this system revolutioned a lot the consciousness of the coach, which is able to develop tactics which are *suited on both his player and the opponent ones* [9] [10].

Having this highly reliable quantity of data, the step into machine learning was short. Indeed, since 2013 many artificial intelligence engineers have been hired by the company to produce better results, finding patterns among players (comparing for example a rookie with a top player, to understand the potential) or measuring the consistency

in the actions, making it possibile to develop tactics and countertactics centered on one player. Starting from 2019 many of these data has been captured also on some NCAA (i.e *National Collegiate Athletic Association*) matches, in order to produce the same set of data to determine the *NBA draft* [11].

Even if the product is proprietary, its patent can be found easily online and is quite rich of information [12]. Attached to the patent there are some useful schemas which describe the pipeline that the entire system uses. Firstly, after capturing the images a simple **background subtraction** is applied and adjusted, in order to separate the foreground (i.e. the players, the referees and the ball) from the background. At this point a **segmentatation algorithm** is run on the foreground blobs, producing a consistent representation of all the actors moving on the field and determining the bounding boxes. Here picking the ball out of the humans is quite straightforward, as its shape is well known. For distinguishing the humans instead the color is the key: through **histogram correlation** and player's precedent history the player and the officials are identified, computing also the trajectory and some direct informations (speed, spatial positioning).

Although being still partnered, the brand lost its exclusive in player tracking back in 2018, sharing the field with a similar company called *Second Spectrum* [13]. They operate in the same environment, but they are more centered on the broadcast development, confirming that American sports are purely based on entertainment.

B. Second Spectrum - Court Vision

Other than working on the data analysis, *Second Spectrum* is very active in the field of broadcast. For this reason, it stole the exclusive on optical player tracking from *Stats Perform* a few years ago. Their main product on this domain is purely entertaining (so it didn't impact the sport in the strict meaning) but it is strongly computer vision based, so I decided to write a small section about it. Using almost the same technology described above (the 6 camera setup of SportVU) plus some additional side cameras to integrate the data feed, they are able to produce a "real-time" overlay on the live broadcasting. In this way the data gathered for evaluating the players after the match are also used to spice up the broadcasting area [9] [14].

Other than some purely visual effects (like thunders hitting the hoop during a dunk or a flaming "3" after a successful *three-pointer*) available in **mascot mode**, the user can select from the streaming application the **player mode**, which superimposes over the ball owner and his teammates the likelihood to pass or take a shot or the **coach mode**, which displays the tactics adopted while the match goes on, highlighting positioning and weak spots [15].

During an interview [16], the CEO of the company said: "there will be a day, when you look back and say: I can't image we used to watch the same thing at the same time, that seems silly", which underlines perfectly the impact of the product. Even if the augmented reality is really cool and powerful, I feel that the key revolution of this application is the user-centric streaming, giving the power of deciding what to watch and how. In this case is funny, as *computer vision application* (the overlays part) brought a big impact in a *totally unrelated way* (the whole sport streaming concept).

C. Intel - True View

The technologies which interested me most while picking the papers/product to study was the Intel True View [17]. Born in 2011 as **FreeD**, the only product of *Replay Technologies*, it has been acquired by Intel in 2016 (together with the whole company) [18]. The original concept was designed for long field sports like soccer and american football, but has rapidily shown useful also in basketball. Many NBA stadiums adopted it (at least Chicago Bulls, Atalanta Hawks, Sacramento Kings, Washington Wizards, Cleveland Cavaliers, Milwaukee Bucks and Dallas Mavericks [19] [20] [21] [22]).

The name of the product is pretty suggestive of what is the final goal. The analogy of *FreeD* with *3-D* has the evocative meaning of movement freedom. Using from 20 to 40 fixed cameras (28 or 36 in the case of basketball) placed all around the stadium, they are able to produce a quite accurate 3D representation of game, having as a final result a volumetric video, which can be seen by any position (as shown in Fig. 1).

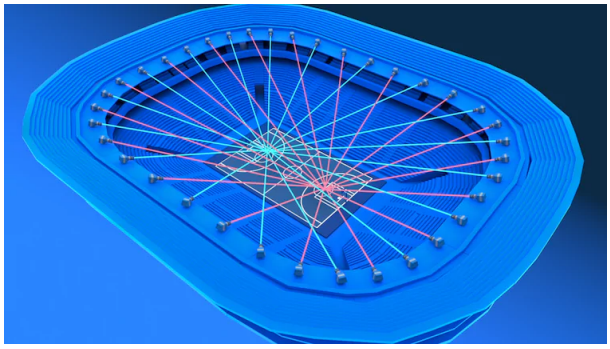

Fig. 1. The 36 cameras dual-focus system used in basketball stadiums [23]. Their fixed position and orientation is the key that enables the reconstruction of the volumetric video

Also the technology is proprietary, therefore there is no way to know how it really works, but some conclusion might be drawn. The cameras are oriented to point at the center of each half-court (18 each in the image above). During the installation (or eventually before the match) they must be calibrated in order to allow the system to know the mutual displacement (of both distance and rotation) of all the cameras. At this point, thanks to the different angulation, it should be possibile to apply a mesh reconstruction algorithm starting from the fundamental matrix of the system (which is a transformation function reconstructing the depth of a point from at least two camera). This is feasible only

using the previous fixed calibration, as there is the need for many *equations* in order to produce some *unknowns*. The complexity **should be linear** with the number of cameras, therefore once solved with two of them, it should be easy[1] to translate the solution to more cameras.

The infrastructure is obviously fully powered by Intel Machines, with a dedicated server (Xeon Series) for each camera in the stadium itself, using optic fibers connections (which now is pretty common, but wasn't back in 2013). Each server works to preprocess the images and extract relevant data, which will be merged and stored in the cloud, building a video of *voxels* (i.e. volumetric pixel), developing de facto a 3D reconstruction of the match. At this point flying a virtual camera inside the field is straightforward. Once the data are up in the cloud, they are also available to streaming platform, using the most common standard encodings (e.g. H264, H265, M-JPEG) [24] [25].

Although being developed mainly for broadcast applications, the True View system didn't only introduced a completely immersive game experience, but also became a great tool for coaches and trainers, who could improve their teams' tactics. Indeed, another big opportunity offered by the volumetric representation of the match is to get a first person view (FPV) of a specific action/moment and analyze the player behaviour. This last feature is used for broadcasting too and also kicked off some research on the FPV game analysis.

While scraping Google Scholar and the CVPR documents I found some FPV analysis and I did not even consider those as useful, since weareable device (as Action Cams) are quite unpopular in sports. After learning about Intel True View, i changed my mind. The guidelines are the use of Neural Networks to process and analyze FPV gameplays, finding errors or similairties with other players and since they are just experimental I will just leave the references [26] [27] [28].

The True View solution has been quite a revolution, improving the user experience a lot, both at the stadium and at home. The volumetric data produced are usable in many different context (Virtual and Augmented Reality, or during video game development) so the challenge to face now is probabily different: It is possible to obtain something good as True View using less resources? With the installation cost of more or less 1.100.000$ and all the other costs coming from the infrastructure (True View produces more or less 85GB of data each second [20] [24] [29]), this technology is only available to major teams, penalizing and widening the gap with minor leagues.

IV. Conclusions

In this essay there is a journey through some of the technologies that nowadays are integrated in the basketball world. Some of them are strictly related to this sport, while other are broader. In different ways, these three products impacted the game, making relevant changes to the tissue of the sport and opening up many possibilities.

I feel that the common thread of the three solutions presented above is the high performance and accuracy developed by big enterprises that have a huge quantity of resources. Since the target is the NBA (which obviously can allocate lots of money) the super expensive and performing *high end product* is natural, but this set-up clips the wings of many researcher/little companies which might have revolutionary ideas but can not compete with the big corporations.

To me, the future of the research should concentrate a lot on recreating the existing result with less resources and in a cheaper but reliable way, instead of trying to improve what already exist (which, however, could clearly be improved).

[1] At least I suppose that

References

[1] Ben Dowsett. The nba's referee 'problem' will last as long as we want it to. *Forbes*, mar 2019.
[2] https://en.wikipedia.org/wiki/List_of_National_Basketball_Association_referees.
[3] Spencer Garfield. Robot referees in basketball. *PSU Vanguard*, apr 2020.
[4] Kyle Bonn. Fifa defends refereeing, var criticism in press conference. *NBC sports*, jun 2019.
[5] https://official.nba.com/replay/.
[6] https://www.nbastuffer.com/analytics101/sportvu-data/.
[7] https://www.statsperform.com/.
[8] Andy Bailey. 5 technological advances that are changing the nba. *Bleacher Report*, may 2016.
[9] Marcus Woo. Artificial intelligence in nba basketball. *Inside Science*, dec 2018.
[10] https://www.statsperform.com/team-performance/basketball/optical-tracking/.
[11] Press Report. Magic become first nba team to use ai player tracking data. *Real GM*, feb 2019.
[12] Gal Oz Michael Tamir. Real-time objects tracking and motion capture in sports events, aug 2008. US Patent App. 11/909,080.
[13] https://www.secondspectrum.com/index.html.
[14] Kevin 0'Connor. Pictures at a basketball revolution. *The ringer*, oct 2018.
[15] https://www.clipperscourtvision.com/.
[16] https://www.youtube.com/watch?v=9TxFKz1hVcA.
[17] https://www.intel.it/content/www/it/it/sports/technology/true-view.html1.
[18] https://www.intel.it/content/www/it/it/sports/technology/true-view.html.
[19] Sam Carp. Atlanta hawks bringing intel true view to state farm arena. *SportsPro*, nov 2019.
[20] Press Release. Sacramento kings launch immersive media experiences with intel® true view. *NBA*, feb 2021.
[21] Steven Impey. Washington wizards and intel bring true view to capital one arena. *SportsPro*, oct 2019.
[22] Press Release. Bucks partner with intel sports to bring true view to fiserv forum. *NBA*, oct 2019.
[23] https://www.nba.com/bucks/news/bucks-partner-intel-sports-bring-true-view-fiserv-forum.
[24] Harleen Gill and Ritesh Kale. Svg tech insight: Immersive media experiences with intel true view delivers new reality for sports, 2020.
[25] https://www.intel.it/content/www/it/it/sports/technology/true-view.html.
[26] Jianbo Shi Hyun Soo Park Shan Su, Jung Pyo Hong. Predicting behaviors of basketball players from first person videos, 2017.
[27] Stella X. Yu Jianbo Shi Gedas Bertasius, Hyun Soo Park. Am i a baller? basketball performance assessment from first-person videos, 2017.
[28] Jianbo Shi Gedas Bertasius, Aaron Chan. Egocentric basketball motion planning from a single first-person image, 2018.
[29] The Stadium Business Press. Intel brings true view tech to ligue 1 stadia, 2020.