Comparing Free-Throw Forms Among NBA Players Through 3D Similarity Measures

Paul Ibrahim[1]
[1] Cary Academy

Abbreviated abstract: In this paper, we propose a method to compare free-throw forms of various NBA players. To characterize each player's form, we apply a multivariate kernel density estimation to the sample of the player's free throw attempts. We then proceed to apply a variety of three-dimensional similarity measures between both individual player and clustered kernel density estimates, therein providing a variety of metrics by which we can assess free-throw form similarity among NBA players.

Tracking Data Source: linouk23. "linouk23/NBA-Player-Movements." *GitHub*, 19 Sept. 2016, github.com/linouk23/NBA-Player-Movements/tree/master/data/2016.NBA.Raw.SportVU.Game.Logs

UCSAS 2020

Modeling the Data

In our analysis, we use publicly available SportVU tracking data from the 2015-16 NBA season [1]. Over the course of a basketball game, a player's shooting form is evident in two distinct situations: jump shots in the midst of gameplay, and free-throw attempts.

To afford the maximum possible approaches of spatial comparison, we modeled each player's shooting form with two methods: trivariate local regression (LOESS) and three-dimensional kernel density estimation (KDE). An example is shown below in Figure 1 .

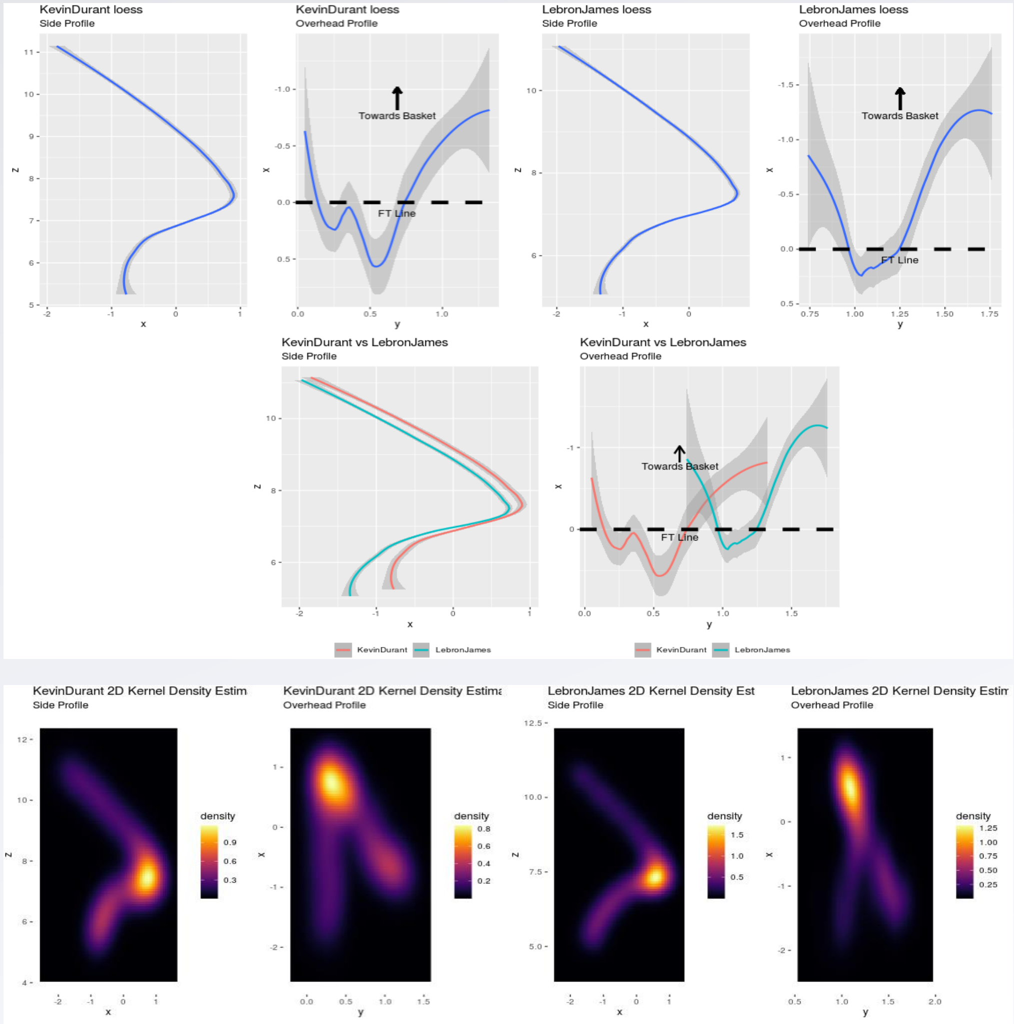

Figure 1: LOESS and KDE Comparative Representations

UCSAS 2020

Three-Dimensional Methods of Comparison

Method 1: LOESS Curve Frechet Distance

To overcome the obstacle posed by temporal sampling, we can use the LOESS model we previously constructed .Systemically sampling 500 points from each player's LOESS fit, we can use the Frechet Distance algorithm to map each point to its approximate corresponding point on the other player's form curve and calculate a similarity measure between the two player's curves.

Method 2: Bin Difference

We can find the absolute difference in density estimate per corresponding cell between the two players' 3D KDE arrays, and sum these absolute differences together to represent a holistic similarity measure.

Method 3: Kullback-Leibler Divergence

Using the formula given on the left we can calculate a three-dimensional Kullback-Leibler Divergence. We can substitute our determined three-dimensional kernel density estimates (integrates to 1) for two given players in lieu of probability distributions $p(t)$ and $q(t)$ while maintaining the integrity of the Kullback-Leibler Divergence.

$$KL(p,q) = \iiint_{-\infty}^{\infty} p(t) \log\left(\frac{p(t)}{q(t)}\right) dt$$

Where $t = (x, y, z)$, $dt = (dx, dy, dz)$, $p(t)$ represents the target probability distribution and $q(t)$ represents the competing probability distribution

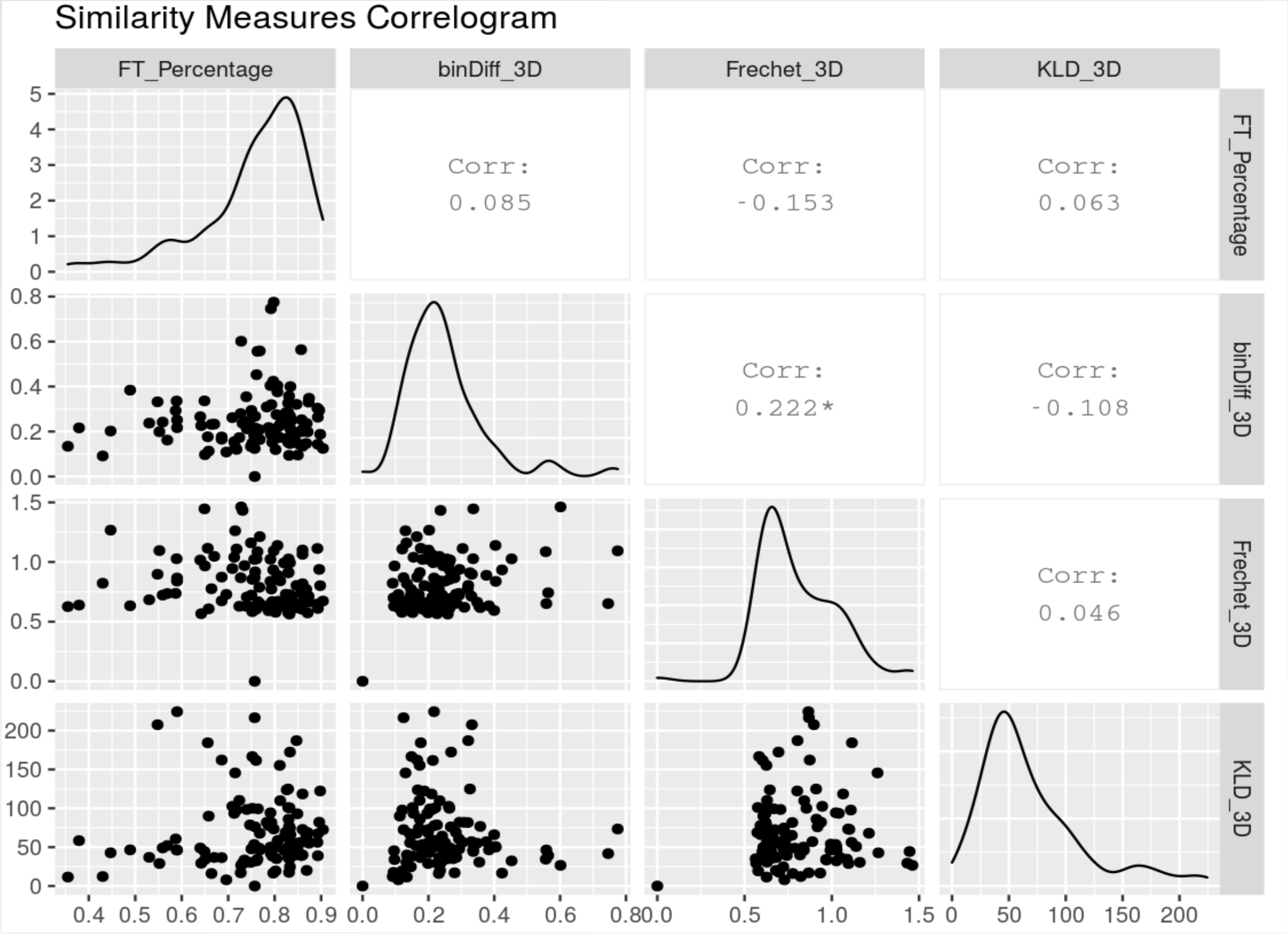

Figure 2: Similarity Measures Correlogram

Paul_Ibrahim@caryacademy.org - 3

UCSAS 2020

Clustering Shot Forms

In order to conduct a more generalized analysis on player shot forms, we can sort them into clusters using the k-means clustering algorithm. We identify that the sum of squares is minimized for our sample of players at 7 clusters. We characterize each cluster in Figure 3 by the cluster mean.

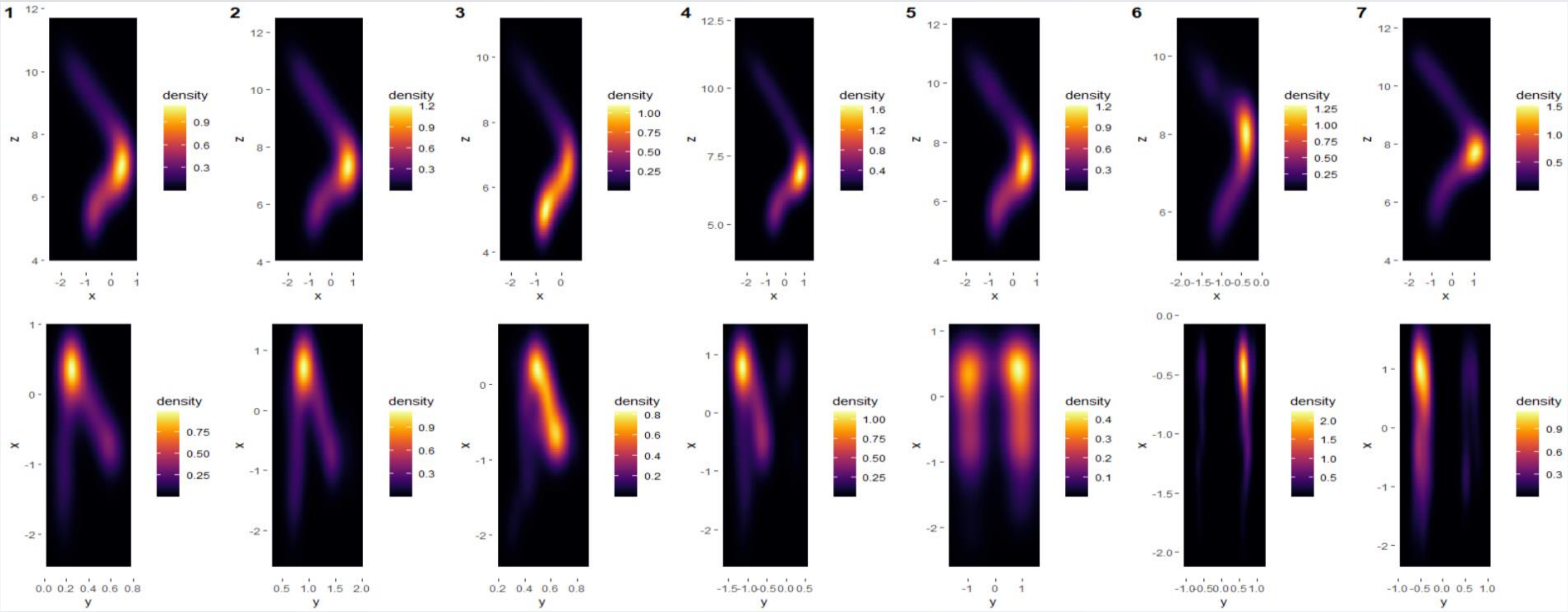

Figure 3: Clustered Shot Forms from Side and Overhead Vantage

Cluster	FT%	3D Frechet Distance	3D Bin Difference	3D Kullbach-Leibler Divergence
1	0.593	0.52732	1.67749	10.17620
2	0.685	0.89270	1.58114	7.41660
3	0.779	0.63914	1.60710	4.67237
4	0.548	0.74580	1.74398	9.94691
5	0.752	0.48673	1.79881	12.38330
6	0.682	0.66081	1.92376	17.48540
7	0.764	0.53964	1.83774	14.67780

Table 1: Cluster Comparative Table

Repeating the methodology outlined in section 3, we compare each cluster to the composite average of all player shooting forms with each of the outlined similarity measures.