

PERBAIKAN INTERPOLASI GERAKKAN MODEL SKELETON 3D DARI DATASET HASIL KAMERA KINECT

Darma Rusjdi^{(1)*}, Dewi Arianti Wulandari⁽²⁾, Efy Yosrita⁽³⁾

⁽¹⁾Program Studi Informatika, Institut Teknologi PLN, Jakarta, Indonesia
Jl. Lkr. Luar Barat, Duri Kosambi, Cengkareng, Kota Jakarta Barat, Jakarta
e-mail : darma@itpln.ac.i⁽¹⁾

Abstract

The Kinect camera's dataset can be used for training and testing of human movement recognition using a deep learning approach, in addition to tracking and estimating the position and position of the human body. Improvement of human movement which is covered by other body parts is still a challenge. The research objective: to design a repair application to move the 3D skeleton model from the Kinect camera dataset using the 3D interpolation histogram smoothing approach on the human body. The study was in the form of a simulation of a skeleton model movement improvement in the context of developing a key-frame based animation application through Kinect camera recording. The prototype development method goes through the cycle stages of analysis, design, implementation. The initial analysis stage selects the kinect camera dataset by examining the file format and data structure. Furthermore, the development for data improvement through the 3D interpolation refinement approach. Movement improvement in terms of measurement of interpolated resultant errors from each time t_0 , t_1 , t_2 and t_3 histogram using the RMSE method and visual observation. showed significant results close to normal movement.

Keywords : *dataset, kinect camera, interpolation, skeleton model*

Dataset hasil kamera Kinect dapat digunakan untuk pelatihan dan pengujian pengenalan gerakan manusia dengan pendekatan deep learning, selain tracking dan estimasi letak dan posisi tubuh manusia. Perbaikan gerakan manusia yang tertutup bagian tubuh lainnya masih menjadi tantangan. Tujuan penelitian: membuat rancang bangun aplikasi perbaikan gerakan model skeleton 3D dari dataset hasil kamera Kinect menggunakan pendekatan penghalusan histogram interpolasi 3D pada bagian tubuh manusia. Kajian berupa simulasi perbaikan gerakan model skeleton dalam rangka pengembangan aplikasi animasi berbasis key-frame melalui perekaman kamera Kinect. Metode pengembangan prototipe melalui tahapan siklus analisis, disain, implementasi. Tahap analisis awal memilih dataset hasil kamera kinect dengan mengkaji format file dan struktur data. Kemudian mengekstraksi beberapa kategori gerakan kedalam kedalam format file untuk eksperimen. Melalui disain awal pengembangan program dilakukan penyesuaian skala angka untuk menghasilkan histogram yang dapat memperlihatkan bagian kesalahan gerakan secara signifikan menggunakan Root Mean Squared Error (RMSE). Selanjutnya pengembangan untuk perbaikan data melalui pendekatan penghalusan interpolasi 3D. Perbaikan gerakan ditinjau dari pengukuran kesalahan resultante interpolasi dari setiap waktu t_0 , t_1 , t_2 dan t_3 histogram dengan metode RMSE dan pengamatan secara visual. menunjukkan hasil cukup signifikan mendekati gerakan yang normal.

Kata Kunci : *dataset, kamera kinect, interpolasi, model skeleton*

1. PENDAHULUAN

Model obyek rangka dalam aplikasi animasi keyframe memiliki kekurangan untuk membuat key-pose untuk gerakan yang natural. Model obyek rangka merupakan salah satu pembentukan model obyek grafika komputer. Dibandingkan model lainnya seperti model bidang dan model padat, model rangka menjadi grafik dasar pengembangan suatu model 3dimensi. Implementasi ke dalam aplikasi animasi keyframe juga menjadi tahapan awal untuk pengembangan ke model bidang atau model padat untuk animasi 3 dimensi. Pembuatan key-pose suatu obyek model menjadi acuan dasar didalam pembuatan animasi berbasis keyframe. Oleh karena keterbatasan

didalam pembuatan keypose ini menyebabkan kualitas gerakan obyek dalam animasi keyframe menjadi kurang luwes bergerak secara natural.

Key-pose tubuh manusia dimodelkan dengan titik-titik sambungan bagian tubuh (manusia) dan grafik penghubung dua titik dapat membentuk model animasi 3D (Etemad, 2016; Rusjdi, 2016). Titik-titik bagian tubuh yaitu: kepala, leher, bahu, dada, perut, pinggul, dengkul/lutut, engkel/tumit, dan kaki. Selanjutnya dari bahu ada titik siku, pergelangan tangan, dan jari. Untuk detail jari dilengkapi dari jempol, dan empat jari. Pembentukan tubuh terjadi dengan menghubungkan titik-titik bagian tubuh sehingga dapat menghasilkan bentuk skeleton. Bentuk skeleton adalah bentuk yang paling sederhana menggambarkan pose atau posisi, arah dari bagian tubuh.

Pada kajian penentuan key-pose berdasarkan dataset hasil kamera kinect menunjukkan ada kelemahan berupa gerakan yang kurang alami berupa noise. Kelemahan kamera karena keterbatasan jarak atau pencahayaan yang mengganggu kualitas gambar dan menimbulkan noise gerakan. Noise gerakan yang diakibatkan oleh terhalangnya dari bagian tubuh lain maupun kelemahan dari kamera kinect. Keterbatasan jarak penangkapan gambar oleh kamera kinect disebabkan karena teknologi kamera dalam menangkap menggunakan semacam sensor jarak. Jarak pada setiap permukaan benda dihasilkan dalam bentuk perbedaan warna. Batas kemampuan menangkap jarak dari kurang 15 cm sampai dengan 3-4 meter. Namun keterbatasan kamera menangkap warna benda dipengaruhi oleh pancaran atau pantulan dari sumber pencahayaan. Semakin terang maka warna cenderung ke arah putih sedangkan berkurangnya cahaya akan menghasilkan warna cenderung ke arah warna hitam.

Namun untuk berbagai penelitian tersebut perlu dilakukan implementasi metode yang lebih simple yang belum membutuhkan tingkat akurasi yang tinggi yang dipengaruhi titik sambung tubuh yang mempengaruhi posisi titik sambung di tingkat bawahnya secara alami. Misalnya ketika titik siku diangkat mempengaruhi gerakan titik pergelangan tangan, atau ketika titik dengkul diangkat akan mempengaruhi titik pergelangan kaki. Dan Ruas antara titik sambung bersifat tetap, misalnya panjang lengan antara titik bahu dan titik siku, antara titik siku dan titik pergelangan tangan relatif tetap panjangnya. Sehingga perubahan atau perbaikan gerakan titik bahu mempunyai pengaruh pada perubahan titik siku selain perubahan hasil pengolahan dengan kamera (kinect).

Noise gerakan tergambar dari histogram berupa perbedaan perubahan vector (dv_{12}). Perbedaan nilai vektor berupa resultante dari proyeksi vektor x searah sumbu x , vektor y searah sumbu y , vektor z searah sumbu z . Dua perubahan vector minimal dihasilkan oleh 4 buah titik sambung (tubuh) (p_1, p_2, p_3, p_4). Titik pertama (p_1) dan titik kedua (p_2) menghasilkan arah vektor kesatu (v_1). Titik kedua (p_2) dan titik ketiga (p_3) menghasilkan arah vektor kedua (v_2). Dari kedua vektor ini menghasilkan perbedaan arah vektor kesatu (dv_1) yaitu berupa perbedaan vektor pada sumbu x , perbedaan vektor pada sumbu y , dan perbedaan vektor pada sumbu z . Dan perbedaan arah vektor kedua (dv_2) didapat dari perbedaan antara vektor ketiga (v_3) dihasilkan dari titik kedua (p_2) dan titik ketiga (p_3) dan vektor keempat (v_4) dihasilkan dari titik ketiga (p_3) dan titik keempat (p_4).

$$v_1 = p_1 - p_2 \quad (1)$$

$$v_2 = p_2 - p_3 \quad (2)$$

$$v_3 = p_3 - p_4 \quad (3)$$

$$dv_1 = v_2 - v_1 \quad (4)$$

$$dv_2 = v_3 - v_2 \quad (5)$$

$$dv_{12} = dv_2 - dv_1 \quad (6)$$

Pada perubahan gerakan yang halus atau natural ditunjukkan oleh 2 atau lebih perubahan vector yang tetap atau sedikit perubahan. Jika perbedaan perubahan vector berbeda secara signifikan menunjukkan adanya noise dari gerakan.

Untuk mengatasi hal tersebut solusi yang diusulkan dalam kajian ini dengan menggunakan penghalusan setiap posisi gerakan dengan pendekatan metode kurva bezier (minimal dengan tiga titik). Persamaan metode bezier 3 titik pendukung sebagai berikut:

$$P = P_0.(1-u)^2 + 2.P_1.(1-u)(u) + P_2.u^2 \quad (7)$$

Dimana u adalah parameter posisi titik berada lintasan garis yang dibentuk oleh persamaan Bezier. Dan nilai u diantara nilai 0 sampai dengan 1 [$0 \leq u \leq 1$]. Jika $u=0$ adalah P_0 dan $u=1$ adalah P_2

Untuk persamaan metode Bezier dengan 4 titik pendukung yaitu:

$$P = P_0.(1-u)^3 + 3.P_1.(1-u)^2(u) + 3.P_2.(1-u)(u)^2 + P_3.u^3 \quad (8)$$

Dengan persamaan ini dapat menghasilkan posisi (u) perbaikan yang berbeda dengan pengaruh titik keempat menyebabkan perubahan vektor yang berbeda.

Untuk menentukan perubahan dari titik awal, nilai u bisa didapat dari perbandingan resultante vektor yang dibentuk oleh titik sasaran dengan total Panjang keseluruhan vector yang dibentuk oleh keempat titik. Secara umum hal ini dapat dilakukan untuk semua titik dimulai dari titik pertama ke titik kedua, titik ketiga, dan titik keempat. Selanjutnya mulai dari titik kedua ke titik ketiga, titik keempat, dan titik kelima, dan seterusnya.

2. METODE PENELITIAN

Metode penelitian dilakukan secara eksperimen dengan mengimplementasikan algoritma kurva Bezier untuk memperbaiki noise gerakan model animasi 3D. Pengumpulan data menggunakan dataset hasil Kinect (Xia, 2012; Kao, 2019) yaitu Skeletal joint Locations.txt yang berisi. Sepuluh aktivitas gerakan. Titik-titik data tersebut diketahui ada 20 titik sambung tubuh (j) masing-masing dengan koordinat tiga dimensi (x,y,z). Banyaknya pose gerakan (frame skeleton) adalah m baris data. Struktur data terdiri dari kode frame, $x_1,y_1,z_1,x_2,y_2,z_2,\dots,x_{20},y_{20},z_{20}$. Dari masing-masing titik-titik dilakukan sambungan proses gerakan untuk mencari perubahan arah vektor dengan menggunakan persamaan (6).

Pada pembacaan data pecahan harus dikonversi nilai yang dapat ditampilkan ke layer dalam ukuran piksel sehingga perlu disesuaikan dengan ukuran layer tampilan. Variabel yang akan diukur perubahan arah vector berdasarkan perubahan dari ketiga sumbu x , y , dan z . Dengan titik awal setiap titik $P(m,j)$ dilakukan pencarian perubahan arah vector tersebut dv_{mn} yaitu: $dv(j,1)$, $dv(j,2)$, ..., $dv(j, n)$, dimana n adalah $m-3$ titik akhir disesuaikan menggunakan persamaan (8).

Tabel 1. Perhitungan nilai $abs(dv_{12}) \leq 2$ (posisi Gerakan normal)

frame	X	Y	Vx	Dvx	Dv12x
1	0	0			
2	20	1	2.861253573		
3	40	2	2.861253573	0	
4	60	2	0	-2.861253573	-1.430626787

Tabel 2. Perhitungan nilai $\text{abs}(\text{dv}12) > 2$ (posisi gerakan kurang wajar)

frame	x	Y	vx	dvx	Dv12x
1	0	0			
2	20	5	14.03059616		
3	40	7	5.708295552	-8.322300609	
4	60	5	-5.708295552	-11.4165911	-9.869445856

Perbedaan perubahan arah mendekati atau sama dengan nol menunjukkan gerakan yang normal. Jika perubahan berbeda secara signifikan >2 menunjukkan ada kesalahan (noise) pada gerakan sehingga harus diperbaiki dengan persamaan kurva bezier. Sehingga Langkah berikut mendata setiap perubahan dari awal gerakan sampai dengan pose terakhir. Pendataan ini ditampilkan dalam diagram histogram yang secara visual terlihat titik-titik sambung yang memiliki noise gerakan. Langkah berikutnya mengimplementasi persamaan kurva Bezier dan mengulangi Langkah penghitungan perbedaan perubahan arah vector.

Secara keseluruhan dibuat algoritma dan dibuatkan program aplikasi menggunakan Bahasa pemrograman visual basic. Implementasi pengembangan aplikasi dapat berupa gerakan animasi berbasis keyframe (Rusjdi, 2016) dengan metode perancangan berbasis prototipe berupa persiapan, siklus analisis, pengembangan, dan pengujian. Aplikasi selesai jika hasil dinilai mencukupi kebutuhan. Implementasi aplikasi dengan melihat perubahan yang optimal.

Setelah diimplementasi dibandingkan hasilnya berdasarkan histogram dan pengamatan gerakan model dari awal dan akhir. Analisis dan model statistik yang digunakan menghasilkan histogram yang dapat memperlihatkan bagian kesalahan gerakan secara signifikan menggunakan Root Mean Squared Error (RMSE).

$$\text{RMSE} = \left(\frac{\sum (y_i - \hat{y}_i)^2}{n} \right)^{1/2} \quad (9)$$

Keterangan:

RMSE = nilai root mean square error

y = nilai hasil observasi

\hat{y} = nilai hasil prediksi

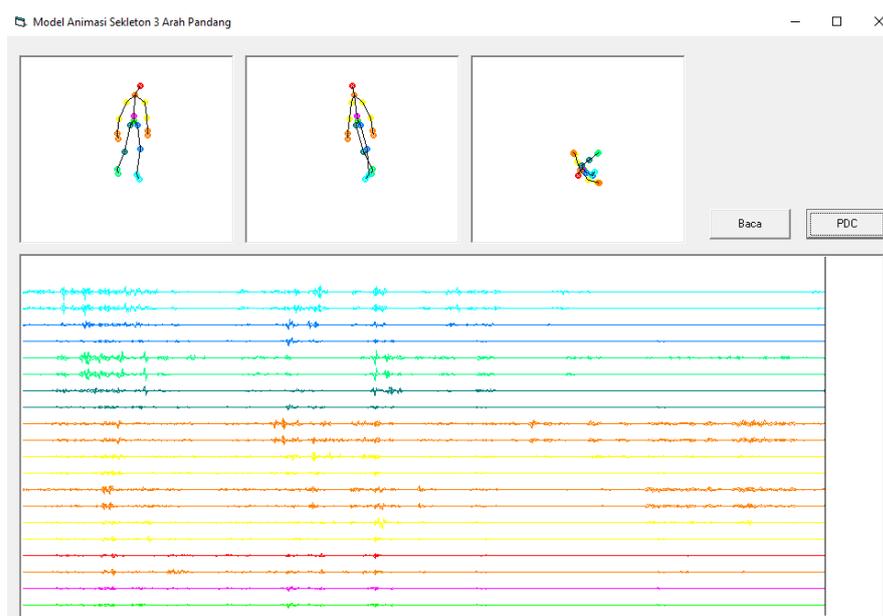
i = urutan data pada database

n = jumlah data

Selanjutnya pengembangan untuk perbaikan data melalui pendekatan penghalusan interpolasi 3D. Proses perbaikan dilakukan sampai menghasilkan gerakan yang cukup baik. Perbaikan gerakan ditinjau dari pengukuran kesalahan resultante interpolasi dari setiap waktu t_0 , t_1 , t_2 dan t_3 histogram dengan metode RMSE dan pengamatan secara visual.

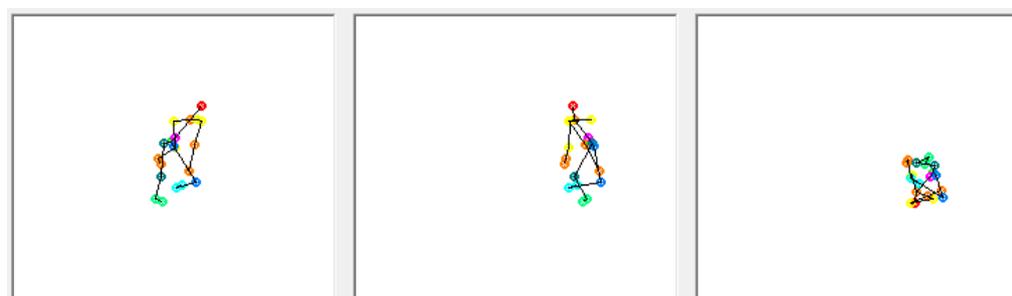
3. HASIL DAN PEMBAHASAN

Tampilan antarmuka aplikasi terdiri dari histogram perubahan gerakan setiap titik sambung (join) dan tiga buah pandangan dari depan, samping dan dari atas yang menampilkan skeleton. Serta dua buah tombol untuk menampilkan gerakan skeleton berdasarkan data hasil kinect dan tampilan gerakan skeleton yang sudah diperbaiki.

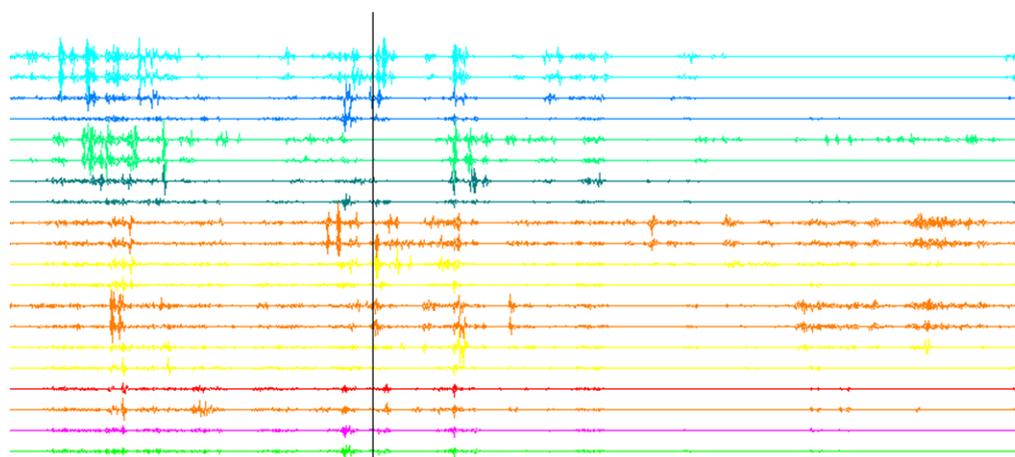


Gambar 1. Tampilan antarmuka aplikasi.

Tampilan gerakan skeleton yang terdiri dari 20 bagian sambungan tubuh (join) ditampilkan melalui tiga bidang gambar yang menunjukkan pandangan dari depan, dari samping dan dari atas. Bagian histogram menunjukkan tampilan yang sudah diperbaiki agar terlihat perubahan grafik pada bagian yang kurang baik.

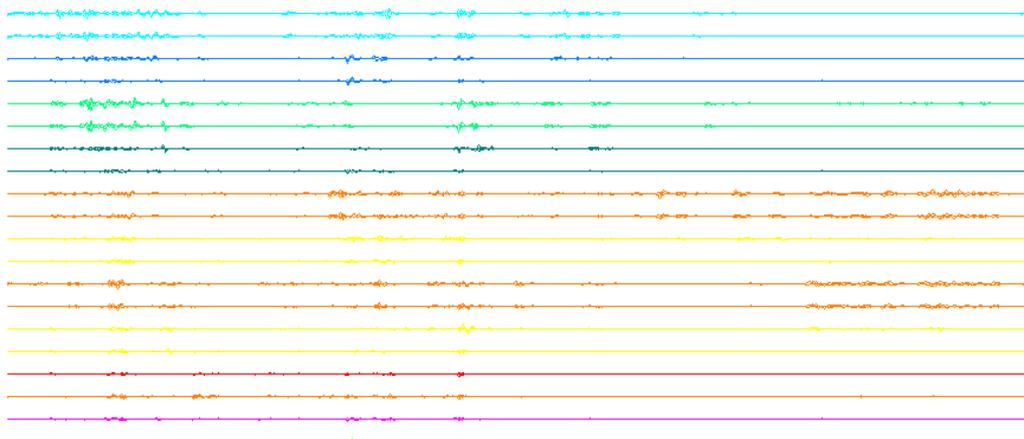


Gambar 2. Tampilan skeleton dari tiga arah



Gambar 3. Histogram gerakan 20 titik sambung (joint) asli

Setelah perbaikan histogram terlihat lebih halus, menunjukkan gerakan sudah semakin halus dan wajar.



Gambar 4. Histogram perbaikan gerakan 20 titik sambung (joint)

Hasil yang didapat berupa gerakan yang lebih natural dengan penghalusan perbedaan perubahan vector. Perbaikan gerakan ditinjau dari pengukuran kesalahan resultante interpolasi dari setiap waktu t_0 , t_1 , t_2 dan t_3 histogram dengan metode RMSE dari 650 menghasilkan kesalahan sebesar 0,56273 cukup kecil namun beberapa bagian terdapat kesalahan yang cukup terlihat secara signifikan. Dan pengamatan secara visual menunjukkan hasil cukup signifikan mendekati gerakan yang normal. Gerakan terlihat cukup membaik namun perubahan pada bagian yang tertutup kurang sesuai dengan harapan. Hal ini dikarenakan perbaikan bersifat global. Untuk kedepannya dapat dikembangkan metode yang memperbaiki bagian yang secara khusus diperbaiki. Hal ini disebabkan oleh teknologi kamera RGBD Kinect yang masih memiliki beberapa kelemahan yang digunakan.

KESIMPULAN DAN SARAN

Gerakan model animasi dari hasil kamera Kinect terdapat noise gerakan. Noise gerakan ditunjukkan berdasarkan perbedaan arah gerakan masing-masing titik sambung divisualisasikan didalam histogram. Pendekatan perbaikan dilakukan melalui persamaan kurva Bezier melalui empat titik masing-masing bagian tubuh. Dengan menampilkan histogram dan pengamatan gerakan didapatkan perbaikan berupa penghalusan noise gerakan menjadi lebih kecil. Untuk kajian masa mendatang dapat dilakukan pendataan skeleton melalui kamera kinect dan kamera RGB secara langsung dan mengembangkan perbaikan gerakan dengan pendekatan yang lebih baik.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada LPPM ITPLN yang telah memberi dukungan yang membantu pelaksanaan penelitian dan penulisan artikel.

DAFTAR PUSTAKA

Etamad, S. Ali, Ali Arya, Avi Parush and Steve DiPaola (2016), Perceptual validity in animation of human motion, *Computer Animation And Virtual Worlds*,; 27:58–71
Han, F.; B. Reily; W. Hoff; and H. Zhang.(2016). Space-time representation of people based on

- 3d skeletal data: a review. *arXiv preprint arXiv:1601.01006*, 2016.
- Kao, J.-Y.; Ortega, A.; Tian, D.; Mansour, H.; Vetro, A. (2019). Graph Based Skeleton Modeling for Human Activity Analysis. Mitsubishi Electric Research Laboratories. <http://www.merl.com/TR2019-037> June 28, 2019
- Papadopoulos, G.T.; Axenopoulos, A.; Daras, P. Real-time skeleton-tracking-based human action recognition using kinect data. In Proceedings of the International Conference on Multimedia Modeling, Dublin, Ireland, 6–10 January 2014; pp. 473–483.
- Paul, S.N.; Singh, Y.J. Survey on Video Analysis of Human Walking Motion. *Int. J. Signal Process. Image Process. Pattern Recognit.* 2014, 7, 99–122. [CrossRef]
- Presti, L.L.; Cascia, M.L. 3D Skeleton-based Human Action Classification: A Survey. *Pattern Recognit.* 2016, 53, 130–147.
- Roanna Lun. 2015, A Survey of Applications and Human Motion Recognition with Microsoft Kinect. Electrical Engineering & Computer Science Faculty Publications
- Rusjdi, D, Pengembangan AMORIK Menggunakan Metode Garis Singgung Terhadap Dua Lingkaran Dan Persamaan Kurva Bezier Orde Dua, *Jurnal Petir*, ISSN: 1978-9262, vol.9 no.1 Januari 2016
- S. Song, C. Lan, J. Xing, W. Zeng, and J. Liu. An End-to-End Spatio-Temporal Attention Model for Human Action Recognition from Skeleton Data. In *AAAI*, volume 1, page 7, 2017.
- Sarafianos, N., Boteanu, B., Ionescu, B., & Kakadiaris, I. A. (2016). *3D Human pose estimation: A review of the literature and analysis of covariates. Computer Vision and Image Understanding, 152, 1–20.*
- Xia, L., Chen, C.C., and Aggarwal, JK.. View Invariant Human Action Recognition Using Histograms of 3D Joints. *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2012