(12) PATENT APPLICATION PUBLICATION
(19) INDIA
(22) Date of filing of Application :08/12/2020

(21) Application No.202011053368 A
(43) Publication Date : 11/12/2020

(54) Title of the invention : ROBUST REAL-TIME ACTION OR POSE RECOGNITION TRAINING MODEL USING 3D GRAPHICS ENGINE

(51) International classification :G06K 9/62 G06T 7/73 G06N 3/04
(31) Priority Document No :NA
(32) Priority Date :NA
(33) Name of priority country :NA
(86) International Application No :NA
Filing Date :NA
(87) International Publication No : NA
(61) Patent of Addition to Application Number :NA
Filing Date :NA
(62) Divisional to Application Number :NA
Filing Date :NA

(71)Name of Applicant :
1)Mr. Pradeep Bedi
Address of Applicant :Associate Professor Department of Computer Science & Engineering Graphic Era Hill University Bell Road, Clement Town, Dehradun, Uttarakhand, India- 248002 Uttarakhand India
2)Dr. Piyush Choudhary
3)Mr. Awanit Kumar
4)Dr. Durga Prasad Gupta
5)Dr. Geetanjali Amarawat
6)Mr. Shrawan Kumar Sharma
7)Mr. Vijay Kumar Chhipa
8)Mr. Amol Laxman Mangrulkar
9)Dr. Ashutosh Priya
10)Prof. (Dr.) Tulika Saxena
(72)Name of Inventor :
1)Mr. Pradeep Bedi
2)Dr. Piyush Choudhary
3)Mr. Awanit Kumar
4)Dr. Durga Prasad Gupta
5)Dr. Geetanjali Amarawat
6)Mr. Shrawan Kumar Sharma
7)Mr. Vijay Kumar Chhipa
8)Mr. Amol Laxman Mangrulkar
9)Dr. Ashutosh Priya
10)Prof. (Dr.) Tulika Saxena

(57) Abstract :
Human pose recognition is a crucial one. The 3D images are divided into two categories namely 3D single and 3D multiple. The 3D single is further classified into 3D model free and model based. In model free the estimation is done in two stages namely: single stage and 2D to 3D. In this single stage using direct prediction as well as body structure constraints are used. The assumptions which are made are the joints are named as parts. When the parts are connecting it is known as pair or limb. Using Dilated Residual Network (DRN) the semantic segmentation was done. This will suits small amount of tasks. To accommodate more tasks mask R-CNN is used. The parameters used for the position are x, y and z. The orientation parameters are qo, qx, qy and q z. The rotation was given by two parameters w and . The features are extracted using ResNet feature extractor. Then the multilayer perceptron is used. This contains 256 nodes including the information about position and orientation. The 20 output layer from the orientation and the 15 output layer from the position are normalized to form a unit magnitude. The rendering of the objects are carried out using the blender rendering software. The 3D model is presented to a blender and the synthetic dataset was matched. Then the prediction was made in three categories namely bounding box, location and rotation prediction. The other method is top-down, bottom-up and real time. For the multi person pose estimation, Top down as well as bottom up approach is made. In top down approach the parts are identified then the pose. But in bottom up approach it is vice-versa. The input is first fed into the transformation, projection and rotation. Then the transformed image is fed into the Key Point Estimation (KPE). This produces a confidence map using the gestures and then it is fed into the Transformation Parameter Estimation (TPE). The EP-PP transformation takes place. For the wide angle images are trained well before and the EP-PP transformation takes place. The learning and prediction was carried out using the radio frequency tracker. If the movement is fast then it is difficult to track it. In such scenario, it is using convolutional neural network based compensation. Then it is fed into the 3D graphics software for the projection.

No. of Pages : 15 No. of Claims : 8