

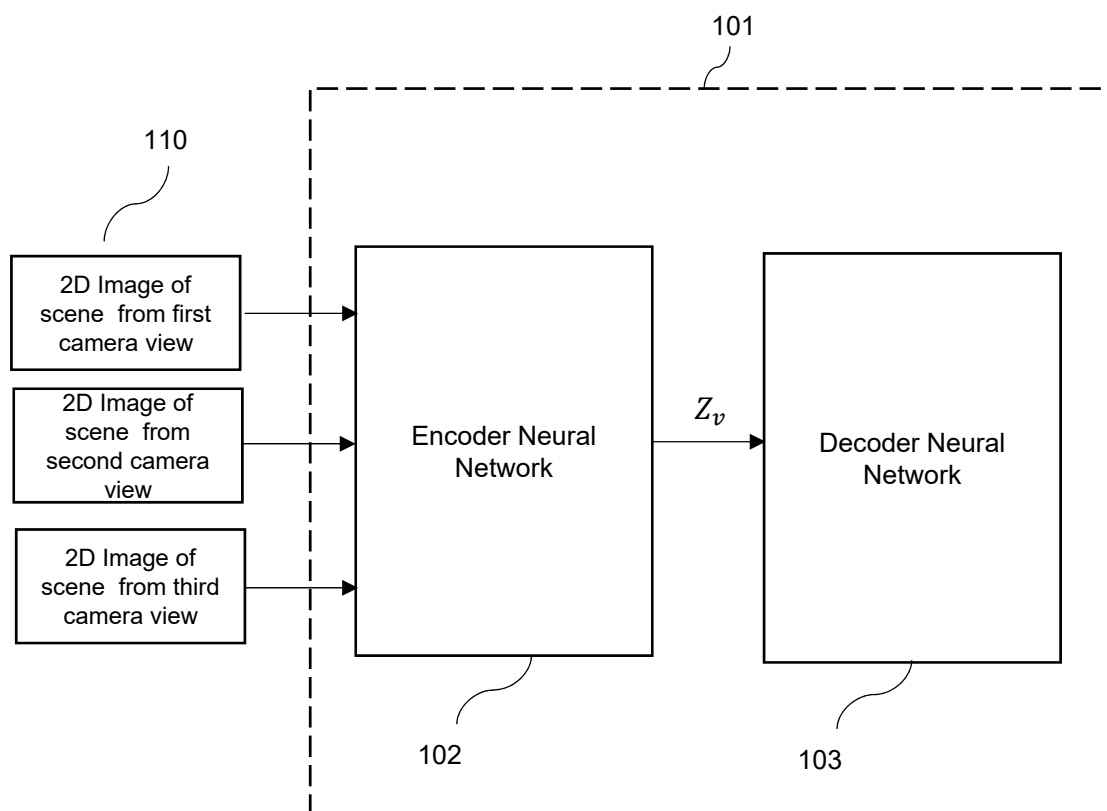


FIGURE 1

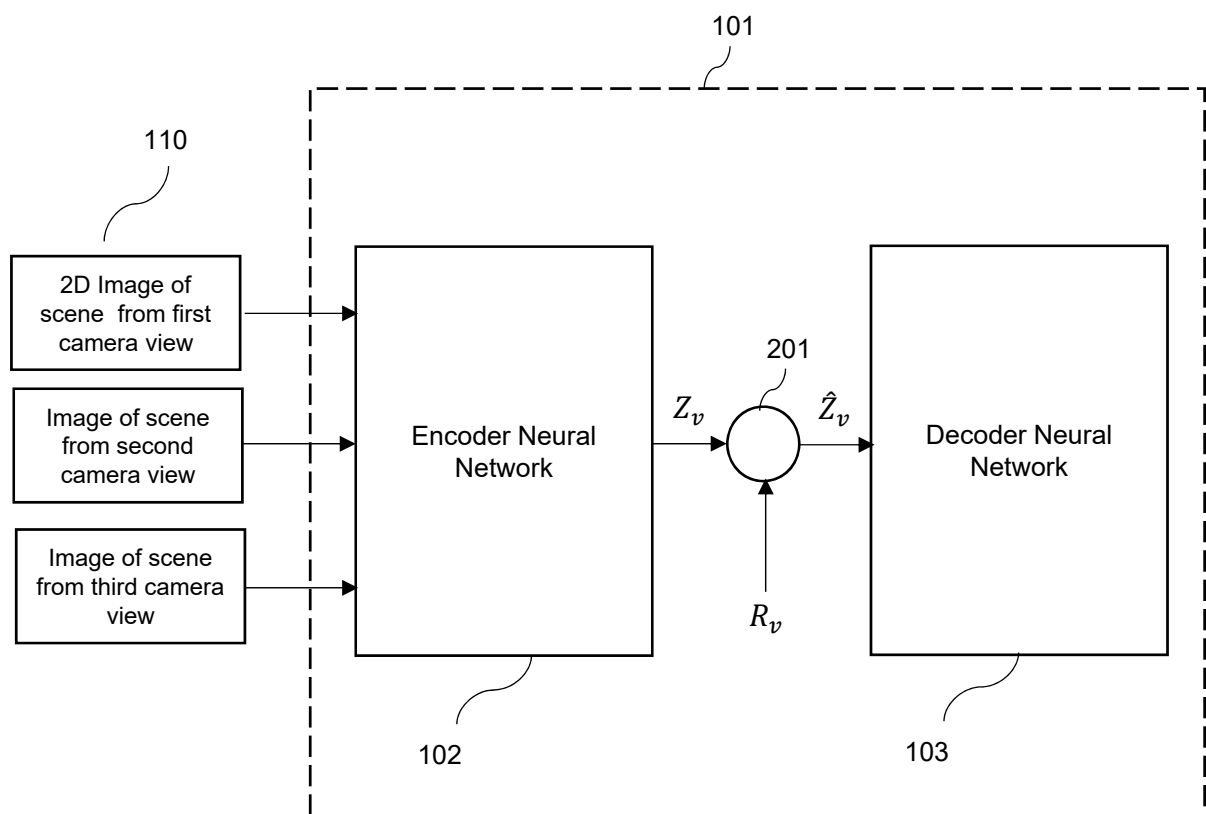


FIGURE 2

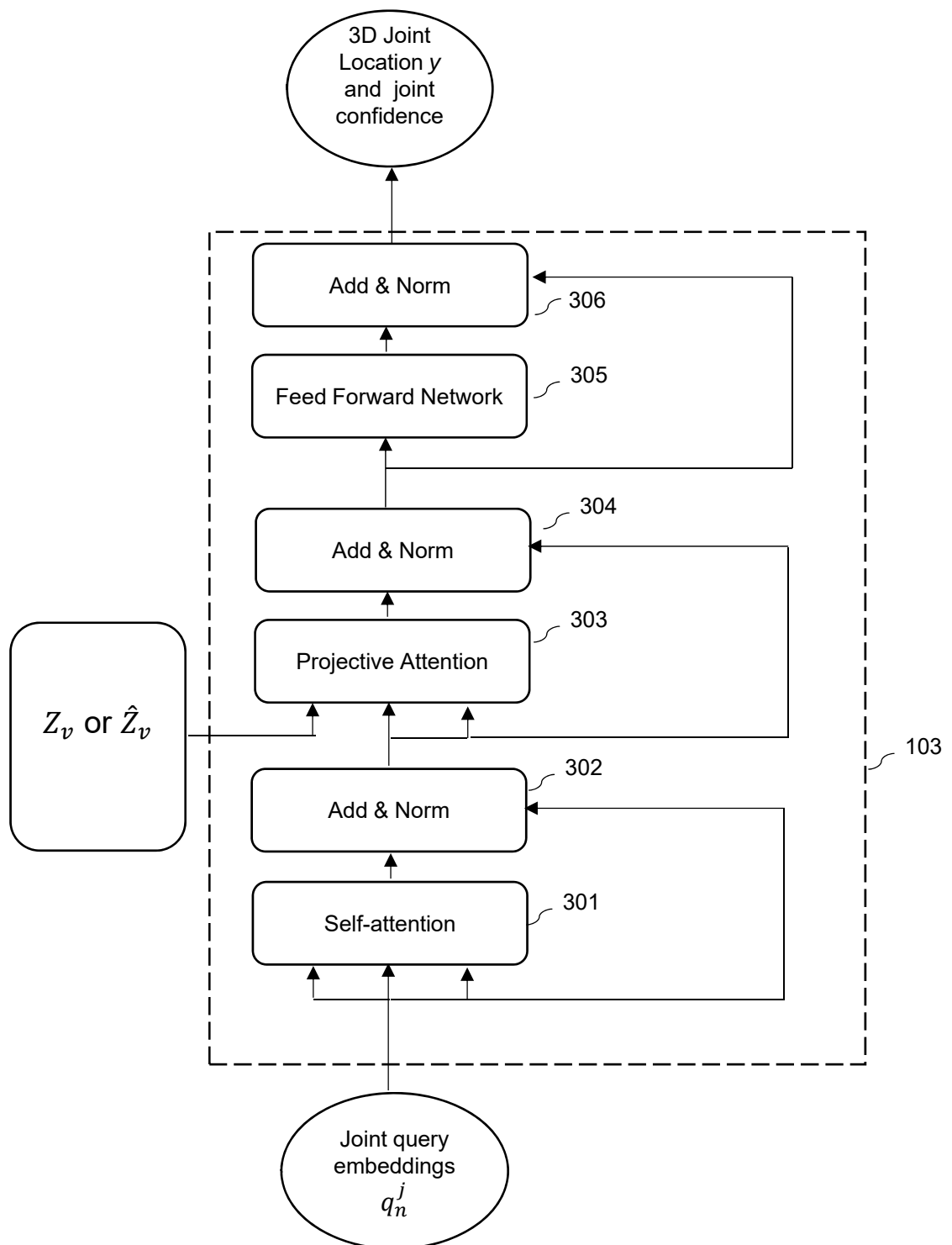


FIGURE 3

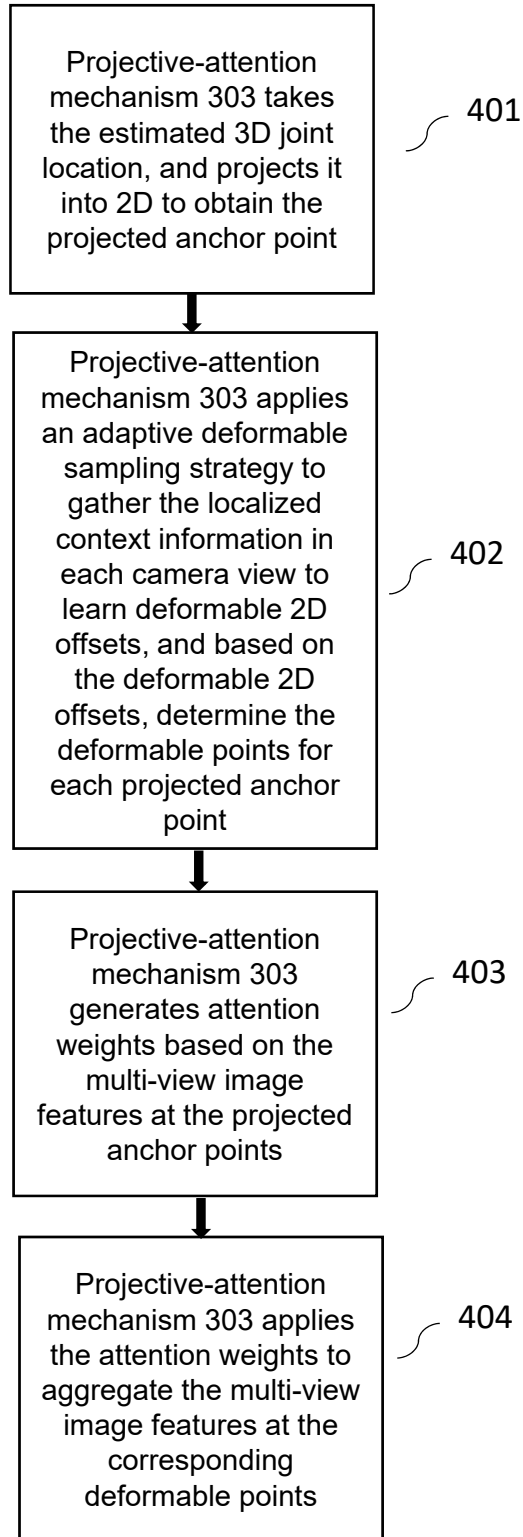


FIGURE 4

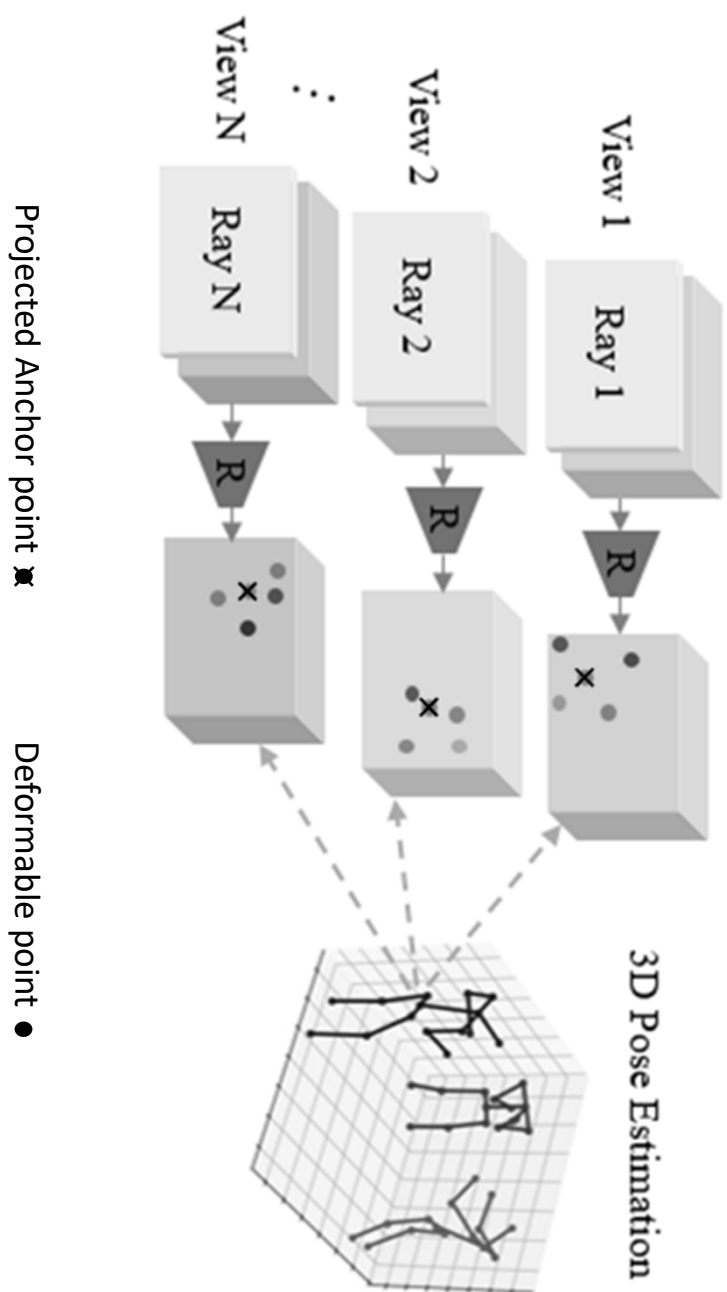


FIGURE 5



Projected Anchor point x Deformable point ●

FIGURE 6

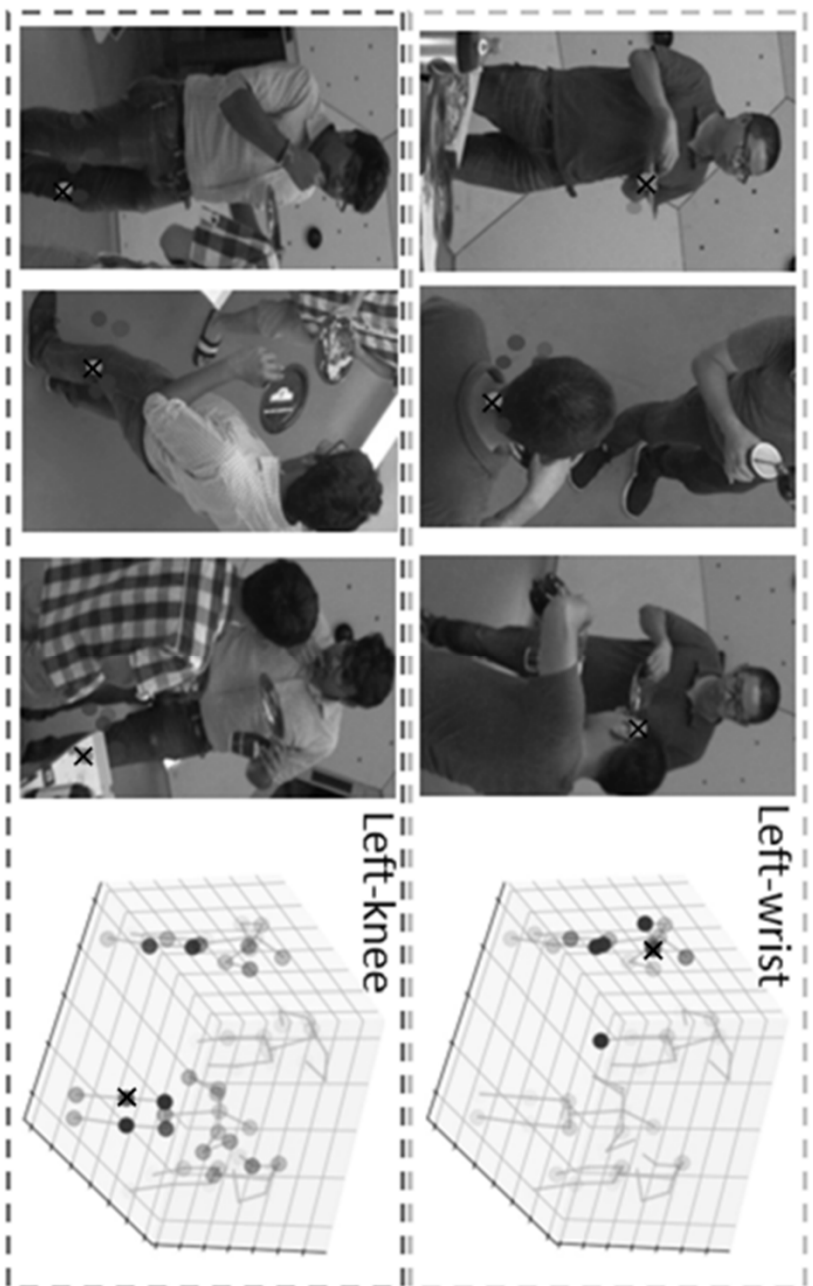


FIGURE 7

Table 1: Result on the Panoptic dataset. MvP is more accurate and faster than VoxelPose.

Methods	AP ₂₅	AP ₅₀	AP ₁₀₀	AP ₁₅₀	Recall@500	MPJPE[mm]	Time[ms]
VoxelPose [35]	84.0	96.4	97.5	97.8	98.1	17.8	320
MvP (Ours)	92.3	96.6	97.5	97.7	98.2	15.8	170

FIGURE 8

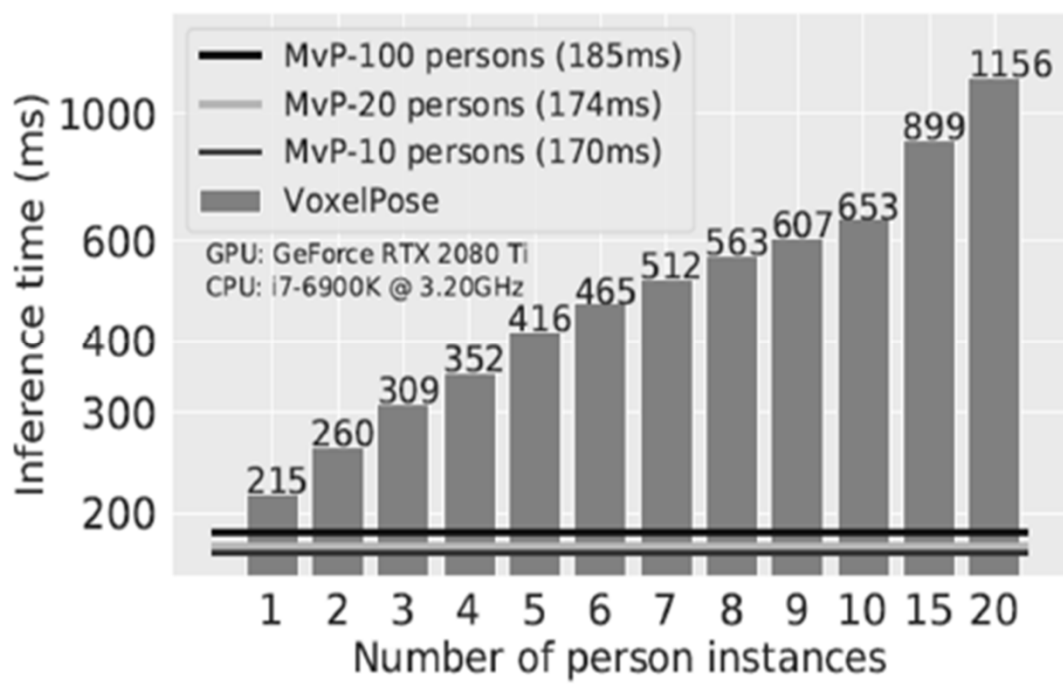


FIGURE 9

Table 2: Results (in PCP) on Shelf and Campus datasets.

Methods	Shelf					Campus				
	Actor 1	Actor 2	Actor 3	Average		Actor 1	Actor 2	Actor 3	Average	
Belagiannis <i>et al.</i> [2]	75.3	69.7	87.6	77.5		93.5	75.7	84.4	84.5	
Ershadi <i>et al.</i> [8]	93.3	75.9	94.8	88.0		94.2	92.9	84.6	90.6	
Dong <i>et al.</i> [6]	98.8	94.1	97.8	96.9		97.6	93.3	98.0	96.3	
VoxelPose [35]	99.3	94.1	97.6	97.0		97.6	93.8	98.8	96.7	
MvP (Ours)	99.3	95.1	97.8	97.4		98.2	94.1	97.4	96.6	

FIGURE 10

Table 3

<i>RConv</i> AP ₂₅ AP ₁₀₀ MPJPE			
w/	92.3	97.5	15.8
w/o	87.5	96.2	17.4

(a) The effect of *RdyConv*. w/o means removing RayConv.

<i>Dec.</i> AP ₂₅ AP ₁₀₀ MPJPE			
2	6.3	92.5	49.6
3	63.4	95.6	22.8
4	86.8	96.8	17.5
5	91.8	97.6	16.2
6	92.3	97.5	15.8
7	92.0	97.5	15.9

(d) Number of decoder layers.

<i>Query</i> AP ₂₅ AP ₁₀₀ MPJPE			
Per-joint	67.4	84.7	41.2
Hier.	82.5	93.2	19.5
Hier.+ad.	92.3	97.5	15.8

(b) Different joint query embedding schemes.

<i>Cum.</i> AP ₂₅ AP ₁₀₀ MPJPE			
1	4.7	61.0	93.8
2	37.7	93.0	34.8
3	71.8	95.1	21.1
4	84.1	96.7	19.3
5	92.3	97.5	15.8

(e) Number of camera views.

<i>Thr.</i> AP ₂₅ AP ₁₀₀ MPJPE			
0.0	93.1	98.5	16.3
0.1	92.3	97.5	15.8
0.2	91.1	96.2	15.5
0.4	89.2	93.7	15.0

(c) Different confidence threshold during evaluation.

<i>K</i> AP ₂₅ AP ₁₀₀ MPJPE			
1	88.6	96.3	18.2
2	89.3	97.5	17.4
4	92.3	97.7	15.8
8	84.4	91.1	20.3

(f) Number of deformable points *K*.

FIGURE 11

Table 4 : Results of replacing camera ray directions with 2D coordinates in RayConv

Positional Input	AP ₂₅	AP ₁₀₀	MPJPE
Camera Ray Directions	92.3	97.5	15.8
2D Spatial Coordinates	83.3	93.0	18.1

FIGURE 12

Table 5 : Comparison between Dense Attention and Projective Attention

Models	AP ₂₅	AP ₁₀₀	MPJPE	GPU Memory[G]
MvP-Dense	0.0	16.1	114.5	31.0
MvP	92.3	97.5	15.8	6.1

FIGURE 13