Type of action	Scenarios	Train	Validation	Test
Upper body movement	Directions	83,856	50,808	114,080
	Discussion	154,392	68,640	140,764
Full body upright variations	Greeting	69,984	33,096	84,980
	Posing	70,948	25,800	85,912
	Purchases	49,096	33,268	48,496
	Taking Photo	67,152	38,216	89,608
	Waiting	98,232	54,928	123,432
Walking variations	Walking	114,468	47,540	93,320
	Walking Dog	77,068	30,648	59,032
	Walking Pair	76,620	36,876	52,724
Variations while seated on a chair	Eating	109,360	39,372	97,192
	Phone Talk	132,612	39,308	92,036
	Sitting	110,228	46,520	89,616
	Smoking	138,028	50,776	85,520
Sitting on the floor	Sitting Down	112,172	50,384	105,396
Various Movements	Miscellaneous	-	-	105,576
Total		1,464,216	646,180	1,467,684

(a) The number of 3D human poses in Human3.6M in training, validation and testing aggregated over each scenario. We used 5 subjects for training (2 female and 3 male), 2 for validation (1 female and 1 male) and 4 subjects for testing (2 female and 2 male). The number of video frames is the same as the number of poses (4 cameras capturing at 50Hz). The number of TOF frames can be obtained by dividing the table entries by 8 (1 sensor capturing at 25Hz).

MoCap System		DV System	
No x Sensor	10 x Vicon T40	No x Sensor	4 x Basler piA1000
Resolution	4 Megapixels	Resolution	1000x1000
Freq.	200Hz	Freq.	50Hz
Sync	hardware	Sync	hardware

TOF System		Body Scanner	
No x Sensor	1 x Mesa SR4000	Sensor	Vitus Smart LC3
Resolution	176x144	No. Lasers	3
Freq.	25Hz	Point Density	7dots/cm3
Sync	software	Tolerance	< 1mm

(b) Technical summary of our different sensors.
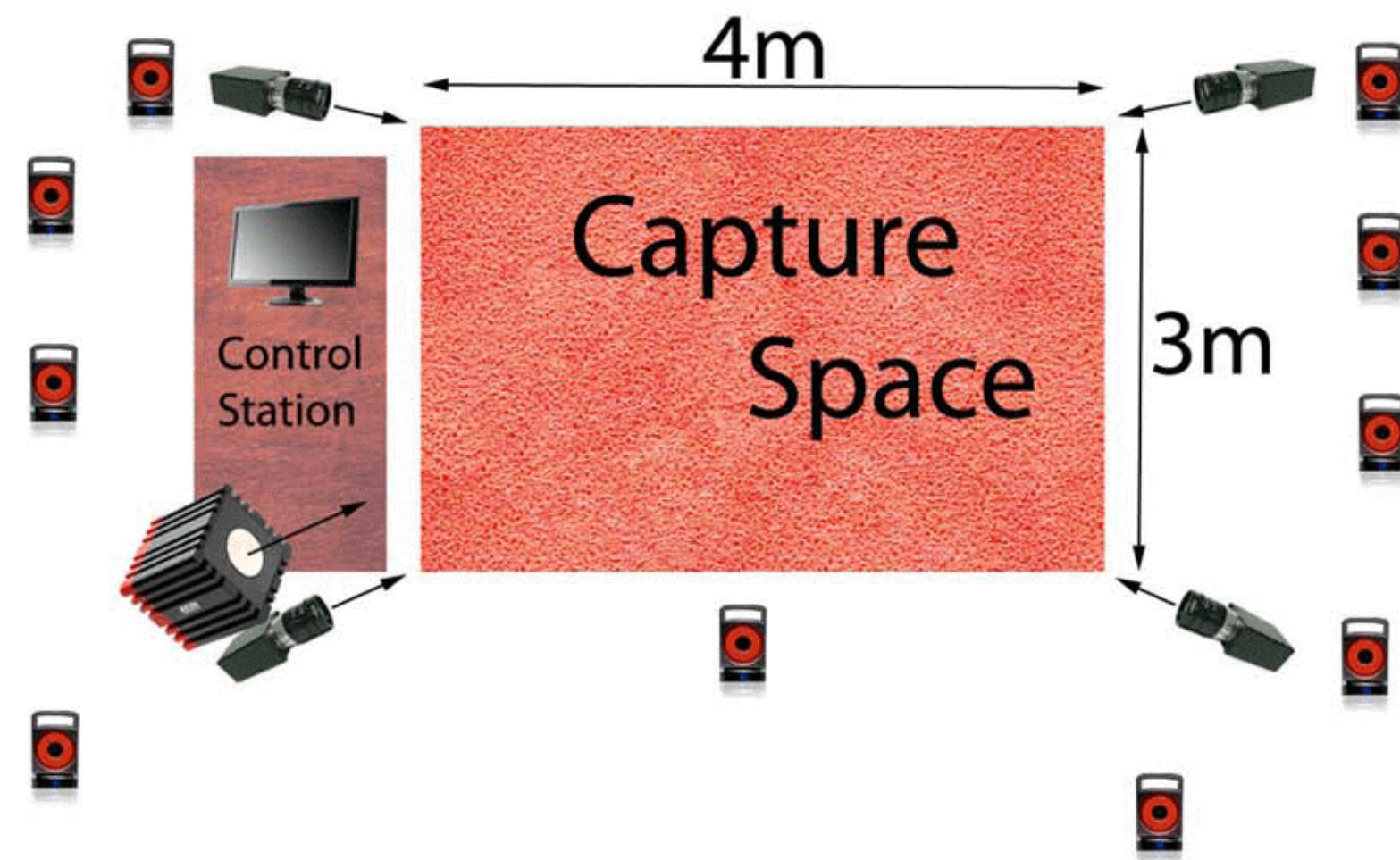
(c) Floor plan showing the capture region and the placement of the video, MoCap and TOF cameras.