Capturing Hand Motion with an RGB-D Sensor, Fusing a Generative Model with Salient Points

Sequences

Table 1. Sequences. *Set A* is used for evaluation of the components of the presented pipeline, while *Set B* is used as a comparison benchmark with the FORTH tracker [1]. All frames of *Set A* are used for evaluation, while for the sequences of *Set B* the evaluation starts at the noted *starting frame* ("ID Start"), since initialization of the compared trackers is different, while the last frame is rejected, since the public software of [1] failed for the last frame of one sequence. The number of the hands in each scene is noted, as well as the characterization of the collisions that take place in the scene: *some*, *severe* and *no apparent* collision. Only two hand sequences can be characterized by *severe* collisions. The public software of [1] can handle tracking of only one hand

	Sequence	ID	Hands	Total	ID Start	ID End	Collision
Set A	Walk	1	2	231	0	$total - 1$	Severe
	Cross	2	2	153	0	$total - 1$	Severe
	Cross & Twist	3	2	155	0	$total - 1$	Severe
	Helix - Tips	4	2	173	0	$total - 1$	Some
	Dance	5	2	265	0	$total - 1$	Severe
	Helix - Blend	6	2	136	0	$total - 1$	No
	Hug	7	2	194	0	$total - 1$	Severe
	Grasp	8	1	106	0	$total - 1$	No
	Fly	9	1	138	0	$total - 1$	No
	Rock	10	1	139	0	$total - 1$	Some
	Bunny	11	1	134	0	$total - 1$	Some
Set B	Bunny	12	1	727	420	$total - 2$	Some
	Fly	13	1	778	480	$total - 2$	No
	Rock	14	1	378	250	$total - 2$	Some

References

1. Oikonomidis, I., Kyriazis, N., Argyros, A.: Efficient model-based 3d tracking of hand articulations using kinect. In: BMVC. pp. 101.1–101.11 (2011)