Kinder-Gator: The UF Kinect Database of Child and Adult Motion

Aishat Aloba, Gianne Flores, Julia Woodward, Alex Shaw, Amanda Castonguay, Isabella Cuba, Yuzhu Dong, Eakta Jain, and Lisa Anthony

Department of CISE, University of Florida,USA

Abstract

Research has suggested that children's whole-body motions are different from those of adults. However, research on children's motions, and how these motions differ from those of adults, is limited. One possible reason for this limited research is that there are few motion capture (mocap) datasets for children, with most datasets focusing on adults instead. Furthermore, there are few datasets that have both children's and adults' motions to allow for comparison between them. To address these problems, we present Kinder-Gator, a new dataset of 10 children and 10 adults performing whole-body motions in front of the Kinect v1.0. The data contains 3D joint positions for 58 motions, such as walk, run, and wave, which have been manually labeled according to the category of the participant (child vs. adult), and the motion being performed. We believe this dataset will be useful in supporting research and applications in animation and whole-body motion recognition and interaction.

CCS Concepts

•***Computing methodologies*** → ***Motion capture;*** •***Human-centered computing*** → *User studies;*

1. Introduction

Motion-based applications for children have become more prevalent. For example, animated movies include child characters that depend on motion datasets to enable movement. Existing motion datasets are composed of either adult motion (performed by adults) [SBB10, KCV*15] or adult actors performing "child-like" motion [XWCH15, PSCO13]. However, character animators [Ani13] have noted that children's motions (e.g., movements and poses) are different from those of adults. Thus, there is a need for child motion datasets to help understand children's motions and how they differ from motions of adults to aid character animation. We found only one publicly available dataset [GFB12] containing both child and adult motions but in this dataset, participants performed motions as demonstrated by the researchers. Clearly, there is a need for more datasets containing the motions of children performing actions as naturally as possible to support realistic avatars.

We introduce Kinder-Gator, a novel Kinect motion dataset of child and adult motion. The dataset contains 58 motions, such as wave your hand, walk in place, kick a ball, point at the camera, etc., from 10 children (ages 5 to 9) and 10 adults, tracked using a Microsoft Kinect 1.0. The Kinect tracks 3D positions (x:horizontal, y:vertical, z:depth) of twenty joints such as the hip, spine, elbows, and knees, along with their respective timestamps.

The motions in the Kinder-Gator dataset have been performed naturally and are manually labeled according to the category of the participant (child and adult) and the motion being performed (one of 58 motions). Analysis of a subset of the dataset made available to Jain et al. [JAA*16], including 6 actions, showed that children perform motions faster than adults and that naïve participants can perceive differences between the motion of a child and that of an adult. These results are an example of how our dataset can facilitate research into understanding the differences between children's and adults' motions, which will have implications for future research in animation, whole-body recognition, and interaction.

Kinder-Gator can be used in character animation to generate avatars tailored to children. Furthermore, because the dataset contains both child and adult motion, Kinder-Gator can be used to perform cross-generational morphing to transform adult motion to child motion and vice versa. The dataset can also be used to detect deviations from normal motor development in children since it includes natural motions of children across a wide agegroup. Because of the actions included in Kinder-Gator, it can be also be applied to gait recognition, hand pose recognition, and stroke gesture recognition. Furthermore, Kinder-Gator can be used in character animation to generate avatars tailored to children. This dataset is publicly available (temporary URL: https://github.com/hayeesha/Kinder-Gator).

2. Related Work

There have been a number of mocap datasets that target whole-body motions. Gross and Shi [GS01] created the Motion of Body (MOBO) dataset which contains RGB video images of 25 adult actors walking with variationson the treadmill (slow, fast, inclined, and with a ball). Sigal and Black [SBB10] created the HumanEva dataset which contains mocap data for 4 adults performing 6 motions: walking, jogging, hello/goodbye motion, throw/catch, make a box, combo (i.e., balancing on feet) captured using a Vicon [vic]. However, these datasets include only a small number of motions. Furthermore, the MOBO dataset uses image sequencing to track motions, requiring a separate processing step to extract the positions of the joints, which makes motion analysis more complex than using mocap devices in which the positions of the joints have already been identified by the device. Kolykhalova et al. [KCV*15] created the MADS dataset which contains mocap data for 5 adult actors performing martial-arts, dance, and sports actions captured using a custom motion tracking device. The motion capture devices used in the HumanEva and MADS datasets provide high motion tracking accuracy, but the devices are difficult to obtain due to their high cost and require people with experience in how to operate the devices. With the availability of low-cost technology such as the Microsoft Kinect, however, mocap data containing more accurate prediction of joint motions has become easier.

Bloom et al. [BMA12] created the G3D dataset which includes video, depth, and skeleton data of ten actors performing twenty gaming actions (e.g., walk, run, jump, and climb), captured using the Kinect. Leightley et al. [LYJM15] also created the K3Da dataset, which contains Kinect data for fifty-four adults performing clinically-relevant motions (e.g., balancing).

A limitation of these mocap datasets is that they only consider adult motions (a comprehensive list of adult mocap datasets can be found in [vB]), which could be because recruiting children for research studies or mocap sessions is difficult [FDG13]. A workaround adopted by researchers has been to recruit adult actors to perform "child-like motions." Volkova et al. [VdlRBM14] created the MPI database which contains mocap data of 8 adult actors performing motions to express the emotions of a child (e.g., anger, disgust) while listening to fairy-tale stories. However, Jain et al. [JAA*16] have shown that children's movements are different from those of adults, so it is likely that even well trained actors are still not the same as actual child actors. Kinder-Gator addresses the limitations of existing datasets by providing a motion dataset that includes motions from both children and adults performing a large diverse set of 58 motions captured using the Kinect.

It is important to note that Guerra-Filho and Biswas [GFB12] created the Human Motion Database (HMD) which contains mocap data of 50 child and adult actors performing seventy different actions. However, in HMD, the actions were demonstrated to participants to maintain consistency in performance among the range of actions while in Kinder-Gator, we aimed to ensure that the participants performed the motions as naturally as possible. This approach will allow analysis of natural realistic motions as people actually perform them.

3. Dataset Collection

We collected a total of 58 motions in our dataset. The motions were chosen by reviewing studies involving whole-body motions [HHTR05, HKL06, NWL10]. Based on the review, we selected motions that were used in prior work, that people would be familiar with, and that we hypothesized would show differences between children and adults. We also had a set of simple "warm up" motions like waving to help get participants into the study. The motions collected from this review were classified into four categories: (a) **warm-up motions**: these motions are easy to perform and are used in day-to-day activities (9 motions). (b) **exercise motions**: these motions induce exertion when performed and are commonly used in exercise and fitness activities (14 motions). (c) **mime motions**: these motions involve the conceptualization of imaginary objects (16 motions). (d) **communication motions**: these motions are used to convey information to other people (19 motions) (Table 1).

3.1. Study Setup

Motions in the Kinder-Gator dataset were collected using Kinect v1.0 hardware and its accompanying Kinect for Windows SDK v1.8 software. Two researchers were responsible for prompting for the next motion to be performed and controlling the Kinect software. In each study session, a participant stood within an area denoted by a square, forward facing the Kinect, and movements were only allowed within that specific area. The denotation of a specific area was to ensure that people did not move outside the tracking range of the Kinect. Participants performed all motions from a standing position. Before the start of each motion, participants stood with their arms outstretched in the form of a T-pose and then counted down from 3 to 1 while lowering their arms to their sides. The T-pose was required to get an accurate demarcation of the intended start of a motion.

The duration of each motion was dependent on the motion being performed. For wave your hand, walk in place, walk in place as fast as you can, run in place, run in place as fast as you can, fly like a bird, swim, climb an imaginary ladder, and do 5 jumping jacks, the duration was typically about 5 cycles (10 steps or repetitions). For motions involving making poses with the body, the duration was 3 seconds, since the experiment staff required participants to hold the pose for 3 seconds. For all other motions, the duration of the motion varied depending on the participant. Participants always returned their hands down to their sides to demarcate the end of the motion. In each session, a participant performed 58 motions. In order to ensure that participants were performing the motions as naturally as possible, participants were allowed to perform the motion free-form. That is, we did not require that the motions be performed in any predefined manner. However, when a participant forgot how a motion was to be performed, one of the researchers showed an example (this occurred for 4 different children on 2 to 6 actions, and 2 different adults on 1 to 2 actions; details are provided in the dataset).

3.2. Participants

We recruited 10 children and 10 adults as participants via flyers, emails, and advertisements on social platforms. Recruitment and study protocol procedures were approved by our institutional review board. Child participant ages ranged from 5 to 9 (mean = 6.70, SD = 1.42). Five children were female. Two children were ambidextrous and none were left-handed. We focus on ages 5 to 9 to capture the variability within children's motions since children in this age group are still growing in terms of their motor development []. The adult participant ages ranged from 19 to 32 (mean = 23.40, SD = 4.33). Five adults were female and only one adult was

Warm-up	Exercise	Mime	Communication
Raise your hand	Put your hands on your hip and lean to the side	Push an imaginary button in front of you	Point at the camera
Raise your other hand	Put your hands on your hips and lean to the other side	Swipe across an imaginary screen in front of you	Motion someone to stop
Wave your hand	Put your hands on your hips and twist back and forth	Swipe across an imaginary screen in front of you with your other hand	Motion someone to come here
Wave your other hand	Touch your toes	Fly like a bird	Draw a [circle, square, triangle] in the air
Bow	Do a forward lunge	Fly like an airplane	Draw the letter [A, C, K, M, X] in the air
Raise your arm to one side	Lift your leg to one side	Swim	Make the letter [Y, M, C, A, K, P, T, X] with your body
Raise your other arm to the other side	Lift your other leg to the other side	Kick a ball	–
Bend your knee	Walk in place	Kick a ball as hard as you can	
Bend your other knee	Walk in place as fast as you can	Kick a ball with the other leg	–
–	Run in place	Kick a ball as hard as you can with that leg	–
–	Run in place as fast as you can	Throw a ball	–
–	Jump	Throw a ball as far as you can	–
–	Jump as high as you can	Throw a ball with your other arm	–
–	Do five jumping jacks	Throw a ball as far as you can with that arm	–
–	–	Punch	–
–	–	Climb an imaginary ladder	–

Table 1: *A list of the 58 motions in the Kinder-Gator dataset.*

left-handed (Table 2). All participants were familiar with motion interaction systems such as the Microsoft Kinect or Sony EyeToy. Participants each received a $10 gift card to a local grocery store as compensation.

3.3. Data Collection

The 58 motions in our dataset were collected using Kinect v1.0. The Kinect tracks 3D positions of twenty joints along three dimensions (x:horizontal, y:vertical, z:depth) with the corresponding timestamps, at 30 frames per second. The joints tracked by the Kinect are shown in Figure 1a. The joint positions are recorded in meters and the timestamp is recorded in milliseconds. A total of 1159 motion trials (58 motions x 20 participants) are included in our dataset; motion trial for jump high for one adult (pID: 565) is missing due to a software error. The total time it took to perform the motions ranged from 248s to 343s (M = 301s, SD = 37.2) for children and from 302s to 424s (M = 344s, SD = 35.3) for adults. To ensure that the data collected is in a format that facilitates easy retrieval and analysis, we performed some data post-processing. In the post-processing stage, the data was re-organized such that each row corresponds to the positions of all the joints at one frame. The first column has the timestamp such that the difference between the last row and the second row gives the duration of the motion in milliseconds; the first row is the header. Subsequent columns have the x, y, and z positions of each joint as recorded by the Kinect. The last two columns have the ID of the participant and the motion label; the values are always the same in each row for one motion. Then, each motion's data has been saved as a .csv file.

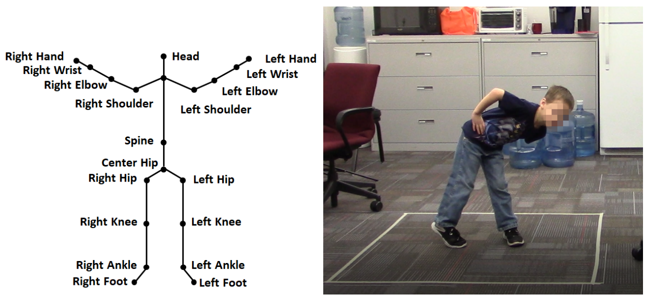

Figure 1: *(a) Joints tracked by the Kinect. (b) A child performing "put your hands on your hips and lean to the side" motion.*

4. Discussion

Since Kinder-Gator contains many different types of motions, we expect that researchers will extract different motion subsets for analysis. As an example, a subset of this dataset was made available to Jain et al. [JAA*16] who conducted a perception study to identify whether naïve viewers can perceive the difference between child and adult motion when presented with point-light displays of the motions. The study was conducted using 4 children (Child IDs: 290, 337, 644, 723) and 4 adults (Adult IDs: 734, 921, 934, 970) and 6 motions from our dataset: walk in place, run in place as fast as you can, jump as high as you can, wave your hand, fly like a bird, and do 5 jumping jacks. Evaluations of the survey responses found that naïve viewers could perceive the differences at levels significantly above chance and with 70% accuracy for dynamic motions, such as walk and run. ??add dataset with most variation?? This find-

CID	Sex	Age	Hand	Grade	AID	Sex	Age	Hand	Edu. Compl.
337	F	5	B	Pre-K	565	F	19	R	High school
595	M	5	B	Pre-K	577	F	19	R	Some college
106	M	6	R	K	604	F	20	R	Some college
290	M	6	R	K	976	M	20	R	Some college
342	F	6	R	K	734	M	22	R	Undergrad
474	F	6	R	1	876	F	23	R	Undergrad
103	F	8	R	3	888	F	25	R	Grad
169	M	8	R	2	921	M	26	L	Undergrad
723	M	8	R	3	970	M	28	R	Grad
644	F	9	R	4	934	M	32	R	Grad

Table 2: *Demographics for the children and adults in our dataset.*

ing suggest that there are indeed quantifiable variations between child and adult motion. These types of results show that Kinder-Gator can be an effective dataset for understanding the differences between child and adult motion.

Dong et al. [DPR*17] used the same subset of motions to generate androgynous characters representing child motions, adult motions, and dynamically scaled motions (adult motions transformed to child motions using dynamic scaling laws). Observations of these androgynous characters showed that dynamically scaled adult avatars are similar to adults in terms of coordination, but similar to children in terms of the pace of the motion. In a different work, Dong et al. [DAAJ18] also created 'child-like' motions from adult motions using a style translation algorithm. The algorithm trains a model using child and adult motions such that given an adult motion as a test input, the model returns a child motion. Rendered characters of the motions used for training, test, and output child motion showed that the output child motion is similar to that of a child than an adult. Taken together, these findings imply that Kinder-Gator enables cross-generational morphing of motions which can be used to create characters that show observable differences between child and adult motion.

5. Conclusion

Kinder-Gator is a dataset of 58 motions performed by 10 children (ages 5 to 9) and 10 adults (ages 19 to 32) recorded using the Kinect v1.0. Kinder-Gator can be used in research fields such as animation, whole-body interaction, and recognition. One of the main intended applications for Kinder-Gator is creating believable child characters for games and animated movies. Style translation methods [HPP05, XWCH15] can be leveraged along with this dataset to enable cross-generational morphing and create compelling avatars. Furthermore, since our dataset contains natural motions of children and covers an age group where children's motor skills are still developing, it can be used to create realistic avatars that model this progression in children's motor skills . Additionally, it can be used for early detection of motor disabilities in children since its motions provides a model for normal motor development in children.In additon to animation, our dataset can also be used to create more robust recognizers that can recognize both child and adult motions. Due to the diverse set of motions in the dataset, subsets of the dataset can be employed in: (a) gait recognition to quantify gait features that are different between children and adults (e.g., walk in place), (b) body recognition to facilitate research into recognition of more diverse sets of motions used to convey information (e.g., motion someone to stop), and (c) stroke gesture recognition to create recognizers that can recognize both 2D and 3D stroke motions (e.g., draw the letter A in the air). The goal of our dataset is to increase the number of motion datasets for research and to encourage research investigating the differences and similarities between child and adult motions.

References

[Ani13] ANIMATIONADDICTS: The challenge of animating underaged adults. http://www.animation-addicts.com/2013/01/30/animating-kids/, 2013. Accessed: 2017-11-12. 1

[BMA12] BLOOM V., MAKRIS D., ARGYRIOU V.: G3D: A gaming action dataset and real time action recognition evaluation framework. *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (2012), 7–12. doi:10.1109/CVPRW.2012.6239175. 2

[DAAJ18] DONG Y., ALOBA A., ANTHONY L., JAIN E.: Style translation to create child-like motion. *Eurographics* (2018), [to appear]. 4

[DPR*17] DONG Y., PARYANI S., RANA N., ALOBA A., ANTHONY L., JAIN E.: Adult2child: Dynamic scaling laws to create child-like motion. In *Proceedings of the Tenth International Conference on Motion in Games* (New York, NY, USA, 2017), MIG '17, ACM, pp. 13:1–13:10. URL: http://doi.acm.org/10.1145/3136457.3136460, doi:10.1145/3136457.3136460. 4

[FDG13] FOSS E., DRUIN A., GUHA M. L.: Recruiting and retaining young participants: Strategies from five years of field research. In *Proceedings of the 12th International Conference on Interaction Design and Children* (New York, NY, USA, 2013), IDC '13, ACM, pp. 313–316. URL: http://doi.acm.org/10.1145/2485760.2485798, doi:10.1145/2485760.2485798. 2

[GFB12] GUERRA-FILHO G., BISWAS A.: The human motion database: A cognitive and parametric sampling of human motion. *Image Vision Comput. 30*, 3 (Mar. 2012), 251–261. URL: http://dx.doi.org/10.1016/j.imavis.2011.12.002, doi:10.1016/j.imavis.2011.12.002. 1, 2

[GS01] GROSS R., SHI J.: *The CMU Motion of Body (MoBo) Database*. Tech. Rep. 1, 2001. doi:10.1109/MP.2008.4430762. 1

[HHTR05] HÖYSNIEMI J., HÄMÄLÄINEN P., TURKKI L., ROUVI T.: Children's intuitive gestures in vision-based action games. *Commun. ACM 48*, 1 (Jan. 2005), 44–50. URL: http://doi.acm.org/10.1145/1039539.1039568, doi:10.1145/1039539.1039568. 2

[HKL06] HWANG B.-W., KIM S., LEE S.-W.: A full-body gesture database for automatic gesture recognition. In *7th International Conference on Automatic Face and Gesture Recognition (FGR06)* (April 2006), pp. 243–248. doi:10.1109/FGR.2006.8. 2

[HPP05] HSU E., PULLI K., POPOVIĆ J.: Style translation for human motion. *ACM Trans. Graph. 24*, 3 (July 2005), 1082–1089. URL: http://doi.acm.org/10.1145/1073204.1073315, doi:10.1145/1073204.1073315. 4

[JAA*16] JAIN E., ANTHONY L., ALOBA A., CASTONGUAY A., CUBA I., SHAW A., WOODWARD J.: Is the motion of a child perceivably different from the motion of an adult? *ACM Trans. Appl. Percept. 13*, 4 (July 2016), 22:1–22:17. URL: http://doi.acm.org/10.1145/2947616, doi:10.1145/2947616. 1, 2, 3

[KCV*15] KOLYKHALOVA K., CAMURRI A., VOLPE G., SANGUINETI M., PUPPO E., NIEWIADOMSKI R.: A Multimodal Dataset for the Analysis of Movement Qualities in Karate Martial Art. In *Proceedings of the 7th International Conference on Intelligent Technologies for Interactive Entertainment* (2015). doi:10.4108/icst.intetain.2015.260039. 1, 2

[LYJM15] LEIGHTLEY D., YAP M. H., J. COULSON Y. B., MCPHEE J. S.: Benchmarking Human Motion Analysis Using Kinect One: an open source dataset. *IEEE Conference of sia-Pacific Signal and Information Processing Association*, December (2015), 1–7. doi:10.1109/APSIPA.2015.7415438. 2

[NWL10] NORTON J., WINGRAVE C. A., LAVIOLA JR. J. J.: Exploring strategies and guidelines for developing full body video game interfaces. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games* (New York, NY, USA, 2010), FDG '10, ACM, pp. 155–162. URL: http://doi.acm.org/10.1145/1822348.1822369, doi:10.1145/1822348.1822369. 2

[PSCO13] PIANA S., STAGLIANÒ A., CAMURRI A., ODONE F.: A set of Full-Body Movement Features for Emotion Recognition to Help Children affected by Autism Spectrum Condition. *IDGEI 1st International workshop* (2013), 1–7. 1

[SBB10] SIGAL L., BALAN A. O., BLACK M. J.: HumanEva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion. *International Journal of Computer Vision 87*, 1-2 (2010), 4–27. doi:10.1007/s11263-009-0273-6. 1

[vB] VAN BOXTEL J. J.: Free mocap databases. http://jeroenvanboxtel.com/MocapDatabases.html. Accessed:2017-12-28. 2

[VdlRBM14] VOLKOVA E., DE LA ROSA S., BULTHOFF H. H., MOHLER B.: The MPI emotional body expressions database for narrative scenarios. *PLoS ONE 9*, 12 (2014), e113647. 2

[vic] Vicon. https://www.vicon.com. Accessed: 2017-12-27. 2

[XWCH15] XIA S., WANG C., CHAI J., HODGINS J.: Realtime style transfer for unlabeled heterogeneous human motion. *ACM Trans. Graph. 34*, 4 (July 2015), 119:1–119:10. URL: http://doi.acm.org/10.1145/2766999, doi:10.1145/2766999. 1, 4