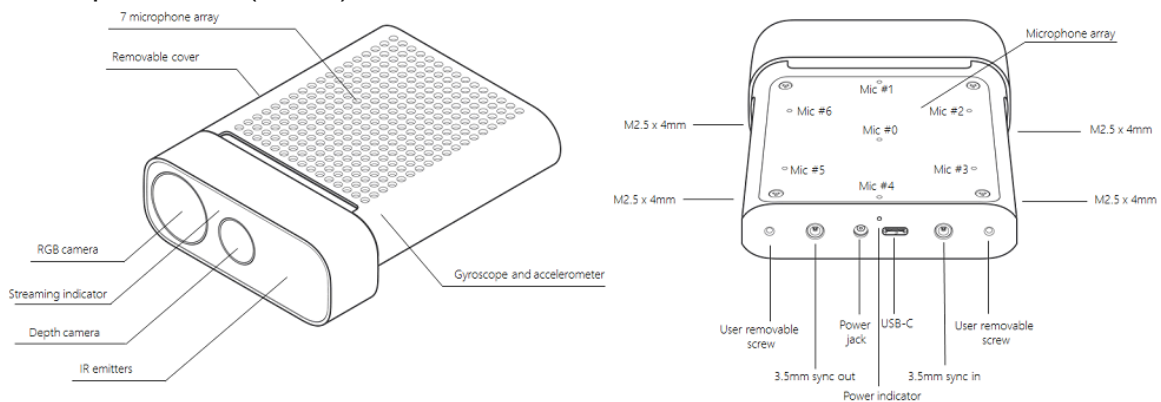


## Azure Kinect DK Specifications (Skeletal Tracking)

### Which features can be collected and extracted with this device?

Azure Kinect DK is a developer kit with advanced AI sensors that provide sophisticated computer vision and speech models. Kinect contains a depth sensor, spatial microphone array with a video camera (color camera), and orientation sensor as an all-in-one small device with multiple modes, options, and software development kits (SDKs).



- Sensor SDK for low-level sensor and device access.
- Body Tracking SDK for tracking bodies in 3D.
  - Provides body segmentation.
  - Contains an anatomically correct skeleton for each partial or full body in FOV.
  - Offers a unique identity for each body.
  - Can track bodies over time.
- Speech Cognitive Services SDK for enabling microphone access and Azure cloud-based speech services.

## Depth camera

The **depth camera** supports the modes indicated below:

Mode	Resolution	FoI	FPS	Operating range*	Exposure time
<b>NFOV unbinned</b>	640x576	75°x65°	0, 5, 15, 30	0.5 - 3.86 m	12.8 ms
<b>NFOV 2x2 binned (SW)</b>	320x288	75°x65°	0, 5, 15, 30	0.5 - 5.46 m	12.8 ms
<b>WFOV 2x2 binned</b>	512x512	120°x120°	0, 5, 15, 30	0.25 - 2.88 m	12.8 ms
<b>WFOV unbinned</b>	1024x1024	120°x120°	0, 5, 15	0.25 - 2.21 m	20.3 ms
<b>Passive IR</b>	1024x1024	N/A	0, 5, 15, 30	N/A	1.6 ms

\*15% to 95% reflectivity at 850nm, 2.2  $\mu\text{W}/\text{cm}^2/\text{nm}$ , random error std. dev.  $\leq 17$  mm, typical systematic error  $< 11$  mm + 0.1% of distance without multi-path interference. Depth may be provided outside of the operating range indicated above. It depends on an object's reflectivity.

## Color camera

Azure Kinect DK includes an OV12A10 12MP CMOS sensor rolling shutter sensor. The native operating modes are listed below:

RGB Resolution (HxV)	Camera Aspect Ratio	Format Options	Frame Rates (FPS)	Nominal FOV (HxV)(post-processed)
3840x2160	16:9	MJPEG	0, 5, 15, 30	90°x59°
2560x1440	16:9	MJPEG	0, 5, 15, 30	90°x59°
1920x1080	16:9	MJPEG	0, 5, 15, 30	90°x59°
1280x720	16:9	MJPEG/YUY2/NV12	0, 5, 15, 30	90°x59°

4096x3072	4:3	MJPEG	0, 5, 15	90°x74.3°
2048x1536	4:3	MJPEG	0, 5, 15, 30	90°x74.3°

The RGB camera is USB Video class-compatible and can be used without the Sensor SDK. The RGB camera color space: BT.601 full range [0..255]. The MJPEG chroma sub-sampling is 4:2:2.

Acceptable RGB camera manual exposure values:

<b>exp</b>	<b>2<sup>exp</sup></b>	<b>50Hz</b>	<b>60Hz</b>
-11	488	500	500
-10	977	1250	1250
-9	1953	2500	2500
-8	3906	10000	8330
-7	7813	20000	16670
-6	15625	30000	33330
-5	31250	40000	41670
-4	62500	50000	50000
-3	125000	60000	66670
-2	250000	80000	83330
-1	500000	100000	100000
0	1000000	120000	116670
1	2000000	130000	133330

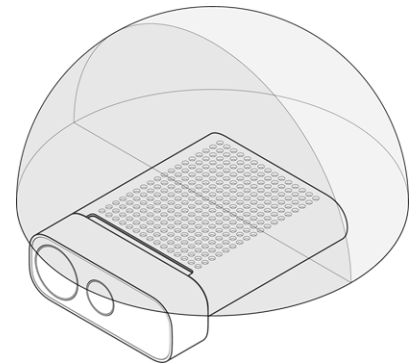
### **Motion sensor (IMU)**

The embedded Inertial Measurement Unit (IMU) is an LSM6DSMUS and includes both an accelerometer and a gyroscope. The accelerometer and gyroscope are simultaneously sampled at 1.6 kHz. The samples are reported to the host at a 208 Hz.

### **Microphone array**

Azure Kinect DK embeds a high-quality, seven microphone circular array that identifies as a standard USB audio class 2.0 device. All 7 channels can be accessed. The performance specifications are:

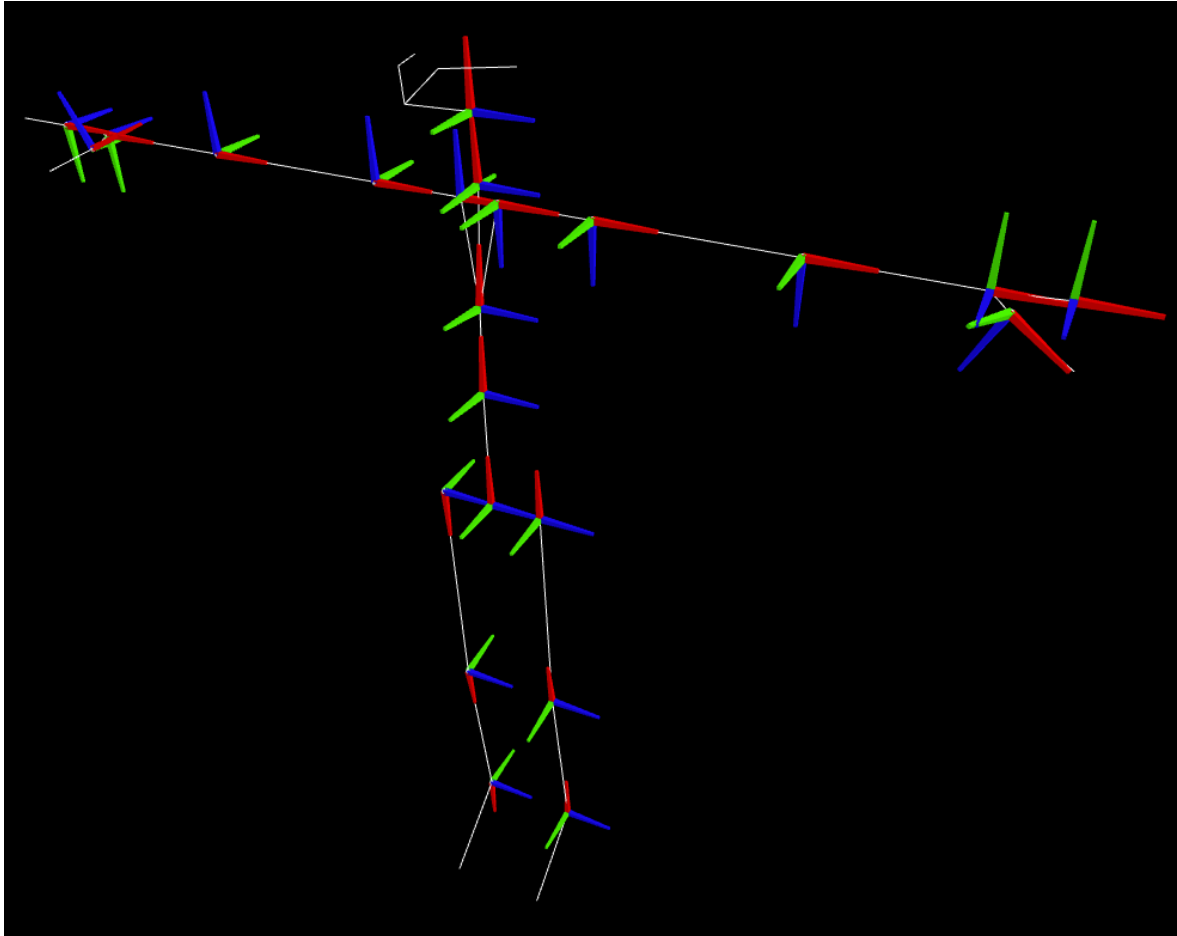
- Sensitivity: -22 dBFS (94 dB SPL, 1 kHz)
- Signal to noise ratio > 65 dB
- Acoustic overload point: 116 dB



### **Body tracking joints**

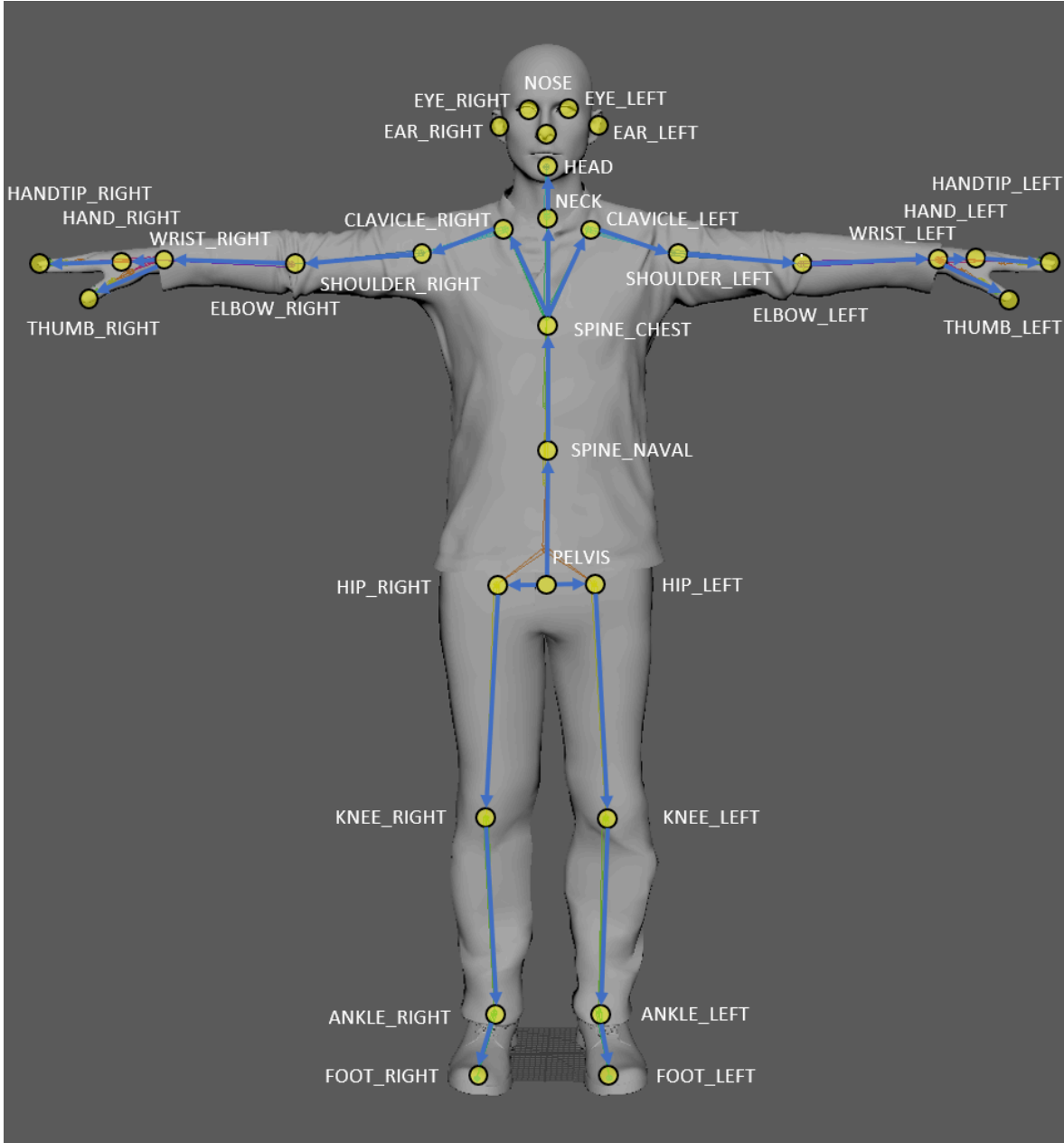
The position and orientation of each joint form its own right-handed joint coordinate system. All joint coordinate systems are absolute coordinate systems in the depth

camera 3D coordinate system.



| x-axis = red | y-axis = green | z-axis = blue |

A skeleton includes 32 joints with the joint hierarchy flowing from the center of the body to the extremities. Each connection (bone) links the parent joint with a child joint. The figure illustrates the joint locations and connection relative to the human body.



The following table enumerates the standard joint connections.

<b>Index</b>	<b>Joint name</b>	<b>Parent joint</b>
0	PELVIS	-
1	SPINE_NAVAL	PELVIS
2	SPINE_CHEST	SPINE_NAVAL
3	NECK	SPINE_CHEST
4	CLAVICLE_LEFT	SPINE_CHEST
5	SHOULDER_LEFT	CLAVICLE_LEFT
6	ELBOW_LEFT	SHOULDER_LEFT
7	WRIST_LEFT	ELBOW_LEFT
8	HAND_LEFT	WRIST_LEFT
9	HANDTIP_LEFT	HAND_LEFT
10	THUMB_LEFT	WRIST_LEFT
11	CLAVICLE_RIGHT	SPINE_CHEST
12	SHOULDER_RIGHT	CLAVICLE_RIGHT
13	ELBOW_RIGHT	SHOULDER_RIGHT
14	WRIST_RIGHT	ELBOW_RIGHT
15	HAND_RIGHT	WRIST_RIGHT
16	HANDTIP_RIGHT	HAND_RIGHT
17	THUMB_RIGHT	WRIST_RIGHT
18	HIP_LEFT	PELVIS

Index	Joint name	Parent joint
19	KNEE_LEFT	HIP_LEFT
20	ANKLE_LEFT	KNEE_LEFT
21	FOOT_LEFT	ANKLE_LEFT
22	HIP_RIGHT	PELVIS
23	KNEE_RIGHT	HIP_RIGHT
24	ANKLE_RIGHT	KNEE_RIGHT
25	FOOT_RIGHT	ANKLE_RIGHT
26	HEAD	NECK
27	NOSE	HEAD
28	EYE_LEFT	HEAD
29	EAR_LEFT	HEAD
30	EYE_RIGHT	HEAD
31	EAR_RIGHT	HEAD

#### References:

- Sych, T., McClister, C., Assaf, W., Urban, E., Coulter, D., Allen, B. (2023, July 28). *About Azure Kinect DK*.  
<https://learn.microsoft.com/en-us/azure/kinect-dk/about-azure-kinect-dk>
- Sych, T., Meadow, P., Allen, B. (2023, June 01). *Cámara de profundidad de Azure Kinect DK*  
<https://learn.microsoft.com/es-es/azure/kinect-dk/depth-camera>
- Sych, T., Allen, B. (2022, September 21). *Azure Kinect DK coordinate systems*  
<https://learn.microsoft.com/en-us/azure/kinect-dk/coordinate-systems>
- Jenks, A., Coulter, D., Sych, T., de los Santos, M., Motiv, T., Borsechnik, J., Meadow, P., Towle, E., Duan, A., Allen, B. (2022, February 09). *Azure Kinect DK hardware specifications*  
<https://learn.microsoft.com/en-us/azure/Kinect-dk/hardware-specification>



---

## How are such features related to Skeletal tracking?

Skeletal tracking is the process of detecting and tracking the human skeleton and body joints in real-time. Skeletal tracking is an important application of Azure Kinect DK, as it allows for a variety of applications.

The depth map captured provides the 3D spatial information required to accurately track the human body in real-time. The depth map is used to create a 3D point cloud of the environment, which is then analyzed to detect and track the human body.

Once the body is detected, Azure Kinect DK can track the movement of the body joints in real-time, using algorithms that are based on machine learning and computer vision techniques. The tracked body joints can be used to create a skeletal model of the human body.

Some videos about it:

- Moving Digits. (2020, October 8). *A Short Introduction to Motion Capture and Showcase of Different Systems* [Video]. YouTube. <https://www.youtube.com/watch?v=NJ7-27xRv6w&feature=youtu.be>
- CSharp TV. (2020, September 3). *Developing with Kinect for Azure - Understanding the human body by Andreas Erben* [Video]. YouTube. <https://www.youtube.com/watch?v=BwSvYnrVzhw>

## References:

- Tölgyessy, M., Dekan, M., & Chovanec, L. (2021). Skeleton tracking accuracy and precision evaluation of Kinect V1, Kinect V2, and the Azure Kinect. *Applied Sciences (Basel, Switzerland)*, 11(12), 5756. <https://doi.org/10.3390/app11125756>.
- Büker, L., Quinten, V., Hackbarth, M., Hellmers, S., Diekmann, R., & Hein, A. (2023). How the processing mode influences Azure Kinect Body Tracking results. *Sensors (Basel, Switzerland)*, 23(2), 878. <https://doi.org/10.3390/s23020878>.
- Clark, R. A., Mentiplay, B. F., Hough, E., & Pua, Y. H. (2019). Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and Kinect alternatives. *Gait & Posture*, 68, 193–200. <https://doi.org/10.1016/j.gaitpost.2018.11.029>.
- Ripic, Z., Kuenze, C., Andersen, M. S., Theodorakos, I., Signorile, J., & Eltoukhy, M. (2022). Ground reaction force and joint moment estimation during gait using an Azure Kinect-driven musculoskeletal modeling approach. *Gait & Posture*, 95, 49–55. <https://doi.org/10.1016/j.gaitpost.2022.04.005>.

---

## How could we relate such features to learning?

There are multiple studies in the literature which relate skeletal tracking and RGB camera, from devices such as the Azure Kinect, to different learning activities and environments. Such studies seek to analyze skeletal tracking and facial tracking to gain a deeper understanding on:

- Kinesthetic Learning
- Multimodal sensing
- Facial recognition.

### References:

- Ocampo, H., Silva, G., & Salinas, P. (2015). Kinect TEAM: Kinesthetic Learning Applied to Mathematics Using Kinect. *Procedia Computer Science*, 75, 169–172. <https://doi.org/10.1016/j.procs.2015.12.234>
- Lin, L., & Shi, L. (2022a). Research on student classroom attention recognition based on multimodality. *Proc.SPIE*, 12506, 125063D. <https://doi.org/10.1117/12.2662889>
- Fakhar, S., Baber, J., Bazai, S. U., Marjan, S., Jasinski, M., Jasinska, E., Chaudhry, M. U., Leonowicz, Z., & Hussain, S. (2022). Smart classroom monitoring using novel real-time facial expression recognition system. *Applied Sciences (Basel, Switzerland)*, 12(23), 12134. <https://doi.org/10.3390/app122312134>.

---

## How to cite this document

Morán-Mirabal, L.F., Güemes-Frese, L.E., Ruiz-Ramirez, J.A. (2024). *Azure Kinect DK Specifications - Skeletal Tracking* (Report No. IFE-EC-DR006-01) Experiential Classroom Living Lab, Institute for the Future of Education. <https://hdl.handle.net/112867656561476>

### Document Change Control

<b>Date</b>	<b>Version</b>	<b>Description of Change</b>	<b>Made By</b>
09/Dec/2022	1.0	Document Creation	Moran, L.
05/Apr/2024	2.0	The document handle link is included	Ruiz, A.