AI Audio-Visual Tracking Box MT100



© 2025 AISPEECH and/or its affiliates. All rights reserved

CATALOG

1 Product Introduction

2 Main Features

3 Product Specifications

4 External Dimension

5 Component Instructions

© 2025 AISPEECH and/or its affiliates. All rights reserved

Product Introduction

The AIVIS-MT100 is a powerful AI Audio-Visual Tracking Box launched by AISpeech, featuring a compact design and easy deployment. It is ideal for enterprises, conference rooms, training rooms, and university lecture halls. This box integrates advanced image algorithms such as AI posture recognition and image stitching. Combined with AISpeech's high-end ceiling microphones and their precise DOA (Direction of Arrival) capabilities in large conference spaces, the MT100's multi-PTZ camera scheduling strategy and image algorithms enable real-time tracking of speakers and movements, delivering a high-definition, smooth audiovisual experience. The MT100 supports binding up to 6 cameras and an unlimited number of ceiling microphones simultaneously, making it suitable for various large and extra-large meeting spaces and presentation environments. It meets the needs of both remote and in-person hybrid meetings.

© 2025 AISPEECH and/or its affiliates. All rights reserved

Main Features

Speaker Tracking

MT100 uses real-time DOA (Direction of Arrival) data from ceiling microphones to precisely locate the active speaker. By calling PTZ camera presets, it drives the corresponding lens for speaker tracking; when the speaker adjusts posture near a preset point, the PTZ camera auto-calibrates to the preset to keep the subject centered at all times.

Scene Policy Configuration

Based on a B/S (browser–server) architecture, MT100 provides a visual configuration interface for fine-grained management of ceiling microphones, cameras, and linkage zones. It supports real-time device-status monitoring, PTZ camera preview, and on/off control of linkage zones, allowing adjustments to match on-site requirements.

Multi-Camera Orchestration & Image Integration

MT100 connects to up to six PTZ cameras, while AISpeech's high-end ceiling microphone systems support unlimited cascading and can synchronize DOA data to MT100 in real time. Leveraging a high-performance video codec engine,MT100 dispatches the appropriate PTZ camera to configure preset zones, together with posture recognition, optimizes framing. It also supports Picture-in-Picture (PiP) with adjustable sub-window position and scale to accommodate more use cases. During close-up transitions, MT100 applies image transition strategies to make remote-meeting visuals more natural.

Multi-Scenario Image Algorithm Scheduling

MT100 works with the C40T 4K conference camera and the C60 dual-lens AI voice-tracking camera to deliver precise talker-area response for diverse meeting scenarios; camera roles can be configured per scenario to flexibly orchestrate view response, presenter tracking, and whiteboard close-ups, and a Discussion Mode is available. At the same time, MT100 provides configurable switching policies, audio-trigger timing windows, and scene-management strategies for flexible adaptation.

© 2025 AISPEECH and/or its affiliates. All rights reserved

Product Specifications

Function	Specification
Resolution	• Maximum Support for 4K @ 30fpsOutput
Video compression	• H.265/H.264/MJPEG
Audio interface	• 1x LINE IN (3.5MM)（Reserved） • 1x LINE OUT（3.5MM）(Reserved)
Video input	• Maximum support for 4K @ 30fps input
Video output	• Maximum support for 4K @ 30fps output
Video switching	• Built-in AI algorithm, supports automatic video switching
HDMI interface	• 1x HDMI-IN • 1x HDMI-OUT (supports audio output) • Output video resolution：4KP30；1080P30/25；720P30/25；640*480P30/25
USB interface	• 1x USB2.0 Host • 1 USB3.0 Device，USB3.0 compatible with USB2.0： • UVC supports UVC1.1protocol； • UVC video formats support H.265/H.264/MJPEG，video resolution 4K30 (without second encoding)； • 1080P30/25；720P30/25；360P30/25； • Supports UAC, audio format supports PCM.

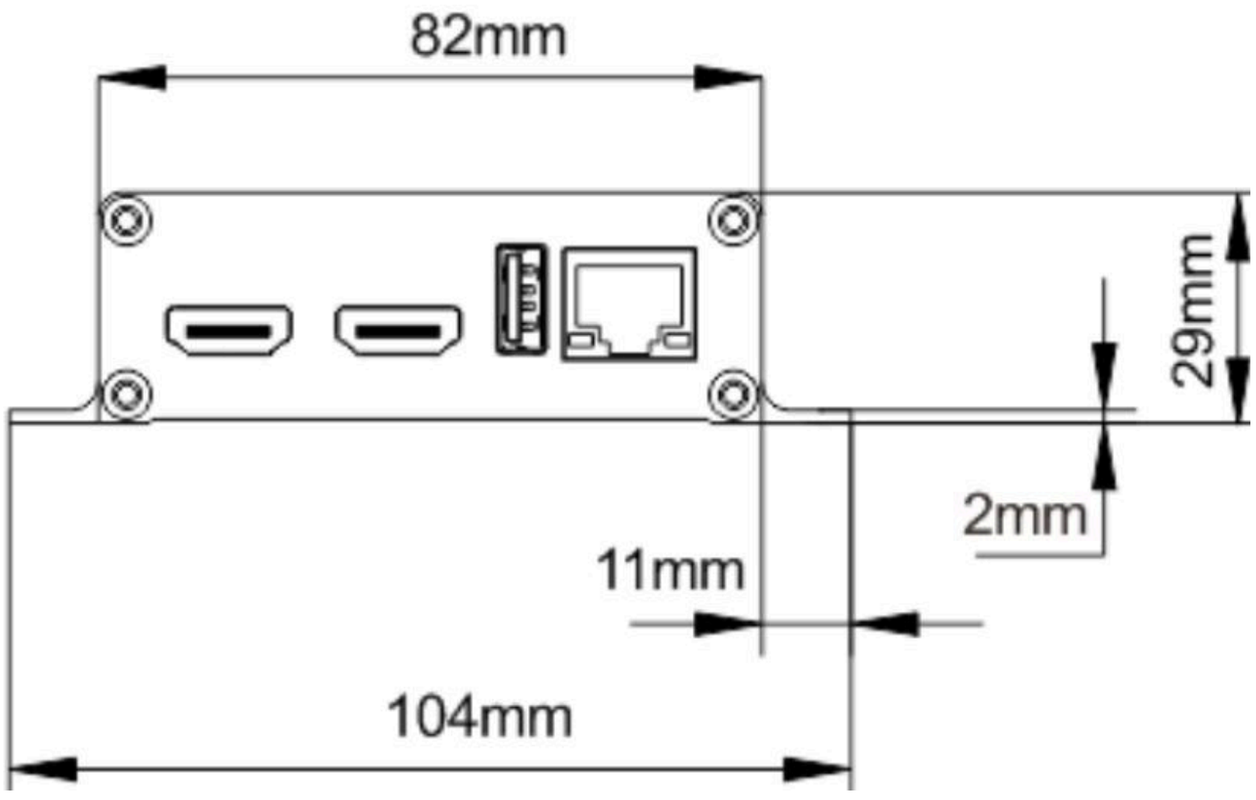
© 2025 AISPEECH and/or its affiliates. All rights reserved

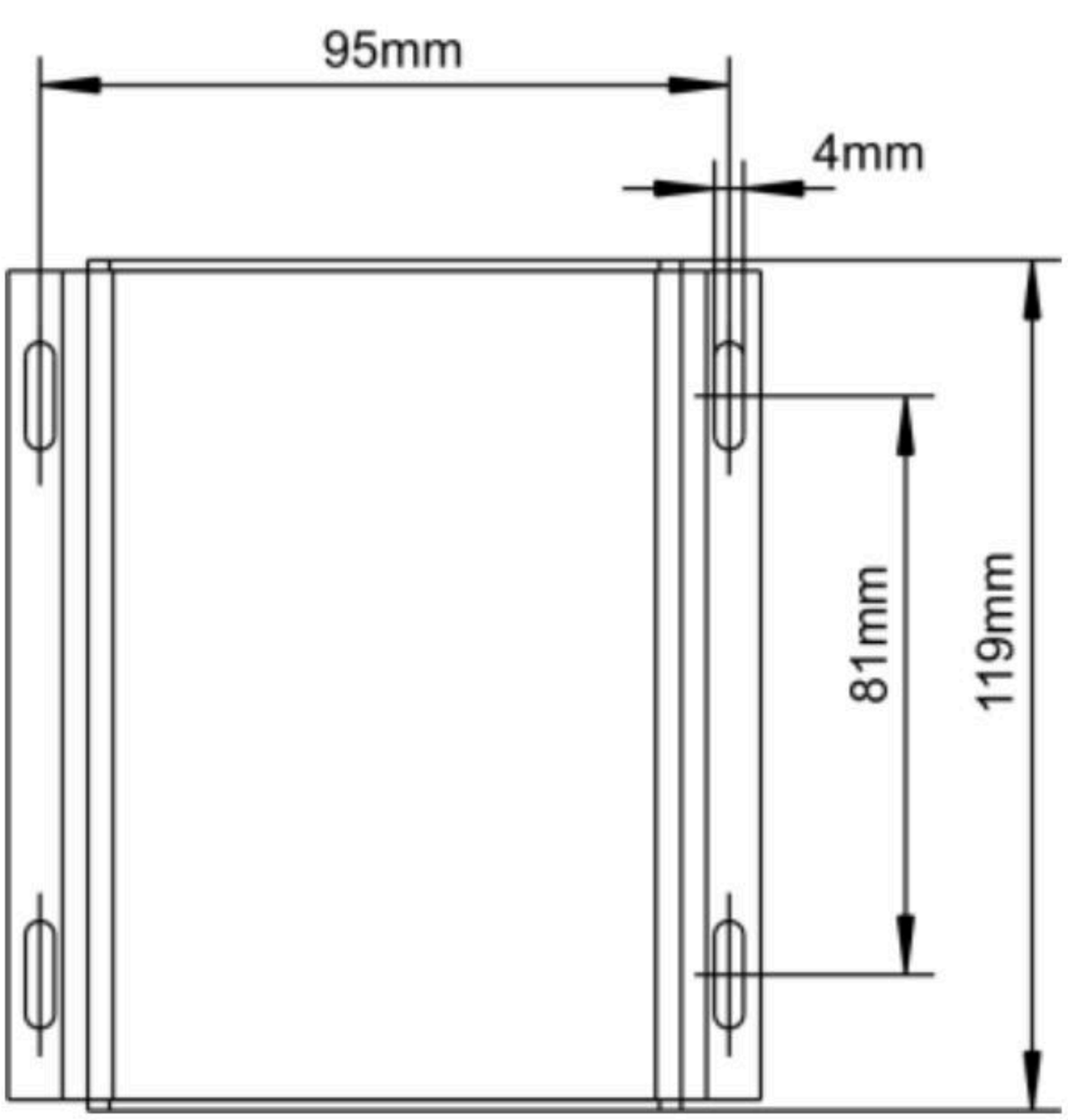
Product Specifications

Function	Specification
TF card	• 1 external Micro SD card, up to 256GB
Power	• DC 12V
Operating temperature	• 0°C ~ + 40°C
Storage temperature	• -20°C ~ +40°C
Dimensions	• 119 mm×104 mm×29mm
Weight	• 0.2KG
HDMI interface	• 1x HDMI-IN • 1x HDMI-OUT (supports audio output) • Output video resolution：4KP30；1080P30/25；720P30/25；640*480P30/25

© 2025 AISPEECH and/or its affiliates. All rights reserved

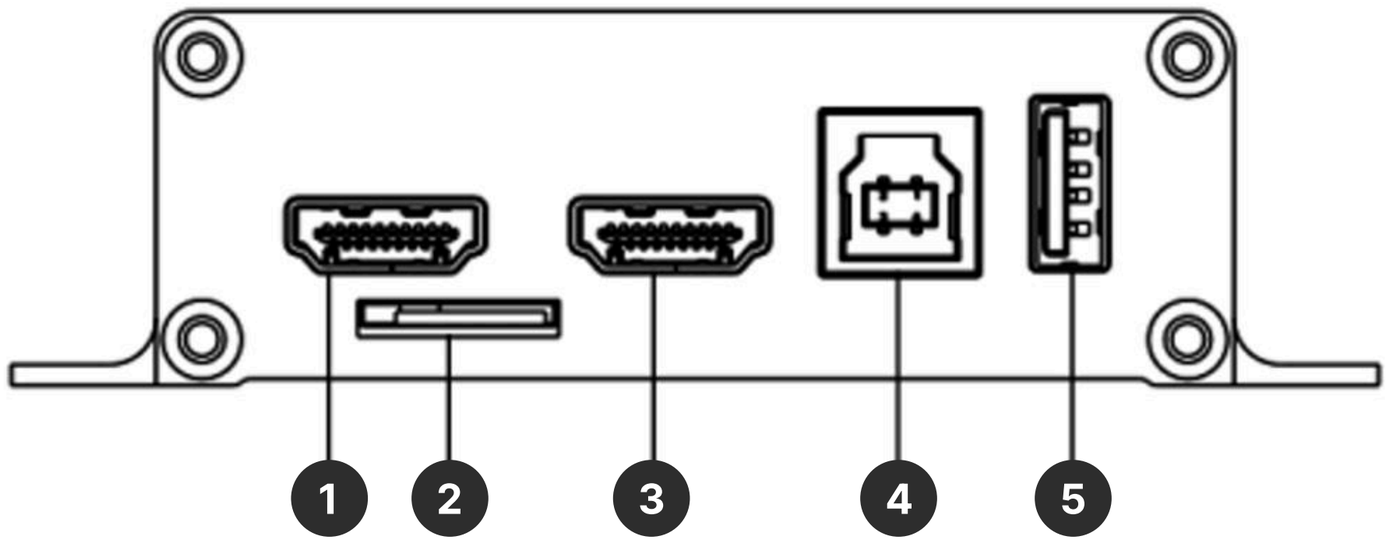
External Dimension





© 2025 AISPEECH and/or its affiliates. All rights reserved

Component Instructions



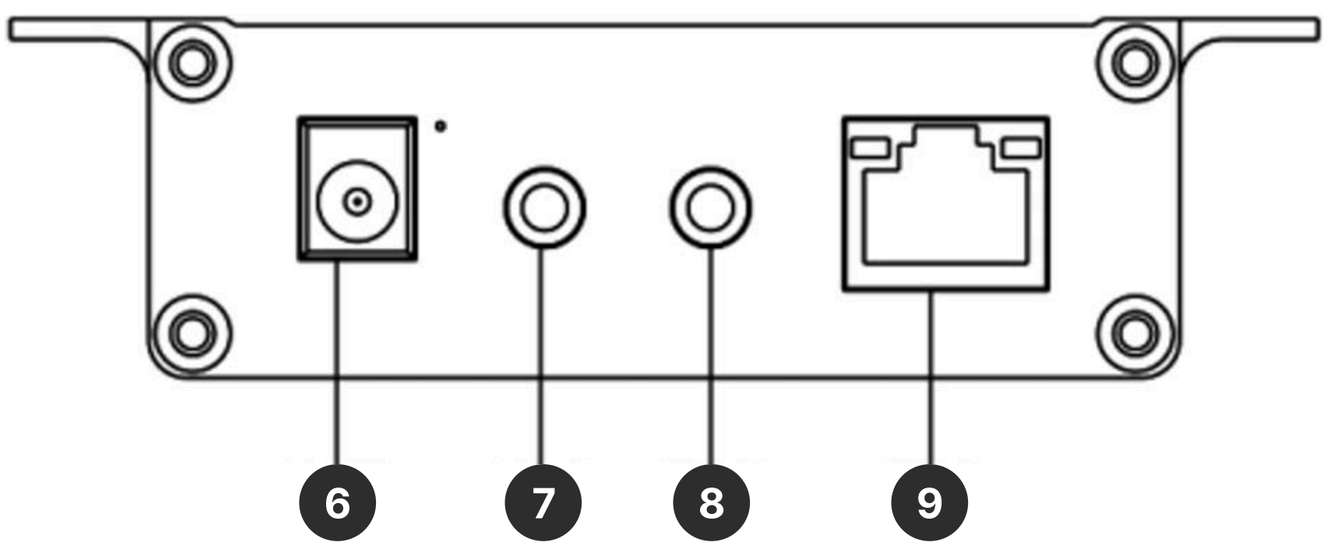
1 HDMI IN

2 TF Card Slot (reserved function)

3 HDMI OUT

4 USB 3.0

5 USB 2.0



6 Power (DC12V) and PoE (802.3af)

7 LINE OUT (reserved function)

8 LINE IN (reserved function)

9 LAN

Learn more : https://meeting.aispeech.com/

© 2025 AISPEECH and/or its affiliates. All rights reserved