Product Brief

Product brief

OVERVIEW

With Yobe software, voice simply works - and not just in the lab, but in the real world. Yobe software enables your voice platform to confidently and consistently work in noisy, multi-voice, and ever-changing real-world use cases through the advanced extraction of key biological, linguistic, and acoustic markers. Yobe software is compact, runs on the edge, and is flexible for your needs. For more information on the product and use cases, see below. For supporting technical information, see our solution brief.

Yobe software delivers distinct voice benefits across a range of industry verticals, resulting in a broad and impactful set of use cases.

Core Industry Verticals:
Consumer Electronics / Smart Devices, Manufacturing / Industrial Automation, Retail & E-commerce, Automotive, Government & Public Sector, DoD & Surveillance, & Healthcare
Additional Verticals: Hospitality & Tourism, Gaming & Virtual Reality, Sports & Fitness, Energy & Utilities, Entertainment & Media, Transportation & Logistics, & Education.

Platforms enabled by Yobe Software*
-Voice interface (VUI) -Voice ID (Biometrics) -Voice Data Analytics -Voice Prompt for Generative AI (LLMs)

Industry Example: AUTOMOTIVE

Benefits of Yobe Software to Voice Platforms

Use Case: Creating Privacy Zones for Rideshare Passengers

Car manufacturers have begun incorporating voice control technology into their vehicles, aiming to revolutionize the driving and riding experience. This example highlights the successful implementation of Yobe's enhanced voice controls, transforming the user experience, while accelerating productivity during the ride-share experience.

CUSTOMER NEED:
To allow rideshare passengers to use personalized voice assistants and voice commands in the cabin of a vehicle.

CHALLENGE:
Accurate voice capture for STT/transcription and user identification in the presence of road noise, wind noise, music interference, and high levels of cross-talk (other passengers talking at the same time).

YOBE SOLUTION:
Yobe Near-Voice/Speaker-Dependent SDKs integrated into the customer's Android application.

RESULTS:
The solution enabled passengers in a 6-seat minivan to talk to a personalized voice assistant (embedded on a tablet mounted on the back of the seat in front of them) and only have their voice transcribed for accurate voice queries and commands in the presence of road noise, music, and other passenger's speech. By creating a personalized voice setting, the speech-to-text accuracy was improved by over 90% in this use case while adding a layer of user functionality not previously seen.

*Voice platfom benefits by Industry Vertical.

Product brief

OFFERING

Our product is offered as a Software Development Kit (SDK) and is available for Android platforms, Windows platforms, Linux platforms, and Embedded ARM-64 platforms. The SDK includes a library implementing the Yobe C++ API, full documentation and example code that illustrates how to use the library.

The Yobe SDK is tailored to handle two specific types of acoustical configurations –
NEAR-VOICE and FAR-VOICE.

For NEAR-VOICE, the target talker is closer to the device than the interfering noise. This situation arises in use-cases like voice-enabled apps requiring phone-in-hand operation, Public kiosks with voice input, and in-hand voice remote controls. In these use-cases, Yobe software can exspand the range of environments where reliable operation is possible.

For FAR-VOICE, the interfering noise is much closer to the device than the target talker (within two feet). This situation typically arises in use-cases where noise is generated from the device itself or a nearby device. Examples are voice input for appliances, industrial machines and robots. The technology can also be applied to devices with audio output like smart speakers. In these use-cases, Yobe software can extend the effective range of devices by over 20 feet.

For either acoustical configuration, Yobe's algorithms can also focus attention on a specific talker if necessary using SPEAKER-DEPENDENT processing. This can significantly improve performance in the presence of interfering talkers. In SPEAKER-DEPENDENT operation, a target talker is established by a manual or automatic enrollment process causing other interfering voices to be suppressed. This can be applied, in public settings like Kiosk clusters where many customers are talking at the same time, or for user personalization like apps running on a cell phone in which the device should only respond to the owner.

When this extra functionality is unnecessary or undesirable, SPEAKER-INDEPENDENT operation can be specified.

Note: our automatic range adaptive feature (the ability to change between Near and Far voice scenarios) is currently under development.

Speaker-Independent Requirements	
Data Memory	50 kbytes
Library Size	75 kbytes
1-CPU Usage	15%
Platform	Linux, Windows, ARM (32-bit, 64-bit)

Metric	Results
Processing Latency	less than 50 ms
Rate of Environmental Adaptation	Every 128ms
ASR/STT Accuracy Enhancement	From 65% to 90%+ in -SNRs
SNR Improvement (Far Voice)	15 dB+

Product brief

Our executables generate two outputs that our customers use to build on:

- Enhanced Voice (Voice-extracting output): voice (in buffered form) that has been optimized for machine transcription, machine extraction & machine matching as well as other data analytics
- Identified Voice (Voice-identifying output): Real-time (frame-by-frame) Speaker Identified Voice
- ID Indexing (in development) biomarker template matching with confidence score output

Output	Core Benefit	Example 1	Example 2
ENHANCED VOICE	Clean voice extracted froms a noisy environmrnt that has been optimized for Machine listening (20% bounce in accuracy in low SNRs).	Hands-free voice commands for devices that need to operate in noisy environments or that make their own noise.	Accurate speech-to-text (regardless of the noise) for transcription and translation platforms that operate in the cloud or on the edge.
IDENTIFIED VOICE	User/Target Identified Voice extracted from a noisy/multi-voice environment that has been optimized for Machine or Human listening	Speaker verified speech/ text output as an input for private devices (cell phones), digital humans, kiosk ordering, and personalized document recording.	Biometric-linked mute and unmute function for video/voice calls in high noise and cross-talk environments.
* ID INDEXING	Voice biometric template matching and confidence scoring in noisy environments	Biometric verification and confirmation for user profile retrieval.	Biometric verification and confirmation for application access

Current Integration Options supported by Yobe:

Microphone Capture	Accessing Yobe Output
Yobe's SDK requires two microphones input data (interleaved) . The format can be either 16-bit PCM data or normalized data. The SDK also provides Microphone configurations for Broad-side or End-Fire form-factors.	For Enhanced Voice YOBE SDK provides API access to mono-channel 16-bit PCM.
Yobe can provide real-time audio capture support for Windows/Linux and Embedded (ARM-32/64) platforms.	For Identified Voice YOBE SDK provides API access to mono-channel 16-bit PCM.
On an Android platform, Yobe's SDK has integrated capability to capture device microphones	For ID Indexing, YOBE SDK provides API access to confidence levels every 128ms that the voice captured belongs to enrolled user.

Next Steps

Learn how Yobe can make your voice interface robust to the real world! Visit www.yobeinc.com or contact us at contact.us@yobeinc.com