



Protect your communications against voice deception

CandyVoice

- Expertise in digital voice processing
- Innovative Voice Technologies provider

Envisioned Modes of Collaboration

- Provider of innovative voice-technology components
- Specialized subcontracting
- Collaboration with operational units
- Partnership within the ecosystem

Military Credentials

- TRT: development of a military digital radio modem with frequency hopping
- Thales: integration of the TETRA vocoder
- Past clearances: *Secret Défense* and NATO security clearance

Common features of the technological building blocks

- Proprietary technologies with high TRL (Technology Readiness Level)
- Operate locally (no cloud dependency)
- Real-time performance
- Low computational requirements
- Can operate independently or in synergy
- Simple and intuitive interfaces
- Technologies adaptable to specific needs or use cases (customization options available)
- ITAR-free technology
- Technology demonstrations available upon request

Integration options

- Locally or on a server
- In a Docker container
- Cross-platform support (all major operating systems)
- Compatible with ARM processors (easily integrable into any hardware, e.g., digital radios)
- Runs on embedded processors (no need for GPU cards)

Definition of Voice Imitation

Reproduction of a person's voice (Target Voice), clearly defined and recognizable, by another person (Source Voice) in real time. (≠ voice conversion).

Value Proposition

Model the human voice—even from a small amount of audio data—and imitate that voice in real time using the generated vocal model of the Target Voice.

Innovation / Competitive Advantages

- World-unique technology (real time!)
- Simple and fast process:
  - Does not require prior modelling of the Source Voice
  - Requires very few recordings of the Target Voice
  - Target Voice model generated automatically and rapidly (2 minutes)
- Adaptable to all languages

State of the art – Existing voice imitation approaches

TECHNOLOGICAL APPROACH	GENERAL PRINCIPLE	STRENGTHS	STRUCTURAL LIMITATIONS
CandyVoice real-time Speech-to-Speech (STS)	Frame-level voice imitation without buffering	Ultra-low latency with full preservation of source voice prosody	No structural limitations
Text-to-Speech (TTS)-based voice cloning	Transcription of the source speech followed by resynthesis in the target voice	High audio quality for post-production	High latency, loss of spontaneity, incompatible with real-time interaction
Speech-to-Speech (STS) with delayed processing	Direct imitation of the source voice into the target voice using temporal buffering	Better preservation of prosody	Multi-second delay, disruption of conversational flow
Cloud-based solutions	Voice processing executed on remote servers (GPU-based)	Scalability, high computational power	Network dependency, variable latency, unsuitable for critical or tactical contexts

- End-to-end real-time voice imitation latency: ~70 ms
- Enables smooth and fully interactive conversations
  - Ensures immediate responsiveness with no perceptible delay
  - Allows seamless real-time integration into communication systems

TRL (Technology Readiness Level) : 7