

Azure AI 视频翻译动手实验

视频翻译是 Azure AI 语音中的一项功能，可让你自动无缝翻译和生成多种语言的视频。此功能旨在帮助你本地化视频内容，以满足全球不同受众的需求。你可以跨各种用例（如 vlog、教育、新闻、企业培训、广告、电影、电视节目等）高效创建沉浸式本地化的视频。

通常依赖于使用不同语言录制的音频替换视频的原始语言来迎合不同的受众。传统上，翻译是通过人工录制和手动后期制作来实现的，对于确保受众能够用母语欣赏视频内容至关重要。但是，这一过程存在关键痛点，包括成本高、持续时间长、无法准确复制原始说话者的声音。Azure AI 语音中的视频翻译通过提供用于创建本地化视频的自动化、经济高效的解决方案来解决这些挑战。

Azure AI 语音提供的视频翻译在各个行业和内容类型中具有广泛的用例。下面是一些关键应用：

- 新闻 + 采访：新闻机构可以翻译和配音新闻片段和采访，以便向全世界受众提供准确、及时的信息。
- 广告 + 营销：企业可以本地化其广告和营销视频，以造成不同市场目标受众的共鸣，从而提高品牌知名度和客户参与度。
- 教育 + 学习：教育机构和电子学习平台可以将其教学视频和讲座配制成不同的语言，使学习变得更加便捷和包容。
- 电影 + 电视节目：电影工作室和制作公司可为其电影和电视节目配音以供国际发行，从而吸引更多广泛的观众并最大程度地提高收入潜力。

- Vlog + 短视频：内容所有者可以轻松为其 vlog 和短视频进行翻译和配音，以吸引国际观众，扩大观看人数和参与度。
- 企业培训：企业可为不同区域的员工本地化培训视频，确保员工之间保持一致、有效的沟通。

目前，视频翻译服务仅在East US 区域可用。

① 备注

语音服务不会在语音资源区域之外存储或处理数据。数据仅在创建资源的区域中进行存储或处理。例如，如果在 `westus` 区域中创建语音资源，则数据仅位于 `westus` 区域中。

视频翻译服务的价格参考：

视频翻译预览

批处理: 每输出视频分钟 \$1.00

内容编辑: 每输出视频分钟 \$0.20

个人声音: 每输出视频分钟 \$0.20

创建Azure 语音资源

要开始使用 Azure AI 语音的视频翻译功能，首先需要创建一个 Azure 语音资源。这个资源将为你提供访问 Azure AI 语音服务所需的密钥和终结点。让我们一步步了解如何在 Azure 门户中创建和配置语音资源。以下是具体步骤：

1. 登录到 Azure 门户 (portal.azure.com)，使用你的 Azure 账号凭据。
2. 在门户主页点击"创建资源"按钮，或使用搜索栏查找"语音服务"。
3. 在搜索结果中选择"语音服务"，然后点击"创建"开始配置新的语音资源。

×


定价：全部 ×

操作系统：全部 ×

添加筛选器

☐ 仅有资格享受的 Azure 权益 ⓘ ☐ 仅 Azure 服务

正在显示“语音”的 1 到 4 个结果(共 4 个)。 [清除搜索](#)





语音


Microsoft

Azure Service

使你能够将类人语音功能嵌入到你的应用中的 AI 服务

创建 ▾ 







Azure 通信网关

Microsoft

Azure Service

部署 Azure 通信网关。

创建 ▾ 





Azure Bot

Microsoft

Azure Service

使用 Bot Framework Composer 或 SDK 构建企业级人工智能对话体验。

创建 ▾ 




通信服务

Microsoft

Azure Service

创建用于开发连接 Web 的应用程序的通信服务资源

创建 ▾ 

4. 指定资源区域为East US，设置定价层为S0，如下图所示：

创建语音服务

[Basics](#)
[网络](#)
[Identity](#)
[Tags](#)
[审阅并创建](#)

将可听到的语音转录为可读且可搜索的文本。将实时语音翻译添加到应用和服务。准实时地将文本转换为音频。使用已经在使用的编程语言来快速构建已启用语音的应用和服务。自定义语音系统，以优化特定方案的质量。

[了解详细信息](#)

项目详细信息

订阅 * ⓘ

MICL-New-Non-Prod-Env
 ▼

资源组 * ⓘ

(新项) 语音翻译Lab
 ▼

[新建](#)

实例详细信息

区域 ⓘ

East US
 ▼

名称 * ⓘ

vt-lab-eastus
 ✓

定价层 * ⓘ

Standard S0
 ▼

Free F0

Standard S0

[查看全部定价详细信息](#)

5. 点击创建按钮创建语音资源。

创建视频翻译项目

延续上面的操作步骤，点击“转到资源”按钮，跳转到创建好的语音资源

✓ 部署完成



部署名：Microsoft.CognitiveServicesSpeechServices-20250114140711

订阅：MICL-New-Non-Prod-Env

资源组：语音翻译Lab

> 部署详细信息

✓ 后续步骤

转到资源

再点击“转到Speech Studio”按钮。

开始使用 监视

开始使用 Speech Studio 中的资源



试用所有用例并查看其他用于生成语音 AI 模型的自定义工具

转到 Speech Studio

下面正式开始创建视频翻译项目。

1. 在Speech Studio 界面上点击视频翻译

按场景显示的语音功能

使用语音转文本和文本转语音等 Azure 语音服务功能浏览、试用和查看一些常见用例的示例代码。



使用语音转文本生成字幕

将电视广播、网播、电影、视频、实时事件或其他制作内容的音频内容转换为文本，使观众可以更轻松地访问你的内容。

[试用字幕](#)



通话后听录和分析

批量操作转录呼叫中心录制并提取有价值的信息，如个人身份信息(PII)、情绪和通话摘要。

[试用通话后听录](#)



实时聊天头像

使用虚拟形象进行自然对话，该虚拟形象可识别用户的语音输入，并流畅地使用逼真的 AI 语音做出响应。

[试用实时聊天虚拟形象](#)



语言学习 [预览](#)

从聊天体验中获取有关发音准确性、流畅性、韵律、语法和词汇的即时反馈。

[试用语言学习](#)



视频翻译 [预览](#)

轻松翻译 AI 语音配音并应用到 100 多种语言的视频，可选择 400 多种预生成语音或跨语言使用个人语音。

[启动视频翻译](#)

2. 点击创建一个项目按钮，弹出创建项目界面

新建项目



声音类型 *

预生成神经语音

自动匹配类似于视频中语音的预生成神经语音。

个人声音

你没有访问权限

在视频中创建说话人的个人语音，并应用于翻译后的版本。

观看[个人语音示例视频](#)并[了解详细信息](#)。

视频文件 *



在此处拖放文件
(.mp4 文件 < 500MB, 小于 60 分钟)

[浏览文件...](#)

项目名称 *

输入文件名称

发言者数量 *

自动检测 ☒ 开

1

视频语言 *

英语(美国)

目标语言 *

德语(德国)

词典文件 ①

无

☐ 添加字幕文件

☐ * 我确认视频翻译将自 2024 年 6 月起向我的帐户收取使用费。[查看定价](#)

☐ * 接受 [Azure AI 语音文本转语音的行为准则](#)

创建

取消

3. 声音类型选择“预生成神经语音”

4. 将预先准备好的视频文件拖入文件上传界面。

视频文件1：[Introducing the New and Improved Mercedes-Benz App](#)

视频文件2：[Microsoft Copilot Pro AI Walkthrough](#)

以上文件二选一即可

5. 设置目标语言为“中文(普通话，简体)”
6. 勾选“我确认视频翻译将自2024年6月起向我的账户收取使用费”和“接受Azure AI语音文本转语音的行为准则”两个选项，点击创建按钮。
7. 等待3~5分钟，即可查看初步翻译结果
8. 翻译结束后，点击播放按钮查看翻译效果

修改翻译设置

1. 在翻译界面上点击“语音设置”



2. 尝试在语音设置中，把声音设置为“云健”

语音设置



选择已翻译的视频的语音。语音语言: 中文(普通话, 简体)

说话人姓名

Speaker1

声音类型 *

性别

声音

预生成神经语音

男

云健

添加说话人

更新

取消

3. 点击依次点击“更新”，“保存”和“应用更改”，略等片刻后，播放视频，听听效果。
4. 在播放界面右侧，是对应的字幕块。可以修改字幕。尝试修改一部分内容，点击“保存”和“应用更改”，看看结果。

视频

已翻译

原始

00:00:06.880 / 00:00:59.562

features you love with a fresh look and a new name. 旧，开观焕然一新，名字也变了。

Speaker1

00:00:06.880 - 00:00:18.800

The app's intuitive redesign and enhanced functionality now allow you to have quick access to your most used features like lock and unlock, pre-entry climate control, lights and horn. 应用程序的直观重新设计和增强的功能现在允许您快速访问最常用的功能，如锁车、解锁、预设置车内环境控制、灯光和喇叭。

Speaker1

00:00:19.760 - 00:00:30.012

The updated home screen also has shortcuts like the all-new fuel bar where you can easily view your trip data and the battery bar, where you can view your vehicle's charging status. 更新后的主屏幕还增加了快捷方式，比如全新的燃油条，您可以轻松查看行程数据，还有电池条，您可以查看车辆的充电状态。

5. 在播放界面下部，是音轨编辑界面，可以尝试拉动每句话的语言块控制语速。



查看编程文档

视频翻译功能支持通过REST API 的方式调用，通过编程的方式批量地对视频进行翻译，点击链接查看开发文档：<https://learn.microsoft.com/zh-cn/azure/ai-services/speech-service/video-translation-get-started?tabs=webvtt&%3Bpivots=rest-api&pivots=rest-api>