

AI最前线：声音画像，只闻其声也能“看透”你

人工智能 (<https://www.secrss.com/articles?tag=人工智能>) · 机器之能
(<https://www.secrss.com/articles?author=机器之能>) · 2019-04-21



(<https://www.secrss.com/login>)

基于 AI 算法与声音分析学的融合，「声音画像师」正在帮助医院可以提早发现神疾病及身体疾病，让呼叫中心进行个性化营销，为公司筛选出更为合适的求职者……



你感到抑郁吗？会在开车时打瞌睡吗？有被骗保经历吗？有心脏病隐患吗？AI 或许能够过倾听你的声音来解决这些问题。

人类的声音中蕴藏着比我们想象中还要丰富的信息，你的年龄、身高、体重、精神、情绪、甚至情感、气质、进取心、创造力……与此同时，声音似乎比图像更加安全可控。

基于 AI 算法与声音分析学的融合，「声音画像师」正在帮助医院可以提早发现神疾病及身体疾病，让呼叫中心进行个性化营销，为公司筛选出更为合适的求职者……

作者 | John McCormick

编译 | 张玺

日益强大的机器学习方法，使一切变得皆有可能。

AI 系统可以测量音调、节奏及其他语音特征，并将其与存储的语音模式，包括快乐、悲伤、疯狂及其他情绪进行比较。

虽然声音分析学已经发展了几十年，但成本更低廉的计算力与新一代的 AI 工具，如谷歌的 TensorFlow，才使那些雄心勃勃的项目成为可能。

在情感 AI 或情感计算领域，该技术与计算机视觉结合后，功能将变得愈加强大。例如，车载语音系统可判断驾驶员是否哈欠不断，视觉系统则能判断驾驶员是否在打瞌睡。

咨询公司 Gartner 认为，情感 AI 甚至可能进入消费品市场，Gartner 预测，到 2022 年，10% 的个人设备（目前这一比例少于 1%）将具备情感 AI 功能，如监测个人精神状态的可穿戴设备，和基于玩家情绪自动调整的视频游戏。

但在情感 AI 成为主流之前，必须先克服一个巨大的阻碍：人们的不适应情绪。Gartner 在去年发布的一份调查报告指出，来自美国和英国的 4000 多名受访者中，有 52% 的人表示不希望 AI 分析自己的面部表情，63% 的人表示不希望 AI 为了解她/他们而经常「入侵」其工作与生活。

此外，消费者还特别关心自身隐私，约有 65% 的人坚信，AI 会侵犯其隐私。

Gartner 分析师 Annette Zimmermann 撰写了关于情感 AI 的报告，她表示：「人们普遍不相信 AI。至于大家对 AI 的感受，我认为这是仁者见仁、智者见智的问题，但的确有许多理由令人生疑。」

Zimmermann 表示，系统都不完美，最佳系统的精确率也不超过 85%。

CMU (Carnegie Mellon University, 美国卡耐基梅隆大学) 语音学家 Rita Singh 说，「这不完全正确。而且我们也不知道系统何时才能变得精确。但是，快了。」

带着上述防止误解的说明，我们来了解下 AI 语音分析技术已经影响以及即将颠覆的领域。

01 医学领域：疾病监控师

2017 年，根据美国国家精神卫生研究所的数据，2017 年，全美有五分之一的人（约 4660 万）饱受精神疾病的困扰。据估计，接受治疗的人仅占需要接受治疗的人数的一半，而新兴的语音技术可能使该问题更容易被发现。

2018 年底，从行为分析公司 Cogito 衍生出的 CompanionMx 公司发布了一款名为 Companion 的移动心理健康监测系统。该系统由美国国防高级研究计划局、美国退伍军人事务部及美国国家精神卫生研究所资助研发。

正接受抑郁症、双相情感障碍和其他疾病治疗的患者，可通过 Companion 下载一个 APP，在智能手机上创建音频日志。系统要求患者定期描述个人感受，并将信息自动传输到 AI 模块进行分析。

运用 Cogito 研发的情感 AI 技术，CompanionMX 可以分析患者的音频及某些行为数据，以了解其情感、情绪或行为的变化。譬如，CompanionMX 通过监测智能手机活动，看患者是否减少与外界的联系。如果护理人员看到有问题的迹象，他们会主动联系患者。

美国国家精神卫生研究所资助了一项关于该 APP 的研究项目，自 2015 年 5 月至 2017 年 8 月。

「研究结果十分令人鼓舞。」项目联合首席研究员、布莱根妇女医院 (Brigham and Women's Hospital) 与哈佛医学院数字行为健康与信息学研究项目负责人 David Ahern 说。

Ahern 表示，这款 APP 可以作为护理人员的早期监测系统。因为多数需要治疗的患者直到病情严重才会寻求治疗，所以，此款 APP 是一个非常必要的工具。

根据疾病控制和预防中心 (Centers for disease Control and Prevention) 的数据，全美每年有 60 多万人死于心脏病。研究人员正尝试使用语音 AI 发现警告信号，帮助人们快速获得治疗。

梅奥诊所进行了一项为期两年的研究项目（2017 年 2 月结束），旨在验证语音分析是否能够检测出冠心病。梅奥医学中心心血管研究中心主任 Amir Lerman 解释说，每个人的声音频率不尽相同，它们可以被分析。

梅奥与语音 AI 公司 Beyond Verbal 合作，采用机器学习识别冠心病的语音生物特征判定标识，而后，在计划进行血管造影的患者人群中展开测试。

所有研究参与者都在手机 APP 上录制了自己的声音，而后，Beyond Verbal 对这些声音进行了分析。研究发现：在血管造影上有冠心病病症的患者也有该疾病的语音生物特征判定标识。

Lerman 博士表示，梅奥希望在不久的将来部署这项技术。「我认为这是一个令人惊叹的新领域，为我们如何治疗患者打开了新的大门。」他说。

02 智能驾驶：评估驾驶员精神状态

根据国家公路交通安全管理局 2017 年 10 月的统计数据，在 2015 年，有 800 多名美国人因为疲劳驾驶而丧生，超过 3 万人在司机疲劳驾驶导致的撞车事故中受伤。

现如今，许多大型汽车公司与 AI 公司正在设计一种 AI，运用语音分析与面部识别技术评估驾驶员的警觉性及情绪状态。

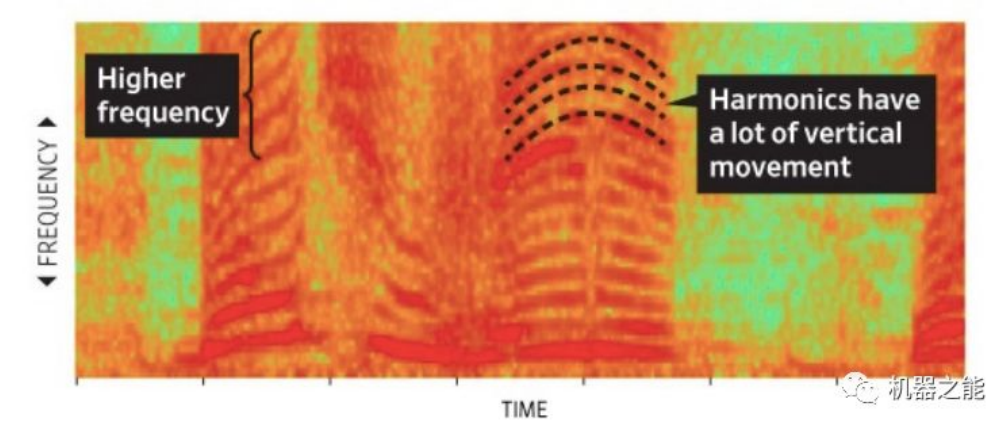


你能看出我生气了吗？

声谱图可用于分析说话人的声音特征。谐波（深色平行线）表示说话人音调和语调的变化。

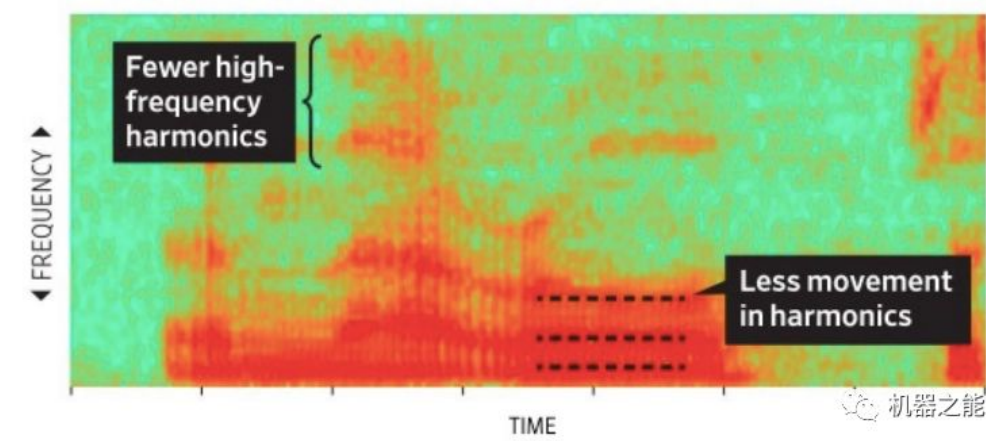
高情绪活动

线条的剧烈运动反应了语调的快速变化。在较高频率下，线条颜色越深，其代表语音质量越高。



低情绪活动

声音更柔和、平静，谐波的强度要小的多，特别是在高频率下。



来源：Cogito

去年的消费电子展（CES）上，丰田发布了一款概念车 Concept-i，它可以识别面部表情及语音声调。该车在转向柱上安装了一个红外摄像头，在仪表盘上安装了一对 3D 传感器，还搭载了语音识别与对话系统。

系统协同工作，以评估驾驶员的状态。例如，头部下垂、弯腰驼背的姿势、困倦或低沉的声音（或者打呵欠的声音）都表示驾驶人员处于疲惫状态。如果系统注意到驾驶人的疲惫状态，会迅速做出反应。

例如，车载语音助手可以与驾驶员进行对话，提高其警觉级别。并且，随着时间点的推移，对话系统会知道哪些话题最可能吸引驾驶员的注意。

去年 9 月，两家 AI 公司——Affectiva 与 Nuance Communications 表示，他们将携手合作，把情感智能融入 Nuance 的对话型汽车助手中，理解并回应驾驶员需求。

Nuance 声称，目前在奥迪、宝马、戴姆勒、菲亚特、福特、通用、现代和丰田等 2 亿多辆汽车上都可以找到这款名为 Dragon Drive 的对话型汽车助手。

该项由 Iovetiva 和 Nuance 公司研发的新技术将使用摄像头来检测面部表情（如微笑），采用麦克风识别声音情绪（如愤怒）。他们的算法采用深度学习、计算机视觉与语音技术，来识别情绪和疲惫的特征指标。

如果驾驶人员已显露疲态，语音助手可通过一些简单的话语提示驾驶员，如「你看起来很疲倦，你想停车休息一下吗？」

目前，上述技术仍处于研发阶段，但是据 Nuance CEO Joe Petro 透露，这些技术或将于几年内正式投入使用。

03 智能营销：个性化的呼叫中心

房地产咨询服务商 Site Selection Group 的数据显示，尽管许多公司将客服业务转移到了海外，但全美仍有 7400 座呼叫中心雇佣了 300 多万名员工。

包括哈门那（Humana，美国大的医疗保险公司）、美国大都会人寿保险公司（MetLife）在内的许多公司，已经部署 Cogito 的 AI 软件，以此来保持保险经纪人的敏锐度及客户满意度。

该系统分析保险经纪人与客户之间的对话，实时跟踪其交互方式。

当呼叫中心接到电话后，会将其流处理至 Cogito 的系统，该系统可评估成百上千个数据点，包括语速、音调等。如果保险经纪人在回答问题前停顿，那可能表明其分心了。如果顾客提高嗓门，这可能是一种沮丧的迹象。

当 Cogito 的系统检测到呼叫过程中可能出现问题时，它将以图标或短信息的形式推送通知到员工的屏幕上，帮助保险经纪人识别、了解客户的情绪。

Cogito CEO Joshua Feast 表示，该系统的主要目标是训练保险经纪人，让她/他们更自信、更投入，更有同情心，「学会与各式各样的客户交谈是一项真正的技能，并非与生俱来，必须后天习得。」

Cogito 表示，其呼叫中心产品的准确性因各类使用场景（如客户服务中心、销售部或理赔管理部）和在各领域中的监控行为而异。

总体上，Cogito 称，该产品的平均准确率为 82%。该公司表示，他们通过人工评审呼叫效果、客户反馈及机器学习分析来验证结果。

大都会保险全球客户解决方案负责人 Kristine Poznanski 表示，大约 15 个月前，大都会在其客户服务中心部署了 Cogito 的系统。

该系统除了为客户服务代表提供呼叫即时反馈及实时指导，还能向经理展示呼叫实时状态。呼叫中心经理可通过数据监控正在进行的呼叫，或者在呼叫结束后与保险经纪人一起复盘呼叫过程。

Poznanski 说道，自从部署该系统以来，公司可追踪客户情绪以评估其接受推荐品牌的可能性，而呼叫中心的首次呼叫解决率和净推荐值也增加了 10%。

04 人事招聘：寻找合适的求职者

根据人事公司 Robert Half International 的数据，超过八成以上的企业老板和经理表示他们碰到过招错人的情况。通常，新员工难以适应企业文化是问题关键所在。

Voicesense 是一款基于语音的 AI 系统，可改善筛选求职者的效率。

Voicesense CEO Yoav Degani 表示，用人方可将视频、音频面试资料上传至 Voicesense 的云端，该公司的系统会分析 200 多个语音参数，包括语调和节奏等。该系统创建了一个囊括求职者气质、进取心、可靠性、创造力等特征的行为模型。



用人方可以根据系统评分判断求职者是否适合这份工作。Degani 说道，例如，一个组织希望聘用一名销售员，系统就会将一个在对话中表现极其主动且参与度极高的人列为可能的匹配对象。但 Degani 也承认，该公司的模型提供的仅仅是可能性，而非确定性。

Degani 表示，在隐私保护方面，Voicesense 不会存储任何数据，并且其工具只分析语音模式，不分析具体的对话内容。AdventHealth Orlando 是 AdventHealth 医疗系统的一部分，它正采用另外一套分析系统 HireVue 辅助招聘工作。该公司在佛罗里达州拥有 8 家医院，员工数超过 25000 人，每年招聘 8000 人。AdventHealth 的人力资源总监 Karla Muniz 表示，上述数据意味着要审核 35000 多份申请。

每一位符合基本工作要求的求职者都会被邀请参加 HireVue 线上面试，系统算法会评价求职者的面试应答情况，如语音和词组，并结合视觉分析，快速检测面部表情（即微表情）。

而后，评估信息会与每项工作相对应的数据点进行匹配，得分最高的求职者会被邀请参加面试。

Muniz 说，从采用 HireVue 以来，AdventHealth 已将招聘一个人的周期从 42 天降低到 36 天。

05 保险业：反欺诈的第一步

根据行业贸易集团保险信息研究所发布的数据，每年财产险和意外伤害险的欺诈金额超过 300 亿美元。

安联集团 (Allianz group) 子公司 Allianz-SP Slovakia (保险公司) 运用 Nemesysco 的语音压力分析技术处理索赔。该技术可获取针对理赔员提出一系列流程问题时人们的反应情况。根据该公司的说法，系统会寻找标识组合，如说话时的短暂停顿（这可能代表发言人正在提供虚假信息）。

Allianz-SP Slovakia 控制和特殊活动负责人 Jaroslava Zemanová 表示，「我们的目标是在没有任何问题的情况下立即赔付，并防止任何类似欺诈的夸大索赔。」

Allianz-SP Slovakia 指出，语音分析并不能证明存在任何不当行为，这只是发现可能欺诈行为的第一步。公司的调查团队需要额外的证据才能拒绝索赔。尽管如此，公司表示，该系统帮他们节省了时间和金钱。

06 刑事侦查：模拟还原罪犯

在某些情况下，语音分析不仅可以提供关于人们健康或情感状态的信息，还能提供关于其外貌的信息。

2014 年，美国海岸警卫队试图追踪一名发出 28 次虚假求救信号的人，对这些电话的紧急反响应费用估计为 50 万美元。

海岸警卫队调查服务中心切萨皮克区域负责人 Marty Martinez 说，这不仅仅是经济上的损失，「它占用了那些真正陷入困境的人的资源。」

除了紧急呼叫录音之外，海岸警卫队调查员基本没有其他线索。然后，她/他们去见了 CMU 的 Singh，Singh 一直致力于计算机语音识别研究。

Singh 仅凭录音就能够确认恶作剧报警人年龄、身高和体重。Martinez 表示，案件目前正在进行中。

Martinez 补充道，该项技术已应用于十多起案件，「它帮助我们缩小调查范围，聚焦调查工作。」

原理是什么？Singh 解释道，人类声音所包含的信息与其身体、生理、人口统计学、医学、环境等其他特征有关。研究人员正在发现这些微特征，并利用它们进行分析。

Singh 说，「我称之为声音画像学。」

Singh 承认，该项技术并不完美，譬如无法精确预测年龄：只能在三年范围内预测。但是技术研究可提高其精度，并将其纳入新领域。



AI最前线：声音画像，只闻其声也能“看透”你 - 安全内参 | 决策者的网络安全知识库
Singh 及其团队最近展示了一个系统，该系统可根据声音重构说话人 60% 至 70% 的面部信息。

Singh 表示，语音分析技术任重道远，但潜力巨大，「它将帮助机器比人类更好地了解人类。」

原文链接：<https://www.wsj.com/articles/what-ai-can-tell-from-listening-to-you-11554169408>

声明：本文来自机器之能，版权归作者所有。文章内容仅代表作者独立观点，不代表安全内参立场，转载目的在于传递更多信息。如有侵权，请联系 anquanneican@163.com。

人工智能 (<https://www.secrss.com/articles?tag=人工智能>)

相关资讯

- AI项目开发中最容易犯的10种错误 (<https://www.secrss.com/articles/56834>)
安全运营 (<https://www.secrss.com/articles?tag=安全运营>) · 安全牛
(<https://www.secrss.com/articles?author=安全牛>) · 2023-07-19
www.secrss.com/articles/56834
- 《认知智能技术与应用研究报告》正式发布
(<https://www.secrss.com/articles/56789>)
人工智能 (<https://www.secrss.com/articles?tag=人工智能>) · 可信AI评测
www.secrss.com/articles/56789 (<https://www.secrss.com/articles?author=可信AI评测>) · 2023-07-18
- 被重塑的世界？ChatGPT崛起下人工智能与国家安全新特征
(<https://www.secrss.com/articles/56768>)
国家安全 (<https://www.secrss.com/articles?tag=国家安全>) · 国际安全研究ISS
www.secrss.com/articles/56768 (<https://www.secrss.com/articles?author=国际安全研究ISS>) · 2023-07-18

评论 (0)

登录后才能发表评论，请先 [登录 / 注册](https://www.secrss.com/login) (<https://www.secrss.com/login>)