

# 基于语音情感能识别的情感化产品设计应用研究

田鑫鑫, 胡志刚

陕西科技大学设计与艺术学院, 陕西 西安

收稿日期: 2023年8月31日; 录用日期: 2023年11月21日; 发布日期: 2023年11月29日

---

## 摘要

情感化产品设计是有目的、有意识地激发人们的某种情感, 比较注重产品对人的心理影响, 遵循情感至上的原则。心理学对情感的定义为: “人对客观现实的一种特殊反应形式”, 语音情感能识别旨在设计创造一种可以识别、数字化和应用人的情感, 同时能够尽可能地反应该真实人类情感的系统。本文将从语音情绪识别国内外发展历史出发, 关注语音识别对用户交互行为的影响, 同时将对机器学习方法以及深度学习方法进行简单介绍与分析。

---

## 关键词

语音情感能识别, 系统设计, 交互设计, 情感分析

---

# Research on the Design Application of Emotional Product Design Based on Voice Emotional Recognition

Xinxin Tian, Zhigang Hu

College of Art and Design, Shaanxi University of Science & Technology, Xi'an Shaanxi

Received: Aug. 31<sup>st</sup>, 2023; accepted: Nov. 21<sup>st</sup>, 2023; published: Nov. 29<sup>th</sup>, 2023

---

## Abstract

Emotional product design is a purposeful and conscious inspiration to people's emotions. It pays more attention to the psychological impact of products on people, and follows the principles of emotional supremacy. The definition of emotions in psychology is "a special reaction form of people to objective reality", speech emotion recognition aims to design and create a system that can recognize, digitize and apply human emotions while reflecting real human emotions as much as possible. This article will start with the history of the development of voice emotional identifi-

cation at home and abroad, pay attention to the impact of voice recognition on user interaction behavior, and simply introduce and analyze machine learning methods and deep learning methods.

## Keywords

**Speech Emotional Recognition, System Design, Interactive Design, Emotional Analysis**

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在当今竞争激烈的市场环境中，产品的功能和性能已经不再是唯一的卖点。消费者越来越注重产品与自身情感的连接，他们渴望与产品产生情感共鸣，寻找能够满足其情感需求的体验[1]。因此，情感化产品设计成为了产品开发和市场营销中不可或缺的一部分。情感化产品设计不仅仅关注产品的功能性，更强调如何在产品中融入情感元素，以引发用户的情感共鸣和情感体验[2]。一个成功的情感化产品能够在消费者心中建立起情感纽带，使其产生忠诚度和情感依赖，从而实现品牌的长期影响力和市场份额的增长。本文将探讨情感化产品设计的重要性，深入分析情感化设计在不同领域的应用案例，探讨如何通过情感化产品设计创造积极的用户体验和情感连接[3]。同时，我们还将讨论情感化产品设计的挑战和策略[4]，以及未来情感化产品设计可能的发展方向。通过本文的阐述，读者将能够更好地理解情感化产品设计的核心概念，并在实际工作中运用这些概念来创造引人入胜的产品，与用户建立更深层次的情感关系。

## 2. 情感化设计在产品设计的重要性

在当今充斥着各类相似产品的市场中，消费者不再仅仅追求产品的功能性能，更加关注产品是否能够触动他们的情感。情感是人类行为的重要驱动力之一，而情感化产品设计能够将产品与消费者的情感需求紧密联系起来，从而在市场竞争中脱颖而出。情感化产品设计有助于建立品牌与消费者之间的情感纽带[4]。当消费者在使用产品时能够体验到愉悦、温馨或是其他积极的情感，他们更有可能与品牌产生情感连接。这种情感连接不仅能够增加用户的忠诚度，还能够促使他们将自己的情感体验与他人分享，从而为品牌口碑的扩散提供助力。情感化产品设计还能够影响消费者的购买决策。情感因素在购买过程中扮演着重要角色，人们倾向于购买那些能够满足其情感需求的产品。一个精心设计的情感化产品能够在消费者心中引发共鸣，使其更倾向于选择这款产品，而不仅仅是基于技术指标或价格[5]。

## 3. 情感化产品设计发展现状

许多知名品牌，如苹果、谷歌和迪士尼，都在产品中融入了情感元素。苹果公司以其简约优雅的设计风格以及用户友好的界面，创造出了与用户之间深刻的情感连接[6]。谷歌的虚拟助手 Google Assistant 不仅能够回答问题，还能够理解用户的情感并作出相应回应，增强了用户与技术之间的情感互动[7]。

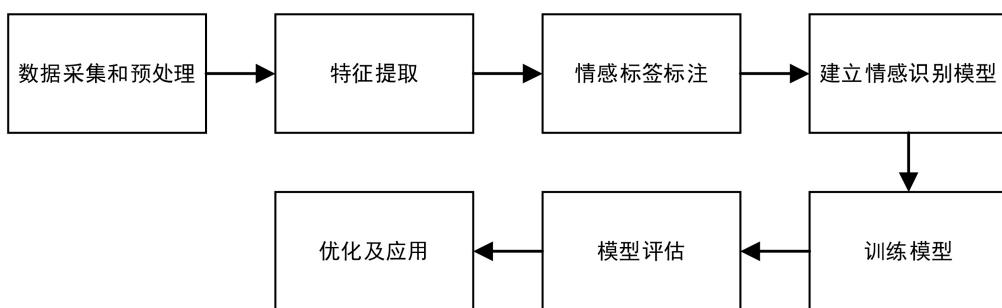
在中国，情感化产品设计也取得了显著进展。互联网巨头腾讯、阿里巴巴等公司在其产品中融入情感元素，创造出与用户情感互动的体验。例如，微信的表情包和语音消息，以及支付宝的红包等功能都

能够触发用户的情感共鸣[8]。中国智能硬件领域也在情感化产品设计方面积极探索。智能音箱、智能家居设备等产品通过语音识别技术和情感传递，使用户与产品之间建立更紧密的情感联系。此外，一些科技公司还推出了情感识别技术，通过分析用户的语音和表情，判断其情感状态，为产品提供更加智能的回应。

#### 4. 语音情感识别的重要性

近年来，语音情感识别(Speech Emotional Recognition, SER)在自然语言处理领域中受到越来越多的关注。语音情感识别是指从语音信号中分析和推断说话人的情感状态或情感体验的过程。这一任务的主要目标是通过分析说话人的声音特征、语音特征、声调、语速等信息，将其情感状态分类或回归成一组情感类别或情感维度。语音情感识别的目标是让计算机系统能够理解和感知说话人的情感状态，这对于许多应用领域如自动客服、情感分析、用户体验评估等都具有重要价值[9]。

语音情绪识别是指利用计算机技术对人的语音信号进行分析和处理，从中提取出说话人的情绪状态，进而实现情感交互和情感诊断等应用。首先情感是抽象的很难被区分出来，同时可能只存在于一段语音中的某些部分[10]，SER 经过数十年的发展，在总结过往的发展过程中发现，深度学习的使用已经成为研究 SER 的一个必要过程。语音情感识别的大致流程分为如下几个步骤数据采集和预处理、特征提取、情感标签标注、建立情感识别模型、训练模型、模型评估以及优化及使用如图 1 所示。



**Figure 1.** Speech emotion recognition flowchart

**图 1.** 语音情感识别流程图

#### 4.1. 语音情感描述模型分类

语音情感描述模型是一种人工智能模型或算法，旨在识别和描述人的语音中包含的情感状态。这些模型可以将情感信息从语音中提取出来，并将其转化为可理解的文本或数字表示，主要可分为离散情感模型与维度情感模型如表 1 所示。

**Table 1.** Discrete-dimensional models compare differences  
**表 1.** 离散维度模型比较区别

比较点	离散情感模型	维度情感模型
情感类别数量	固定的离散类别、明确的情感状态	连续的数值、情感在不同维度上的程度或强度
表达丰富性	粗略分类、缺少对情感细节的详细描述	更丰富的情感表达、可捕捉情感的多样性和复杂性
适用性	简单的情感分类任务	更深入的情感理解和情感描述
优点	简单直观、易于理解	捕捉情感多样性和丰富性
缺点	信息丢失、无法提供详细细节和描述	模型构建和训练较为复杂、计算成本高

#### 4.1.1. 离散情感模型

在对情感进行量化描述时,通常从离散情感模型和情感维度两个方面来进行情感的量化。离散情感模型的优点是简单直接[11],最早由Paul Ekman和Wallace V. Friesen在1971年的研究中提出了基本情感的离散模型,适用于特定场景下对情感进行粗略分类,即将情感表达为预定义的有限个离散情感类别,这些类别通常是人为设定的。这种模型通常使用监督学习算法来进行训练,通过给定标注好情感类别的样本数据来学习情感分类模型,然而,离散情感模型也有一些局限性,它可能无法处理复杂的情感表达,无法提供精细的情感描述,也不能捕捉情感的连续性。

#### 4.1.2. 维度情感模型

维度情感是一种对情感进行连续性描述的方法,与离散情感模型不同,它将情感视为连续的多维空间,而不是简单的离散类别。最早由两位心理学家Russell和Mehrabian在1977年提出。该模型将情感分为两个主要维度:情感价值(valence)和情感唤起(arousal)[12]。在维度情感模型中,情感被描述为在多个情感维度上的连续值,每个维度代表一种情感的程度或强度。通过将情感表示为在这些维度上的连续值,维度情感模型可以提供更丰富、细粒度的情感表达,适用于更复杂的情感分析任务。这种模型对于情感表达的多样性和丰富性非常有帮助。在最近的研究中,深度学习技术(如使用循环神经网络)已经被广泛应用于维度情感模型,以提高模型的性能和泛化能力,具体区别如表1所示。

### 4.2. 语音情感能识别方法

语音情感能识别是一个复杂的任务,由于情感的主观性和多样性,算法的准确性和泛化能力可能会收到影响。因此,一般会使用大量的标注数据来训练和评估算法,并不断改进算法的性能。随着深度学习和多模态方法的不断发展,语音情感能识别算法将继续进步,为情感计算和人机交互领域带来更多可能性。

在过去的数十年中,关于语音情感能识别方法的研究一直以传统的机器学习算法为主,取得了很多的成果,在早期,传统的语音情感能识别算法主要依赖于手工设计的声学特征,如基频、能量、频谱特征等。随着深度学习的兴起,特别是卷积神经网络(CNN)和循环神经网络(RNN)的引入,算法可以自动学习更高级的特征表示,不再依赖于手工设计的特征。同时多模态情感能识别成为研究热点,除了语音数据外,结合视频、文本、面部表情、生理信号等多模态信息,可以提高情感能识别的性能,并可以更全面的理解情感表达。

### 4.3. 语音情感能特征提取

语音情感能识别的关键之一是语音信号的特征提取。在语音情感能识别中,一般采用基于声学和语调特征[13]的方法来提取情感信息。其中,语调特征是比较重要的一类特征,它主要反映了语音信号中的音调、音高、音量等方面的信息。

语调特征的提取主要包括基频、能量、时域声学特征和频域声学特征等。其中,基频是指声音的频率,通常用于表达情感的高低、激动程度等信息。能量是指声音的能量大小,可以用于表达情感的强度、焦虑程度等信息。时域声学特征是指语音信号在时间上的变化,通常包括零交叉率、过零率、短时能量等。频域声学特征是指语音信号在频率上的变化,通常包括梅尔频率倒谱系数(MFCC)、线性预测编码(LPC)等。

#### 4.3.1. 基于时域的特征

基于时域的特征[14]是语音信号处理中的一类重要特征,用于描述语音信号在时间上的变化特性。其中包括短时能量,反映了语音信号在短时间内的能量变化,过零率,用于估计语音信号的频率特性,短

时自相关函数和自相关系数，帮助检测语音信号的周期性特征，以及声道特性，反映了语音信号的共振特性。

#### 4.3.2. 基于频域的特征

语音情感处理中的频域的特征[15]可用于捕捉语音信号在频率上的变化特性。其中最常见的特征是梅尔频谱系数(MFCC) [16]，它通过将语音信号转换到梅尔频率刻度上，并提取其倒谱系数，反映了语音信号的音调、音色和共振等信息。

#### 4.3.3. 基于声学参数的特征

基于声学参数的特征对于语音信号的情感表达和说话者识别等任务具有重要意义。例如，不同情感状态可能导致声带振动频率的变化，从而影响基频特征。不同说话者的声道长度和形状差异较大，因此声道特征可以用于说话者识别。在语音情感识别中，基于声学参数的特征能够捕捉到语音信号中细微的声音变化，帮助识别不同情感状态的特征。综合运用基于声学参数的特征和其他类型的特征，可以提高语音识别和情感识别的准确性和稳定性。

### 5. 语音识别在情感化产品设计中的应用现状

在当今竞争激烈的市场中，产品的功能和性能已经不再是唯一的卖点。消费者越来越注重产品与自身情感的连接，他们渴望与产品产生情感共鸣，寻找能够满足其情感需求的体验。而在这一需求背后，情感化产品与语音识别技术的融合正成为引领产品设计和市场营销的重要趋势。随着人工智能技术的迅速发展，语音识别技术逐渐成为与产品交互的重要方式之一。人们可以通过语音与产品进行沟通和操作，这使得产品不再仅仅是静态的工具，而更像是能够理解和回应用户情感的互动伙伴[17]。在这个背景下，情感化产品设计不仅关注外观和功能，还需要通过语音识别技术传递情感，使产品与用户之间建立更加深入的情感纽带。情感化产品设计通过语音识别技术可以实现更加智能和个性化的交互体验。例如，一款智能音箱不仅能够回答用户的问题，还能够根据用户的语调和情感状态调整回应，让用户感受到与产品的情感互动。语音识别技术还可以在情感化产品设计中加入更丰富的情感元素。产品可以通过语音合成技术模拟不同情感状态下的语调和语速，从而传达愉悦、温馨等情感。此外，产品可以根据用户的语音情感，调整界面的颜色、声音效果等，创造出与用户情感匹配的交互体验。

情感化产品与语音识别技术的融合也面临一些挑战。首先是技术的稳定性和准确性。语音识别技术需要能够准确地理解用户的语音，包括语速、口音等变化，以确保情感传递的准确性。其次是隐私和安全问题。情感化产品通过语音识别技术获取用户情感信息，需要保证用户数据的隐私和安全，避免泄露和滥用。情感化产品与语音识别的结合为产品设计带来了新的可能性和挑战。通过智能化的语音交互，产品能够更好地理解和回应用户情感，创造出与用户更紧密的情感连接。然而，需要注意技术和隐私等方面的问题，以确保用户体验的同时保护用户权益。随着技术的进一步发展，情感化产品与语音识别技术的融合将会在未来持续演进，为用户带来更加丰富和深入的情感体验。

### 6. 结语

随着技术的不断发展，情感化产品设计将更加多样化和个性化。人工智能技术能够分析用户情感，根据其喜好和情感状态调整产品的展示和互动方式。虚拟现实和增强现实技术也能够创造更加身临其境的情感体验。情感化产品设计已经成为品牌成功的重要因素之一。通过在产品中融入情感元素，创造积极的用户体验和情感连接，企业能够赢得消费者的心，并在激烈的市场竞争中取得优势地位。未来，情感化产品设计将继续发展，为用户带来更加丰富的情感体验。

技术方面，情感的主观性和多样性[18]使得构建普适性和准确性较高的情感识别模型成为挑战。数据不平衡、模型的泛化能力以及多模态融合等问题也需要解决。另一方面，随着深度学习技术的不断发展，可以期待更加高效和准确的模型用于语音情感识别。通过引入注意力机制、迁移学习和自监督学习等方法，进一步提升模型性能。其次，小样本学习和迁移学习的方法可以帮助解决数据稀缺的问题，使模型在少量标注数据的情况下实现较好的情感识别效果。此外，跨语言和跨文化情感识别的研究[19]将提高模型的普适性和泛化能力，为不同语言和文化背景下的情感识别提供更好的支持。虽然当前的语音情感识别面临一些挑战，但随着技术的不断进步和应用场景的拓展，可以期待更加准确、普适和可靠的情感识别模型的出现。这将为智能客服、市场调研、心理健康辅助、教育和培训，以及虚拟现实和游戏等领域带来更多应用和突破。通过克服当前的挑战并利用新的技术和方法，可以进一步提升语音情感识别的性能和应用范围，为人们提供更加智能、个性化和情感化的服务和体验。

## 注 释

文中所有图片均为作者自绘。

## 参考文献

- [1] Fishwick, M. (2004) Emotional Design: Why We Love (or Hate) Everyday Things. *The Journal of American Culture*, **27**, 234. [https://doi.org/10.1111/j.1537-4726.2004.133\\_10.x](https://doi.org/10.1111/j.1537-4726.2004.133_10.x)
- [2] Hassenzahl, M. and Tractinsky, N. (2006) User Experience—A Research Agenda. *Behaviour & Information Technology*, **25**, 91-97. <https://doi.org/10.1080/01449290500330331>
- [3] Forlizzi, J. and Ford, S. (2000) The Building Blocks of Experience: An Early Framework for Interaction Designers. *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, New York, 17-19 August 2000, 419-423. <https://doi.org/10.1145/347642.347800>
- [4] Jordan, P.W. (2000) Designing Pleasurable Products. CRC Press, London. <https://doi.org/10.4324/9780203305683>
- [5] Hassenzahl, M., Diefenbach, S. and Göritz, A. (2010) Needs, Affect, and Interactive Products—Facets of User Experience. *Interacting with Computers*, **22**, 353-362. <https://doi.org/10.1016/j.intcom.2010.04.002>
- [6] Martinludwig, H. (2013) Design Thinking ... oder was wir von Apple & Co. lernen können. Brill | Fink, 2013: 13-31. <https://www.th-owl.de/elsa/record/7737>
- [7] Hamamura, T. and Chan, C.S. (2020) Anxious? Just Google It: Social Ecological Factors of Internet Search Records on Anxiety. *Emotion*, **20**, 1475-1484.
- [8] Zhu, Y. (2019) Study on the Strategy Patterns of WeChat Business Advertising in Undergraduates' WeChat Moments. *Proceedings of the 2nd International Symposium on Social Science and Management Innovation (SSMI 2019)*, Xi'an, 29-30 November 2019, 27-31. <https://doi.org/10.2991/ssmi-19.2019.27>
- [9] Konangi, U.M.Y., Katredy, V.R., Rasula, S.K., Marisa, G. and Thakur, T. (2022) Emotion Recognition through Speech: A Review. *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, Salem, 9-11 May 2022, 1150-1153.
- [10] 高利军, 薛雷. 语音情感识别综述[J]. 工业控制计算机, 2022, 35(10): 115-116, 120.
- [11] Ekman, P. and Friesen, W.V. (1971) Constants across Cultures in the Face and Emotion. *Journal of personality and social psychology*, **17**, 124-129. <https://doi.org/10.1037/h0030377>
- [12] Russell, J.A. (1980) A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, **39**, 1161-1178. <https://doi.org/10.1037/h0077714>
- [13] Abdel-Hamid, L. (2020) Egyptian Arabic Speech Emotion Recognition Using Prosodic, Spectral and Wavelet Features. *Speech Communication*, **122**, 19-30. <https://doi.org/10.1016/j.specom.2020.04.005>
- [14] Xiang, S.J., Huang, J.W. and Yang, R. (2007) International Workshop on Information Hiding. Springer, Berlin, 93-108.
- [15] Ghai, M., Lal, S., Duggal, S. and Manik, S. (2017) Emotion Recognition on Speech Signals Using Machine Learning. *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)*, Andhra Pradesh, 23-25 March 2017, 34-39. <https://doi.org/10.1109/ICBDACI.2017.8070805>
- [16] Dahake, P.P., Shaw, K. and Malathi, P. (2016) Speaker Dependent Speech Emotion Recognition Using MFCC and Support Vector Machine. *2016 International Conference on Automatic Control and Dynamic Optimization Techniques*

- (ICACDOT), Pune, 9-10 September 2016, 1080-1084. <https://doi.org/10.1109/ICACDOT.2016.7877753>
- [17] 黄孟卉. 情感化设计在智能家居设计中的应用[J]. 黑河学院学报, 2023, 14(7): 137-139, 148.
- [18] Zhou, S.J., Huang, D.M., Liu, C. and Jiang, D.Z. (2022) Objectivity Meets Subjectivity: A Subjective and Objective Feature Fused Neural Network for Emotion Recognition. *Applied Soft Computing*, **122**, Article ID: 108889. <https://doi.org/10.1016/j.asoc.2022.108889>
- [19] Desplanques, B. and Demuynck, K. (2018) Cross-Lingual Speech Emotion Recognition through Factor Analysis. *Interspeech* 2018, Hyderabad, 2-6 September 2018, 3648-3652. <https://doi.org/10.21437/Interspeech.2018-1778>