**1、介绍**

**音乐元创造(MuMe)**

生成模型与创作模型(**人文主义要素**)

**行为功能学**、**认知科学**、**相关的理论结构**和**神经科学**

**创作=生成+反射+评价**

**系统的反馈**

**评估创造性系统（内部、外部）**

**2、创新理论视角**

小结：

Boden三种**创造性**：**组合创意、探索性创造力、转换创造**

CSF**创作系统框架，描述和比较创新系统**

评估表述为四个参数的函数：**人工制品、创造者、受众和上下文。**

**3、语境化评价：MuMe和音乐CC系统的兴趣领域**

共同目标：

**旋律、和声和节奏内容的生成**；在观众中产生**情感反应**；以及在创作过程中考虑到**表演者或观众的互动系统**。

**1、旋律、和声：**

风格仿制品

**音乐图灵测试**

**区分计算机生成与人类生成**

**超越音乐学和文体学，将目光投向支持听力和音乐认知的心理学理论**

**2、节奏生成**

**Meter和Rhythm**是非常**开放**的问题

**要使mume的发展超越这些有限的表达手段，就需要更多的理论发展**

**3、在听众中创造情感反应**

**作曲家的目标是让听众自己的大脑状态与自己的大脑状态保持一致**

**在人类听众中实现特定的情感反应对人类专家作曲家来说是一个挑战**

**在情感反应方面，本文着重探讨了如何利用Mume系统不仅可以唤起听众的某些情绪，而且可以利用听众的主观和生理反应作为对生成系统的反馈。**

**4、交互系统：包括作曲家、表演者和观众反馈**

**人类互动者**的影响：

**直接参与、仪器控制、明确反馈、内隐反馈**

**“情感表现框架”**，它提供了一个**听众态度模型**和一种**用情感元数据注释作品**的方法

**在明确的计算创造环境中，似乎很少有工作能真正与计算机共享创造责任。**

**4、外部方法：通过测试观众的感知、认知和情感反应来评估创作过程和人工制品**

我们的目标**不一定**是提倡**类人系统**（或指定任何特定的mume系统设置），而是研究人类的**感知、认知和情感反应**，作为评估系统创造性输出影响的一种手段。

**在系统内提供反馈**

**4.1 行为测试与人类创造力评估**

**创造力的行为测试**

**发散思维、收敛思维和艺术能力或依靠自我评估**

可以**应用于计算创新系统的开发或改进**

托伦斯创造性思维测试

**流畅性、灵活性、独创性和精化 =>**

**流畅性、独创性、精雕细琢、标题抽象性和抗过早闭合**

这些测试中的大多数都是检查**个人对可能的解决方案的阐述，测量细节、数量、新颖性和各种想法**。

**聚合思维任务:** 需要创造性思维来重新评估一个问题，确定一个有洞察力的解决问题的观点，或者改变策略来寻求正确的解决方案。

**4.2 共识评估技术**

**共识评估技术**（CAT）

**相关领域的专家小组对艺术品、理论、产品或性能的创造力进行判断**。

合理的**可靠性和可重复性**

**不考虑创造艺术品产生的过程**

图灵测试评估方法

在一个任务中，所有健康的人都会有很高的表现水平：一般对话。无论出于什么原因，音乐创作都不是这样一项任务；**这使得大多数玩家没有能力对“游戏”做出明智的判断**。第二，在图灵的模拟游戏中，**隐藏的人类试图欺骗玩家做出错误的选择**；

因为它只专注于一个**创造性的系统**，它没有解决**欺骗的企图**。

需要**定性的反馈**，可以用来增强所讨论的创造性系统

**4.3 计算创造力的扩展**

第4.1节中列出的一些行为方法和策略已被用于评估计算创造力,

**有助于消除判据对人工系统生成的人工制品的偏见**

**类似于人类创造力评估中使用的评估指标的标准:**

**新颖性**（不典型性、创新性或阶级成员）和**质量**（价值等级）

**关注生成的人工制品**，而不是**评估创建该人工制品的内部过程**

**创造性三脚架**

**描述和评估创造性系统行为**的一种技术提供，它是基于**评估技能**（技术能力）、**欣赏**（在领域中有价值）和**想象**（超越了过去的适当新颖性）

**既关注生成的人工制品，也关注系统的创造性行为**

**创造性过程的知识**不仅影响观察者对**创造性的判断、人类创造力的价值判断**，而且还影响**计算系统的上下文**

如果这个系统被认为是**熟练**的、**有鉴赏力**的和**富有想象力**的，那么这个软件应该被认为是有创造力的。

一个模型侧重于**创造性生成**的行为，称为FACE（**框架、美学、概念、表达**）(Framing, Aesthetic, Concept, Expression)

另一个**测试创造性系统可能对观察者产生的影响**，称为IDEA（**迭代开发-执行-欣赏**）(Iterative Development-Execution-Appreciation)

FACE框架可以用来**描述mume模型中旋律/和声产生和节奏产生的基本过程**，IDEA方法可以用来**融合观众的背景知识和偏好**。

**4.4 问卷、相关研究和评分量表**

检验听者**对系统或系统输出的主观反应**的一种有价值的方法

（1）解决特定的感兴趣的研究问题

（2）基于上面的行为和心理测量测试

（3）勾勒出一套创意系统的关键组成部分

使用**离散响应标准**（例如，要求听者选择多个选项中的一个，如选择多个问题）或**评分量表**进行调查

尽管相关研究**不能提供因果关系的证据**，但它们可以成为一个有价值的指标，并**激发能够检验因果关系的实证研究**。

CAT还采用了**可靠性**（评分者之间一致性）和**有效性**的衡量标准，以确保调查结果的**可靠性和可复制性**。

**4.5 测量运动和生理反应**

生理测量可以用来捕捉心理和情绪状态的物理表现

**这些技术已被用于评估感知的紧张、刺激强度或在一系列领域中的诱发降级，并且最近也被应用于测量音乐聆听过程中观众的生理和情感反应**

**4.5.1 情感捕捉**

动作捕捉是另一种**测量情绪、唤醒和具体认知状态**实时指标的方法

**4.5.2 视线追踪**

眼睛跟踪是另一种间接测量**在线、实时注意、感知和认知方面**的技术。眼睛跟踪可以作为**注意力、信息处理和决策以及探索性行为（通过测量图像或视觉场景周围的步态）**的一种度量。

**不需要公开的响应**

评估**视觉注意力和观众参与度**。

**4.5.3 反应时间**

反应时间（RT）可作为**感知或认知过程的间接测量**，而不是要求参与者直接做出判断或评级

**当一个节奏或旋律模式改变时，可能会要求一个听者作出反应**

**4.6 用脑电图测量神经反应**

**听者的神经活动**纳入他们的系统或表现

ERP组件可用于**评估受众或合作者的期望机制和语义处理**

**检查大脑皮层广袤区域的时空活动**

**创造性与非创造性任务之间检查振荡带活动**

**振荡活动可以在实时表演期间与音乐家或共同创作者一起使用或捕捉观众的情感状态并将其纳入音乐创作**

**4.7 特殊方法的好处和局限性**

外部方法的选择可能取决于几个因素，包括研究者对创新系统的**目标、系统的架构，甚至预算和时间限制**。

**调查问卷数据**对于mume研究人员来说是一种有吸引力的资源，因为这些数据相对**容易获取和分析**，而且**不需要昂贵的专用设备来收集数据**。**要求观察者对他们的观点进行评级的行为实际上会改变观察者的观点**。从计算创意中提取的方法，如**创意三脚架**，也可能特别有用，因为这些方法是专门为测试计算创意工艺品而设计的。

**通过感知和认知的间接测量来评估创造力**，例如**眼睛跟踪和反应时间测量**。必须使用**专用设备以及用于数据分析的专用软件。**

**5、内部评价：为自我反省建立人类创造行为和观众感知、认知和情感状态的模型**

**人类感知、认知和情感状态的建模方法**及其相关的评估方法可以为开发具有自我评估能力的更复杂的创造性系统提供信息。**自我评价可能被视为创造性行为的必要组成部分**。**自我评价可以形成反思的基础。**

**自我反思因子**(Self-reflective agents)

数据本身不足以进行反射：**系统必须具有对数据进行推理的方法**

人类感知和认知建模的突出技术，这些技术可用于**使创造性系统将受众的知识纳入他们的推理和元推理过程**

**5.1 基于行为测试的自我评估**

**发散性和收敛性思维测试**也可以作为自我反思的内部测试

自我评估测试提供了另一种从人类行为测试中获得自我反思的清晰方法。**这些测试要求测试者反思自己的创造力，或报告个人经历信息，例如参与者参与艺术追求（例如参加视觉艺术课或写短篇小说的经验）**

**准备、孵化、启发和验证（preparation, incubation, illumination, veriﬁcation）**

**5.2 基于观众感知、认知和情感状态外部测试的反射模型**

**感知和生理外部评价方法的数据**作为**内部重建的基础**

如果一个系统试图在听者中调用某种**特定类型的情感反应**，那么它应该拥有某种**人类情感反应的内部模型**（例如，**基于Valence和Arousal的二维模型**）

如果系统被赋予了一种**预测听者可能如何响应**的方法，那么它可以将听者的**实际响应与其预测进行比较**。

与**表现后反思(post-performance reﬂection)**不同，生理反应的实时测量可用于**持续、在线自我评估和反思**。创意因子可以使用**眼睛跟踪测量**、**反应时间数据**、**心率数据和连续脑电图**来评估系统的预期目标是否已实现，以及是否（如何）**更新创意过程**以支持未来的预期结果。

**5.3 感知和情感反应模型的预测和期望度量**

**因为期望对于学习是至关重要的**

***听演讲时，人们不会被动地“听”单词；相反，听者会对下一步会发生什么做出含蓄的预测。***

**正确的预测是一个有价值的进化特征，因为我们不能总是依靠直接的经验来形成预测**

**将听者期望模型纳入生成过程可能对mume系统的表现力和创造力产生强大的影响**

**非常可预测的刺激被认为是无聊的，而极其复杂或强烈的刺激被认为是不可接近或过度刺激的。**因此，曲线的中心产生**最佳复杂性**的**“最佳点”**，这通常是刺激复杂性/强度的函数，也是观察者的背景和经验。

随着时间的推移，整个序列的信息理论特性对感知和记忆产生了**动态影响**，复杂的刺激通常会增加**识别记忆性能**差的影响。

**即为什么人类喜欢有创造力**，即使这样做对他们没有任何生物学上的好处。

**预测有助于大脑学习和编码一个领域的信息，而（违背）期望则与美学和影响有关。**

**心理表征或概念空间**的网络可以被认为是一个复杂的先验分布，具有**统计上定义的共现性和相关特征**。

探索性创造力涉及到根**据已知的概率分布或概念表示生成新的想法**，而对创造艺术品（价值判断）的偏好**通常属于熟悉、新颖或可预测和复杂的首选范围内。**

**5.4 概念表示**

**概念之间的相似性是通过多维空间中点或区域之间的距离来隐式表示的**

在实际的音乐术语中，潜在的任务可以包括构建高层次概念的概念空间表示，如**类型或情绪**

**重大挑战：构建这种具有几何意义的空间**；其次，**建立它们的感知有效性**。

**概念空间理论可以与已建立的建模技术（如统计模型或深层神经网络）相结合**

将**观众或共同创作者的概念空间模型结合**起来，可以为mume系统提供一种**自我反思的方法和思想理论**

**人类为中心**的创造力方法，但因为音乐（和音乐创造力）是一种人类构造，**人类对音乐概念化的方式可以为人工智能和机器创造力提供重要的见解。**

**将系统的探索行为限定到更高价值的空间区域， 概念空间的探索受到质量维度的限制**，如情感、兴趣和紧张

**6、总结和结论**

本文概述了可用于或适用于**计算创造力**研究的**理论和经验方法**，特别是针对**mume群体**。

从***科学的角度***，我们讨论了一系列**提供客观评价创造力的方法**。这种形式的评估称为**外部评估**，*因为判断或测量的来源是来自系统本身的外部*。

***创造性系统本身****及其****内部自我评价能力***的角度来考虑评价。

**创造力和计算创造力**理论的联系

如何**对创造性人工制品和产生它们的过程作出判断**，从根本上与**系统设计问题**有关。

讨论范围限定在Mume感兴趣的四个主要领域，包**括旋律和和声的产生**、**在听众中产生情感反应的节奏的产生，以及结合表演者和观众反馈**。

本文提倡**科学的方法**研究创造力和发展的Mume系统。**精确的评估方法，连同明确陈述的假设**，是任何科学或科学上一致的学科进步的基本基石。不仅为个人研究人员提供了好处，也为整个科研环境提供了好处。

创造性评价的主题更普遍地为更先进的创造性系统的发展提供了见解，在这种系统中，**评价能力**成为一个基本组成部分。