



(12) 发明专利申请

(10) 申请公布号 CN 116847123 A

(43) 申请公布日 2023. 10. 03

(21) 申请号 202310958973.5

H04N 21/2387 (2011.01)

(22) 申请日 2023.08.01

H04N 21/44 (2011.01)

H04N 21/439 (2011.01)

(71) 申请人 南拳互娱(武汉)文化传媒有限公司

地址 430000 湖北省武汉市东湖开发区珞  
狮南路517号明泽大厦3楼B2、B3区

(72) 发明人 余骐楠 潘安 陈思远 李倩

汪琥 张梦晗 孔玺铉 温泉  
罗昕美

(74) 专利代理机构 北京曼京知识产权代理事务  
所(普通合伙) 11965

专利代理师 尹昌浩

(51) Int. Cl.

H04N 21/233 (2011.01)

H04N 21/234 (2011.01)

H04N 21/2368 (2011.01)

权利要求书2页 说明书6页 附图1页

(54) 发明名称

一种视频后期剪辑与视频合成优化方法

(57) 摘要

本发明涉及影视传媒技术领域,具体涉及一种视频后期剪辑与视频合成优化方法,通过内容分析和特征提取,自动识别关键场景和特征,并提取视觉特征、语义信息和情感内容,借助语义关联和场景划分,将视频分割为不同场景,并从音频、字幕等提取语义信息,实现更准确的场景划分,智能剪辑和剧情推进自动拆分视频为段落或镜头,并保持视觉流畅性和情感连贯性,动态音视频合成确保音频与视频的节奏、情感和情绪相匹配,提供更具感染力的观看体验。智能特效与过渡根据场景和情感特征自动生成合适的视觉特效,增强视频的视觉吸引力,多模态数据融合将视频中的多模态数据融合起来进行联合优化,提高视频质量。

1. 一种视频后期剪辑与视频合成优化方法,其特征在于,具体包括以下步骤:

S1、对原始视频进行内容分析,自动识别关键场景和特征,并从视频中提取视觉特征、语义信息和情感内容;

S2、基于视频内容分析结果,将视频分割为不同的场景或话题;

S3、根据语义关联和场景划分结果,智能剪辑算法自动拆分视频为段落或镜头,并根据剧情需要进行顺序调整;

S4、将视频剪辑后的不同段落或镜头与动态音频进行智能合成,采用音频分析技术,确保音频与视频的节奏、情感和情绪相匹配;

S5、引入智能特效生成算法,根据场景和情感特征自动生成合适的视觉特效;

S6、将视频中的多模态数据融合起来进行联合优化;

S7、引入用户交互机制,允许用户根据自己的喜好和需求定制视频剪辑风格;

S8、结合增强现实技术,将虚拟场景融合到视频中;

S9、引入自动质量评估算法,对生成的优化视频进行评估;

S10、采用分布式计算技术,加速视频剪辑与合成过程。

2. 根据权利要求1所述的一种视频后期剪辑与视频合成优化方法,其特征在于,所述智能剪辑与剧情推进步骤进一步包括:

a. 根据语义关联和场景划分结果,智能剪辑算法根据视频的内容进行镜头拆分,并自动选择最佳镜头顺序。

b. 在进行镜头拆分和顺序调整的过程中,智能剪辑算法确保镜头之间的视觉过渡平滑。

3. 根据权利要求1所述的一种视频后期剪辑与视频合成优化方法,其特征在于,所述动态音视频合成步骤进一步包括:

a. 通过音频分析技术,智能音频合成算法根据视频内容的节奏、情感和情绪,自动选择合适的音频素材进行智能合成。

b. 智能音频合成算法根据视频的节奏和情感动态调整音频素材的混合比例和音量。

4. 根据权利要求1所述的一种视频后期剪辑与视频合成优化方法,其特征在于,所述智能特效与过渡步骤进一步包括:

a. 引入智能特效生成算法,根据场景和情感特征自动生成合适的视觉特效。

b. 智能特效生成算法根据视频内容的情感变化和视觉效果需求,动态调整特效的程度和持续时间。

5. 根据权利要求1所述的一种视频后期剪辑与视频合成优化方法,其特征在于,所述多模态数据融合步骤进一步包括:

a. 利用跨模态关联技术,将视频中的视觉特征、音频特征和文本特征多模态数据进行融合,并形成综合特征表示。

b. 在视频合成过程中,智能算法根据综合特征表示进行优化调整。

6. 根据权利要求1所述的一种视频后期剪辑与视频合成优化方法,其特征在于,所述用户交互与个性化步骤进一步包括:

a. 用户交互模块提供界面供用户选择个性化选项。

b. 根据用户选择的个性化选项,智能算法对视频后期剪辑与合成参数进行自动调整。

7. 根据权利要求1所述的一种视频后期剪辑与视频合成优化方法,其特征在于,所述增强现实与虚拟场景融合步骤进一步包括:

- a. 引入增强现实技术,将虚拟场景与实际视频内容进行融合。
- b. 智能算法根据视频内容和虚拟场景的特征,动态调整增强现实效果。

8. 根据权利要求1所述的一种视频后期剪辑与视频合成优化方法,其特征在于,所述自动质量评估与优化步骤进一步包括:

- a. 引入自动质量评估算法,对生成的优化视频进行评估。
- b. 根据自动质量评估的结果,智能算法自动调整视频后期处理参数。

9. 根据权利要求1所述的一种视频后期剪辑与视频合成优化方法,其特征在于,所述分布式计算与快速处理步骤进一步包括:

- a. 利用分布式计算技术,将视频后期处理任务分解为多个子任务,并在多个计算节点上并行处理,加速视频剪辑与合成过程。
- b. 通过快速处理技术,提高视频后期处理的响应速度和处理效率。

## 一种视频后期剪辑与视频合成优化方法

### 技术领域

[0001] 本发明涉及影视传媒技术领域,尤其涉及一种视频后期剪辑与视频合成优化方法。

### 背景技术

[0002] 随着影视行业的发展以及自媒体的兴起,影视剪辑逐渐成为一项重要的技术服务于制片工作中,但是传统的影视剪辑流程通常采用人工剪辑为主,软件剪辑为辅的剪辑方式,对于体量较大或数量较多的影片在同期剪辑的过程中剪辑效率较低,同时最终成片的剪辑质量也参差不齐,针对这一问题,自动化影视剪辑技术应运而生,但是现有的自动剪辑技术对于特定的视频特征还是需要人为进行特征的干预与选择,无法在选定待剪辑特征后,自动进行特征视频的摘录与选择,从根本上并未进一步实现视频的自动化剪辑处理。

[0003] 经检索CN202211150109.4公开了及一种视频后期剪辑与视频合成优化方法,具体包括,S1、将原始视频文件进行视频关键帧拆分,S2、对拆分后的视频关键帧建立动态视频特征提取,S3、将提取的动态视频片段进行视频合成;其中所述的视频合成中建立了动态视频片段自动拼接优化。

[0004] 但是经本发明人探索发现该技术方案仍然存在至少以下缺陷:

[0005] 第一、视频后期剪辑与合成技术虽然能够进行基本的剪辑和特效处理,但在智能化处理方面仍有限,缺乏全面的视频内容分析和语义理解能力,导致剪辑和合成结果可能缺乏情感连贯性和创意。第二、现有技术为用户交互与个性化方面表现不足,缺乏灵活的个性化选项,用户无法根据个人喜好和需求定制视频剪辑风格,限制了观看体验的个性化和定制化。第三、现有技术在增强现实与虚拟场景融合方面还有待改进,在将虚拟场景与实际视频内容无缝融合的能力上仍存在一定局限性,限制了视频内容创意和视觉体验的创新性

### 发明内容

[0006] 有鉴于此,本发明的目的在于提出一种视频后期剪辑与视频合成优化方法解决上述的问题。

[0007] 基于上述目的,本发明提供了一种视频后期剪辑与视频合成优化方法。

[0008] 一种视频后期剪辑与视频合成优化方法,具体包括以下步骤:

[0009] S1.对原始视频进行内容分析,自动识别关键场景和特征,并从视频中提取视觉特征、语义信息和情感内容。

[0010] S2.基于视频内容分析结果,将视频分割为不同的场景或话题。利用自然语言处理技术,从视频的音频、字幕或音频转录中提取语义信息,并与视觉特征关联,实现更准确的场景划分。

[0011] S3.根据语义关联和场景划分结果,智能剪辑算法自动拆分视频为段落或镜头,并根据剧情需要进行顺序调整。在剪辑过程中,保持视觉流畅性和情感连贯性,以推进视频的剧情发展。

- [0012] S4. 将视频剪辑后的不同段落或镜头与动态音频进行智能合成。采用音频分析技术,确保音频与视频的节奏、情感和情绪相匹配,提供更具感染力的观看体验。
- [0013] S5. 引入智能特效生成算法,根据场景和情感特征自动生成合适的视觉特效,如转场、滤镜、色彩调整,以增强视频的视觉吸引力。
- [0014] S6. 将视频中的多模态数据(视觉、音频、文本)融合起来进行联合优化。通过跨模态关联,实现更准确的特征提取和合成,提高视频质量。
- [0015] S7. 引入用户交互机制,允许用户根据自己的喜好和需求定制视频剪辑风格。个性化选项可包括音乐选择、特效定制、剧情推进速度,以满足不同用户的需求。
- [0016] S8. 结合增强现实技术,将虚拟场景融合到视频中,实现更加丰富和创新的视觉体验。
- [0017] S9. 引入自动质量评估算法,对生成的优化视频进行评估,并根据评估结果自动调整剪辑和合成参数,不断优化视频质量。
- [0018] S10. 采用分布式计算技术,加速视频剪辑与合成过程,保证高效率 and 快速响应。
- [0019] 进一步的,智能剪辑与剧情推进步骤进一步包括:
- [0020] a. 根据语义关联和场景划分结果,智能剪辑算法根据视频的内容进行镜头拆分,并自动选择最佳镜头顺序,以推进视频剧情的发展。
- [0021] b. 在进行镜头拆分和顺序调整的过程中,智能剪辑算法确保镜头之间的视觉过渡平滑,以保持观看的连贯性和流畅性。
- [0022] 进一步的,动态音视频合成步骤进一步包括:
- [0023] a. 通过音频分析技术,智能音频合成算法根据视频内容的节奏、情感和情绪,自动选择合适的音频素材进行智能合成。
- [0024] b. 智能音频合成算法根据视频的节奏和情感动态调整音频素材的混合比例和音量,以确保音频与视频内容的高度匹配。
- [0025] 进一步的,智能特效与过渡步骤进一步包括:
- [0026] a. 引入智能特效生成算法,根据场景和情感特征自动生成合适的视觉特效,如转场、滤镜、色彩调整。
- [0027] b. 智能特效生成算法根据视频内容的情感变化和视觉效果需求,动态调整特效的程度和持续时间,以增强视频的视觉吸引力。
- [0028] 进一步的,多模态数据融合步骤进一步包括:
- [0029] a. 利用跨模态关联技术,将视频中的视觉特征、音频特征和文本特征多模态数据进行融合,并形成综合特征表示。
- [0030] b. 在视频合成过程中,智能算法根据综合特征表示进行优化调整,以提高视频质量和观看体验。
- [0031] 进一步的,用户交互与个性化步骤进一步包括:
- [0032] a. 用户交互模块提供界面供用户选择个性化选项,包括音乐选择、特效定制、剧情推进速度。
- [0033] b. 根据用户选择的个性化选项,智能算法对视频后期剪辑与合成参数进行自动调整,以满足用户的个性化需求。
- [0034] 进一步的,增强现实与虚拟场景融合步骤进一步包括:

[0035] a. 引入增强现实技术,将虚拟场景与实际视频内容进行融合,创造更加丰富和创新的视觉体验。

[0036] b. 智能算法根据视频内容和虚拟场景的特征,动态调整增强现实效果,以提供逼真和吸引人的视觉体验。

[0037] 进一步的,自动质量评估与优化步骤进一步包括:

[0038] a. 引入自动质量评估算法,对生成的优化视频进行评估,包括视频流畅性、视觉吸引力、音频质量方面。

[0039] b. 根据自动质量评估的结果,智能算法自动调整视频后期处理参数,不断优化视频质量。

[0040] 进一步的,分布式计算与快速处理步骤进一步包括:

[0041] a. 利用分布式计算技术,将视频后期处理任务分解为多个子任务,并在多个计算节点上并行处理,加速视频剪辑与合成过程。

[0042] b. 通过快速处理技术,提高视频后期处理的响应速度和处理效率,保证用户体验。

[0043] 本发明的有益效果:

[0044] 1. 通过视频内容分析与特征提取、语义关联与场景划分等技术,该方法能够自动识别关键场景和特征,实现智能剪辑和合成。这将保证视频剧情的连贯性和情感连贯性,同时利用智能特效和音频合成,增强视频的视觉吸引力和感染力,从而显著提升视频质量。

[0045] 2. 通过引入用户交互与个性化模块,用户可以根据个人喜好和需求定制视频剪辑风格,包括音乐选择、特效定制和剧情推进速度等。这样的个性化选项将满足不同用户的需求,提供更具定制化和个性化的观看体验。

[0046] 3. 结合增强现实技术,将虚拟场景融合到实际视频中,为观众创造更加丰富、创新和逼真的视觉体验。这将带来全新的观看感受,为视频内容的创意和表现提供更大的空间,使观众沉浸于影视作品的世界之中。

## 附图说明

[0047] 为了更清楚地说明本发明或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0048] 图1为本发明实施例的运行逻辑结构示意图;。

## 具体实施方式

[0049] 为使本发明的目的、技术方案和优点更加清楚明白,以下结合具体实施例,对本发明进一步详细说明。

[0050] 如图1所示,一种视频后期剪辑与视频合成优化方法,具体包括以下步骤:

[0051] S1. 对原始视频进行内容分析,自动识别关键场景和特征,并从视频中提取视觉特征、语义信息和情感内容。

[0052] S2. 基于视频内容分析结果,将视频分割为不同的场景或话题。利用自然语言处理技术,从视频的音频、字幕或音频转录中提取语义信息,并与视觉特征关联,实现更准确的

场景划分。

[0053] S3. 根据语义关联和场景划分结果,智能剪辑算法自动拆分视频为段落或镜头,并根据剧情需要进行顺序调整。在剪辑过程中,保持视觉流畅性和情感连贯性,以推进视频的剧情发展。

[0054] S4. 将视频剪辑后的不同段落或镜头与动态音频进行智能合成。采用音频分析技术,确保音频与视频的节奏、情感和情绪相匹配,提供更具感染力的观看体验。

[0055] S5. 引入智能特效生成算法,根据场景和情感特征自动生成合适的视觉特效,如转场、滤镜、色彩调整,以增强视频的视觉吸引力。

[0056] S6. 将视频中的多模态数据(视觉、音频、文本)融合起来进行联合优化。通过跨模态关联,实现更准确的特征提取和合成,提高视频质量。

[0057] S7. 引入用户交互机制,允许用户根据自己的喜好和需求定制视频剪辑风格。个性化选项可包括音乐选择、特效定制、剧情推进速度,以满足不同用户的需求。

[0058] S8. 结合增强现实技术,将虚拟场景融合到视频中,实现更加丰富和创新的视觉体验。

[0059] S9. 引入自动质量评估算法,对生成的优化视频进行评估,并根据评估结果自动调整剪辑和合成参数,不断优化视频质量。

[0060] S10. 采用分布式计算技术,加速视频剪辑与合成过程,保证高效率 and 快速响应。

[0061] 进一步的,智能剪辑与剧情推进步骤进一步包括:

[0062] a. 根据语义关联和场景划分结果,智能剪辑算法根据视频的内容进行镜头拆分,并自动选择最佳镜头顺序,以推进视频剧情的发展。

[0063] b. 在进行镜头拆分和顺序调整的过程中,智能剪辑算法确保镜头之间的视觉过渡平滑,以保持观看的连贯性和流畅性。

[0064] 进一步的,动态音视频合成步骤进一步包括:

[0065] a. 通过音频分析技术,智能音频合成算法根据视频内容的节奏、情感和情绪,自动选择合适的音频素材进行智能合成。

[0066] b. 智能音频合成算法根据视频的节奏和情感动态调整音频素材的混合比例和音量,以确保音频与视频内容的高度匹配。

[0067] 进一步的,智能特效与过渡步骤进一步包括:

[0068] a. 引入智能特效生成算法,根据场景和情感特征自动生成合适的视觉特效,如转场、滤镜、色彩调整。

[0069] b. 智能特效生成算法根据视频内容的情感变化和视觉效果需求,动态调整特效的程度和持续时间,以增强视频的视觉吸引力。

[0070] 进一步的,多模态数据融合步骤进一步包括:

[0071] a. 利用跨模态关联技术,将视频中的视觉特征、音频特征和文本特征多模态数据进行融合,并形成综合特征表示。

[0072] b. 在视频合成过程中,智能算法根据综合特征表示进行优化调整,以提高视频质量和观看体验。

[0073] 进一步的,用户交互与个性化步骤进一步包括:

[0074] a. 用户交互模块提供界面供用户选择个性化选项,包括音乐选择、特效定制、剧情

推进速度。

[0075] b. 根据用户选择的个性化选项,智能算法对视频后期剪辑与合成参数进行自动调整,以满足用户的个性化需求。

[0076] 进一步的,增强现实与虚拟场景融合步骤进一步包括:

[0077] a. 引入增强现实技术,将虚拟场景与实际视频内容进行融合,创造更加丰富和创新的视觉体验。

[0078] b. 智能算法根据视频内容和虚拟场景的特征,动态调整增强现实效果,以提供逼真和吸引人的视觉体验。

[0079] 进一步的,自动质量评估与优化步骤进一步包括:

[0080] a. 引入自动质量评估算法,对生成的优化视频进行评估,包括视频流畅性、视觉吸引力、音频质量方面。

[0081] b. 根据自动质量评估的结果,智能算法自动调整视频后期处理参数,不断优化视频质量。

[0082] 进一步的,分布式计算与快速处理步骤进一步包括:

[0083] a. 利用分布式计算技术,将视频后期处理任务分解为多个子任务,并在多个计算节点上并行处理,加速视频剪辑与合成过程。

[0084] b. 通过快速处理技术,提高视频后期处理的响应速度和处理效率,保证用户体验。

[0085] 通过上述的视频后期剪辑与视频合成优化方法,其具体运行逻辑结构流程如下:首先,我们需要对原始视频进行内容分析。这就是让计算机“看懂”视频,自动识别出视频中的关键场景和特征,比如哪些部分是重要的,哪些是情感丰富的,哪些是需要强调的。然后,我们使用计算机视觉和深度学习技术,从视频中提取出视觉特征,语义信息和情感内容等,这些将成为后续处理的基础,有了内容分析和特征提取的结果,我们接下来将视频分割为不同的场景或话题。这样做有助于我们更好地理解视频的结构和故事情节。我们还需要利用自然语言处理技术,从视频的音频、字幕或音频转录中提取语义信息。这些语义信息会与之前的视觉特征关联起来,帮助我们更准确地划分场景和理解视频的含义,有了场景划分和语义关联的结果,我们就可以开始进行智能剪辑。智能剪辑算法会自动将视频拆分成段落或镜头,并根据剧情需要进行顺序调整。这样就能够保持视频剧情的连贯性和情感连贯性,让整个视频的故事更加流畅和引人入胜,接下来,我们需要将视频剪辑后的不同段落或镜头与动态音频进行智能合成。这就是让视频的画面和音频完美匹配,让声音与画面的情感和节奏相符。音频分析技术会帮助我们选择合适的音频素材,以确保音频与视频内容高度匹配,让观众有更加感染力的观看体验,引入智能特效生成算法,我们可以根据场景和情感特征自动生成合适的视觉特效。这些特效可以是转场、滤镜、色彩调整等,让视频更具吸引力和视觉冲击力。我们还会根据视频内容的情感变化和视觉效果需求,动态调整特效的程度和持续时间,以增强视频的观赏效果,接下来,我们将视频中的多模态数据(视觉、音频、文本等)融合起来进行联合优化。通过跨模态关联,我们能够实现更准确的特征提取和合成,进一步提高视频的质量和观看体验,我们引入用户交互机制,让用户可以根据自己的喜好和需求定制视频剪辑风格。用户可以选择自己喜欢的音乐,定制特效,调整剧情推进速度等。这样的个性化选项会满足不同用户的需求,让观看体验更具个性和定制性,我们还结合增强现实技术,将虚拟场景融合到视频中,创造更加丰富和创新的视觉体验。通过虚拟

场景的融合,视频内容可以更加生动、有趣,让观众沉浸其中,提升观看体验,引入自动质量评估算法,对生成的优化视频进行评估。通过自动评估结果,我们可以自动调整剪辑和合成参数,不断优化视频质量,让视频效果更加出色,最后,为了应对大规模视频处理,我们采用分布式计算技术,加速视频剪辑与合成过程,保证高效率和快速响应。这样能够让视频处理更加高效和流畅。

[0086] 本发明旨在涵盖落入所附权利要求的宽泛范围之内的所有这样的替换、修改和变型。因此,凡在本发明的精神和原则之内,所做的任何省略、修改、等同替换、改进等,均应包含在本发明的保护范围之内。

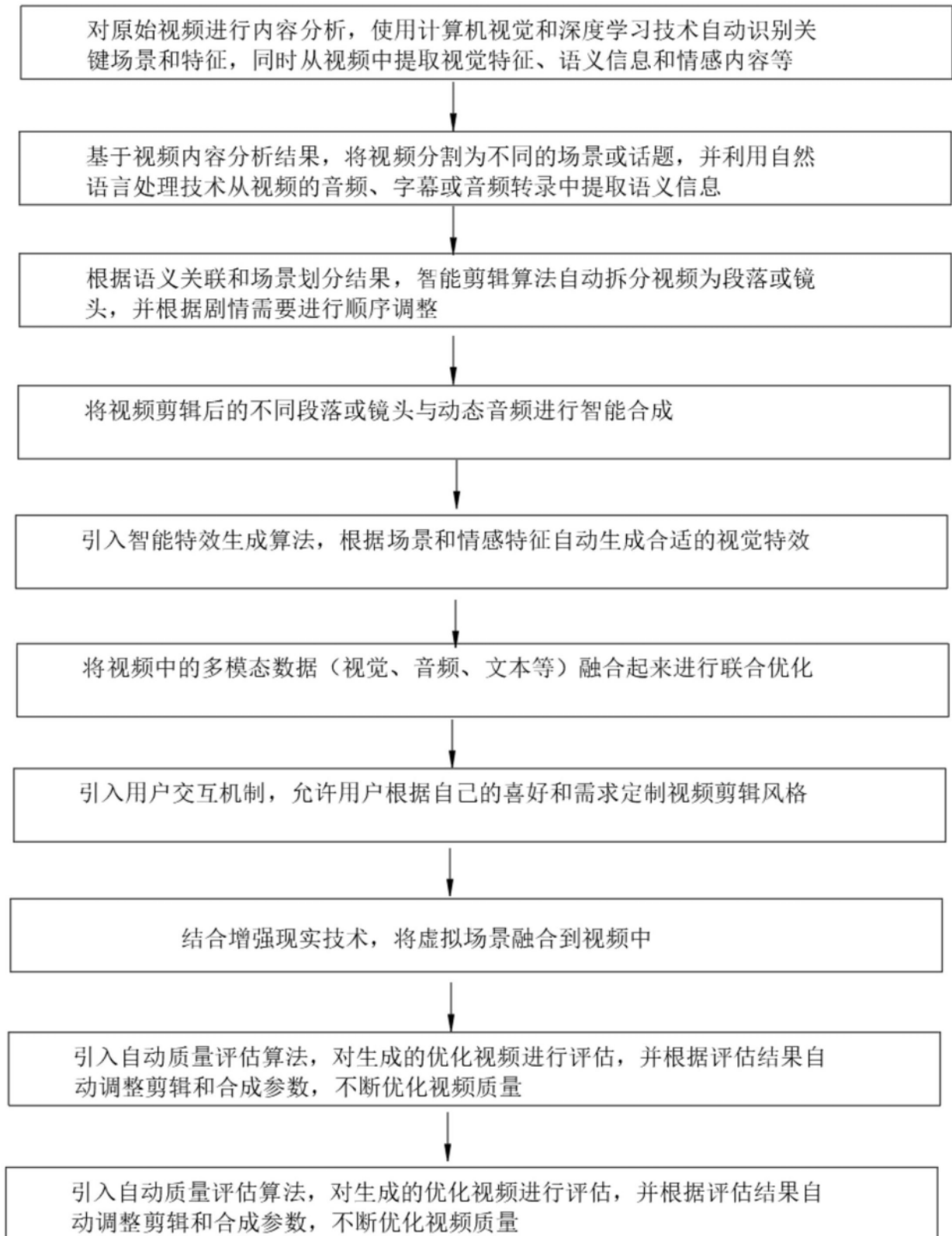


图1