

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号
WO 2024/056022 A1

(43) 国际公布日
2024 年 3 月 21 日 (21.03.2024)

(51) 国际专利分类号:
H04N 5/278 (2006.01)

(21) 国际申请号: PCT/CN2023/118772

(22) 国际申请日: 2023 年 9 月 14 日 (14.09.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202211117721.1 2022年9月14日 (14.09.2022) CN

(71) 申请人: 北京字跳网络技术有限公司
(**BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区紫金数码园4号楼2层0207, Beijing 100190 (CN)。

(72) 发明人: 黄雪航(**HUANG, Xuehang**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局, Beijing 100028 (CN)。 黄展鹏(**HUANG, Zhanpeng**); 中国

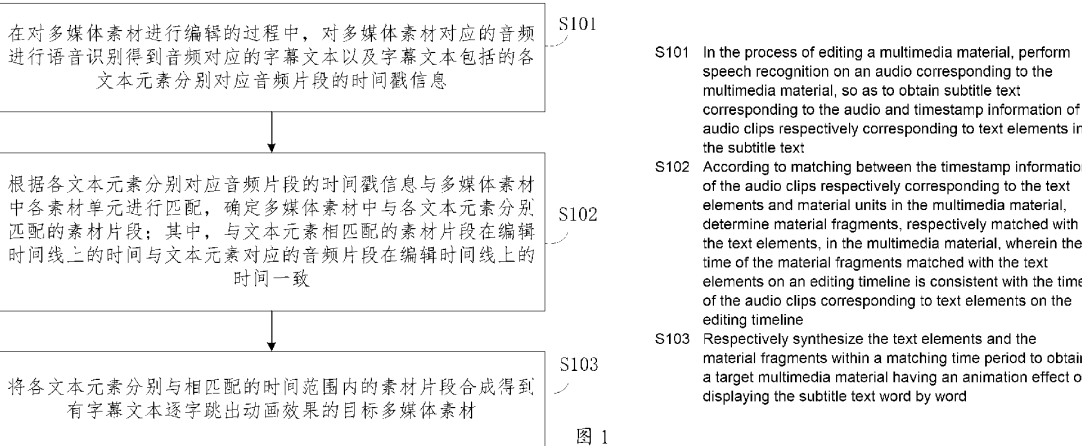
北京市朝阳区七圣中街12号院融中心B1小邮局, Beijing 100028 (CN)。 俞志云(**YU, Zhiyun**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局, Beijing 100028 (CN)。

(74) 代理人: 北京市柳沈律师事务所 (**LIU, SHEN & ASSOCIATES**); 中国北京市海淀区彩和坊路10号1号楼10层, Beijing 100080 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD,

(54) **Title:** SUBTITLE PROCESSING METHOD AND DEVICE

(54) 发明名称: 字幕处理方法及装置



(57) **Abstract:** The present invention relates to a subtitle processing method and device. The method comprises: in the process of editing multimedia material fragments, performing speech recognition on an audio corresponding to a multimedia material, so as to obtain subtitle text corresponding to the audio and timestamp information of audio clips corresponding to text elements in the subtitle text; according to the timestamp information of the audio clips corresponding to the text elements, determining material fragments, matched with the text elements, among the multimedia material fragments; and then respectively synthesizing the text elements and the material fragments within a matching time period to obtain a target multimedia material having an animation effect of displaying the subtitle text word by word. The solution of the present invention can achieve a subtitle animation effect that a corresponding text subtitle appears upon a certain word is spoken; in addition, dynamic subtitles can be automatically generated by means of an instruction inputted by a user, user's operation is simple, and user experience is improved.

[见续页]

SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚
(AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE,
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR,
HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO,
PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN,
TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本公开涉及一种字幕处理方法及装置, 其中, 该方法包括: 在对多媒体素材片段进行编辑的过程中, 通过对多媒体素材对应的音频进行语音识别得到音频对应的字幕文本以及字幕文本中各文本元素对应的音频片段的时间戳信息; 根据各文本元素对应的音频片段的时间戳信息, 确定多媒体素材片段中与文本元素相匹配的素材片段; 再将各文本元素分别与相匹配的时间内的素材片段合成得到有字幕文本逐字跳出动画效果的目标多媒体素材。本公开的方案能够实现说到某个词时相应文本字幕出现的字幕动画效果; 此外, 用户输入指令可实现动态字幕自动生成, 用户操作简单, 有利于提升用户体验。