(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 4 307 299 A3**

(12) **EUROPEAN PATENT APPLICATION**

(88) Date of publication A3:
01.05.2024 Bulletin 2024/18

(43) Date of publication A2:
17.01.2024 Bulletin 2024/03

(21) Application number: **23213490.8**

(22) Date of filing: **17.03.2021**

(51) International Patent Classification (IPC):
G10L 15/26 *(2006.01)* ***G10L 25/87*** *(2013.01)*
G10L 15/04 *(2013.01)* ***G10L 15/16*** *(2006.01)*
G10L 15/06 *(2013.01)* ***G06N 3/04*** *(2023.01)*
G06N 3/044 *(2023.01)* ***G06N 3/045*** *(2023.01)*
G06N 3/08 *(2023.01)*

(52) Cooperative Patent Classification (CPC):
G10L 15/26; G06N 3/044; G06N 3/045; G06N 3/08; G10L 15/04; G10L 15/063; G10L 15/16; G10L 25/87

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **07.05.2020 US 202063021660 P**

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
21717715.3 / 4 128 219

(71) Applicant: **Google LLC**
Mountain View, CA 94043 (US)

(72) Inventors:
- **SAINATH, Tara N.**
  Mountain View, 94043 (US)
- **GARCIA, Basilio Castillo**
  Mountain View, 94043 (US)
- **RYBACH, David**
  Mountain View, 94043 (US)
- **STROHMAN, Trevor**
  Mountain View, 94043 (US)
- **PANG, Ruoming**
  Mountain View, 94043 (US)

(74) Representative: **Lewis, Stefanie Janneke et al**
Kilburn & Strode LLP
Lacon London
84 Theobalds Road
London WC1X 8NL (GB)

(54) **EMITTING WORD TIMINGS WITH END-TO-END MODELS**

(57) A method (400) includes receiving a training example (302) that includes audio data (202) representing a spoken utterance (12) and a ground truth transcription (204). For each word (310) in the utterance, the method also includes inserting a placeholder symbol before the word identifying a respective ground truth alignment (312, 314) for a beginning and an end of the word, and generating a first constrained alignment (330) for a beginning word piece and a second constrained alignment for an ending word piece. The first constrained alignment is aligned with the ground truth alignment for the beginning of the respective word and the second constrained alignment is aligned with the ground truth alignment for the ending of the respective word. The method also includes constraining an attention head of a second pass decoder (230) by applying the first and second constrained alignments.

EP 4 307 299 A3

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 23 21 3490

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	TARA N SAINATH ET AL: "Two-Pass End-to-End Speech Recognition", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 29 August 2019 (2019-08-29), XP081489070, * page 1, paragraph 1 - page 2, paragraph 3 * * page 3, left-hand column, paragraph 3.2 * * page 3, paragraph 4.2 - page 4, paragraph 4.3 * -----	1-11	INV. G10L15/26 G10L25/87 G10L15/04 G10L15/16 G10L15/06 G06N3/04 G06N3/044 G06N3/045 G06N3/08
A	LI BO ET AL: "Towards Fast and Accurate Streaming End-To-End ASR", ICASSP 2020 - 2020 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP), IEEE, 4 May 2020 (2020-05-04), pages 6069-6073, XP033794369, DOI: 10.1109/ICASSP40776.2020.9054715 * abstract * * page 6069, left-hand column, paragraph 1 - right-hand column, paragraph 5 * ----- -/--	1-11	

TECHNICAL FIELDS SEARCHED (IPC)
G10L
G06N

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	15 March 2024	Ebbinghaus, Stefanie

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons
& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 23 21 3490

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	SAINATH TARA N ET AL: "An Attention-Based Joint Acoustic and Text on-Device End-To-End Model", ICASSP 2020 - 2020 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP), IEEE, 4 May 2020 (2020-05-04), pages 7039-7043, XP033793174, DOI: 10.1109/ICASSP40776.2020.9053510 * page 6070, paragraph 2 - right-hand column, paragraph 3 * * page 6071, paragraph 3.2 * * page 6072, paragraph 4.4 - page 6073, paragraph 4.5 * -----	1-11	
X,P	SAINATH TARA N. ET AL: "Emitting Word Timings with End-to-End Models", INTERSPEECH 2020 , 25 October 2020 (2020-10-25), pages 3615-3619, XP055817592, ISCA DOI: 10.21437/Interspeech.2020-1059 Retrieved from the Internet: URL:https://isca-speech.org/archive/Interspeech_2020/pdfs/1059.pdf [retrieved on 2021-06-24] * the whole document * -----	1-11	TECHNICAL FIELDS SEARCHED (IPC)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	15 March 2024	Ebbinghaus, Stefanie

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)