

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2025-512659  
(P2025-512659A)

(43)公表日 令和7年4月22日(2025.4.22)

(51)国際特許分類	F I	テーマコード(参考)
G 0 6 F 16/783(2019.01)	G 0 6 F 16/783	5 B 1 7 5
G 0 6 N 3/02(2006.01)	G 0 6 N 3/02	
G 0 6 N 3/0455(2023.01)	G 0 6 N 3/0455	

審査請求 未請求 予備審査請求 未請求 (全29頁)

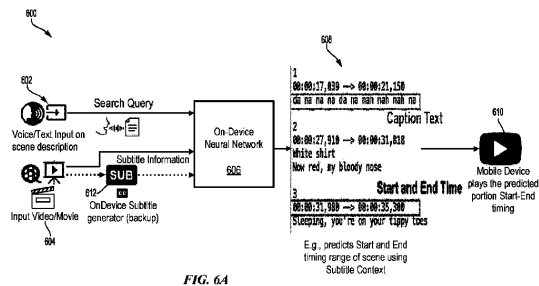
(21)出願番号	特願2024-547596(P2024-547596)	(71)出願人	595020643 クアルコム・インコーポレイテッド QUALCOMM INCORPORATED
(86)(22)出願日	令和5年2月16日(2023.2.16)	(74)代理人	110003708 弁理士法人鈴榮特許総合事務所
(85)翻訳文提出日	令和6年8月9日(2024.8.9)	(72)発明者	パテル、シュバム・ディーバック アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、 モアハウス・ドライブ 5 7 7 5
(86)国際出願番号	PCT/US2023/013252	(72)発明者	ブドワニ、パワン・アースダラム アメリカ合衆国、カリフォルニア州 9
(87)国際公開番号	WO2023/167791		
(87)国際公開日	令和5年9月7日(2023.9.7)		
(31)優先権主張番号	202241011422		
(32)優先日	令和4年3月3日(2022.3.3)		
(33)優先権主張国・地域又は機関	インド(IN)		
(81)指定国・地域	AP(BW,CV,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV)		

最終頁に続く

(54)【発明の名称】 オンデバイス人工知能ビデオ検索

(57)【要約】

人工ニューラルネットワーク(ANN)を使用するオンデバイスビデオクエリ及び検索のためのコンピュータで実施する方法は、ANNによって、ビデオと検索クエリとを受信することを含む。ビデオは、フレームのシーケンス及び関連付けられたサブタイトル情報を含む。モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とが生成される。検索クエリとサブタイトル情報との間の相関は、第1の表現及び第2の表現に基づいて、ANNによってモバイルデバイスにおいて判定される。ANNは、モバイルデバイスにおいて、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測する。



**【特許請求の範囲】****【請求項 1】**

人工ニューラルネットワーク（ANN）を使用してモバイルデバイス上でビデオを検索するためのコンピュータで実施する方法であって、

前記 ANN によって、前記ビデオと検索クエリとを受信することであって、前記ビデオが、フレームのシーケンス及び関連付けられたサブタイトル情報を含む、ことと、

前記モバイルデバイスにおいて、前記 ANN によって、前記検索クエリ内の単語の第 1 のセットについての第 1 の表現と、前記サブタイトル情報内の単語の第 2 のセットについての第 2 の表現とを生成することと、

前記モバイルデバイスにおいて、前記 ANN によって、前記第 1 の表現及び前記第 2 の表現に基づいて相関を判定することと、

前記モバイルデバイスにおいて、前記 ANN によって、前記相関に基づいて、前記検索クエリに応じて、コンテンツを含む前記ビデオの一部を予測することと、

を含む、コンピュータで実施する方法。

**【請求項 2】**

前記予測することが、前記検索クエリに応じて、前記コンテンツを含む前記ビデオの前記一部分についての開始時間及び終了時間を更に示す、請求項 1 に記載のコンピュータで実施する方法。

**【請求項 3】**

前記開始時間に含まれているビデオの前記一部分を前記終了時間まで表示することを更に含む、請求項 2 に記載のコンピュータで実施する方法。

**【請求項 4】**

前記 ANN が、トランスフォーマニューラルネットワークを含む、請求項 1 に記載のコンピュータで実施する方法。

**【請求項 5】**

前記ビデオとともに含まれるクロードキャプション情報に基づいて、前記関連付けられたサブタイトル情報を生成することを更に含む、請求項 1 に記載のコンピュータで実施する方法。

**【請求項 6】**

前記検索クエリが、シーン、イベント、単語、又は句の記述のうちの 1 つ又は複数を含む、請求項 1 に記載のコンピュータで実施する方法。

**【請求項 7】**

前記検索クエリが、前記モバイルデバイスの発話入力テキスト入力を介して供給される、請求項 1 に記載のコンピュータで実施する方法。

**【請求項 8】**

人工ニューラルネットワーク（ANN）を使用してモバイルデバイス上でビデオを検索するための装置であって、

メモリと、

前記メモリに結合された少なくとも 1 つのプロセッサと、を備え、前記少なくとも 1 つのプロセッサが、

前記 ANN によって、前記ビデオと検索クエリとを受信し、前記ビデオが、フレームのシーケンス及び関連付けられたサブタイトル情報を含み、

前記モバイルデバイスにおいて、前記 ANN によって、前記検索クエリ内の単語の第 1 のセットについての第 1 の表現と、前記サブタイトル情報内の単語の第 2 のセットについての第 2 の表現とを生成し、

前記モバイルデバイスにおいて、前記 ANN によって、前記第 1 の表現及び前記第 2 の表現に基づいて相関を判定し、

前記モバイルデバイスにおいて、前記 ANN によって、前記相関に基づいて、前記検索クエリに応じて、コンテンツを含む前記ビデオの一部を予測する、

ように構成されている、装置。

10

20

30

40

50

## 【請求項 9】

前記少なくとも 1 つのプロセッサが、前記検索クエリに応じて、前記コンテンツを含む前記ビデオの前記一部分についての開始時間及び終了時間を示す予測を生成するように更に構成されている、請求項 8 に記載の装置。

## 【請求項 10】

前記少なくとも 1 つのプロセッサが、前記開始時間に含まれているビデオの前記一部分を前記終了時間まで表示するように更に構成されている、請求項 9 に記載の装置。

## 【請求項 11】

前記 ANN が、トランスフォーマニューラルネットワークを含む、請求項 8 に記載の装置。

10

## 【請求項 12】

前記少なくとも 1 つのプロセッサが、前記ビデオとともに含まれるクローズドキャプション情報に基づいて、前記関連付けられたサブタイトル情報を生成するように更に構成されている、請求項 8 に記載の装置。

## 【請求項 13】

前記検索クエリが、シーン、イベント、単語、又は句の記述のうちの 1 つ又は複数を含む、請求項 8 に記載の装置。

## 【請求項 14】

前記検索クエリが、前記モバイルデバイスの発話入力テキスト入力を介して供給される、請求項 8 に記載の装置。

20

## 【請求項 15】

人工ニューラルネットワーク (ANN) を使用してモバイルデバイス上でビデオを検索するためのプログラムコードを記録した非一時的コンピュータ可読媒体であって、前記プログラムコードが、プロセッサによって実行され、

前記 ANN によって、前記ビデオと検索クエリとを受信するためのプログラムコードであって、前記ビデオが、フレームのシーケンス及び関連付けられたサブタイトル情報を含む、プログラムコードと、

前記モバイルデバイスにおいて、前記 ANN によって、前記検索クエリ内の単語の第 1 のセットについての第 1 の表現と、前記サブタイトル情報内の単語の第 2 のセットについての第 2 の表現とを生成するためのプログラムコードと、

30

前記モバイルデバイスにおいて、前記 ANN によって、前記第 1 の表現及び前記第 2 の表現に基づいて相関を決定するためのプログラムコードと、

前記モバイルデバイスにおいて、前記 ANN によって、前記相関に基づいて、前記検索クエリに応じて、コンテンツを含む前記ビデオの一部を予測するためのプログラムコードと、

を含む、非一時的コンピュータ可読媒体。

## 【請求項 16】

前記プログラムコードが、前記検索クエリに応じて、前記コンテンツを含む前記ビデオの前記一部分についての開始時間及び終了時間を示す予測を生成するためのプログラムコードを更に含む、請求項 15 に記載の非一時的コンピュータ可読媒体。

40

## 【請求項 17】

前記プログラムコードが、前記開始時間に含まれているビデオの前記一部分を前記終了時間まで表示するためのプログラムコードを更に含む、請求項 16 に記載の非一時的コンピュータ可読媒体。

## 【請求項 18】

前記 ANN が、トランスフォーマニューラルネットワークを含む、請求項 15 に記載の非一時的コンピュータ可読媒体。

## 【請求項 19】

前記プログラムコードが、前記ビデオとともに含まれるクローズドキャプション情報に基づいて、前記関連付けられたサブタイトル情報を生成するためのプログラムコード

50

を更に含む、請求項 15 に記載の非一時的コンピュータ可読媒体。

【請求項 20】

前記検索クエリが、シーン、イベント、単語、又は句の記述のうちの 1 つ又は複数を含む、請求項 15 に記載の非一時的コンピュータ可読媒体。

【請求項 21】

前記検索クエリが、前記モバイルデバイスの発話入力テキスト入力を介して供給される、請求項 15 に記載の非一時的コンピュータ可読媒体。

【請求項 22】

人工ニューラルネットワーク (ANN) を使用してモバイルデバイス上でビデオを検索するための装置であって、

前記 ANN によって、前記ビデオと検索クエリとを受信するための手段であって、前記ビデオが、フレームのシーケンス及び関連付けられたサブタイトル情報を含む、手段と、

前記モバイルデバイスにおいて、前記 ANN によって、前記検索クエリ内の単語の第 1 のセットについての第 1 の表現と、前記サブタイトル情報内の単語の第 2 のセットについての第 2 の表現とを生成するための手段と、

前記モバイルデバイスにおいて、前記 ANN によって、前記第 1 の表現及び前記第 2 の表現に基づいて相関を判定するための手段と、

前記モバイルデバイスにおいて、前記 ANN によって、前記相関に基づいて、前記検索クエリに応じて、コンテンツを含む前記ビデオの一部を予測するための手段と、

を備える、装置。

【請求項 23】

前記検索クエリに応じて、前記コンテンツを含む前記ビデオの前記一部分についての開始時間及び終了時間を示す予測を生成するための手段を更に備える、請求項 22 に記載の装置。

【請求項 24】

前記開始時間に含まれているビデオの前記一部分を前記終了時間まで表示するための手段を更に備える、請求項 23 に記載の装置。

【請求項 25】

前記 ANN が、トランスフォーマーニューラルネットワークを含む、請求項 22 に記載の装置。

【請求項 26】

前記ビデオとともに含まれるクローズドキャプション情報に基づいて、前記関連付けられたサブタイトル情報を生成するための手段を更に備える、請求項 22 に記載の装置。

【請求項 27】

前記検索クエリが、シーン、イベント、単語、又は句の記述のうちの 1 つ又は複数を含む、請求項 22 に記載の装置。

【請求項 28】

前記検索クエリが、前記モバイルデバイスの発話入力テキスト入力を介して供給される、請求項 22 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

【0001】 本出願は、その開示全体が参照により本明細書に明確に組み込まれる、2022年3月3日出願の「ON-DEVICE ARTIFICIAL INTELLIGENCE VIDEO SEARCH」と題するインド特許出願第202241011422号の優先権を主張する。

【0002】

【0002】 本開示の態様は、概して、ニューラルネットワークに関し、より具体的には

10

20

30

40

50

、人工ニューラルネットワークを使用するオンデバイスビデオ検索に関する。

【背景技術】

【0003】

【0003】人工ニューラルネットワークは、人工ニューロン（例えば、ニューロンモデル）の相互接続されたグループを備えてもよい。人工ニューラルネットワークは、計算デバイスであってもよく、又は計算デバイスによって実行されることになる方法として表されてもよい。畳み込みニューラルネットワークは、フィードフォワード人工ニューラルネットワークの1つのタイプである。畳み込みニューラルネットワークは、各々が受容野を有し、入力空間をまとめてタイリングする、ニューロンの集合を含み得る。深層畳み込みニューラルネットワーク（DCNs）などの畳み込みニューラルネットワーク（CNNs）は、多数の適用例を有する。特に、これらのニューラルネットワークアーキテクチャは、画像認識、パターン認識、発話認識、自律運転、及び他の分類タスクなどの、様々な技術において使用される。

10

【0004】

【0004】スマートフォン又は他のモバイルデバイスなどのエッジデバイスは、例えば、音楽又はビデオなどのメディアを消費するために広く使用されている。ニューラルネットワークの多くの有用な用途を考えると、エッジデバイス上での使用に対する需要が増大している。ビデオ、歌、又は他のシーケンス内の特定のコンテンツを検索することは、ユーザにとって一般的なタスクである。例えば、しばしば、ユーザは、映画又はビデオ全体を見ることなく、例えば、映画、重要なイベント（例えば、ゴール）ダイアログ、又はビデオ内の会話からのお気に入りの又は印象的なシーンを再生することを望む場合がある。しかしながら、そのようなイベントを自動的に検索することは、面倒であり、時間がかかり、電力の観点から計算コストが高い。これは、モバイルデバイスなどのリソースが限られたデバイスにおいて特に深刻である。

20

【発明の概要】

【0005】

【0005】本開示は、独立請求項にそれぞれ記載される。本開示のいくつかの態様は、従属請求項に記載される。

【0006】

【0006】本開示の一態様では、人工ニューラルネットワーク（artificial neural network、ANN）を使用してモバイルデバイス上でビデオを検索するためのコンピュータで実施する方法は、ANNによって、ビデオと検索クエリとを受信することを含む。ビデオは、フレームのシーケンス及び関連付けられたサブタイトル情報を含む。コンピュータで実施する方法はまた、モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成することを含む。コンピュータで実施する方法は、加えて、モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定することを含む。コンピュータで実施する方法は、モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測することを更に含む。

30

40

【0007】

【0007】本開示の別の態様は、ANNによって、ビデオと検索クエリとを受信するための手段を含む装置を対象とする。ビデオは、フレームのシーケンス及び関連付けられたサブタイトル情報を含む。装置はまた、モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成するための手段を含む。加えて、装置は、モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定するための手段を含む。装置は、モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測するための手段を更に含む。

50

## 【0008】

【0008】 本開示の別の態様では、非一時的コンピュータ可読媒体が提示される。非一時的コンピュータ可読媒体は、人工ニューラルネットワーク（ANN）を使用してモバイルデバイス上でビデオを検索するためのプログラムコードを記録している。プログラムコードは、プロセッサによって実行され、ANNによって、ビデオと検索クエリとを受信するためのプログラムコードを含む。ビデオは、フレームのシーケンス及び関連付けられたサブタイトル情報を含む。プログラムコードはまた、モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成するためのプログラムコードを含む。プログラムコードは、加えて、モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定するためのプログラムコードを含む。プログラムコードは、モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測するためのプログラムコードを更に含む。

10

## 【0009】

【0009】 本開示の別の態様は、人工ニューラルネットワーク（ANN）を使用してモバイルデバイス上でビデオを検索するための装置を対象とする。装置は、メモリと、メモリに結合された1つ又は複数のプロセッサと、を有する。プロセッサ（単数又は複数）は、ANNによって、ビデオと検索クエリとを受信するように構成されている。ビデオは、フレームのシーケンス及び関連付けられたサブタイトル情報を含む。プロセッサ（単数又は複数）はまた、モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成するように構成されている。プロセッサ（単数又は複数）は、加えて、モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定するように構成されている。プロセッサ（単数又は複数）は、モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測するように更に構成されている。

20

## 【0010】

【0010】 本開示の追加の特徴及び利点が以下で説明される。本開示が、本開示の同じ目的を遂行するための他の構造を修正又は設計するための基礎として容易に利用されてもよいことを、当業者は理解されたい。そのような等価な構成が、添付の特許請求の範囲に記載されている本開示の教示から逸脱しないことも、当業者は認識されたい。その編成及び動作方法の両方に関して本開示の特性であると考えられる新規の特徴は、更なる目的及び利点とともに、添付の図に関して検討されると以下の説明からより良く理解されよう。しかしながら、図の各々は例示及び説明のために提供されるにすぎず、本開示の範囲を定めるものではないことを明確に理解されたい。

30

## 【図面の簡単な説明】

## 【0011】

【0011】 本開示の特徴、性質、及び利点は、同様の参照符号が全体にわたって対応して識別する図面と併せて読まれると、以下に記載される詳細な説明からより明らかになる。

40

【図1】【0012】 本開示のいくつかの態様による、汎用プロセッサを含むシステムオンチップ（SOC）を使用するニューラルネットワークの例示的な実装形態を示す図である。

【図2A】【0013】 本開示の態様によるニューラルネットワークを示す図である。

【図2B】 本開示の態様によるニューラルネットワークを示す図である。

【図2C】 本開示の態様によるニューラルネットワークを示す図である。

【図2D】【0014】 本開示の態様による、例示的な深層畳み込みネットワーク（DCN）を示す図である。

【図3】【0015】 本開示の態様による、例示的な深層畳み込みネットワーク（DCN）

50

を示すブロック図である。

【図4】[0016] 本開示の態様による、人工知能(AI)機能をモジュール化し得る例示的なソフトウェアアーキテクチャを示すブロック図である。

【図5】[0017] 本開示の態様による、ビデオのオンデバイスクエリ及び検索のための例示的なプロセスを示す高レベルブロック図である。

【図6A】[0018] 本開示の態様による、ビデオのクエリ及び検索のための例示的なアーキテクチャを示すブロック図である。

【図6B】[0019] 本開示の態様による、例示的なオンデバイスニューラルネットワークを示すブロック図である。

【図7】[0020] 本開示の態様による、人工ニューラルネットワークを使用するビデオのクエリ及び検索のための例示的なプロセスを示すフロー図である。

【発明を実施するための形態】

【0012】

【0021】 添付の図面に関して以下に記載される発明を実施するための形態は、様々な構成を説明するものであり、説明される概念が実施され得る構成のみを表すものではない。「発明を実施するための形態」は、様々な構想の完全な理解をもたらすことを目的とする、具体的な詳細を含む。しかしながら、当業者には、これらの具体的な詳細を伴わずとも、これらの構想を実践することができる点が明らかとなるであろう。場合によっては、そのような構想を不明瞭にすることを回避するために、周知の構造及び構成要素は、ブロック図の形式で示されている。

【0013】

【0022】 教示に基づいて、本開示の範囲は、本開示の任意の他の態様とは無関係に実装されるか、又は本開示の任意の他の態様と組み合わせて実装されるかに関わらず、本開示の任意の態様を包含することを意図することを、当業者は理解されたい。例えば、記載された任意の数の態様を使用して、装置が実装されてもよく、又は方法が実践されてもよい。加えて、本開示の範囲は、記載された本開示の様々な態様に加えて、又はそれらの態様以外の、他の構造、機能、又は構造及び機能を使用して実践される、そのような装置又は方法を包含するものとする。開示される本開示の任意の態様は、請求項の1つ又は複数の要素によって具現化される場合があることを理解されたい。

【0014】

【0023】 「例示的」という単語は、「例、事例、又は例示としての役割を果たす」を意味するために使用される。「例示的」として説明されるいかなる態様も、必ずしも他の態様よりも好ましい又は有利であると解釈されるべきであるとは限らない。

【0015】

【0024】 特定の態様が説明されるが、これらの態様の多くの変形及び置換は本開示の範囲内に入る。好ましい態様のいくつかの利益及び利点が述べられるが、本開示の範囲は、特定の利益、用途、又は目的に限定されるものではない。むしろ、本開示の態様は、異なる技術、システム構成、ネットワーク、及びプロトコルに広く適用可能であるものとし、そのうちのいくつかが例として図及び好ましい態様の以下の説明において示される。詳細な説明及び図面は、限定的ではなく、本開示の例示にすぎず、本開示の範囲は、添付の特許請求の範囲及びその等価物によって定義される。

【0016】

【0025】 説明したように、ビデオ内の特定のコンテンツを検索することは、ユーザにとって一般的なタスクである。しばしば、ユーザは、映画又はビデオ全体を見ることなく、例えば、映画、重要なイベント(例えば、ゴール)ダイアログ、又はビデオ内の会話からのお気に入りの又は印象的なシーンを再生することを望む場合がある。しかしながら、そのようなイベントを自動的に検索することは、面倒であり、時間がかかり、電力の観点から計算コストが高い。これは、モバイルデバイスなどのリソースが限られたデバイスにおいて特に深刻である。

【0017】

10

20

30

40

50

【0026】 ビデオ検索のための従来の技術は、ビデオフレームの意味理解を強化することを含み、検索において示される正しいビデオの瞬間を見つけるためにビデオフレームの各々を検索することを伴う。より高速な結果を生成するために、いくつかの従来の技術はまた、すべてのビデオフレームの前処理と、キャッシュされたコーパス/データベースに学習を記憶することを含む。しかしながら、ビデオフレームの意味理解は、困難なタスクである。加えて、そのような従来の技術は、時間がかかり、ビデオを検索する際にかなりの時間が費やされる。更に、従来の技術は、計算コストが高く、これは、スマートフォン又は他のモバイルデバイスなどのリソースが限られたデバイスにおいて深刻である。

【0018】

【0027】 これら及び他の課題に対処するために、本開示の態様は、ビデオのオンデバイスクエリ及び検索を対象とする。人工ニューラルネットワークは、ユーザからの検索クエリ、ビデオからのサブタイトル情報を処理し、検索されたコンテンツを含む可能性が最も高いビデオの一部を示す予測を生成することができる。いくつかの態様では、人工ニューラルネットワークは、検索クエリに一致する、又は検索されたコンテンツを含む可能性が高い、可能性が高いイベントのN個の予測のリストを生成することができ（例えば、検索クエリが、複数回発生するイベントを記述する場合）、Nは、整数である。

10

【0019】

【0028】 いくつかの態様では、予測は、検索されたコンテンツを含む一部分を識別するための開始時間及び終了時間を含むことができる。予測された部分は、ビデオプレーヤ又はメディアプレーヤにおいてテキスト又はタイムスタンプ（例えば、タイムライン位置）として表示することができ、ユーザが、検索されたコンテンツを再生するために予測された部分に容易にナビゲートすることを可能にする、又はいくつかの態様では、モバイルデバイスは、予測された部分を自動的に再生することができる。処理及び予測は、オンデバイスで行うことができる。オンデバイスは、クラウド又はリモートコンピューティングの支援なしでの処理及び予測を指すことができる。

20

【0020】

【0029】 図1は、人工ニューラルネットワーク（例えば、ニューラルエンドツーエンドネットワーク）を使用してビデオを検索するように構成された中央処理ユニット（central processing unit、CPU）102又はマルチコアCPUを含むことができる、システムオンチップ（system-on-a-chip、SOC）100の例示的な実装形態を示す。変数（例えば、ニューラル信号及びシナプス重み）、計算デバイス（例えば、重みを有するニューラルネットワーク）に関連付けられたシステムパラメータ、遅延、周波数ピン情報、及びタスク情報は、ニューラル処理ユニット（NPU）108に関連付けられたメモリブロックに、CPU102に関連付けられたメモリブロックに、グラフィックス処理ユニット（GPU）104に関連付けられたメモリブロックに、デジタル信号プロセッサ（DSP）106に関連付けられたメモリブロックに、メモリブロック118に記憶されてもよく、又は複数のブロックにわたって分散されてもよい。CPU102において実行される命令は、CPU102に関連付けられたプログラムメモリからロードされてもよく、又はメモリブロック118からロードされてもよい。

30

【0021】

【0030】 SOC100はまた、GPU104、DSP106、第5世代（5G）接続性、第4世代ロングタームエボリューション（4G LTE）接続性、Wi-Fi接続性、USB接続性、Bluetooth接続性などを含んでもよい接続性ブロック110、並びに、例えば、ジェスチャを検出及び認識する場合があるマルチメディアプロセッサ112などの、特定の機能に適合された追加の処理ブロックを含んでもよい。一実装形態では、NPU108は、CPU102、DSP106、及び/又はGPU104において実装される。SOC100はまた、センサプロセッサ114、画像信号プロセッサ（ISP）116、及び/又は全地球測位システムを含んでもよいナビゲーションモジュール120を含んでもよい。

40

【0022】

50

【0031】 SOC100は、ARM命令セットに基づいてもよい。本開示の態様では、汎用プロセッサ102にロードされる命令は、ANNによって、ビデオと検索クエリとを受信するためのコードを含むことができる。ビデオは、フレームのシーケンス及び関連付けられたサブタイトル情報を含む。汎用プロセッサ102はまた、モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成するためのコードを含むことができる。汎用プロセッサ102は、加えて、モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定するためのコードを含むことができる。汎用プロセッサ102は、モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測するためのコードを更に含むことができる。 10

【0023】

【0032】 深層学習アーキテクチャは、各層内の連続的により高い抽象レベルにおいて入力を表現することを学習し、それによって入力データの有用な特徴表現を構築することによって、オブジェクト認識タスクを実行してもよい。このようにして、深層学習は、従来の機械学習の主要なボトルネックに対処する。深層学習の出現より前に、オブジェクト認識問題に対する機械学習手法は、場合によっては、浅い分類器と組み合わせて、人間が設計した特徴に大きく依拠してきた可能性がある。浅い分類器は、例えば、入力がどのクラスに属するかを予測するために特徴ベクトル成分の加重和を閾値と比較することができる、2クラス線形分類器であってもよい。人間が設計した特徴は、領域専門知識を有する技術者によって特定の領域に適合されたテンプレート又はカーネルであってもよい。対照的に、深層学習アーキテクチャは、人間の技術者が設計することができるものと同様である特徴を表現することを、しかし訓練を通して、学習してもよい。更に、深層ネットワークは、人間が考えていなかった可能性がある新しいタイプの特徴を表現及び認識することを学習してもよい。 20

【0024】

【0033】 深層学習アーキテクチャは、特徴の階層を学習してもよい。視覚データが提示された場合、例えば、第1の層は、入力ストリームにおけるエッジなどの比較的単純な特徴を認識することを学習してもよい。別の例では、聴覚データが提示される場合、第1の層は、特定の周波数におけるスペクトル電力を認識することを学習することができる。第1の層の出力を入力として取る第2の層は、視覚データの場合は単純な形状などの特徴の組み合わせ、又は聴覚データの場合は音の組み合わせを認識することを学習してもよい。例えば、上位層は、視覚データにおける複雑な形状又は聴覚データにおける単語を表現することを学習することができる。更に上位の層は、一般的な視覚オブジェクト又は話された句を認識することを学習してもよい。 30

【0025】

【0034】 深層学習アーキテクチャは、自然階層構造を有する問題に適用されたとき、特によく機能する場合がある。例えば、電動ビークルの分類は、車輪、フロントガラス、及び他の特徴を認識することを最初に学習することから恩恵を受ける場合がある。これらの特徴は、自動車、トラック、及び飛行機を認識するために、異なる方法で上位層において組み合わせられてもよい。 40

【0026】

【0035】 ニューラルネットワークは、様々な結合性パターンを用いて設計されてもよい。フィードフォワードネットワークでは、所与の層内の各ニューロンが上位層内のニューロンと通信して、情報が下位層から上位層に渡される。上述したように、フィードフォワードネットワークの連続する層内で階層表現が構築されてもよい。ニューラルネットワークはまた、再帰結合又はフィードバック(トップダウンとも呼ばれる)結合を有してもよい。再帰結合では、所与の層内のニューロンからの出力は、同じ層内の別のニューロンに伝達されてもよい。再帰型アーキテクチャは、シーケンスの中でニューラルネットワークに配信される入力データチャンクのうちの2つ以上にわたるパターンを認識する際に 50

用であり得る。所与の層内のニューロンから下位層内のニューロンへの結合は、フィードバック（又は、トップダウン）結合と呼ばれる。多数のフィードバック結合を有するネットワークは、高レベルの概念の認識が入力の特定の低レベルの特徴を弁別する助けとなり得るときに有用であり得る。

**【 0 0 2 7 】**

【0036】 ニューラルネットワークの層間の結合は、全結合又は局所結合されていてもよい。図 2 A は、全結合ニューラルネットワーク 2 0 2 の一例を示す。全結合ニューラルネットワーク 2 0 2 では、第 1 の層内のニューロンは、その出力を第 2 の層内のあらゆるニューロンに伝達してもよく、その結果、第 2 の層内の各ニューロンは、第 1 の層内のあらゆるニューロンから入力を受信することになる。図 2 B は、局所結合ニューラルネットワーク 2 0 4 の一例を示す。局所結合ニューラルネットワーク 2 0 4 では、第 1 の層内のニューロンは、第 2 の層内の限られた数のニューロンに結合されていてもよい。より一般的には、局所結合ニューラルネットワーク 2 0 4 の局所結合層は、ある層内の各ニューロンが同じ又は類似の結合性パターンを、しかし異なる値（例えば、2 1 0、2 1 2、2 1 4、及び 2 1 6）を有してもよい結合強度を伴って、有するように構成することができる。所与の領域の中の上位層ニューロンは、訓練を通じてネットワークへの全入力のうちの制約された部分の特性に調整される入力を受信することができるので、局所結合の結合性パターンは、上位層内で空間的に別個の受容野を生じることがある。

10

**【 0 0 2 8 】**

【0037】 局所結合ニューラルネットワークの一例は、畳み込みニューラルネットワークである。図 2 C は、畳み込みニューラルネットワーク 2 0 6 の一例を示す。畳み込みニューラルネットワーク 2 0 6 は、第 2 の層内の各ニューロンについての入力に関連付けられた結合強度が共有される（例えば、2 0 8）ように構成されてもよい。畳み込みニューラルネットワークは、入力の空間口ケーションが意味のある問題に好適である場合がある。

20

**【 0 0 2 9 】**

【0038】 畳み込みニューラルネットワークの 1 つのタイプは、深層畳み込みネットワーク（DCN）である。図 2 D は、車載カメラなどの画像キャプチャデバイス 2 3 0 から入力された画像 2 2 6 から視覚特徴を認識するように設計された、DCN 2 0 0 の詳細な例を示す。本例の DCN 2 0 0 は、交通標識及び交通標識上に記載された数字を識別するように訓練されてもよい。もちろん、DCN 2 0 0 は、車線区分線を識別すること又は交通信号を識別することなどの、他のタスクのために訓練されてもよい。

30

**【 0 0 3 0 】**

【0039】 DCN 2 0 0 は、教師あり学習を用いて訓練されてもよい。訓練中に、DCN 2 0 0 は、速度制限標識の画像 2 2 6 などの画像を提示される場合があり、次いで、出力 2 2 2 を生成するために、順方向パスが計算される場合がある。DCN 2 0 0 は、特徴抽出セクション及び分類セクションを含んでよい。画像 2 2 6 を受信すると、畳み込み層 2 3 2 は、畳み込みカーネル（図示せず）を画像 2 2 6 に適用して、特徴マップの第 1 のセット 2 1 8 を生成し得る。例として、畳み込み層 2 3 2 についての畳み込みカーネルは、 $28 \times 28$  の特徴マップを生成する  $5 \times 5$  のカーネルであってもよい。本例では、特徴マップの第 1 のセット 2 1 8 において 4 つの異なる特徴マップが生成されるため、畳み込み層 2 3 2 において 4 つの異なる畳み込みカーネルが画像 2 2 6 に適用された。畳み込みカーネルはまた、フィルタ又は畳み込みフィルタと呼ばれることもある。

40

**【 0 0 3 1 】**

【0040】 特徴マップの第 1 のセット 2 1 8 は、特徴マップの第 2 のセット 2 2 0 を生成するために、最大プーリング層（図示せず）によってサブサンプリングされてもよい。最大プーリング層は、特徴マップの第 1 のセット 2 1 8 のサイズを縮小する。すなわち、 $14 \times 14$  などの特徴マップの第 2 のセット 2 2 0 のサイズは、 $28 \times 28$  などの特徴マップの第 1 のセット 2 1 8 のサイズよりも小さい。縮小されたサイズにより、メモリ消費量を低減しながら、類似の情報を後続の層に提供する。特徴マップの第 2 のセット 2 2 0

50

は、1つ又は複数の後続の畳み込み層（図示せず）を介して更に畳み込まれて、特徴マップの1つ又は複数の後続のセット（図示せず）を生成し得る。

【0032】

【0041】 図2Dの例では、特徴マップの第2のセット220が畳み込まれて、第1の特徴ベクトル224を生成する。更に、第1の特徴ベクトル224が更に畳み込まれて、第2の特徴ベクトル228を生成する。第2の特徴ベクトル228の各特徴は、「標識」、「60」、及び「100」などの画像226の可能性のある特徴に対応する数を含んでもよい。ソフトマックス関数（図示せず）は、第2の特徴ベクトル228における数を確率に変換してもよい。したがって、DCN200の出力222は、画像226が1つ又は複数の特徴を含む確率である。

【0033】

【0042】 本例では、「標識」及び「60」に対する出力222における確率は、「30」、「40」、「50」、「70」、「80」、「90」、及び「100」などの出力222のその他のものの確率よりも高い。訓練の前に、DCN200によって生成された出力222は、不正確である可能性がある。したがって、出力222とターゲット出力との間で誤差が計算されてもよい。ターゲット出力は、画像226のグラウンドトゥース（例えば、「標識」及び「60」）である。次いで、DCN200の重みが調整されてよく、そのため、DCN200の出力222はターゲット出力とより密に整合される。

【0034】

【0043】 重みを調整するために、学習アルゴリズムは重みに対する勾配ベクトルを算出してよい。勾配は、重みが調整された場合に誤差が増加又は減少することになる量を示すことができる。最上層において、勾配は、最後から2番目の層内の活性化されたニューロンと出力層内のニューロンとを結合する重みの値に直接対応することができる。下位層では、勾配は、重みの値、及び上位層の算出された誤差勾配に依存することができる。次いで、誤差を低減するように、重みを調節することができる。重みを調整するこの方式は、ニューラルネットワークを通じた「逆方向パス」を伴うので、「逆伝搬法」と呼ばれることがある。

【0035】

【0044】 実際には、重みの誤差勾配は、算出された勾配が真の誤差勾配に近似するように、少数の例にわたって算出される場合がある。この近似法は、確率的勾配降下法と称される場合がある。確率的勾配降下法は、システム全体の達成可能な誤差率が下げ止まるまで、又は、誤差率がターゲットレベルに達するまで、繰り返すことができる。学習の後、DCNは、新たな画像が提示されてよく、ネットワークを通じた順方向パスが、DCNの推定又は予測と見なされてよい出力222を与え得る。

【0036】

【0045】 深層信念ネットワーク（DBNs）は、隠れノードの複数の層を備える確率的モデルである。DBNは、訓練データセットの階層表現を抽出するために使用されてもよい。DBNは、制限付きボルツマンマシン（RBMs）の層を積層することによって取得されてもよい。RBMは、入力の特徴にわたる確率分布を学習できるタイプの人工ニューラルネットワークである。RBMは、各入力のカテゴリ分類されるべき先のクラスについての情報がない状態で確率分布を学習できるので、RBMは、教師なし学習において頻繁に使用される。教師なしと教師ありとのハイブリッドパラダイムを使用して、DBNの下部RBMは、教師なし方式で訓練することができ、特徴抽出器として機能することができ、上部RBMは、（前の層及びターゲットクラスからの入力の同時分布上で）教師あり方式で訓練することができ、分類器として機能することができる。

【0037】

【0046】 深層畳み込みネットワーク（DCNs）は、追加のプーリング層及び正規化層を用いて構成された、畳み込みネットワークのネットワークである。DCNは、多数のタスクに対して最先端の性能を達成している。DCNは、入力と出力ターゲットの両方が、多数の標本に対して知られており、かつ勾配降下法の使用によってネットワークの重み

10

20

30

40

50

を修正するために使用される、教師あり学習を使用して訓練することができる。

【0038】

[0047] DCNは、フィードフォワードネットワークであってもよい。加えて、上述したように、DCNの第1の層内のニューロンから次の上位層内のニューロンのグループへの結合は、第1の層内のニューロンにわたって共有される。DCNのフィードフォワード結合及び共有される結合は、高速処理のために活用することができる。DCNの計算負担は、例えば、再帰結合又はフィードバック結合を含む同様のサイズのニューラルネットワークの計算負担よりもはるかに小さい場合がある。

【0039】

[0048] 畳み込みネットワークの各層の処理は、空間的に不変のテンプレート又は基底射影と見なされてもよい。入力が、カラー画像の赤色チャンネル、緑色チャンネル、及び青色チャンネルなどの複数のチャンネルに最初に分解される場合、その入力に対して訓練される畳み込みネットワークは、画像の軸に沿った2つの空間次元及び色情報をキャプチャする第3の次元を有する、3次元と見なされてもよい。畳み込み結合の出力は、後続の層内で特徴マップを形成するものと見なされてもよく、特徴マップ(例えば、220)の各要素が、前の層(例えば、特徴マップ218)内のある範囲のニューロンから、かつ複数のチャンネルの各々から、入力を受信する。特徴マップにおける値は、整流、 $\max(0, x)$ などの非線形性を用いて更に処理されてもよい。隣接するニューロンからの値は更に、ダウンサンプリングに対応するプーリングが行われてもよく、追加の局所不変性及び次元数低減をもたらすことができる。白色化に対応する正規化が、特徴マップ内のニューロン間の側方抑制を通して適用される場合もある。

【0040】

[0049] 深層学習アーキテクチャの性能は、より多くのラベル付けされたデータポイントが利用可能になるほど、又は計算能力が高まるほど、向上し得る。現代の深層ニューラルネットワークは、わずか15年前に典型的な研究者が利用可能であったものよりも数千倍大きい計算リソースを用いて、ルーチン的に訓練される。新しいアーキテクチャ及び訓練パラダイムは、深層学習の性能を更に高めることができる。整流線形ユニットは、勾配消失として知られる訓練問題を低減する場合がある。新しい訓練技術は、過剰適合を低減し、したがって、より大きいモデルがより良い汎化を達成することを可能にする場合がある。カプセル化技術は、所与の受容野内のデータを抽出し、全体的な性能を更に高める場合がある。

【0041】

[0050] 図3は、深層畳み込みネットワーク350を示すブロック図である。深層畳み込みネットワーク350は、結合性及び重み共有に基づいて、複数の異なるタイプの層を含み得る。図3に示されているように、深層畳み込みネットワーク350は、畳み込みブロック354A、354Bを含む。畳み込みブロック354A、354Bの各々は、畳み込み層(CONV)356、正規化層(LNorm)358、及び最大プーリング層(MAX POOL)360を用いて構成され得る。

【0042】

[0051] 畳み込み層356は、特徴マップを生成するために入力データに適用され得る、1つ又は複数の畳み込みフィルタを含み得る。畳み込みブロック354A、354Bのうちの2つのみが示されているが、本開示は、そのように限定しておらず、代わりに、設計選好に従って、任意の数の畳み込みブロック354A、354Bが深層畳み込みネットワーク350に含まれてもよい。正規化層358は、畳み込みフィルタの出力を正規化してもよい。例えば、正規化層358は、白色化又は側方抑制をもたらし得る。最大プーリング層360は、局所不変性及び次元数低減のために、空間にわたってダウンサンプリング集約をもたらし得る。

【0043】

[0052] たとえば、深層畳み込みネットワークの並列フィルタバンクは、高性能及び低電力消費を達成するために、SOC100のCPU102又はGPU104上にロード

10

20

30

40

50

されてもよい。代替実施形態では、並列フィルタバンクは、SOC 100のDSP 106又はISP 116上にロードされてもよい。加えて、深層畳み込みネットワーク350は、センサ及びナビゲーションにそれぞれ専用のセンサプロセッサ114及びナビゲーションモジュール120などの、SOC 100上に存在する場合がある他の処理ブロックにアクセスしてもよい。

#### 【0044】

[0053] 深層畳み込みネットワーク350はまた、1つ又は複数の全結合層362(FC1及びFC2)を含み得る。深層畳み込みネットワーク350は、ロジスティック回帰(LR)層364を更に含んでもよい。深層畳み込みネットワーク350の各層356、358、360、362、364の間には、更新されることになる重み(図示せず)がある。層の各々(例えば、356、358、360、362、364)の出力は、畳み込みブロック354Aのうちの最初に供給された入力データ352(例えば、画像、オーディオ、ビデオ、センサデータ、及び/又は他の入力データ)から階層的特徴表現を学習するために、深層畳み込みネットワーク350内の層(例えば、356、358、360、362、364)のうちの後続の1つの層の入力として機能し得る。深層畳み込みネットワーク350の出力は、入力データ352に対する分類スコア366である。分類スコア366は確率のセットであってよく、ここで、各確率は、入力データが特徴のセットからの特徴を含む確率である。

10

#### 【0045】

[0054] 図4は、人工知能(AI)機能をモジュール化し得る例示的なソフトウェアアーキテクチャ400を示すブロック図である。このアーキテクチャを使用して、SOC 420の様々な処理ブロック(たとえば、CPU 422、DSP 424、GPU 426、及び/又はNPU 428)に、本開示の態様による、AIアプリケーション402のためのポストトレーニング量子化のための開示されるような適応丸めをサポートさせることができるアプリケーションが設計され得る。

20

#### 【0046】

[0055] AIアプリケーション402は、例えば、デバイスが現在動作しているロケーションを示すシーンの検出及び認識を行い得る、ユーザ空間404において定義された機能呼び出すように構成されてもよい。AIアプリケーション402は、例えば、認識されるシーンがオフィスであるか、講堂であるか、レストランであるか、又は湖などの屋外設定であるかに応じて、マイクロフォン及びカメラを異なるように構成してもよい。AIアプリケーション402は、AI機能アプリケーションプログラミングインターフェース(API)406において定義されたライブラリに関連付けられたコンパイル済みプログラムコードに対する要求を行ってもよい。この要求は、最終的には、例えば、ビデオ及び測位データに基づいて推論応答を提供するように構成された深層ニューラルネットワークの出力に依拠してもよい。

30

#### 【0047】

[0056] ランタイムフレームワークのコンパイル済みコードであってもよいランタイムエンジン408は、AIアプリケーション402が更にアクセス可能なものであってもよい。AIアプリケーション402は、ランタイムエンジンに、たとえば、特定の時間間隔における、又はアプリケーションのユーザインターフェースによって検出されたイベントによってトリガされる推論を要求させてもよい。次に、ランタイムエンジンは、推論応答を提供させられるとき、SOC 420上で動作しているカーネル412などの、オペレーティングシステム(OS)空間410の中でのオペレーティングシステムに信号を送ってもよい。次に、オペレーティングシステムは、CPU 422、DSP 424、GPU 426、NPU 428、又はそれらの何らかの組み合わせ上で量子化の連続的な緩和を実行させてもよい。CPU 422は、オペレーティングシステムによって直接アクセスされてもよく、他の処理ブロックは、それぞれ、DSP 424、GPU 426、又はNPU 428用のドライバ414、416、又は418などの、ドライバを通じてアクセスされてもよい。例示的な例では、深層ニューラルネットワークは、CPU 422、DSP 424、及

40

50

びGPU426などの処理ブロックの組み合わせ上で動作するように構成されてもよく、又はNPU428上で動作してもよい。

【0048】

[0057] アプリケーション402（たとえば、AIアプリケーション）は、たとえば、デバイスが現在動作しているロケーションを示すシーンの検出及び認識を行い得る、ユーザ空間404において定義された機能呼び出すように構成されてもよい。アプリケーション402は、たとえば、認識されるシーンがオフィスであるか、講堂であるか、レストランであるか、又は湖などの屋外設定であるかに応じて、マイクロフォン及びカメラを異なるように構成してもよい。アプリケーション402は、現在のシーンの推定値を提供するために、SceneDetectアプリケーションプログラミングインターフェース（API）406において定義されたライブラリに関連付けられたコンパイル済みプログラムコードに対する要求を行ってもよい。この要求は、最終的には、たとえば、ビデオ及び測位データに基づいてシーン推定値を提供するように構成された差分ニューラルネットワークの出力に依拠してもよい。

10

【0049】

[0058] ランタイムフレームワークのコンパイル済みコードであってもよいランタイムエンジン408は、アプリケーション402が更にアクセス可能なものであってもよい。アプリケーション402は、ランタイムエンジンに、たとえば、特定の時間間隔における、又はアプリケーションのユーザインターフェースによって検出されたイベントによってトリガされるシーン推定値を要求させてもよい。シーンを推定させられたとき、ランタイムエンジンは、今度は、SOC420上で実行している、カーネル412など、オペレーティングシステム410に信号を送信することができる。次に、オペレーティングシステム410は、CPU422、DSP424、GPU426、NPU428、又はそれらの何らかの組合せ上で計算を実行させてもよい。CPU422は、オペレーティングシステムによって直接アクセスされてもよく、他の処理ブロックは、DSP424用、GPU426用、又はNPU428用のドライバ414～418などの、ドライバを通じてアクセスされてもよい。例示的な例では、差分ニューラルネットワークは、CPU422及びGPU426などの処理ブロックの組合せ上で動作するように構成されてもよく、又はNPU428上で動作してもよい。

20

【0050】

[0059] 説明したように、本開示の態様は、ビデオのオンデバイスクエリ及び検索を対象とする。人工ニューラルネットワークは、ユーザからの検索クエリ、ビデオからのサブタイトル情報を処理し、検索されたコンテンツを含む可能性が最も高いビデオの一部を示す予測を生成することができる。いくつかの態様では、人工ニューラルネットワークは、検索クエリに一致する、又は検索されたコンテンツを含む可能性が高い、可能性が高いイベントのN個の予測のリストを生成することができ（例えば、検索クエリが、複数回発生するイベントを記述する場合）、Nは、整数である。

30

【0051】

[0060] いくつかの態様では、予測は、検索されたコンテンツを含む一部分を識別するための開始時間及び終了時間を含むことができる。予測された部分は、ビデオプレーヤ又はメディアプレーヤにおいてテキスト又はタイムスタンプ（例えば、タイムライン位置）として表示されてもよい。結果として、ユーザは、検索されたコンテンツを再生するために予測された部分により容易にナビゲートすることができ、又はいくつかの態様では、モバイルデバイスは、予測された部分を自動的に再生することができる。処理及び予測は、オンデバイスで行うことができる。オンデバイスは、クラウド又はリモートコンピューティングの支援なしにモバイルデバイスのリソースを使用する処理及び予測を指すことができる。

40

【0052】

[0061] 図5は、本開示の態様による、ビデオのオンデバイスクエリ及び検索のための例示的なプロセス500を示す高レベルブロック図である。図5を参照すると、プロッ

50

ク502で、プロセス500は、モバイルデバイスを介して検索クエリ及びビデオを受信することができる。検索クエリは、例えば、ビデオ内のシーン又はイベントの記述を含むことができる。

【0053】

[0062] ブロック504において、検索クエリとビデオに関連付けられたサブタイトル情報とを処理することができる。人工ニューラルネットワーク（例えば、図3の350）は、検索クエリ及びサブタイトル情報を処理することができる。サブタイトル情報は、ビデオ内のシーン又はアクションの記述を含むことができる。サブタイトル情報は、フレーム単位で段落形式で提供されてもよい。ANNは、検索クエリ及びサブタイトル情報の特徴を判定することができる。いくつかの態様では、ANNは、特徴をマッチングするために、検索クエリの顕著な部分を判定して、ビデオに関連付けられたサブタイトル情報を検索することができる。

10

【0054】

[0063] ブロック506において、ANNは、検索クエリに応じて、対象素材を含むビデオの一部分の予測を生成することができる。予測は、検索クエリに記述されたシーン又はイベントを含む可能性が最も高いフレームのセットを含むことができる。いくつかの態様では、予測は、検索クエリに記述されたシーン又はイベントを含む可能性が最も高いビデオの一部分についての開始時間及び終了時間を含むことができる。

【0055】

[0064] ブロック508において、ビデオの予測された部分を、（例えば、図1のマルチメディアプロセッサ112を介して）モバイルデバイスのディスプレイに表示することができる。

20

【0056】

[0065] 図6Aは、本開示の態様による、ビデオのクエリ及び検索のための例示的なアーキテクチャ600を示すブロック図である。図6Aを参照すると、アーキテクチャ600は、入力デバイス602を含むことができる。入力デバイス602は、モバイルデバイスのマイクロフォン（例えば、図1のセンサ114）、マルチメディアデバイス（例えば、図1のマルチメディアプロセッサ112）、又はテキスト入力を受信するための入力/出力デバイスを含むことができる。いくつかの態様では、入力デバイス602は、発話信号又は他のセンサ入力をテキスト入力に変換するように構成され得る。入力デバイス602は、ユーザから検索クエリを受信することができる。検索クエリは、映画に含まれるシーン又はイベントの記述であってもよい。検索クエリは、単語、句、又は文の形態であってもよい。

30

【0057】

[0066] アーキテクチャ600はまた、ビデオ入力604を受信することができる。ビデオ入力604は、例えば、映画であってもよい。ビデオ入力604は、ストレージ又はストリーミングメディアソースから受信することができる。ビデオ入力604は、フレームの時間的シーケンスを含むことができる。いくつかの態様では、ビデオ入力604は、関連付けられたクローズドキャプション（closed captioning、CC）情報又はサブタイトル情報を含むことができる。サブタイトル情報と呼ばれることもあるクローズドキャプション情報は、ビデオのフレーム内に表示された文字間の映画内のダイアログの時限トランスクリプトを含むことができる。いくつかの態様では、サブタイトル生成器612は、任意選択的に、オンデバイスに含まれ、例えば、そのような情報がビデオ入力604とともに含まれないとき、ビデオ入力604のフレームのためのサブタイトル情報を生成するために使用することができる。サブタイトル情報は、例えば、文又は段落の形態であってもよい。

40

【0058】

[0067] 検索クエリ及びビデオのサブタイトル情報は、オンデバイスニューラルネットワーク606に供給することができる。オンデバイスニューラルネットワーク606は、例えば、トランスフォーマニューラルネットワークであってもよい。トランスフォーマ

50

ニューラルネットワークは、自己注意を使用して入力シーケンス内の任意の位置に関するコンテキスト情報を提供する、深層学習モデルである。いくつかの態様では、オンデバイスニューラルネットワーク606は、トークン置換を正確に分類する効率的に学習するエンコーダ（efficiently learning encoder that classifies token replacements accurately、ELECTRA）スモールモデルであってもよい。もちろん、これは、一例にすぎず、トランスフォーマからの双方向エンコーダ表現（bi-directional encoder representations from transformers、BERT）、ロバストに最適化されたBERT手法（robustly optimized BERT approach、RoBERTa）、XLNet、Transformer-XL、及びトランスフォーマの生成事前訓練トランスフォーマ（generative pre-trained transformer、GPT）ファミリーなどの他のアーキテクチャも採用することができる。ELECTRAスモールモデルは、質問応答自然言語プロセッサ（natural language processor、NLP）である。ELECTRAスモールモデルは、クエリ及び入力パラグラフが与えられた場合に回答を予測するように事前訓練することができる。

10

【0059】

【0068】 オンデバイスニューラルネットワーク606は、検索クエリ内の単語の表現と、サブタイトル情報内の単語の表現とを生成することができる。単語の表現の各々は、検索クエリ及びサブタイトル情報内の単語の各々にそれぞれ基づくコンテキスト情報を含むことができる。オンデバイスニューラルネットワーク606は、単語の生成された表現に基づいて、検索クエリとサブタイトル情報との間の相関を判定することができる。次に、オンデバイスニューラルネットワーク606は、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオ入力604の一部分の予測608を生成することができる。予測608は、検索されたコンテンツを含む可能性が最も高いビデオの一部分を示すことができる。いくつかの態様では、予測608は、そのようなコンテンツを含む1つ又は複数のフレームを示すことができる。例えば、図6Aの例に示すように、オンデバイスニューラルネットワーク606は、3つの候補部分を生成することができ、各候補部分は、検索クエリを満たすことができるサブタイトル情報を含む。いくつかの態様では、予測608はまた、検索クエリに応じて、コンテンツを含むビデオ入力604の一部分又はセグメントについての開始時間及び終了時間を含むことができる。

20

【0060】

30

【0069】 再生デバイス610は、開始時間及び終了時間によって示されるビデオ入力604の予測された部分（単数又は複数）のリストを表示することができる。更に、いくつかの態様では、再生デバイスは、識別された開始時間にナビゲート（例えば、早送り）し、識別された終了時間まで、ビデオ入力604の予測された部分の再生を開始することができる。

【0061】

【0070】 図6Bは、本開示の態様による、例示的なオンデバイスニューラルネットワーク606を示すブロック図である。図6Bを参照すると、オンデバイスニューラルネットワーク606は、検索生成器ネットワーク620及び弁別器ネットワーク630を含むことができる。検索生成器ネットワーク620及び弁別器ネットワーク630は各々、例えば、トランスフォーマネットワークとして構成され得る。検索生成器ネットワーク620は、入力として検索クエリを受信することができ、検索クエリをコンテキストベクトル表現 $h_{sq}$ にマッピングすることができる。コンテキストベクトル表現 $h_{sq}$ は、検索クエリ内の重要な単語（例えば、名詞、動詞）に焦点を当ててもよい。コンテキストベクトル表現 $h_{sq}$ は、弁別器ネットワーク630に供給することができる。

40

【0062】

【0071】 弁別器ネットワーク630は、第1の入力としてサブタイトル情報を受信することができる。弁別器ネットワーク630は、ビデオ入力604の各フレームに関連付けられたサブタイトル情報をコンテキストベクトル表現 $h_{st}$ にマッピングすることができる。いくつかの態様では、弁別器ネットワーク630は、サブタイトル情報内の重要な

50

単語に焦点を当てるように、コンテキストベクトル表現  $h_{S_I}$  を生成することができる。

【0063】

[0072] 弁別器ネットワーク630は、コンテキストベクトル表現  $h_{S_Q}$  を受信することもできる。弁別器ネットワーク630は、コンテキストベクトル表現  $h_{S_Q}$  をコンテキストベクトル表現  $h_{S_I}$  と比較して、コンテキストベクトル表現  $h_{S_Q}$  とコンテキストベクトル表現  $h_{S_I}$  との間の相関を判定することができる。すなわち、弁別器ネットワーク630は、(例えば、検索クエリ内のみの)データの単語をサブタイトル情報に含まれる単語と区別するように訓練することができる。コンテキストベクトル表現  $h_{S_Q}$  とコンテキストベクトル表現  $h_{S_I}$  との間の相関に基づいて、弁別器ネットワーク630は、検索クエリがサブタイトル情報に一致するかどうか、したがって、ビデオ入力604の対応する部分が検索クエリを満たすコンテンツを含むかどうかの予測(例えば、608)を生成することができる。

10

【0064】

[0073] 図7は、本開示の態様による、人工ニューラルネットワーク(ANN)を使用するビデオのクエリ及び検索のための例示的なプロセス700を示すフロー図である。図7に示すように、ブロック702において、プロセス700は、ANNによって、フレームのシーケンス及び関連付けられたサブタイトル情報を含むビデオと検索クエリとを受信する。例えば、図6Aを参照して説明したように、入力デバイス602は、ユーザから検索クエリを受信することができる。検索クエリは、映画に含まれるシーン又はイベントの記述であってもよい。検索クエリは、単語、句、又は文の形態であってもよい。アーキテクチャ600はまた、ビデオ入力604を受信することができる。ビデオ入力604は、例えば、映画であってもよい。ビデオ入力604は、ストレージ又はストリーミングメディアソースから受信することができる。ビデオ入力604は、フレームの時間的シーケンスを含むことができる。いくつかの態様では、ビデオ入力604は、関連付けられたクロードキャプション(CC)情報又はサブタイトル情報を含むことができる。サブタイトル情報と呼ばれることもあるクロードキャプション情報は、ビデオのフレーム内に表示された文字間の映画内のダイアログの時限トランスクリプトを含むことができる。

20

【0065】

[0074] ブロック704において、プロセス700は、モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成する。図6Bを参照して説明したように、検索生成器ネットワーク620は、入力として検索クエリを受信ことができ、検索クエリをコンテキストベクトル表現  $h_{S_Q}$  にマッピングすることができる。コンテキストベクトル表現  $h_{S_Q}$  は、検索クエリ内の重要な単語(例えば、名詞、動詞)に焦点を当ててもよい。同様に、弁別器ネットワーク630は、サブタイトル情報を受信することができる。弁別器ネットワーク630は、ビデオ入力604の各フレームに関連付けられたサブタイトル情報をコンテキストベクトル表現  $h_{S_I}$  にマッピングすることができる。

30

【0066】

[0075] ブロック706において、プロセス700は、モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定する。例えば、図6Bを参照して説明したように、弁別器ネットワーク630は、コンテキストベクトル表現  $h_{S_Q}$  をコンテキストベクトル表現  $h_{S_I}$  と比較して、コンテキストベクトル表現  $h_{S_Q}$  とコンテキストベクトル表現  $h_{S_I}$  との間の相関を判定することができる。すなわち、弁別器ネットワーク630は、(例えば、検索クエリ内のみの)データの単語をサブタイトル情報に含まれる単語と区別するように訓練することができる。加えて、相関は、検索クエリの単語に対するコンテキストと、サブタイトル情報の単語に対するコンテキストとの対応に基づいてもよい。

40

【0067】

50

[0076] ブロック708において、プロセス700は、モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測する。例えば、図6Bを参照して説明したように、コンテキストベクトル表現 $h_{sQ}$ とコンテキストベクトル表現 $h_{sI}$ との間の相関に基づいて、弁別器ネットワーク630は、検索クエリがサブタイトル情報に一致するかどうか、したがって、ビデオ入力604の対応する部分が検索クエリを満たすコンテンツを含むかどうかの予測（例えば、608）を生成することができる。予測608は、検索されたコンテンツを含む可能性が最も高いビデオの一部を示すことができる。いくつかの態様では、予測608は、そのようなコンテンツを含む1つ又は複数のフレームを示すことができる。いくつかの態様では、予測608はまた、検索クエリに応じて、コンテンツを含むビデオ入力604の一部又はセグメントについての開始時間及び終了時間を含むことができる。

10

【0068】

[0077] 以下の番号付きの条項に、実装形態が含まれる。

1. 人工ニューラルネットワーク(ANN)を使用してモバイルデバイス上でビデオを検索するためのコンピュータで実施する方法であって、

ANNによって、ビデオと検索クエリとを受信することであって、ビデオが、フレームのシーケンス及び関連付けられたサブタイトル情報を含む、ことと、

モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成することと、

20

モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定することと、

モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測することと、

を含む、コンピュータで実施する方法。

2. 予測することが、検索クエリに応じて、コンテンツを含むビデオの一部についての開始時間及び終了時間を更に示す、条項1に記載のコンピュータで実施する方法。

3. 開始時間に含まれているビデオの一部を終了時間まで表示することを更に含む、条項1又は2に記載のコンピュータで実施する方法。

4. ANNが、トランスフォーマニューラルネットワークを含む、条項1から3のいずれか一項に記載のコンピュータで実施する方法。

30

5. ビデオとともに含まれるクローズドキャプション情報に基づいて、関連付けられたサブタイトル情報を生成することを更に含む、条項1から4のいずれか一項に記載のコンピュータで実施する方法。

6. 検索クエリが、シーン、イベント、単語、又は句の記述のうちの1つ又は複数を含む、条項1から5のいずれか一項に記載のコンピュータで実施する方法。

7. 検索クエリが、モバイルデバイスの発話入力テキスト入力を介して供給される、条項1から6のいずれか一項に記載のコンピュータで実施する方法。

8. 人工ニューラルネットワーク(ANN)を使用してモバイルデバイス上でビデオを検索するための装置であって、

40

メモリと、

メモリに結合された少なくとも1つのプロセッサと、を備え、少なくとも1つのプロセッサが、

ANNによって、ビデオと検索クエリとを受信し、ビデオが、フレームのシーケンス及び関連付けられたサブタイトル情報を含み、

モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成し、

モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定し、

50

モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部分を予測する、  
ように構成されている、装置。

9．少なくとも1つのプロセッサが、検索クエリに応じて、コンテンツを含むビデオの一部分についての開始時間及び終了時間を示す予測を生成するように更に構成されている、条項8に記載の装置。

10．少なくとも1つのプロセッサが、開始時間に含まれているビデオの一部分を終了時間まで表示するように更に構成されている、条項8又は9に記載の装置。

11．ANNが、トランスフォーマーニューラルネットワークを含む、条項8から10のいずれか一項に記載の装置。

12．少なくとも1つのプロセッサが、ビデオとともに含まれるクローズドキャプション情報に基づいて、関連付けられたサブタイトル情報を生成するように更に構成されている、条項8から11のいずれか一項に記載の装置。

13．検索クエリが、シーン、イベント、単語、又は句の記述のうちの1つ又は複数を含む、条項8から12のいずれか一項に記載の装置。

14．検索クエリが、モバイルデバイスの発話入力テキスト入力を介して供給される、条項8から13のいずれか一項に記載の装置。

15．人工ニューラルネットワーク(ANN)を使用してモバイルデバイス上でビデオを検索するためのプログラムコードを記録した非一時的コンピュータ可読媒体であって、プログラムコードが、プロセッサによって実行され、

ANNによって、ビデオと検索クエリとを受信するためのプログラムコードであって、ビデオが、フレームのシーケンス及び関連付けられたサブタイトル情報を含む、プログラムコードと、

モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成するためのプログラムコードと、

モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定するためのプログラムコードと、

モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部分を予測するためのプログラムコードと、

を含む、非一時的コンピュータ可読媒体。

16．プログラムコードが、検索クエリに応じて、コンテンツを含むビデオの一部分についての開始時間及び終了時間を示す予測を生成するためのプログラムコードを更に含む、条項15に記載の非一時的コンピュータ可読媒体。

17．プログラムコードが、開始時間に含まれているビデオの一部分を終了時間まで表示するためのプログラムコードを更に含む、条項15又は16に記載の非一時的コンピュータ可読媒体。

18．ANNが、トランスフォーマーニューラルネットワークを含む、条項15から17のいずれか一項に記載の非一時的コンピュータ可読媒体。

19．プログラムコードが、ビデオとともに含まれるクローズドキャプション情報に基づいて、関連付けられたサブタイトル情報を生成するためのプログラムコードを更に含む、条項15から18のいずれか一項に記載の非一時的コンピュータ可読媒体。

20．検索クエリが、シーン、イベント、単語、又は句の記述のうちの1つ又は複数を含む、条項15から19のいずれか一項に記載の非一時的コンピュータ可読媒体。

21．検索クエリが、モバイルデバイスの発話入力テキスト入力を介して供給される、条項15から20のいずれか一項に記載の非一時的コンピュータ可読媒体。

22．人工ニューラルネットワーク(ANN)を使用してモバイルデバイス上でビデオを検索するための装置であって、

ANNによって、ビデオと検索クエリとを受信するための手段であって、ビデオが、フレームのシーケンス及び関連付けられたサブタイトル情報を含む、手段と、

10

20

30

40

50

モバイルデバイスにおいて、ANNによって、検索クエリ内の単語の第1のセットについての第1の表現と、サブタイトル情報内の単語の第2のセットについての第2の表現とを生成するための手段と、

モバイルデバイスにおいて、ANNによって、第1の表現及び第2の表現に基づいて相関を判定するための手段と、

モバイルデバイスにおいて、ANNによって、相関に基づいて、検索クエリに応じて、コンテンツを含むビデオの一部を予測するための手段と、

を備える、装置。

23. 検索クエリに応じて、コンテンツを含むビデオの一部についての開始時間及び終了時間を示す予測を生成するための手段を更に備える、条項22に記載の装置。

10

24. 開始時間に含まれているビデオの一部を終了時間まで表示するための手段を更に備える、条項22又は23に記載の装置。

25. ANNが、トランスフォーマニューラルネットワークを含む、条項22から24のいずれか一項に記載の装置。

26. ビデオとともに含まれるクロズドキャプション情報に基づいて、関連付けられたサブタイトル情報を生成するための手段を更に備える、条項22から25のいずれか一項に記載の装置。

27. 検索クエリが、シーン、イベント、単語、又は句の記述のうちの1つ又は複数を含む、条項22から26のいずれか一項に記載の装置。

28. 検索クエリが、モバイルデバイスの発話入力テキスト入力を介して供給される、条項22から27のいずれか一項に記載の装置。

20

【0069】

[0078] 一態様では、受信するための手段、生成するための手段、判定するための手段、予測するための手段、及び/又は表示するための手段は、列挙された機能を実行するように構成された、CPU102、CPU102に関連付けられたプログラムメモリ、専用メモリブロック118、全結合層362、NPU428、及び/又はルーティング接続処理ユニット216であってもよい。別の構成では、上述の手段は、上述の手段によって列挙された機能を実行するように構成された任意のモジュール又は任意の装置であってもよい。

【0070】

30

[0079] 上述の方法の様々な動作は、対応する機能を実行することが可能な、任意の好適な手段によって実行することができる。手段は、限定はしないが、回路、特定用途向け集積回路(ASIC)、又はプロセッサを含む、様々なハードウェア及び/又はソフトウェア構成要素(単数又は複数)及び/又はモジュール(単数又は複数)を含んでもよい。一般に、図に示される動作がある場合、それらの動作は、同様の番号を付された対応する相対物のミーンズプラスファンクション構成要素を有してもよい。

【0071】

[0080] 本明細書で使用される場合、「判定すること(determining)」という用語は、多種多様なアクションを包含する。例えば、「判定すること/判定すること」は、算出すること、計算すること、処理すること、導出すること、調査すること、ルックアップすること(例えば、テーブル、データベース、又は別のデータ構造においてルックアップすること)、確認することなどを含んでもよい。加えて、「判定すること」は、受信すること(例えば、情報を受信すること)、アクセスすること(例えば、メモリ内のデータにアクセスすること)などを含んでもよい。更に、「判定すること」は、解決すること、選択すること、選ぶこと、確立することなどを含んでもよい。

40

【0072】

[0081] 本明細書で使用される際、項目の列挙「のうちの少なくとも1つ」を指す句は、単一のメンバーを含む、それらの項目の任意の組み合わせを指す。一例として、「a、b、又はcのうちの少なくとも1つ」は、a、b、c、a-b、a-c、b-c、及びa-b-cを包含するものとする。

50

## 【 0 0 7 3 】

[0082] 本開示に関して説明される様々な例示的な論理ブロック、モジュール及び回路は、汎用プロセッサ、デジタル信号プロセッサ(DSP)、特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)若しくは他のプログラマブル論理デバイス(PLD)、個別ゲート若しくはトランジスタ論理、個別ハードウェア構成要素、又は説明された機能を実行するように設計されたそれらの任意の組み合わせを用いて実装又は実行されてもよい。汎用プロセッサはマイクロプロセッサであってもよいが、代替として、プロセッサは、任意の市販のプロセッサ、コントローラ、マイクロコントローラ、又はステートマシンであってもよい。プロセッサはまた、コンピューティングデバイスの組み合わせ、例えば、DSPとマイクロプロセッサとの組み合わせ、複数のマイクロプロセッサ、DSPコアと連携した1つ又は複数のマイクロプロセッサ、又は任意の他のそのような構成として実装されてもよい。

10

## 【 0 0 7 4 】

[0083] 本開示に関して説明される方法又はアルゴリズムのステップは、直接ハードウェアにおいて、プロセッサによって実行されるソフトウェアモジュールにおいて、又はこれら2つの組み合わせにおいて具現化されてもよい。ソフトウェアモジュールは、当技術分野で知られている任意の形態の記憶媒体内に存在してもよい。使用され得る記憶媒体のいくつかの例は、ランダムアクセスメモリ(RAM)、読取り専用メモリ(ROM)、フラッシュメモリ、消去可能プログラマブル読取り専用メモリ(EPROM)、電氣的消去可能プログラマブル読取り専用メモリ(EEPROM)、レジスタ、ハードディスク、リムーバブルディスク、CD-ROMなどを含む。ソフトウェアモジュールは、単一の命令又は多数の命令を備えてもよく、いくつかの異なるコードセグメントにわたって、異なるプログラム間で、複数の記憶媒体にわたって分散されてもよい。記憶媒体は、プロセッサが記憶媒体から情報を読み取り、記憶媒体に情報を書き込むことができるように、プロセッサに結合されてもよい。代替として、記憶媒体はプロセッサと一体化され得る。

20

## 【 0 0 7 5 】

[0084] 本明細書で開示される方法は、説明される方法を達成するための1つ若しくは複数のステップ又はアクションを含む。それらの方法のステップ及び/又はアクションは、特許請求の範囲から逸脱することなく、互いに入れ替えることができる。換言すれば、ステップ又はアクションの特定の順序が指定されていない限り、特定のステップ及び/又はアクションの順序、並びに/あるいは、それらのステップ及び/又はアクションの使用は、特許請求の範囲から逸脱することなく、修正することができる。

30

## 【 0 0 7 6 】

[0085] 説明された機能は、ハードウェア、ソフトウェア、ファームウェア、又はそれらの任意の組み合わせにおいて実装されてもよい。ハードウェアにおいて実装される場合、例示的なハードウェア構成は、デバイス内の処理システムを備えてもよい。処理システムは、バスアーキテクチャを用いて実装され得る。バスは、処理システムの特定の適用例、及び全体的な設計制約に応じて、任意の数の相互接続バス及びブリッジを含んでよい。バスは、プロセッサ、機械可読媒体、及びバスインターフェースを含む様々な回路を互いにリンクさせてもよい。バスインターフェースは、とりわけ、バスを介してネットワークアダプタを処理システムに接続するために使用され得る。ネットワークアダプタは、信号処理機能を実装するために使用され得る。いくつかの態様では、ユーザインターフェース(例えば、キーボード、ディスプレイ、マウス、ジョイスティックなど)もバスに接続され得る。バスは、タイミングソース、周辺装置、電圧調整器、電力管理回路などの様々な他の回路もリンクさせ得るが、これらの回路は当技術分野でよく知られており、したがって、これ以上は説明されない。

40

## 【 0 0 7 7 】

[0086] プロセッサは、バスを管理することと、機械可読媒体上に記憶されたソフトウェアの実行を含む一般的な処理とを担う場合がある。プロセッサは、1つ又は複数の汎用プロセッサ及び/又は専用プロセッサを用いて実装されてもよい。例としては、マイク

50

ロプロセッサ、マイクロコントローラ、DSPプロセッサ、及びソフトウェアを実行できる他の回路が挙げられる。ソフトウェアは、ソフトウェア、ファームウェア、ミドルウェア、マイクロコード、ハードウェア記述言語と呼ばれるか、又は他の名称で呼ばれるかにかかわらず、命令、データ、又はそれらの任意の組み合わせを意味するものと広く解釈されるものとする。機械可読媒体は、例として、ランダムアクセスメモリ(RAM)、フラッシュメモリ、読取り専用メモリ(ROM)、プログラマブル読取り専用メモリ(PROM)、消去可能プログラマブル読取り専用メモリ(EPROM)、電氣的消去可能プログラマブル読取り専用メモリ(EEPROM)、レジスタ、磁気ディスク、光ディスク、ハードドライブ、若しくは任意の他の好適な記憶媒体、又はそれらの任意の組み合わせを含んでもよい。機械可読媒体は、コンピュータプログラム製品において具現化されてもよい。コンピュータプログラム製品は、パッケージング材料を備えてもよい。

10

**【0078】**

【0087】ハードウェア実装形態では、機械可読媒体は、プロセッサとは別個の処理システムの一部であってもよい。しかしながら、当業者が容易に理解するように、機械可読媒体又はその任意の部分は処理システムの外部にあってもよい。例として、機械可読媒体は、伝送線路、データによって変調された搬送波、及び/又はデバイスとは別個のコンピュータ製品を含んでもよく、それらのすべては、バスインターフェースを通じてプロセッサによってアクセスされてもよい。代替的に、又は加えて、機械可読媒体又はその任意の部分は、キャッシュ及び/又は汎用レジスタファイルがそうであり得るように、プロセッサに統合されてもよい。論じられた様々な構成要素は、ローカル構成要素など、特定のロケーションを有するものとして説明されることがあるが、これらの構成要素は、分散コンピューティングシステムの一部として構成されるいくつかの構成要素など、様々な方法で構成されることもある。

20

**【0079】**

【0088】処理システムは、すべてが外部バスアーキテクチャを通じて他のサポート回路と互いにリンクされる、プロセッサ機能を提供する1つ又は複数のマイクロプロセッサと、機械可読媒体の少なくとも一部分を提供する外部メモリとを有する汎用処理システムとして構成されてもよい。代替的に、処理システムは、説明されたニューロンモデルとニューラルシステムのモデルとを実装するための1つ又は複数の神経形態学的プロセッサを備え得る。別の代替として、処理システムは、プロセッサを有する特定用途向け集積回路(AASIC)、バスインターフェース、ユーザインターフェース、サポート回路、及び単一のチップに統合された機械可読媒体の少なくとも一部分を用いて、あるいは1つ又は複数のフィールドプログラマブルゲートアレイ(FPGAs)、プログラマブル論理デバイス(PLDs)、コントローラ、ステートマシン、ゲート論理、個別ハードウェア構成要素、若しくは任意の他の好適な回路、又は本開示全体にわたって説明された様々な機能を実行することができる回路の任意の組み合わせを用いて実装されてもよい。当業者は、特定の適用例及びシステム全体に課される全体的な設計制約に応じて、処理システムについて説明された機能を最良に実装する方法を認識されよう。

30

**【0080】**

【0089】機械可読媒体は、いくつかのソフトウェアモジュールを備えてもよい。ソフトウェアモジュールは、プロセッサによって実行されると、処理システムに様々な機能を実行させる命令を含む。ソフトウェアモジュールは、送信モジュール及び受信モジュールを含んでもよい。各ソフトウェアモジュールは、単一の記憶デバイスの中に常駐してよく、又は複数の記憶デバイスにわたって分散されてよい。例として、ソフトウェアモジュールは、トリガイベントが発生したときにハードドライブからRAMの中にロードされてもよい。ソフトウェアモジュールの実行中、プロセッサは、アクセス速度を高めるために、命令のうちのいくつかをキャッシュの中にロードしてもよい。次いで、1つ又は複数のキャッシュラインは、プロセッサによる実行のために汎用レジスタファイルの中にロードされてもよい。以下でソフトウェアモジュールの機能に言及するとき、そのような機能が、そのソフトウェアモジュールからの命令を実行するとプロセッサによって実装されること

40

50

が理解されよう。更に、本開示の態様は、プロセッサ、コンピュータ、機械、又はそのような態様を実装する他のシステムの機能の改善をもたらすことを理解されたい。

【0081】

[0090] ソフトウェアにおいて実装される場合、機能は、1つ又は複数の命令又はコードとしてコンピュータ可読媒体上に記憶されるか、又はコンピュータ可読媒体を介して送信されてもよい。コンピュータ可読媒体は、ある場所から別の場所へのコンピュータプログラムの転送を容易にする任意の媒体を含む、コンピュータ記憶媒体と通信媒体の両方を含む。記憶媒体は、コンピュータによってアクセスされ得る任意の利用可能な媒体であり得る。限定ではなく例として、そのようなコンピュータ可読媒体は、RAM、ROM、EEPROM、CD-ROM若しくは他の光ディスクストレージ、磁気ディスクストレージ若しくは他の磁気記憶デバイス、又は命令若しくはデータ構造の形態の所望のプログラムコードを搬送若しくは記憶するために使用され得るとともに、コンピュータによってアクセスされ得る任意の他の媒体を備えることができる。加えて、いかなる接続もコンピュータ可読媒体と適切に呼ばれる。例えば、ソフトウェアが、同軸ケーブル、光ファイバケーブル、ツイストペア、デジタル加入者回線(DSL)、又は赤外線(IR)、無線、及びマイクロ波などのワイヤレス技術を使用して、ウェブサイト、サーバ、又は他のリモートソースから送信される場合、同軸ケーブル、光ファイバケーブル、ツイストペア、DSL、又は赤外線、無線、及びマイクロ波などのワイヤレス技術は、媒体の定義に含まれる。本明細書で使用されるディスク(disk)及びディスク(disc)は、コンパクトディスク(disc)(CD)、レーザーディスク(disc)、光ディスク(disc)、デジタル多用途ディスク(disc)(DVD)、フロッピーディスク(disk)、及びBlu-ray(登録商標)ディスク(disc)を含み、ディスク(disk)は通常、データを磁氣的に再生し、ディスク(disc)は、レーザーを用いてデータを光学的に再生する。したがって、いくつかの態様では、コンピュータ可読媒体は、非一時的コンピュータ可読媒体(例えば、有形媒体)を備えてよい。加えて、他の態様では、コンピュータ可読媒体は、一時的コンピュータ可読媒体(例えば、信号)を備えてもよい。上記のものの組み合わせもまた、コンピュータ可読媒体の範囲内に含まれるものとする。

【0082】

[0091] したがって、いくつかの態様は、本明細書で提示した動作を実行するためのコンピュータプログラム製品を含み得る。例えば、そのようなコンピュータプログラム製品は、命令を記憶した(及び/又は符号化した)コンピュータ可読媒体を備えてもよく、命令は、説明された動作を実行するように1つ又は複数のプロセッサによって実行可能である。いくつかの態様では、コンピュータプログラム製品は、パッケージング材料を含んでもよい。

【0083】

[0092] 更に、説明された方法及び技法を実行するためのモジュール及び/又は他の適切な手段は、適用可能な場合、ユーザ端末及び/又は基地局によってダウンロード及び/又は別の方法で取得され得ることを理解されたい。例えば、そのようなデバイスは、説明された方法を実行するための手段の転送を容易にするためにサーバに結合され得る。代替的に、説明された様々な方法は、ユーザ端末及び/又は基地局が記憶手段(例えば、RAM、ROM、コンパクトディスク(CD)又はフロッピーディスクなどの物理記憶媒体など)をデバイスに結合又は提供すると様々な方法を取得することができるように、記憶手段を介して提供され得る。更に、説明された方法及び技法をデバイスに提供するための任意の他の好適な技法が利用され得る。

【0084】

[0093] 特許請求の範囲は、上記で例示されている厳密な構成及び構成要素に限定されるものではない点を理解されたい。特許請求の範囲から逸脱することなく、上述した方法及び装置の構成、動作、及び詳細において、様々な修正、変更、及び変形が行われ得る。

。

10

20

30

40

50

【図面】

【図 1】

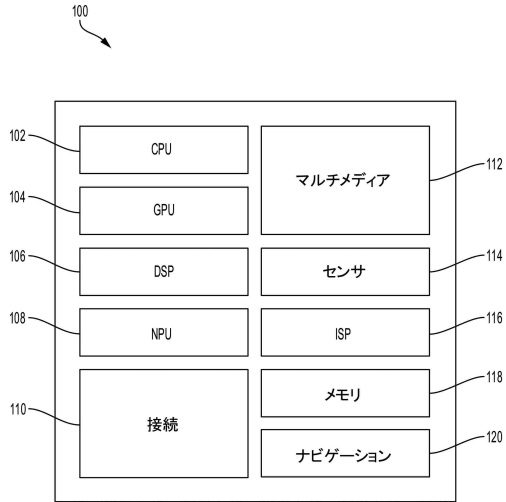


FIG. 1

【図 2 A】

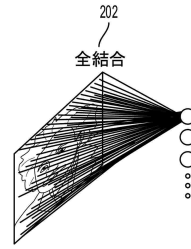


FIG. 2A

10

20

【図 2 B】

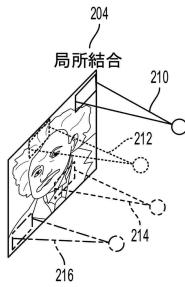


FIG. 2B

【図 2 C】

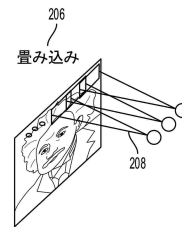


FIG. 2C

30

40

50

【 図 2 D 】

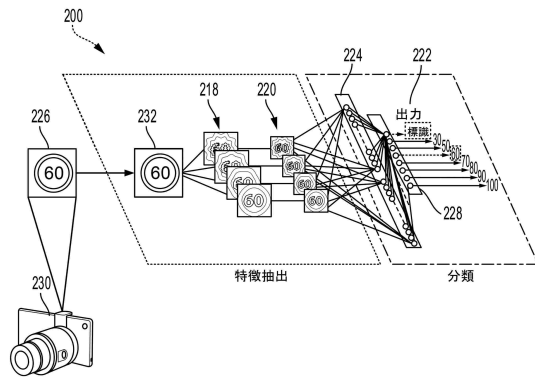


FIG. 2D

【 図 3 】

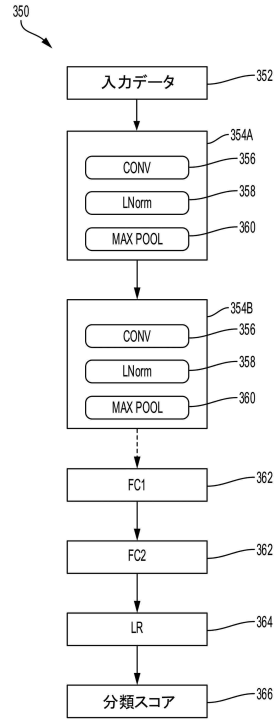


FIG. 3

10

20

【 図 4 】

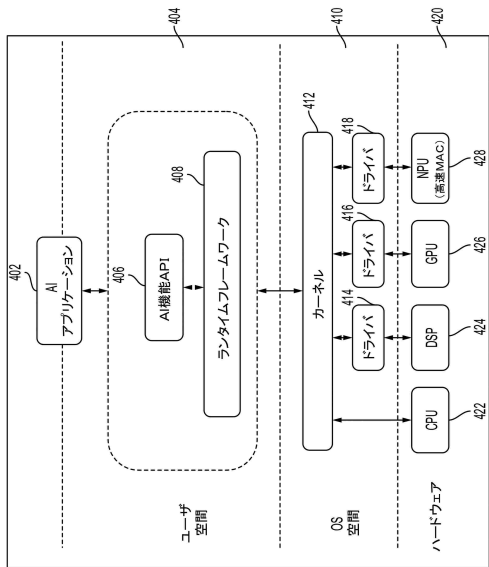


FIG. 4

【 図 5 】

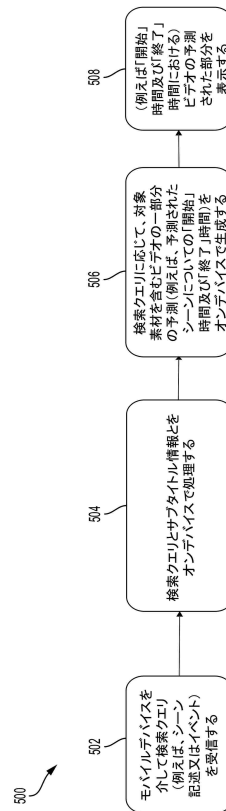


FIG. 5

30

40

50

【 図 6 A 】

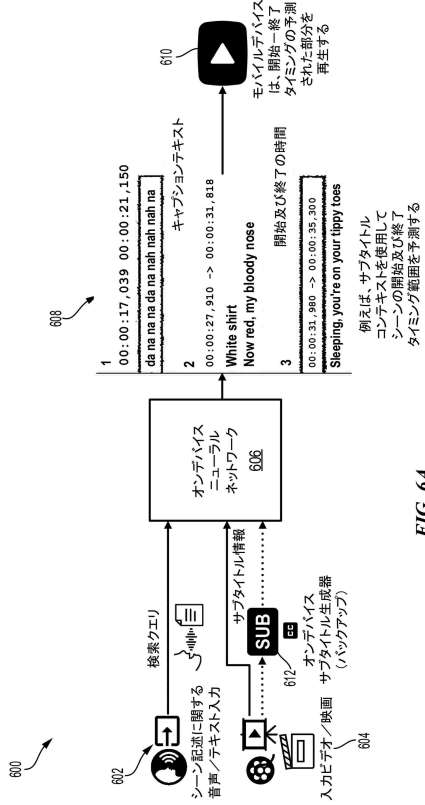


FIG. 6A

【 図 6 B 】

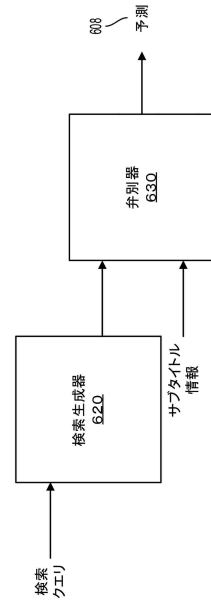


FIG. 6B

【 図 7 】

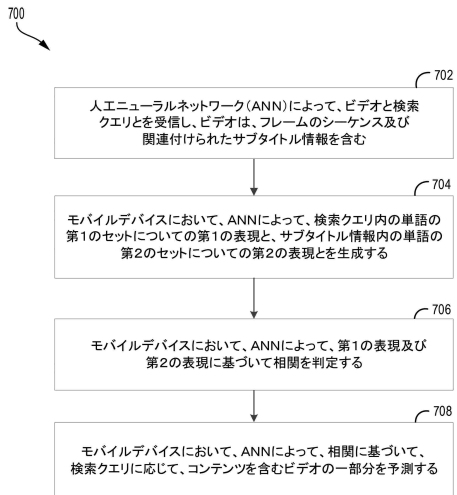


FIG. 7

10

20

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2023/013252

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> <b>INV. G06F16/783</b> <b>ADD.</b>  According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) <b>G06F</b>		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) <b>EPO-Internal, WPI Data</b>		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<b>US 2021/193187 A1 (PHILLIPS CALEB [CA] ET AL) 24 June 2021 (2021-06-24)</b> <b>paragraphs [0003] - [0005]</b> <b>paragraphs [0046], [0067], [0083], [0098]</b>	1-28
X	<b>US 2021/109966 A1 (AYUSH KUMAR [IN] ET AL) 15 April 2021 (2021-04-15)</b> <b>paragraph [0044]</b>	1-28
<input type="checkbox"/> Further documents are listed in the continuation of Box C.		
<input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family	
Date of the actual completion of the international search	Date of mailing of the international search report	
22 May 2023	31/05/2023	
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer  Correia Martins, F	

Form PCT/ISA/210 (second sheet) (April 2005)

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No  
**PCT/US2023/013252**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
<b>US 2021193187 A1</b>	<b>24-06-2021</b>	<b>US 2021193187 A1</b>	<b>24-06-2021</b>
		<b>WO 2021132802 A1</b>	<b>01-07-2021</b>
-----			
<b>US 2021109966 A1</b>	<b>15-04-2021</b>	<b>NONE</b>	
-----			

10

20

30

40

50

## フロントページの続き

,MC,ME,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),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,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

## 1 . B L U E T O O T H

2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5

(72)発明者 ナディパリ、シャラト・チャンドラ

アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5

(72)発明者 コンダパーティ、サイクマー

アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5

Fターム(参考) 5B175 DA04 FB03