(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年6月2日(02.06.2022)

WIPO | PCT

(10) 国際公開番号
WO 2022/113293 A1

(51) 国際特許分類:
G06T 11/60 (2006.01)

(21) 国際出願番号： PCT/JP2020/044317

(22) 国際出願日： 2020年11月27日(27.11.2020)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(71) 出願人: 日本電信電話株式会社 (**NIPPON TELEGRAPH AND TELEPHONE CORPORATION**) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目５番１号 Tokyo (JP).

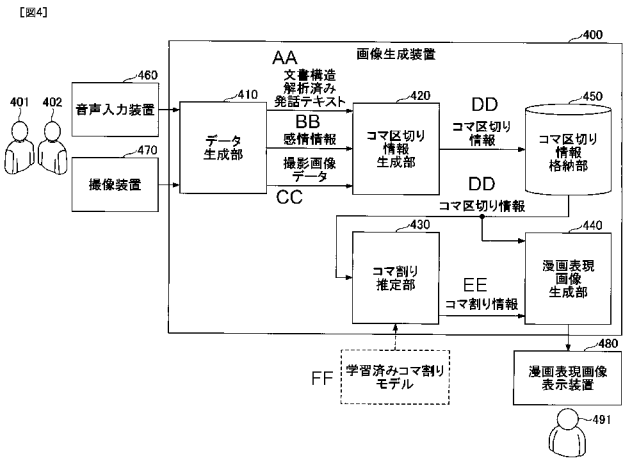
(72) 発明者: 瀬古 俊一(**SEKO, Shunichi**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ知的財産センタ内 Tokyo (JP).

(74) 代理人: 伊東 忠重，外(**ITOH, Tadashige** et al.); 〒1000005 東京都千代田区丸の内二丁目１番１号 丸の内 ＭＹ ＰＬＡＺＡ （明治安田生命ビル） １６階 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, 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,

(54) **Title:** IMAGE GENERATION DEVICE, IMAGE GENERATION METHOD, AND IMAGE GENERATION PROGRAM

(54) 発明の名称： 画像生成装置、画像生成方法及び画像生成プログラム



400 Image generation device
410 Data generation unit
420 Frame break information generation unit
430 Frame split estimation unit
440 Cartoon representation image generation unit
450 Frame break information storage unit
460 Speech input device
470 Image capturing device
480 Cartoon representation image display device
AA Document structure-analyzed utterance text
BB Emotion information
CC Captured image data
DD Frame break information
EE Frame split information
FF Trained frame split model

(57) **Abstract:** The present invention, when an utterance scene is represented in a cartoon style, makes utterance content be easily understood. This image generation device comprises: an estimation unit which inputs, on the basis of an existing cartoon image, information generated on the basis of speech data and captured image data of an utterer in the utterance scene to a trained model that has been trained with frame assignment, and estimates the frame assignment; and an image generation unit which generates an image representing the utterance scene on the basis of the frame assignment estimated by the estimation unit.

(57) 要約：発話シーンを漫画風に表現する際、発話内容を理解しやすくする。画像生成装置は、既存の漫画画像に基づいて、コマ割りを学習した学習済みモデルに、発話シーンにおける話者の音声データと撮影画像データとに基づいて生成した情報を入力し、コマ割りを推定する推定部と、前記推定部により推定されたコマ割りに基づいて、前記発話シーンを示す画像を生成する画像生成部とを有する。

WO 2022/113293 A1

［続葉有］

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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).

添付公開書類:

一　国際調査報告（条約第21条(3)）

明　細　書

発明の名称 ：

画像生成装置、画像生成方法及び画像生成プログラム

技術分野

[0001] 本開示は、画像生成装置、画像生成方法及び画像生成プログラムに関する。

背景技術

[0002] 発話中の話者を撮影した撮影画像データや、発話中の音声を記録した音声データを用いて、発話シーンを漫画風に表現した画像（「漫画表現画像」と称す）を生成する画像生成技術が知られている。当該技術を用いれば、画像の閲覧者（例えば、視覚優位な人、聴覚からでは情報を理解しづらい人等）による発話内容の理解を支援することができる。

先行技術文献

非特許文献

[0003] 非特許文献1：Maciej Pesko, Adam Svystun, Pawel Andruszkiewicz, Przemyslaw Rokita, Tomasz Trzcinski, "Comixify: Transform video into a comics", arXiv:1812.03473v1 [cs.CV], December 11, 2018

発明の概要

発明が解決しようとする課題

[0004] しかしながら、上記画像生成技術の場合、漫画特有の表現方法（コマ割り、吹き出し等）が適用されていないため、発話中の話者の感情が閲覧者に伝わりにくいという問題がある。このため、閲覧者による発話内容の理解が十分に進まない場合がある。

[0005] 本開示は、発話シーンを漫画風に表現する際、発話内容を理解しやすくすることを目的とする。

課題を解決するための手段

[0006] 本開示の一態様によれば、画像生成装置は、

既存の漫画画像に基づいて、コマ割りを学習した学習済みモデルに、発話シーンにおける話者の音声データと撮影画像データとに基づいて生成した情報を入力し、コマ割りを推定する推定部と、

前記推定部により推定されたコマ割りに基づいて、前記発話シーンを示す画像を生成する画像生成部とを有する。

発明の効果

[0007] 本開示によれば、発話シーンを漫画風に表現する際、発話内容を理解しやすくすることができる。

図面の簡単な説明

[0008] [図1]図１は、漫画特有の表現方法による効果を説明するための第１の図である。

[図2]図２は、漫画特有の表現方法による効果を説明するための第２の図である。

[図3]図３は、学習装置の一例を示す図である。

[図4]図４は、画像生成装置の一例を示す図である。

[図5]図５は、学習装置または画像生成装置のハードウェア構成の一例を示す図である。

[図6]図６は、学習用データ生成部の処理の具体例を示す図である。

[図7]図７は、コマ割り学習部の処理の具体例を示す図である。

[図8]図８は、学習処理の流れを示すフローチャートである。

[図9]図９は、データ生成部の処理の具体例を示す図である。

[図10]図１０は、コマ区切り情報生成部の処理の具体例を示す図である。

[図11]図１１は、コマ区切り情報の具体例を示す図である。

[図12]図１２は、コマ割り推定部の処理の具体例を示す図である。

[図13]図１３は、漫画表現画像生成部の処理の具体例を示す図である。

[図14]図１４は、画像生成処理の流れを示すフローチャートである。

発明を実施するための形態

[0009] 以下、各実施形態について添付の図面を参照しながら説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複した説明を省略する。

[0010] ［第１の実施形態］

＜漫画特有の表現方法による効果の説明＞

はじめに、漫画特有の表現方法（コマ割り、吹き出し等）による効果について説明する。図１及び図２は、漫画特有の表現方法による効果を説明するための第１及び第２の図である。

[0011] 図１及び図２に示す漫画画像１１０、漫画画像１２０、漫画画像２１０、漫画画像２２０は、いずれも、二人の話者の発話内容を吹き出しで示した、発話シーンを示す漫画画像である。

[0012] このうち、図１の漫画画像１１０は、１コマの漫画画像である。また、図１の漫画画像２２０は、２コマの漫画コマ画像１２１、１２２を有する漫画画像である。このように、同じ発話シーンを示す漫画画像であっても、コマ割りによって、閲覧者が受ける印象は変わってくる。

[0013] 更に、図２の漫画画像２１０は、３コマの漫画コマ画像２１１～２１３を有し、３コマ目の話者の大きさを２コマ目の話者の大きさよりも大きくした漫画画像である。このように、コマの数を増やし、発話をしていない漫画コマ画像２１２を入れることで、話者が戸惑っている感情を表現することができる。また、３コマ目の話者の大きさを大きくすることで、話者が意を決して発話した感情を表現することができる。

[0014] また、図３の漫画画像２２０は、３コマの漫画コマ画像２２１～２２３を有し、３コマ目のコマの大きさを、他のコマの大きさよりも大きくしてレイアウトし、かつ、効果線を入れた漫画画像である。このように、コマの大きさを変えてレイアウトし、３コマ目のコマに効果線を入れることで、話者が意を決して発話した感情を、より効果的に表現することができるとともに、他方の話者が衝撃を受けた感情を表現することもできる。

[0015] 以上のように、漫画特有の表現方法を適切に用いることで、閲覧者は、話

者の感情を考慮しながら発話内容を理解することができるようになる。そこで、以下に説明する第１の実施形態に係る画像生成装置では、発話シーンから漫画表現画像を生成するにあたり、漫画特有の表現方法が適切に用いられるように構成する。具体的には、学習装置が既存の漫画画像に基づいて漫画特有の表現方法（コマ割り）を学習し、画像生成装置が、当該学習結果（学習済みモデル）を用いて発話シーンから漫画表現画像を生成する。これにより、第１の実施形態に係る画像生成装置によれば、漫画特有の表現方法が適切に用いられた漫画表現画像を生成することができ、発話内容を理解しやすくすることができる。

[0016] ＜学習装置の説明＞

はじめに、漫画特有の表現方法を学習する学習装置について説明する。

[0017] 図３は、学習装置の一例を示す図である。学習装置３００には、学習プログラムがインストールされており、当該プログラムが実行されることで、学習装置３００は、学習用データ生成部３１０、コマ割り学習部３２０として機能する。

[0018] 学習用データ生成部３１０は、既存の漫画画像が複数格納された漫画画像格納部３３０より漫画画像を読み出す。また、学習用データ生成部３１０は、読み出した複数の漫画画像それぞれについて、作品ごと、コマの順番ごとにナンバリングを行い、コマごとに分割した漫画コマ画像を生成する。

[0019] また、学習用データ生成部３１０は、生成した漫画コマ画像それぞれに対して画像認識処理を行うことで、「漫画コマ情報」を抽出する。なお、漫画コマ情報には、対応する漫画コマ画像内の登場人物のセリフ等の発話テキスト、登場人物の感情、シーンを示す情報等が含まれる。

[0020] 更に、学習用データ生成部３１０は、漫画コマ画像及び漫画コマ情報を入力データ、漫画画像を正解データとする学習用データ３４１を生成し、学習用データ格納部３４０に格納する。

[0021] コマ割り学習部３２０は、学習用データ格納部３４０より学習用データ３４１を読み出し、漫画コマ画像及び漫画コマ情報をコマ割りモデルに入力す

る。そして、コマ割り学習部３２０は、コマ割りモデルの出力が、学習用データ３４１の正解データである漫画画像に近づくよう、コマ割りモデルに対して学習処理を行う。

[0022] また、コマ割り学習部３２０は、コマ割りモデルに対する学習処理により、漫画コマ画像及び漫画コマ情報と、漫画画像（コマ割りされたもの）との関係を学習することで、学習済みコマ割りモデルを生成し、後述する画像生成装置に出力する。

[0023] ＜画像生成装置の説明＞

次に、学習済みコマ割りモデルを用いて、発話シーンから漫画表現画像を生成する画像生成装置について説明する。図４は、画像生成装置の一例を示す図である。画像生成装置４００には、画像生成プログラムがインストールされており、当該プログラムが実行されることで、画像生成装置４００は、データ生成部４１０、コマ区切り情報生成部４２０として機能する。また、画像生成装置４００は、コマ割り推定部４３０、漫画表現画像生成部４４０として機能する。

[0024] データ生成部４１０は、画像生成装置４００と接続される音声入力装置４６０（例えば、マイク）より、話者４０１及び話者４０２の発話から検出された音声データを取得する。

[0025] また、データ生成部４１０は、画像生成装置４００と接続される撮像装置４７０（例えば、カメラ）より、話者４０１及び話者４０２を撮影した撮影画像データを取得する。

[0026] また、データ生成部４１０は、音声データから発話テキストを抽出し、文書構造の解析を行うことで、文書構造解析済み発話テキストを生成する。

[0027] また、データ生成部４１０は、音声データと撮影画像データとに基づいて、話者の感情を示す感情情報を生成する。

[0028] 更に、データ生成部４１０は、生成した文書構造解析済み発話テキスト、感情情報、及び、撮影画像データをコマ区切り情報生成部４２０に通知する。

[0029] コマ区切り情報生成部４２０は情報生成部の一例である。コマ区切り情報生成部４２０は、文書構造解析済み発話テキスト、感情情報、撮影画像データに基づいて、１コマとして区切る単位を抽出し、コマ区切り情報を生成する。コマ区切り情報生成部４２０は、生成したコマ区切り情報を、コマ区切り情報格納部４５０に格納する。

[0030] コマ割り推定部４３０は推定部の一例である。コマ割り推定部４３０は、コマ区切り情報格納部４５０より、コマ区切り情報を読み出し、学習済みコマ割りモデルに入力する。また、コマ割り推定部４３０は、コマ区切り情報を入力したことで学習済みコマ割りモデルにより推定された、適切なコマ割りを示すコマ割り情報を、漫画表現画像生成部４４０に出力する。

[0031] 漫画表現画像生成部４４０は画像生成部の一例である。漫画表現画像生成部４４０は、コマ区切り情報格納部４５０より、コマ区切り情報を読み出し、コマ割り情報に基づくコマ割りで漫画表現に変換することで、漫画表現画像を生成する。また、漫画表現画像生成部４４０は、生成した漫画表現画像を漫画表現画像表示装置４８０に送信することで、閲覧者４９１に表示する。

[0032] ＜学習装置及び画像生成装置のハードウェア構成＞

次に、学習装置３００及び画像生成装置４００のハードウェア構成について説明する。なお、学習装置３００と画像生成装置４００とは、同様のハードウェア構成を有することから、ここでは、図５を用いて、まとめて説明する。図５は、学習装置または画像生成装置のハードウェア構成の一例を示す図である。

[0033] 図５に示すように、学習装置３００または画像生成装置４００は、プロセッサ５０１、メモリ５０２、補助記憶装置５０３、操作装置５０４、表示装置５０５、Ｉ／Ｆ（Interface）装置５０６、通信装置５０７、ドライブ装置５０８を有する。なお、学習装置３００または画像生成装置４００の各ハードウェアは、バス５０９を介して相互に接続される。

[0034] プロセッサ５０１は、ＣＰＵ（Central Processing Unit）、ＧＰＵ（Grap

hics Processing Unit）等の各種演算デバイスを有する。プロセッサ５０１は、各種プログラム（例えば、学習プログラムまたは画像生成プログラム等）をメモリ５０２上に読み出して実行する。

[0035] メモリ５０２は、ＲＯＭ（Read Only Memory）、ＲＡＭ（Random Access Memory）等の主記憶デバイスを有する。プロセッサ５０１とメモリ５０２とは、いわゆるコンピュータを形成し、プロセッサ５０１が、メモリ５０２上に読み出した各種プログラムを実行することで、当該コンピュータは各種機能を実現する。なお、ここでいう各種機能とは、学習装置３００の場合、例えば、学習用データ生成部３１０～コマ割り学習部３２０を指し、画像生成装置４００の場合、例えば、データ生成部４１０～漫画表現画像生成部４４０を指す。

[0036] 補助記憶装置５０３は、各種プログラムや、各種プログラムがプロセッサ５０１によって実行される際に用いられる各種データを格納する。例えば、学習装置３００の場合は、漫画画像格納部３３０、学習用データ格納部３４０が、画像生成装置４００の場合は、コマ区切り情報格納部４５０が、それぞれ補助記憶装置５０３において実現される。

[0037] 操作装置５０４は、学習装置３００または画像生成装置４００の管理者が学習装置３００または画像生成装置４００に対して各種指示を入力する際に、管理者の各種操作を受け付ける。表示装置５０５は、学習装置３００または画像生成装置４００の管理者に対して、学習装置３００または画像生成装置４００の内部状態を表示する。

[0038] Ｉ／Ｆ装置５０６は、学習装置３００または画像生成装置４００と、外部装置５２０とを接続する。画像生成装置４００の場合、Ｉ／Ｆ装置５０６を介して、例えば、外部装置５２０である音声入力装置４６０や撮像装置４７０と接続される。

[0039] 通信装置５０７は、学習装置３００または画像生成装置４００がネットワークを介して外部装置と通信可能に接続するための通信デバイスである。画像生成装置４００の場合、通信装置５０７は、漫画表現画像表示装置４８０

と通信可能に接続される。

[0040] ドライブ装置５０８は記録媒体５３０をセットするためのデバイスである。ここでいう記録媒体５３０には、ＣＤ－ＲＯＭ、フレキシブルディスク、光磁気ディスク等のように情報を光学的、電気的あるいは磁気的に記録する媒体が含まれる。また、記録媒体５３０には、ＲＯＭ、フラッシュメモリ等のように情報を電気的に記録する半導体メモリ等が含まれていてもよい。

[0041] なお、補助記憶装置５０３にインストールされる各種プログラムは、例えば、配布された記録媒体５３０がドライブ装置５０８にセットされ、該記録媒体５３０に記録された各種プログラムがドライブ装置５０８により読み出されることでインストールされる。あるいは、補助記憶装置５０３にインストールされる各種プログラムは、不図示のネットワークを介してダウンロードされることで、インストールされてもよい。

[0042] ＜学習装置の各部の処理の具体例＞

次に、学習装置３００の各部（学習用データ生成部３１０、コマ割り学習部３２０）の処理の具体例について説明する。

[0043] （１）学習用データ生成部３１０の処理の具体例

図６は、学習用データ生成部の処理の具体例を示す図である。図６に示すように、学習用データ生成部３１０は、更に、コマ分割処理部６１０とコマ情報抽出部６２０とを有する。

[0044] コマ分割処理部６１０は、漫画画像格納部３３０より漫画画像を読み出す。図６の例は、漫画画像２２０が、作品名＝”Ａ”、ページ＝”１”の漫画画像であることを示している。

[0045] また、コマ分割処理部６１０は、漫画画像２２０内のエッジを検出し、コマとしての区切り線（多角形で囲まれている領域）ごとに漫画画像２２０を分割することで、漫画コマ画像２２１～２２３を生成する。また、コマ分割処理部６１０は、漫画コマ画像２２１～２２３について、上かつ右に位置する漫画コマ画像から順番にナンバリングを行う。

[0046] 図６の例は、コマ分割処理部６１０が、漫画コマ画像２２１に”１”のナン

バリングを行い、漫画コマ画像２２２に”２”のナンバリングを行い、漫画コマ画像２２３に”３”のナンバリングを行った様子を示している。

[0047] コマ情報抽出部６２０は、ナンバリングが行われた漫画コマ画像２２１～２２３に対して画像認識を行い、登場人物のセリフ等の発話テキスト、登場人物の感情、シーンを示す情報を抽出することで、漫画コマ情報６３１～６３３を生成する。

[0048] 図６の例は、コマ情報抽出部６２０が、漫画コマ画像２２１から、テキスト＝”お昼はラーメンにしようか”、感情＝”気楽”、シーン＝”提案”を抽出することで、漫画コマ情報６３１を生成した様子を示している。

[0049] 同様に、図６の例は、コマ情報抽出部６２０が、漫画コマ画像２２２から、テキスト＝”ｎｕｌｌ”、感情＝”ニュートラル”、シーン＝”間”を抽出することで、漫画コマ情報６３２を生成した様子を示している。

[0050] 同様に、図６の例は、コマ情報抽出部６２０が、漫画コマ画像２２３から、テキスト＝”太るから嫌”、感情＝”嫌悪”、シーン＝”提案”を抽出することで、漫画コマ情報６３３を生成した様子を示している。

[0051] なお、学習用データ生成部３１０は、漫画コマ画像２２１～２２３及び漫画コマ情報６３１～６３３を学習用データ３４１の入力データとして、漫画画像２２０を学習用データ３４１の正解データとして、学習用データ格納部３４０に格納する。

[0052] （２）コマ割り学習部３２０の処理の具体例

図７は、コマ割り学習部の処理の具体例を示す図である。図７に示すように、コマ割り学習部３２０は、更に、コマ割りモデル７１０と比較／変更部７２０とを有する。

[0053] コマ割りモデル７１０は、例えば、ＲＮＮ（Recurrent Neural Network）や、ＬＳＴＭ（Long short-term memory）等により構成される。コマ割り学習部３２０は、学習用データ３４１の入力データである漫画コマ画像２２１～２２３及び漫画コマ情報６３１～６３３をコマ割りモデル７１０に入力することで、コマ割りモデル７１０を実行させる。また、コマ割り学習部３２

０は、コマ割りモデル７１０より出力された漫画画像を、比較／変更部７２０に入力する。

[0054] 比較／変更部７２０は、コマ割りモデル７１０より出力された漫画画像が、学習用データ３４１の正解データである漫画画像（例えば、漫画画像２２０）に近づくように、コマ割りモデル７１０のモデルパラメータを更新する。このように、コマ割り学習部３２０では、コマ割りモデル７１０に対して学習処理を行い、漫画コマ画像及び漫画コマ情報と、漫画画像（コマ割りされたもの）との関係を学習することで、学習済みコマ割りモデルを生成する。

[0055] ＜学習処理の流れ＞

次に、学習装置３００による学習処理の流れについて説明する。図８は、学習処理の流れを示すフローチャートである。なお、学習処理を開始するにあたり、漫画画像格納部３３０には、複数の漫画画像が格納されているものとする。

[0056] ステップＳ８０１において、学習用データ生成部３１０は、漫画画像格納部３３０より漫画画像を読み出す。

[0057] ステップＳ８０２において、学習用データ生成部３１０は、読み出した漫画画像をコマごとに分割し、漫画コマ画像を生成する。

[0058] ステップＳ８０３において、学習用データ生成部３１０は、生成した漫画コマ画像に対して画像認識を行い、漫画コマ情報を生成する。

[0059] ステップＳ８０４において、学習用データ生成部３１０は、生成した漫画コマ画像及び漫画コマ情報と漫画画像とを対応付け、学習用データとして、学習用データ格納部３４０に格納する。

[0060] ステップＳ８０５において、コマ割り学習部３２０は、学習用データ格納部３４０より学習用データを読み出し、コマ割りモデルに入力することで、コマ割りモデルに対して学習処理を行う。

[0061] ステップＳ８０６において、コマ割り学習部３２０は、コマ割りモデルに対する学習処理を終了するか否かを判定する。ステップＳ８０６において、

コマ割りモデルに対する学習処理を終了しないと判定した場合には（ステップＳ８０６においてＮＯの場合には）、ステップＳ８０１に戻る。一方、ステップＳ８０６において、コマ割りモデルに対する学習処理を終了すると判定した場合には（ステップＳ８０６においてＹＥＳの場合には）、ステップＳ８０７に進む。

[0062] ステップＳ８０７において、コマ割り学習部３２０は、生成した学習済みコマ割りモデルを出力し、学習処理を終了する。

[0063] ＜画像生成装置の各部の処理の具体例＞

次に、画像生成装置４００の各部（データ生成部４１０、コマ区切り情報生成部４２０、コマ割り推定部４３０、漫画表現画像生成部４４０）の処理の具体例について説明する。

[0064] （１）データ生成部４１０の処理の具体例

図９は、データ生成部４１０の処理の具体例を示す図である。図９に示すように、データ生成部４１０は、更に、音声認識部９１０、文書構造解析部９２０、感情情報推定部９３０を有する。

[0065] 音声認識部９１０は、音声データから発話テキストを抽出し、文書構造解析部９２０に通知する。

[0066] 文書構造解析部９２０は、発話テキストの文書構造を解析し、文書構造解析済み発話テキストを生成する。なお、文書構造解析部９２０による解析は、周知の技術であるSpeech to Text（音声認識機能）や自然文解析機能等を利用する。

[0067] 感情情報推定部９３０は、音声データと撮影画像データとに基づいて、感情情報を生成する。なお、感情情報推定部９３０による感情情報の生成は、周知の技術（例えば、特許第５４３７２９７や特開２０１９－２８４８５号公報等に開示された技術）を利用する。周知の技術を利用するにあたっては、データ生成部４１０は、例えば、撮影画像データより抽出される視線情報を用いて、共感情報を取得するように構成されてもよい。あるいは、データ生成部４１０は、新たに呼吸運動情報を取得し、取得した呼吸運動情報に基

づいて算出される盛り上がり情報を取得するように構成されてもよい。

[0068] （２）コマ区切り情報生成部４２０による処理の具体例

図１０は、コマ区切り情報生成部による処理の具体例を示す図である。図１０に示すように、コマ区切り情報生成部４２０は、更に、発話区間分割部１０１０、感情情報判定部１０２０、発話区間連結部１０３０、コマ区切り情報出力部１０４０を有する。

[0069] 発話区間分割部１０１０は、文書構造解析済み発話テキストに沿って、句点を区切りとして、時刻軸を複数の発話区間に分割する。図１０において、グラフ１０５０は、横軸に時刻を、縦軸に感情情報の変化をとったグラフである。グラフ１０５０において、縦方向の点線は、時刻軸を発話区間ごとに分割した様子を示している。

[0070] 感情情報判定部１０２０は、感情情報の変化が大きい発話区間を判定する。図１０において、グラフ１０５０の実線１０５１は、各時刻における感情情報の変化を示している。

[0071] なお、感情情報の変化は、０～１の範囲の値となるように、予め正規化されているものとする。具体的には、感情情報として共感情報を用いる場合には、共感情報を正規化するために、例えば、従来技術である特許第５４３７２９７号に開示された方法が用いられるものとする。また、感情情報として盛り上がり情報を用いる場合には、盛り上がり情報を正規化するために、例えば、従来技術である特開２０１９－２８４８５号公報に開示された方法が用いられるものとする。

[0072] また、感情情報の変化（Ｍ）は、例えば、下式に基づいて算出されるものとする。

[0073] Ｍ＝（２×ＨＵＰ×ＥＭＴ）／（ＨＵＰ＋ＥＭＴ）

なお、上式において、ＨＵＰは、正規化した盛り上がり度［０≦ＨＵＰ≦１］を表し、ＥＭＴは正規化した共感度［０≦ＥＭＴ≦１］を表す。

[0074] 発話区間連結部１０３０は、発話区間分割部１０１０により時刻軸が分割されることで得た複数の発話区間について、下記の２つの条件（（ｉ）、（

ｉｉ））のうちの少なくともいずれか一方を満たすか否かを判定する。そして、発話区間連結部１０３０は、いずれか一方に該当すると判定した発話区間について、発話内容が同一の発話区間であると判定し、発話区間グループを生成する。

（ｉ）感情情報の変化が大きいと判定され、文書構造的にも関連している発話区間である（例えば、符号１０６２）。

（ｉｉ）発話の開始タイミングや、文書構造的に主張ポイントであるなど、発話内容が重要であると思われる発話区間である（例えば、符号１０６１、１０６３）。

[0075] また、発話区間連結部１０３０は、生成した発話区間グループの中から、以下の３つの条件（（ａ）、（ｂ）、（ｃ））のうちのいずれかを満たす発話区間グループを抽出する。

[0076] （ａ）発話区間グループの位置が最初の位置である

発話の開始時には、話題や発話のきっかけ等のように、重要な情報が含まれている可能性が高い。このため、発話区間連結部１０３０では、最初の位置の発話区間グループを抽出する（例えば、発話区間グループ１０７１参照）。

[0077] （ｂ）感情情報の変化が所定の閾値を超えた発話区間が含まれる

感情情報が変化した際の話者の発話には、主張や心を動かす情報が含まれている可能性が高い。このため、発話区間連結部１０３０では、感情情報の変化が所定の閾値を超えた発話区間が含まれる発話区間グループを抽出する（例えば、発話区間グループ１０７２、１０７３参照）。なお、所定の閾値は、例えば、画像生成装置４００の管理者によって予め設定されているものとする。

[0078] （ｃ）論理構成的な重要度が所定の閾値を超えた発話区間が含まれる。

[0079] 論理構成的に重要な発話には、主張や説明などの情報が含まれている可能性が高い。このため、発話区間連結部１０３０では、発話内容についての論理構成的な重要度が所定の閾値を超えた発話区間を含む発話区間グループを

抽出する（例えば、発話区間グループ１０７４参照）。なお、所定の閾値は、例えば、画像生成装置４００の管理者によって予め設定されているものとする。また、論理構成的な重要度は、従来技術の方法を用いて算出され、０～１の範囲の値となるように、予め正規化されているものとする。

[0080] コマ区切り情報出力部１０４０は、生成された発話区間グループそれぞれに対応する撮影画像データから、漫画表現画像を生成する際に用いる代表画像を抽出する。なお、代表画像を抽出するにあたり、コマ区切り情報出力部１０４０は、既知のハイライト画像抽出技術等を用いる。

[0081] また、コマ区切り情報出力部１０４０は、抽出した代表画像を、対応する発話区間グループの発話テキスト、及び、対応する発話区間グループの感情情報と対応付けることで、コマ区切り情報を生成する。更に、コマ区切り情報出力部１０４０は、生成したコマ区切り情報を、コマ区切り情報格納部４５０に格納する。

[0082] 図１１は、コマ区切り情報の具体例を示す図である。図１１に示すように、コマ区切り情報１１００には、情報の項目として、”発話区間グループＩＤ”、”発話テキスト”、”感情情報”、”代表画像”が含まれる。

[0083] ”発話区間グループＩＤ”には、発話区間グループを識別するための識別子が格納される。”発話テキスト”には、対応する発話区間グループに含まれる発話テキストが格納される。”感情情報”には、対応する発話区間グループにおける感情情報が格納される。”代表画像”には、対応する発話区間グループにおける撮影画像データの中から抽出された代表画像が格納される。

[0084] （３）コマ割り推定部４３０による処理の具体例

図１２は、コマ割り推定部による処理の具体例を示す図である。図１２に示すように、コマ割り推定部４３０は、学習済みコマ割りモデル１２００を有する。

[0085] コマ割り推定部４３０は、コマ区切り情報格納部４５０よりコマ区切り情報（例えば、コマ区切り情報１１００）を読み出し、学習済みコマ割りモデル１２００に入力することで、学習済みコマ割りモデル１２００を実行させ

る。これにより、学習済みコマ割りモデル１２００では、コマ割り情報を推定する。

[0086] 図１２において、コマ割り情報１２１０は、コマ区切り情報１１００に基づいて、学習済みコマ割りモデル１２００を実行させることにより推定されたコマ割り情報の一例である。図１２の例は、

・発話区間グループＩＤ＝”１”の代表画像＝”Ａ－０１．ｊｐｇ”が、推定された大きさ及び形状のコマ１２１１の位置にレイアウトされること、

・発話区間グループＩＤ＝”１”の発話テキストが、コマ１２１１内に吹き出しの位置にレイアウトされること、

・コマ１２１１には効果線が付加されること、

・発話区間グループＩＤ＝”２”の代表画像＝”Ａ－０２．ｊｐｇ”が、推定された大きさ及び形状のコマ１２１２の位置にレイアウトされること、

・発話区間グループＩＤ＝”２”の発話テキストが、コマ１２１２内に吹き出しの位置にレイアウトされること、

・発話区間グループＩＤ＝”３”の代表画像＝”Ａ－０３．ｊｐｇ”が、推定された大きさ及び形状のコマ１２１３の位置にレイアウトされること、

・発話区間グループＩＤ＝”３”の発話テキストが、コマ１２１３内に吹き出しの位置にレイアウトされること、

・コマ１２１３には効果線が付加されること、

等を示すコマ割り情報１２１０が推定された様子を示している。

[0087] コマ割り推定部４３０は、学習済みコマ割りモデル１２００により推定されたコマ割り情報１２１０を、漫画表現画像生成部４４０に出力する。

[0088] なお、コマ割り推定部４３０は、コマ区切り情報を学習済みコマ割りモデル１２００に入力する際、リアルタイム処理かバッチ処理かの処理形態に応じて処理方法を切り替える。

[0089] リアルタイム処理とは、話者４０１または話者４０２が発話中に、リアルタイムに漫画表現画像を生成する処理を指す。リアルタイム処理を行う場合、コマ割り推定部４３０は、新たなコマ区切り情報が追加（新たな行が追加

）されたタイミングで、最新ｎ個のコマ区切り情報を、学習済みコマ割りモデル１２００に入力することで、コマ割り情報を推定する。なお、コマ割り推定部４３０が学習済みコマ割りモデル１２００に入力するコマ区切り情報の数（＝ｎ）は、予め設定されているものとする。

[0090] 一方、バッチ処理とは、話者４０１または話者４０２の過去の発話時に記録された音声データ及び撮影画像データについて生成された、全てのコマ区切り情報を用いて、漫画表現画像を生成する処理を指す。コマ割り推定部４３０は、コマ区切り情報格納部４５０に格納された全てのコマ区切り情報を、学習済みコマ割りモデル１２００に入力することで、コマ割り情報を推定する。

[0091] （４）漫画表現画像生成部４４０の処理の具体例

図１３は、漫画表現画像生成部による処理の具体例を示す図である。図１３に示すように、漫画表現画像生成部４４０は、コマ区切り情報（例えば、コマ区切り情報１１００）と、コマ割り情報（例えば、コマ割り情報１２１０）とに基づいて、漫画表現画像を生成する。

[0092] 図１３の例は、コマ１２１１に、

・発話区間グループＩＤ＝”１”の代表画像＝”Ａ－０１．ｊｐｇ”がレイアウトされ、

・吹き出しの位置に、発話区間グループＩＤ＝”１”の発話テキストがレイアウトされ、

・効果線が付加される、

ことで、漫画コマ画像１３１１が生成された様子を示している。

[0093] 同様に、図１３の例は、コマ１２１２に、

・発話区間グループＩＤ＝”２”の代表画像＝”Ａ－０２．ｊｐｇ”がレイアウトされ、

・吹き出しの位置に、発話区間グループＩＤ＝”２”の発話テキストがレイアウトされる、

ことで、漫画コマ画像１３１２が生成された様子を示している。

[0094] 同様に、図１３の例は、コマ１２１３に、

・発話区間グループＩＤ＝”３”の代表画像＝”Ａ－０３．ｊｐｇ”がレイアウトされ、

・吹き出しの位置に、発話区間グループＩＤ＝”３”の発話テキストがレイアウトされ、

・効果線が付加される、

ことで、漫画コマ画像１３１３が生成された様子を示している。

[0095] ＜画像生成処理の流れ＞

次に、画像生成装置４００による画像生成処理の流れについて説明する。図１４は、画像生成処理の流れを示すフローチャートである。

[0096] ステップＳ１４０１において、データ生成部４１０は、音声入力装置４６０及び撮像装置４７０より、音声データ及び撮影画像データを取得する。

[0097] ステップＳ１４０２において、データ生成部４１０は、音声データに基づいて、発話テキストを生成し、文書構造を解析することで、文書構造解析済み発話テキストを生成する。また、データ生成部４１０は、音声データ及び撮影画像データに基づいて、感情情報を生成する。

[0098] ステップＳ１４０３において、コマ区切り情報生成部４２０は、文書構造解析済み発話テキストに基づいて時刻軸を複数の発話区間に分割したうえで、複数のグループに分類し、発話区間グループを生成する。また、コマ区切り情報生成部４２０は、発話区間グループの位置、感情情報の変化、論理構成的な重要度等に基づいて、所定の発話区間グループを抽出する。更に、コマ区切り情報生成部４２０は、抽出した発話区間グループごとに、撮影画像データから代表画像を抽出し、発話テキスト及び感情情報と対応付けることで、コマ区切り情報を生成する。

[0099] ステップＳ１４０４において、コマ割り推定部４３０は、コマ区切り情報を学習済みコマ割りモデルに入力することで学習済みコマ割りモデルを実行させる。また、コマ割り推定部４３０は、学習済みコマ割りモデルにより推定されたコマ割り情報を、漫画表現画像生成部４４０に出力する。

[0100] ステップＳ１４０５において、漫画表現画像生成部４４０は、コマ区切り情報とコマ割り情報とに基づいて、漫画表現画像を生成する。

[0101] ステップＳ１４０６において、漫画表現画像生成部４４０は、漫画表現画像を漫画表現画像表示装置４８０に表示する。

[0102] ステップＳ１４０７において、画像生成装置４００は画像生成処理を終了するか否かを判定し、終了しないと判定した場合には（ステップＳ１４０７においてＮＯの場合には）、ステップＳ１４０１に戻る。

[0103] 一方、ステップＳ１４０７において、終了すると判定した場合には（ステップＳ１４０７においてＹＥＳの場合には）、画像生成処理を終了する。

[0104] ＜まとめ＞

以上の説明から明らかなように、第１の実施形態に係る画像生成装置４００は、

・既存の漫画画像に基づいて、コマ割りを学習した、学習済みコマ割りモデルを有する。

・発話シーンにおける話者の音声データと撮影画像データとに基づいて生成したコマ区切り情報を、学習済みコマ割りモデルに入力することで、コマ割りを推定する。

・推定したコマ割りに基づいて、発話シーンを示す漫画表現画像を生成する。

[0105] このように、既存の漫画画像を用いて、漫画特有の表現方法（コマ割り）を学習し、発話シーンから漫画表現画像を生成する際に、学習した表現方法を反映させることで、話者の感情が表現された漫画表現画像を生成することができる。

[0106] この結果、第１の実施形態に係る画像生成装置４００によれば、発話シーンを漫画風に表現する際、発話内容を理解しやすくすることができる。

[0107] ［第２の実施形態］

上記第１の実施形態では、学習装置３００と画像生成装置４００とを別体の装置として構成したが、学習装置３００と画像生成装置４００とは、一体

の装置として構成してもよい。

[0108] また、上記第１の実施形態では、画像生成装置４００と、音声入力装置４６０、撮像装置４７０、漫画表現画像表示装置４８０とが、別体の装置として構成されるものとして説明した。しかしながら、画像生成装置４００は、音声入力装置４６０、撮像装置４７０、漫画表現画像表示装置４８０のうちの一部または全部と、一体の装置として構成されてもよい。その場合、画像生成装置４００は、タブレット端末等のように携帯可能な装置により構成されてもよい。

[0109] また、上記第１の実施形態では、発話シーンにおける話者が二人の場合を例に説明したが、話者の人数は二人に限定されず、一人であっても三人以上であってもよい。

[0110] また、上記第１の実施形態では、任意の漫画画像を用いて、コマ割りモデルに対して学習処理を行うものとして説明した。しかしながら、画像生成装置４００の適用先において想定されるシーンに対応する漫画画像を選択的に用いて、コマ割りモデルに対して学習処理を行うように構成してもよい。

[0111] なお、上記実施形態に挙げた構成等に、その他の要素との組み合わせ等、ここで示した構成に本発明が限定されるものではない。これらの点に関しては、本発明の趣旨を逸脱しない範囲で変更することが可能であり、その応用形態に応じて適切に定めることができる。

符号の説明

[0112] ３００ ：学習装置
３１０ ：学習用データ生成部
３２０ ：コマ割り学習部
３４１ ：学習用データ
４００ ：画像生成装置
４１０ ：データ生成部
４２０ ：コマ区切り情報生成部
４３０ ：コマ割り推定部

４４０　：漫画表現画像生成部

４６０　：音声入力装置

４７０　：撮像装置

４８０　：漫画表現画像表示装置

１１００　：コマ区切り情報

１２００　：学習済みコマ割りモデル

１２１０　：コマ割り情報

請求の範囲

[請求項1] 既存の漫画画像に基づいて、コマ割りを学習した学習済みモデルに、発話シーンにおける話者の音声データと撮影画像データとに基づいて生成した情報を入力し、コマ割りを推定する推定部と、

前記推定部により推定されたコマ割りに基づいて、前記発話シーンを示す画像を生成する画像生成部と

を有する画像生成装置。

[請求項2] 前記学習済みモデルは、既存の漫画画像をコマごとに分割し、それぞれの漫画コマ画像と、それぞれの漫画コマ画像におけるセリフ、登場人物の感情を示す情報とを入力として、コマ割りを学習した学習済みモデルである、請求項１に記載の画像生成装置。

[請求項3] 前記推定部は、前記学習済みモデルに、前記発話シーンにおける話者の撮影画像データから抽出した画像と、前記発話シーンにおける話者の音声データから抽出した発話テキストと、前記発話シーンにおける話者の音声データ及び撮影画像データから推定される話者の感情を示す情報とを入力することで、前記コマ割りを推定する、請求項２に記載の画像生成装置。

[請求項4] 前記発話テキストに基づいて生成した複数の発話区間を複数のグループに分け、所定の条件を満たすグループを、コマを区切る単位として抽出し、抽出したグループごとに、前記学習済みモデルに入力する、前記画像と、前記発話テキストと、前記感情を示す情報とを含むコマ区切り情報を生成する情報生成部を更に有する、請求項３に記載の画像生成装置。

[請求項5] 前記所定の条件を満たすグループは、発話の開始時の発話区間を含むグループ、感情を示す情報の変化が閾値を超えた発話区間を含むグループ、発話内容についての論理構成的な重要度が閾値を超えた発話区間を含むグループのいずれかのグループを含む、請求項４に記載の画像生成装置。

[請求項6] 前記推定部は、

前記情報生成部により新たなコマ区切り情報が追加されたタイミングで最新のｎ個のコマ区切り情報を前記学習済みモデルに入力し、コマ割りを推定する、請求項４に記載の画像生成装置。

[請求項7] 既存の漫画画像に基づいて、コマ割りを学習した学習済みモデルに、発話シーンにおける話者の音声データと撮影画像データとに基づいて生成した情報を入力し、コマ割りを推定する推定工程と、

前記推定工程において推定されたコマ割りに基づいて、前記発話シーンを示す画像を生成する画像生成工程と

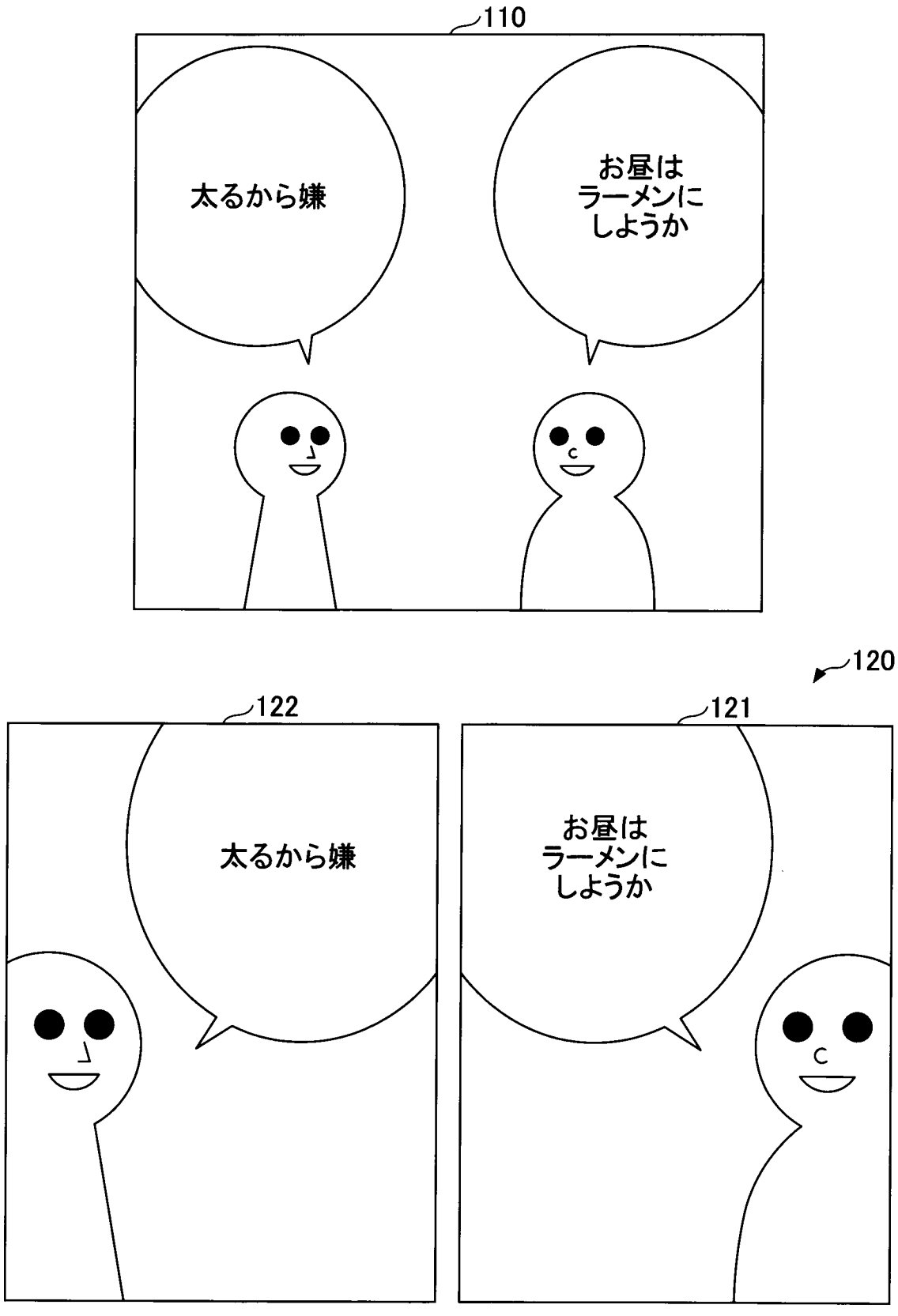
を有する画像生成方法。

[請求項8] 既存の漫画画像に基づいて、コマ割りを学習した学習済みモデルに、発話シーンにおける話者の音声データと撮影画像データとに基づいて生成した情報を入力し、コマ割りを推定する推定工程と、

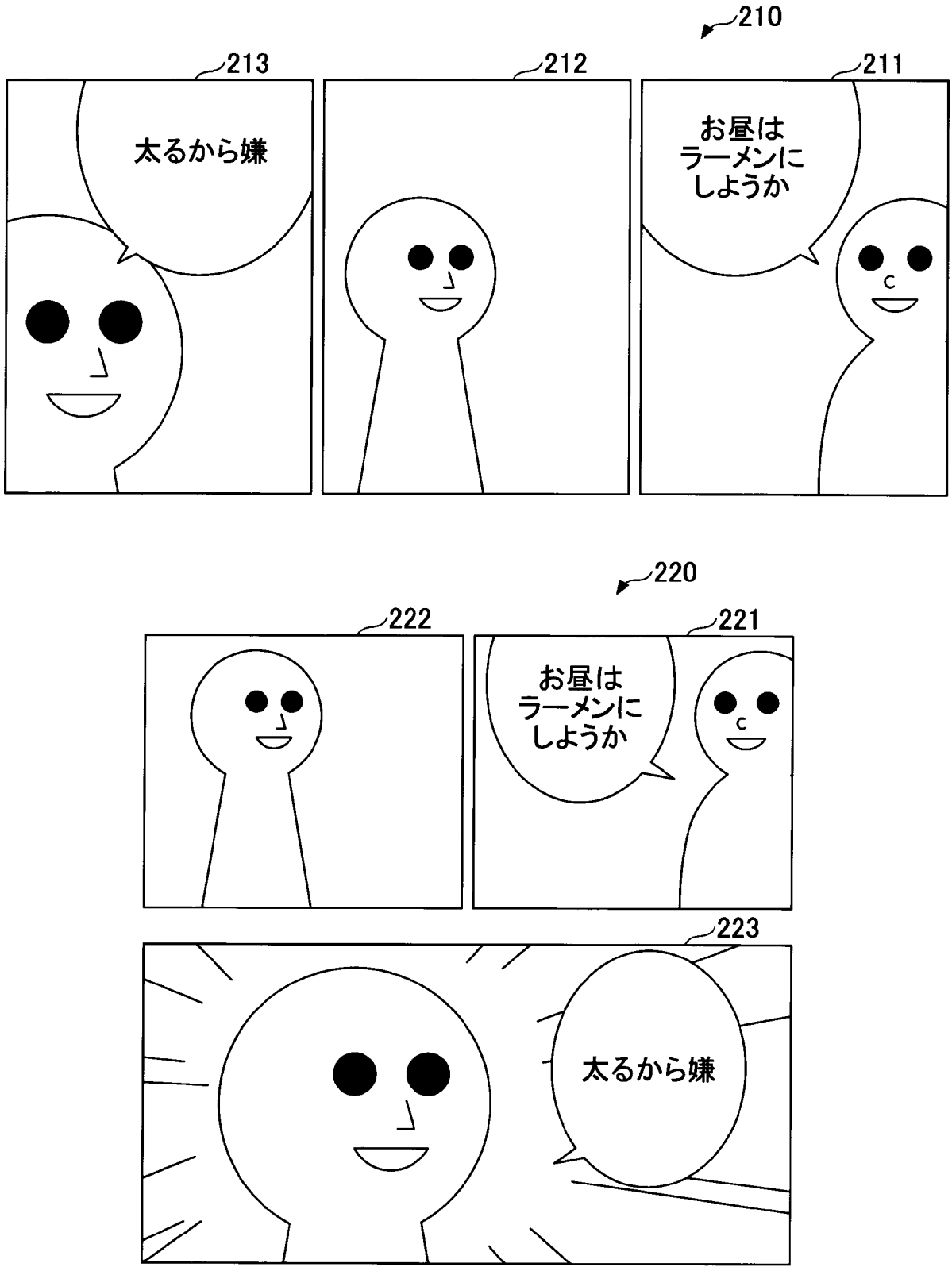
前記推定工程において推定されたコマ割りに基づいて、前記発話シーンを示す画像を生成する画像生成工程と

をコンピュータに実行させるための画像生成プログラム。

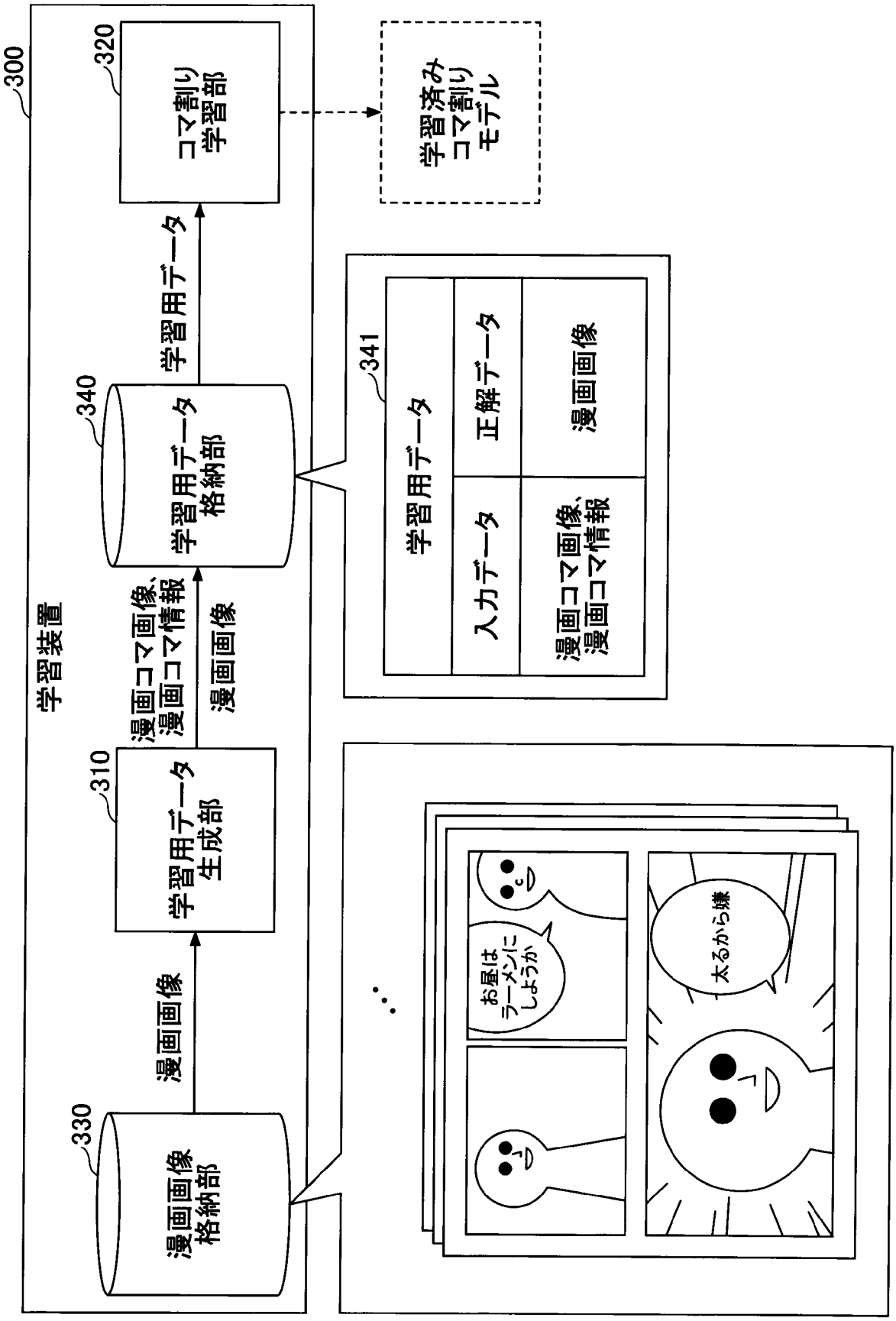
[図1]



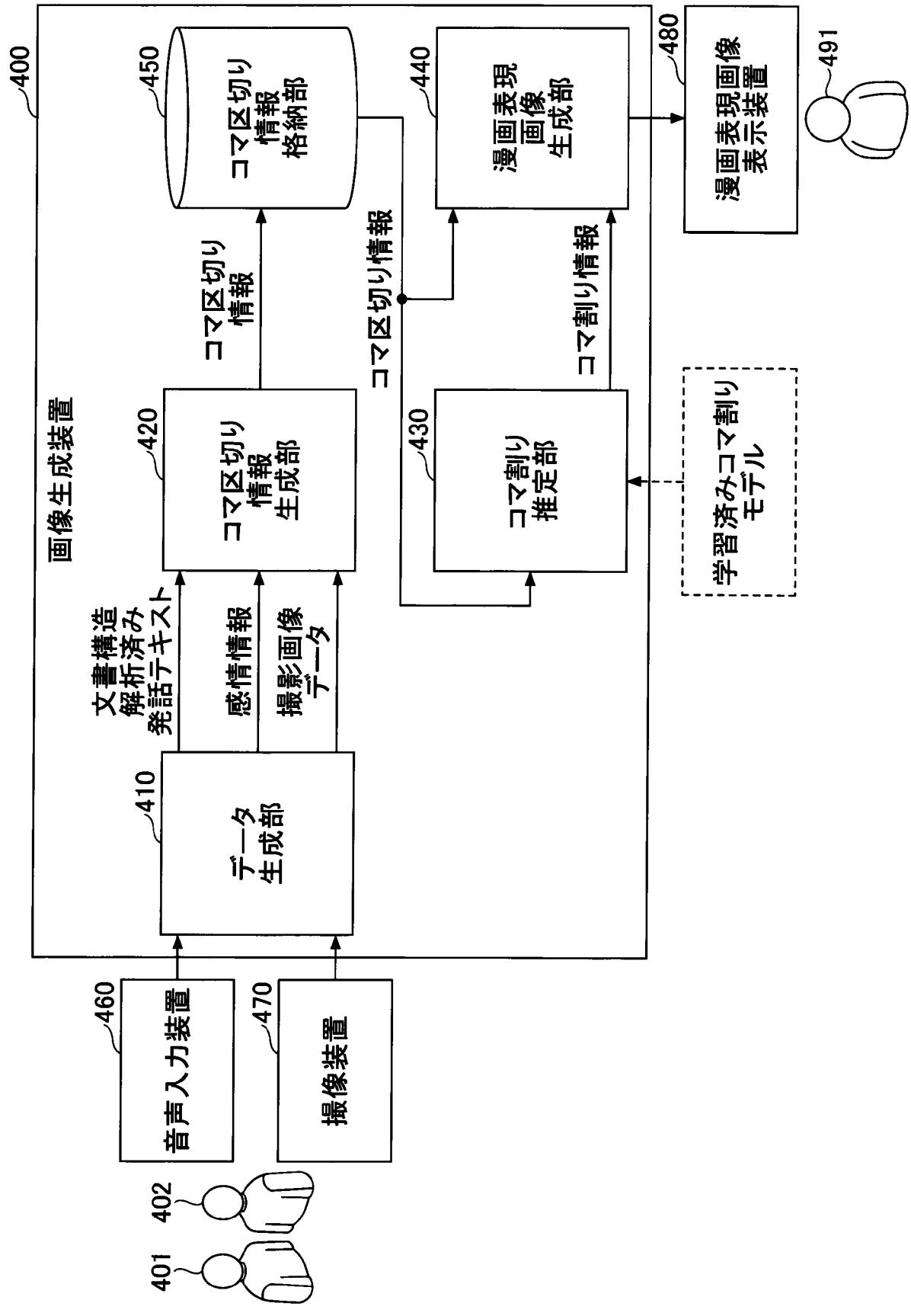
[図2]


210
213
212
211
太るから嫌
お昼は
ラーメンに
しようか
C
220
222
221
お昼は
ラーメンに
しようか
C
223
太るから嫌

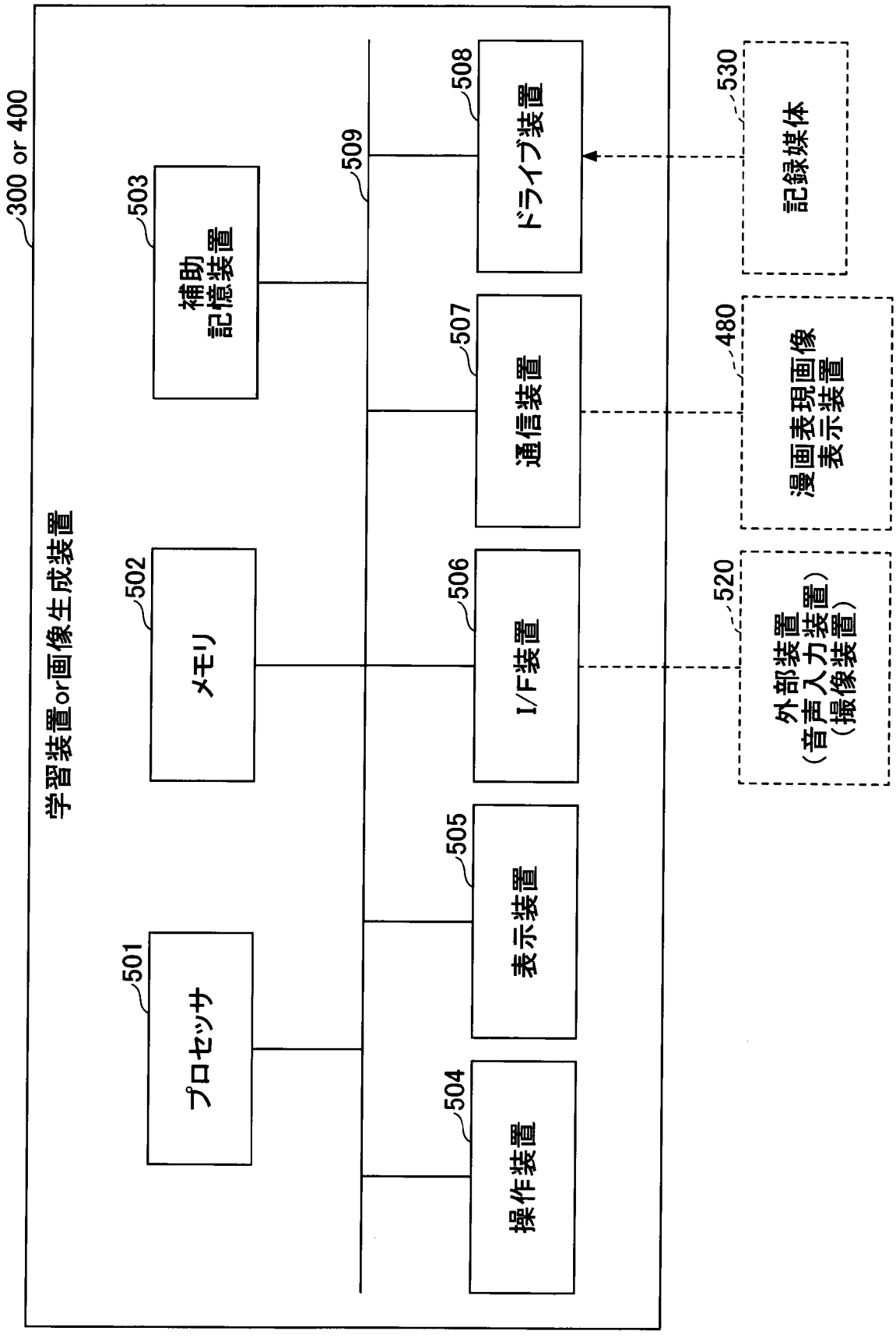
[図3]


300
学習装置
320
コマ割り学習部
学習済み
コマ割り
モデル
学習用データ
340
学習用データ格納部
341
学習用データ
入力データ
正解データ
漫画画像
漫画コマ画像、
漫画コマ情報
漫画コマ画像、漫画コマ情報
漫画画像
310
学習用データ生成部
漫画画像
330
漫画画像格納部
お昼はラーメンにしようか
太るから嫌

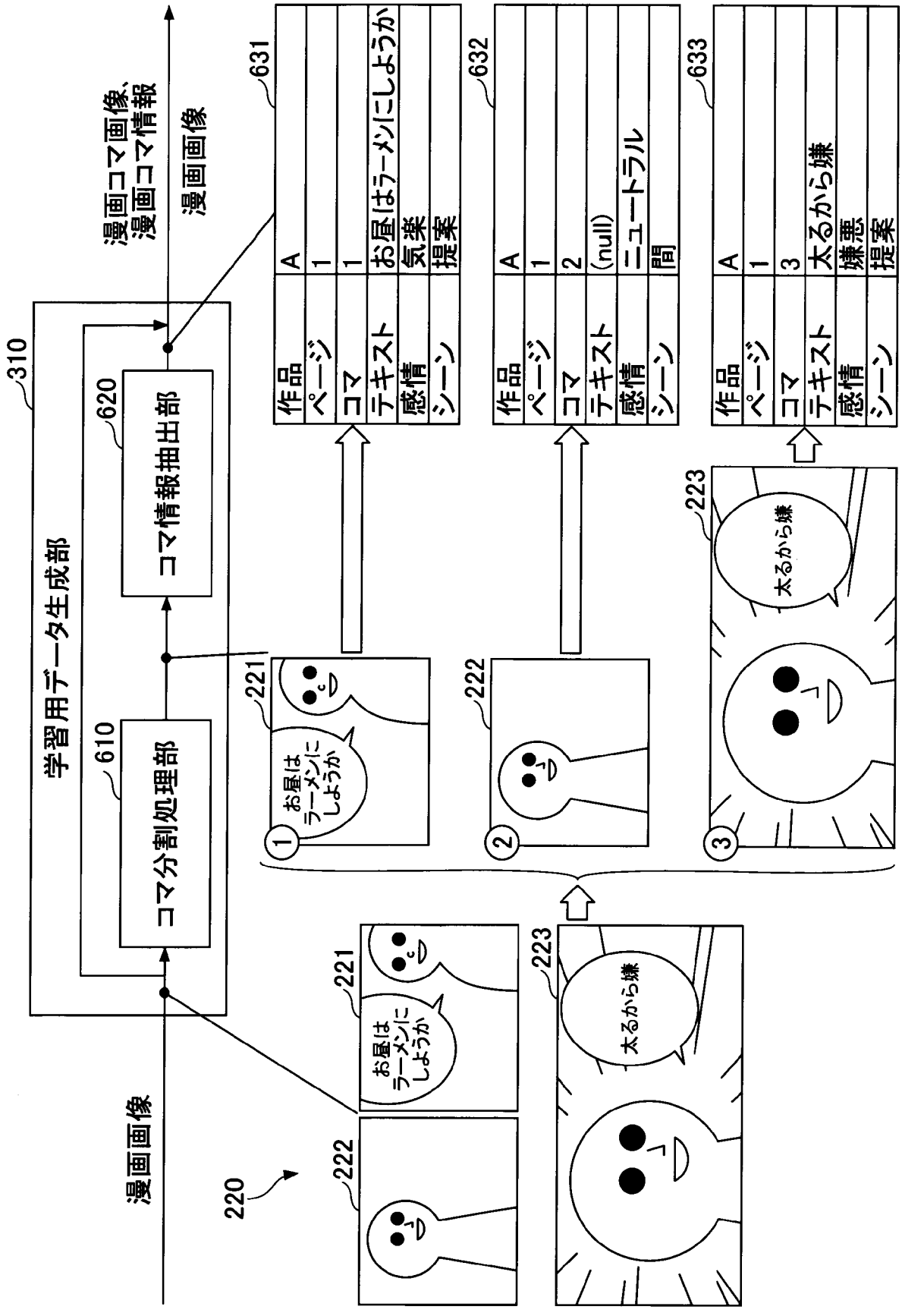
[図4]


400
画像生成装置
410
データ生成部
420
コマ区切り情報生成部
430
コマ割り推定部
440
漫画表現画像生成部
450
コマ区切り情報格納部
460
音声入力装置
470
撮像装置
480
漫画表現画像表示装置
491
401
402
文書構造解析済み発話テキスト
感情情報
撮影画像データ
コマ区切り情報
コマ区切り情報
コマ割り情報
学習済みコマ割りモデル

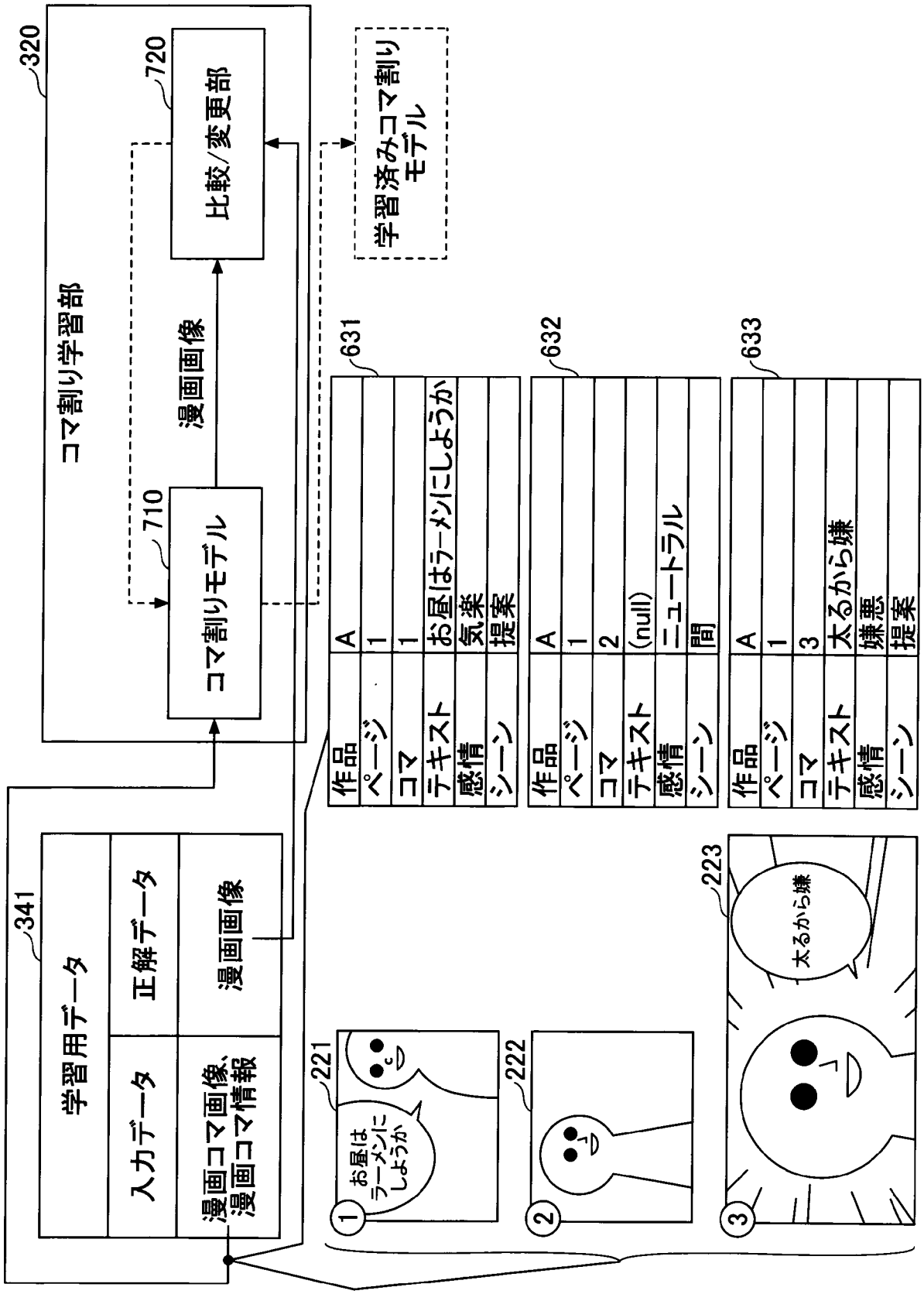
[図5]


300 or 400
学習装置or画像生成装置
501
プロセッサ
502
メモリ
503
補助
記憶装置
504
操作装置
505
表示装置
506
I/F装置
507
通信装置
508
ドライブ装置
509
520
外部装置
（音声入力装置）
（撮像装置）
480
漫画表現画像
表示装置
530
記録媒体

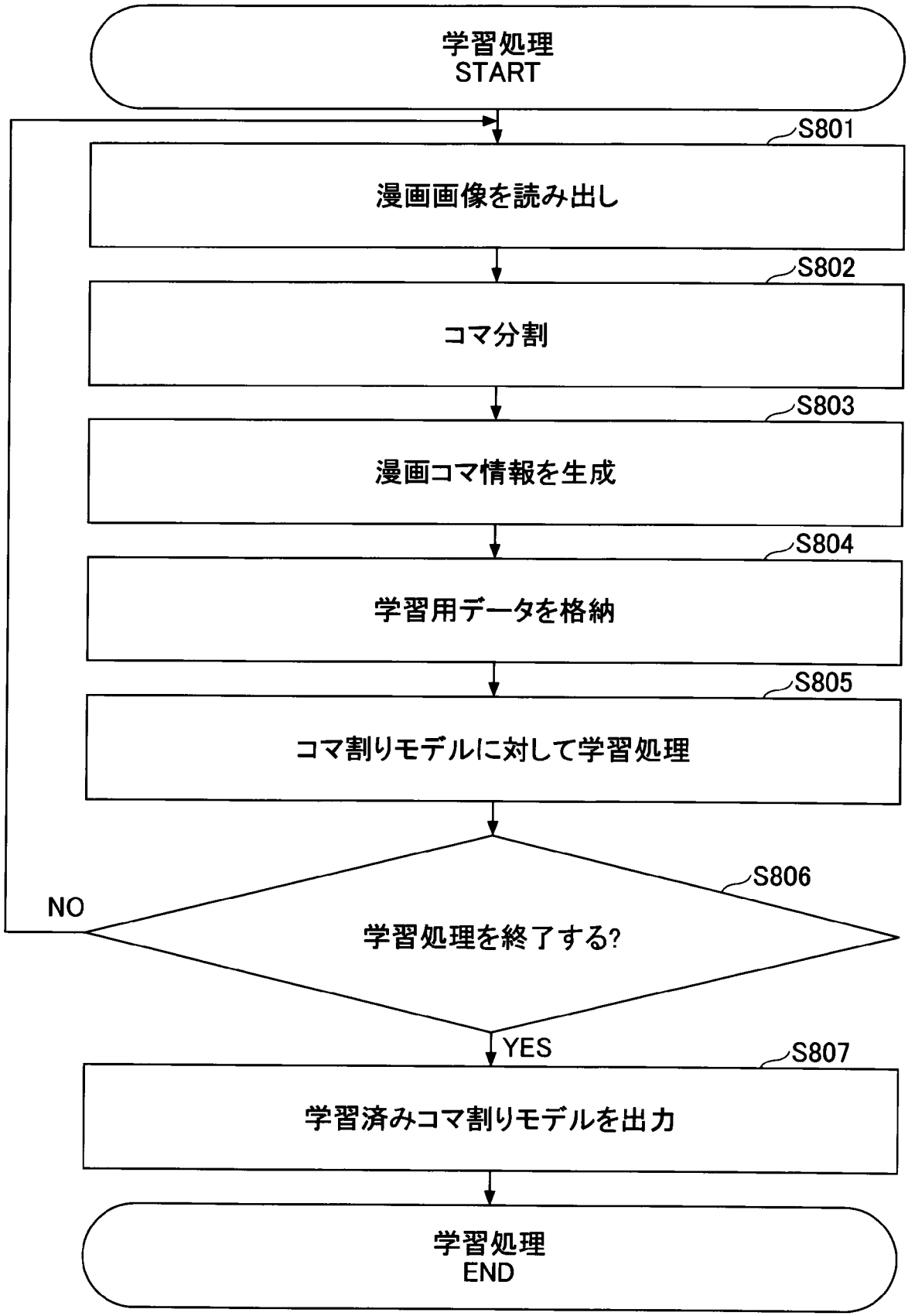
[図6]


漫画画像
220
学習用データ生成部
310
コマ分割処理部
610
コマ情報抽出部
620
漫画コマ画像、漫画コマ情報
漫画画像
221
222
223
お昼はラーメンにしようか
太るから嫌
631
632
633
作品 A
ページ 1
コマ 1
テキスト お昼はラーメンにしようか
感情 気楽
シーン 提案
コマ 2
テキスト (null)
感情 ニュートラル
シーン 間
コマ 3
テキスト 太るから嫌
感情 嫌悪
シーン 提案

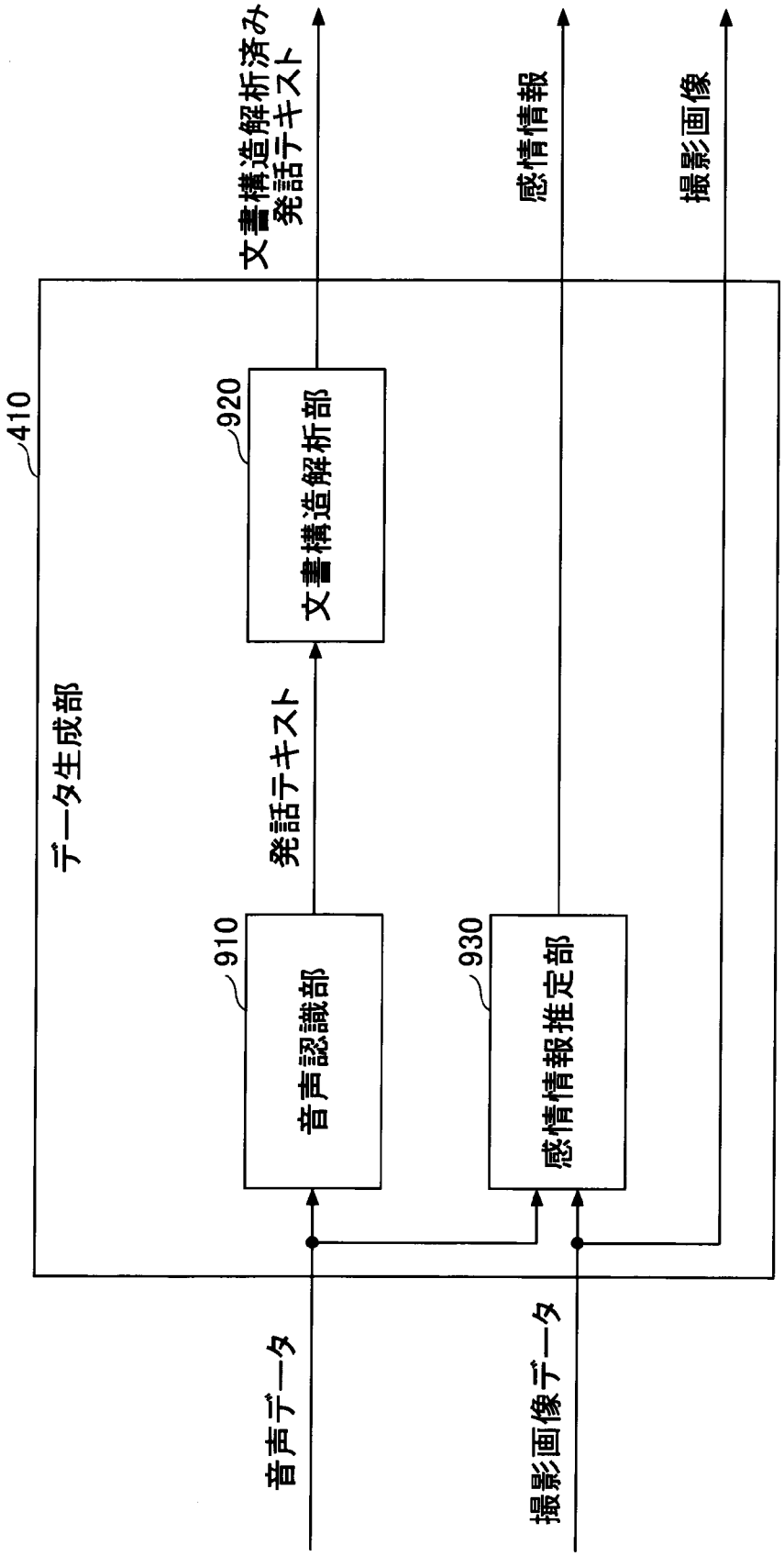
[図7]


320
コマ割り学習部
720
比較/変更部
710
コマ割りモデル
漫画画像
学習済みコマ割りモデル
341
学習用データ
入力データ
正解データ
漫画画像
漫画コマ画像、漫画コマ情報
631
作品 A
ページ 1
コマ 1
テキスト お昼はラーメンにしようか
感情 気楽
シーン 提案
632
作品 A
ページ 1
コマ 2
テキスト (null)
感情 ニュートラル
シーン 間
633
作品 A
ページ 1
コマ 3
テキスト 太るから嫌
感情 嫌悪
シーン 提案
221
お昼はラーメンにしようか
222
223
太るから嫌
①
②
③

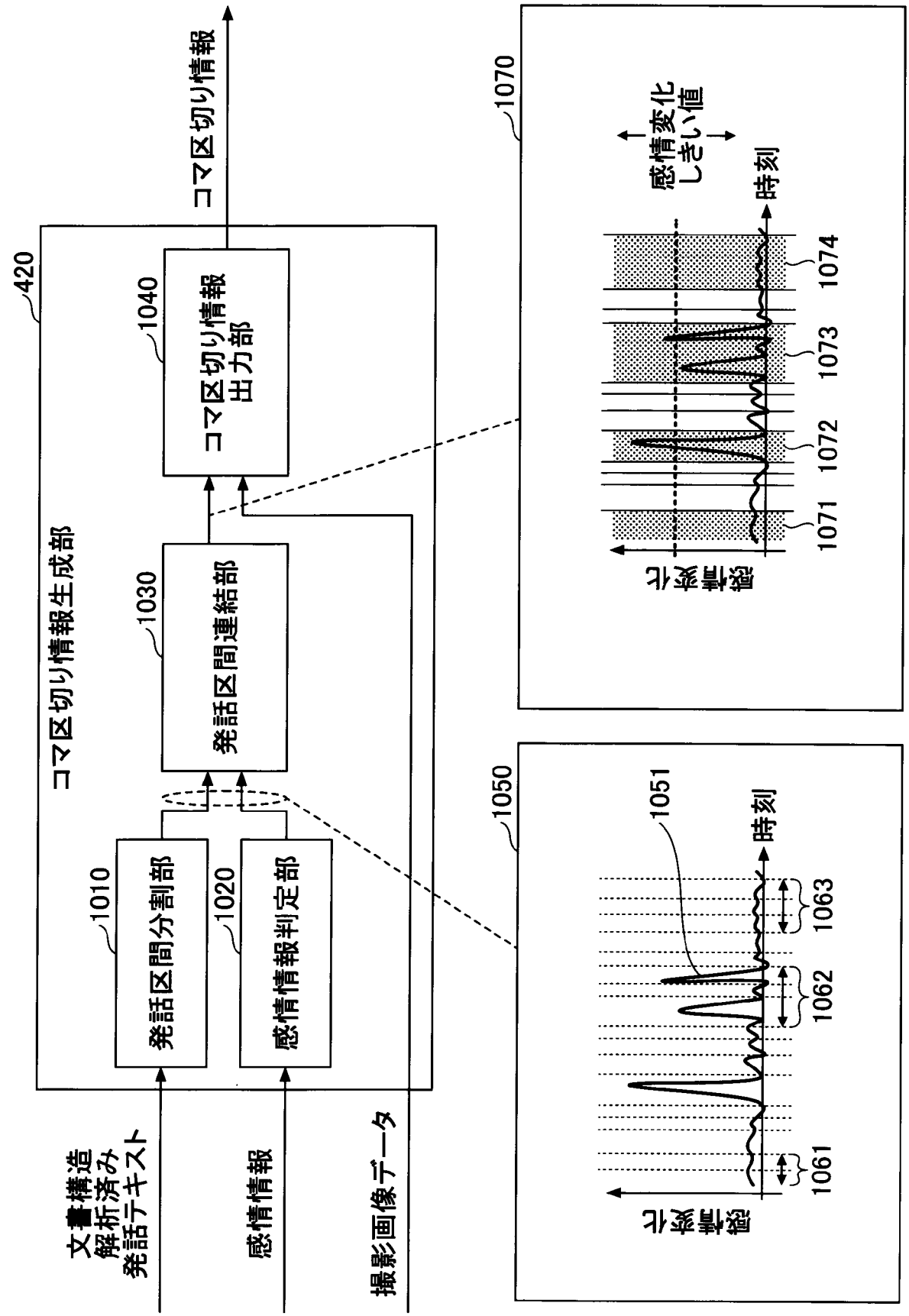
[図8]


学習処理
START
S801
漫画画像を読み出し
S802
コマ分割
S803
漫画コマ情報を生成
S804
学習用データを格納
S805
コマ割りモデルに対して学習処理
S806
学習処理を終了する?
NO
YES
S807
学習済みコマ割りモデルを出力
学習処理
END

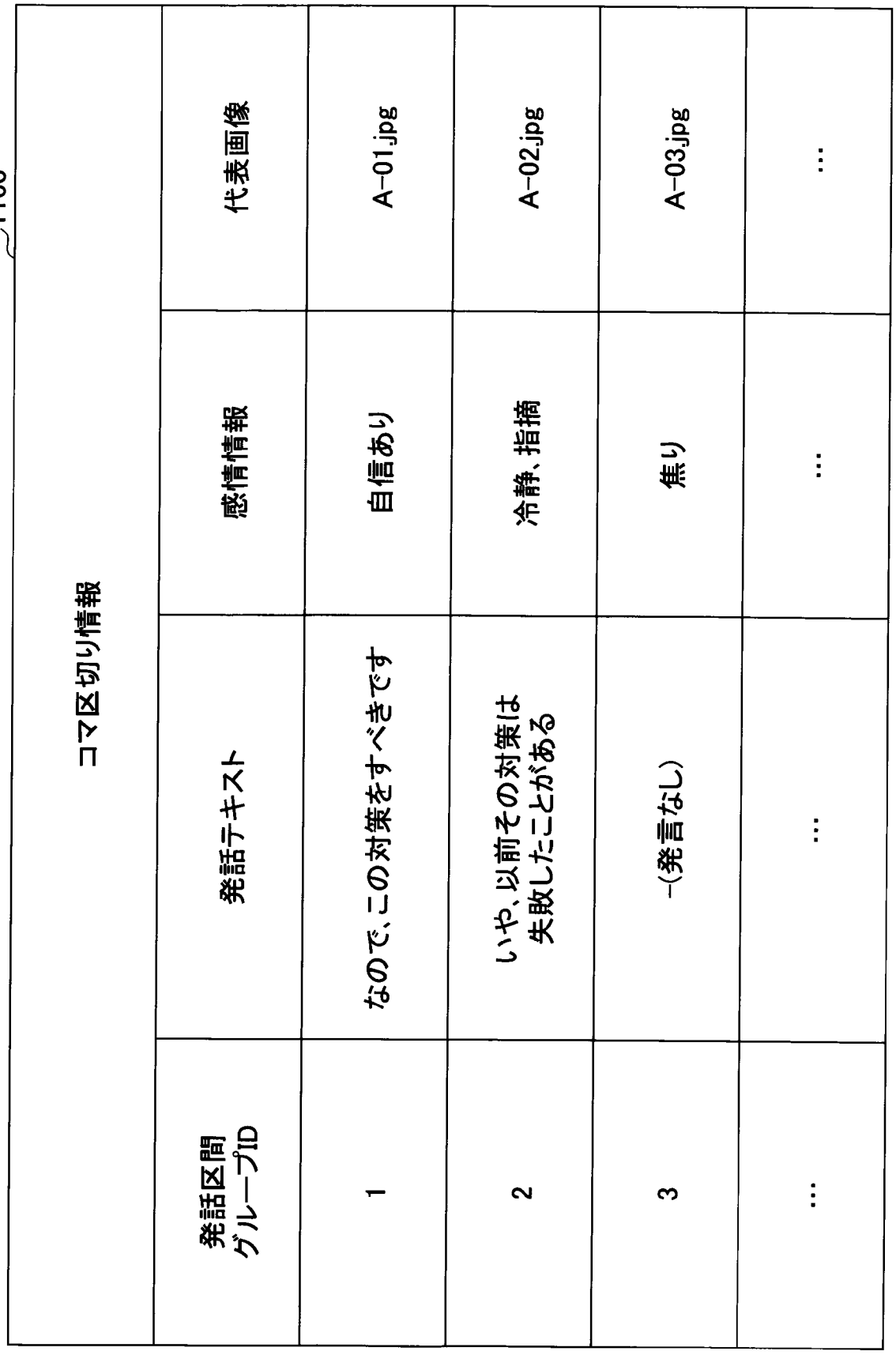
[図9]


410
データ生成部
910
音声認識部
920
文書構造解析部
930
感情情報推定部
発話テキスト
音声データ
撮影画像データ
文書構造解析済み
発話テキスト
感情情報
撮影画像

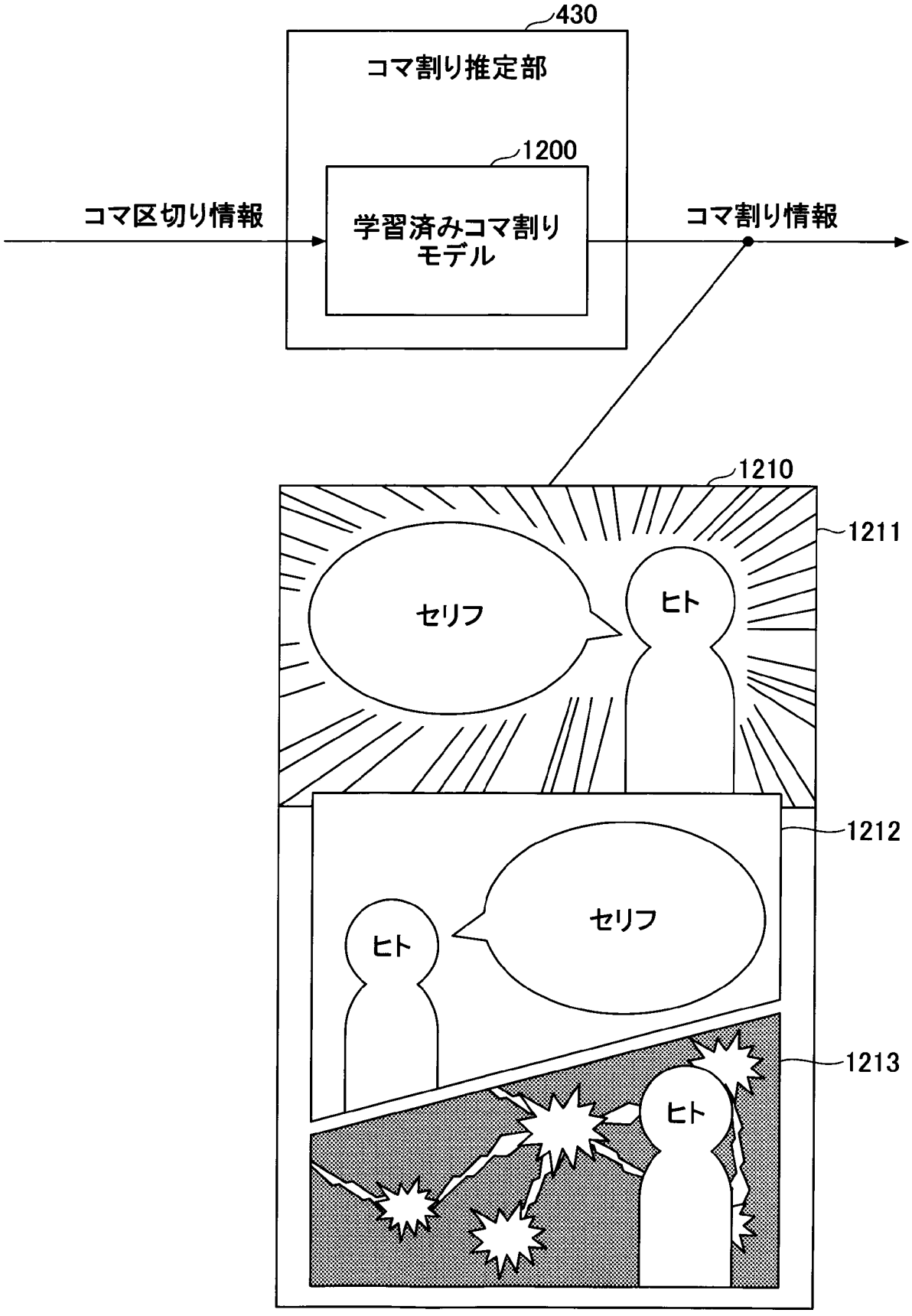
[図10]


420
コマ区切り情報生成部
1010
発話区間分割部
1020
感情情報判定部
1030
発話区間連結部
1040
コマ区切り情報出力部
コマ区切り情報
文書構造解析済み発話テキスト
感情情報
撮影画像データ
1050
1051
時刻
感情変化
1061
1062
1063
1070
感情変化しきい値
時刻
感情変化
1071
1072
1073
1074

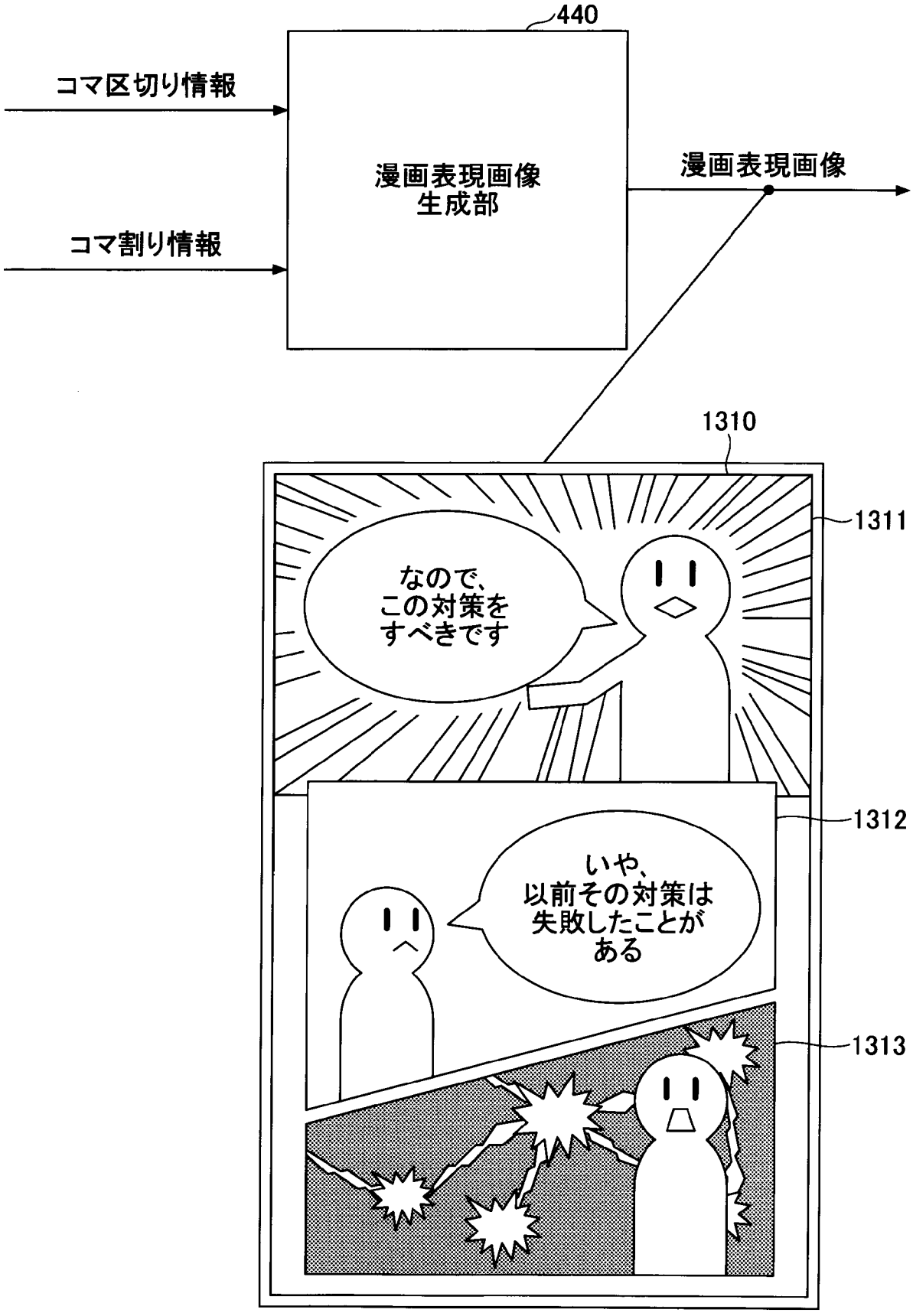
[図11]


1100
コマ区切り情報
発話区間グループID
発話テキスト
感情情報
代表画像
1
なので、この対策をすべきです
自信あり
A-01.jpg
2
いや、以前その対策は失敗したことがある
冷静、指摘
A-02.jpg
3
-(発言なし)
焦り
A-03.jpg
…
…
…
…

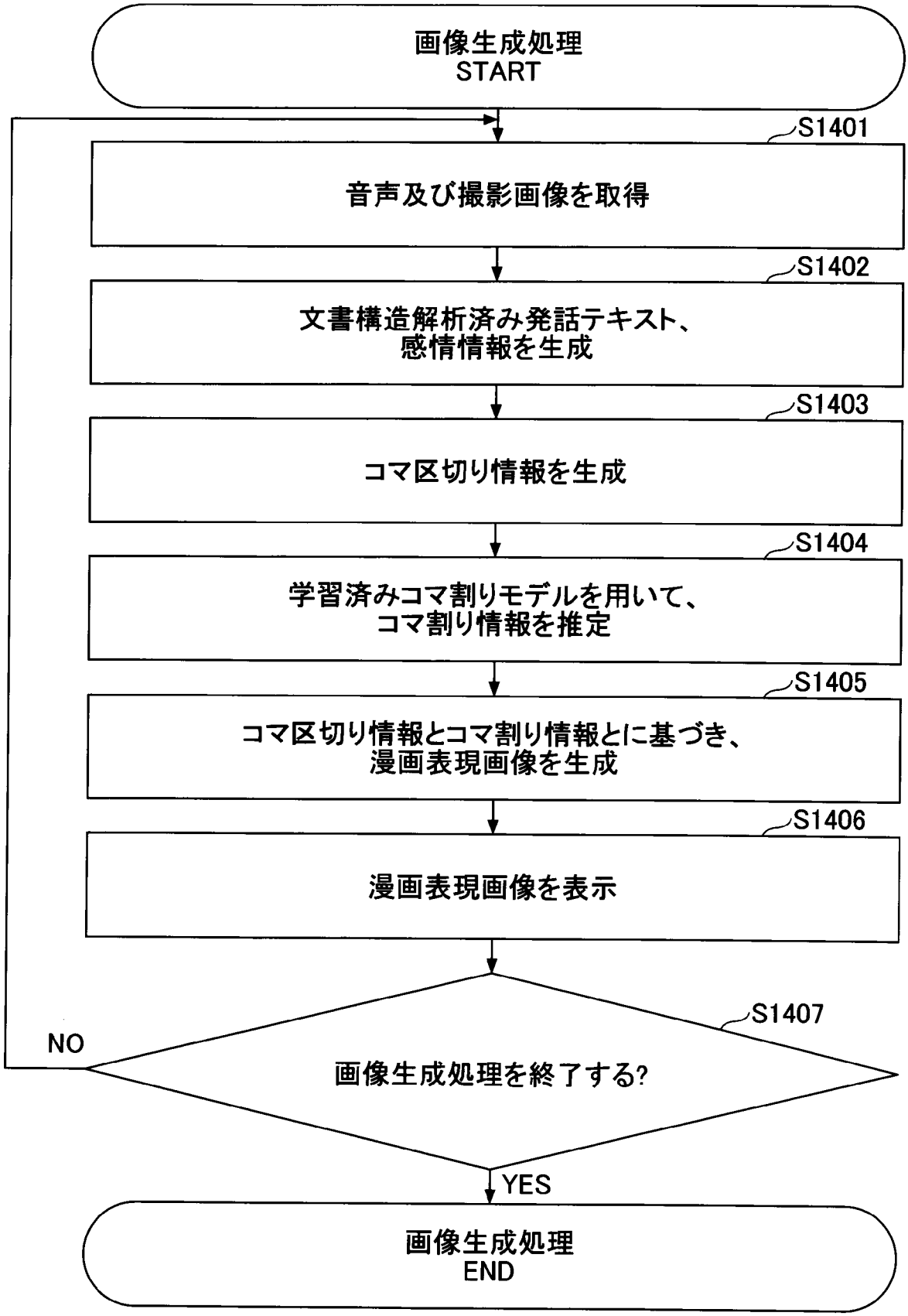
[図12]


430
コマ割り推定部
1200
学習済みコマ割り
モデル
コマ区切り情報
コマ割り情報
1210
1211
セリフ
ヒト
1212
セリフ
ヒト
1213
ヒト

[図13]


440
コマ区切り情報
コマ割り情報
漫画表現画像
生成部
漫画表現画像
1310
1311
なので、
この対策を
すべきです
1312
いや、
以前その対策は
失敗したことが
ある
1313

[図14]


画像生成処理
START
S1401
音声及び撮影画像を取得
S1402
文書構造解析済み発話テキスト、
感情情報を生成
S1403
コマ区切り情報を生成
S1404
学習済みコマ割りモデルを用いて、
コマ割り情報を推定
S1405
コマ区切り情報とコマ割り情報とに基づき、
漫画表現画像を生成
S1406
漫画表現画像を表示
S1407
画像生成処理を終了する?
NO
YES
画像生成処理
END

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2020/044317

A. CLASSIFICATION OF SUBJECT MATTER
G06T 11/60(2006.01)i
FI: G06T11/60 100A

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06T11/60

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2021
Registered utility model specifications of Japan	1996-2021
Published registered utility model applications of Japan	1994-2021

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2020-106910 A (HITACHI SOLUTIONS, LTD.) 09 July 2020 (2020-07-09) entire text, all drawings	1-8
A	JP 2009-65462 A (TOSHIBA CORP.) 26 March 2009 (2009-03-26) entire text, all drawings	1-8
A	JP 2010-191934 A (SONY CORP.) 02 September 2010 (2010-09-02) entire text, all drawings	1-8
A	JP 2007-193824 A (FUJIFILM CORPORATION) 02 August 2007 (2007-08-02) entire text, all drawings	1-8

☐ Further documents are listed in the continuation of Box C.

☒ 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
25 February 2021 (25.02.2021)	16 March 2021 (16.03.2021)

Name and mailing address of the ISA/	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2020/044317

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
JP 2020-106910 A	09 Jul. 2020	(Family: none)	
JP 2009-65462 A	26 Mar. 2009	(Family: none)	
JP 2010-191934 A	02 Sep. 2010	US 2010/0182501 A1 entire text, all drawings CN 101783886 A	
JP 2007-193824 A	02 Aug. 2007	US 2006/0251299 A1 entire text, all drawings	

国際調査報告

国際出願番号
PCT/JP2020/044317

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06T 11/60(2006.01)i
FI: G06T11/60 100A

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G06T11/60

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２１年
日本国実用新案登録公報	１９９６－２０２１年
日本国登録実用新案公報	１９９４－２０２１年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
A	JP 2020-106910 A (株式会社日立ソリューションズ) 09.07.2020 (2020-07-09) 全文,全図	1-8
A	JP 2009-65462 A (株式会社東芝) 26.03.2009 (2009-03-26) 全文,全図	1-8
A	JP 2010-191934 A (ソニー株式会社) 02.09.2010 (2010-09-02) 全文,全図	1-8
A	JP 2007-193824 A (富士フイルム株式会社) 02.08.2007 (2007-08-02) 全文,全図	1-8

☐ Ｃ欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
“O” 口頭による開示、使用、展示等に言及する文献
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
“Y” 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
“&” 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
25.02.2021	16.03.2021

名称及びあて先

日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）

橘　高志 5V 8391

電話番号 03-3581-1101 内線 3571

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2020/044317

引用文献	公表日	パテントファミリー文献	公表日
JP 2020-106910 A	09.07.2020	（ファミリーなし）	
JP 2009-65462 A	26.03.2009	（ファミリーなし）	
JP 2010-191934 A	02.09.2010	US 2010/0182501 A1 全文,全図 CN 101783886 A	
JP 2007-193824 A	02.08.2007	US 2006/0251299 A1 全文,全図	