

(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
2020年7月23日(23.07.2020)



(10) 国際公開番号

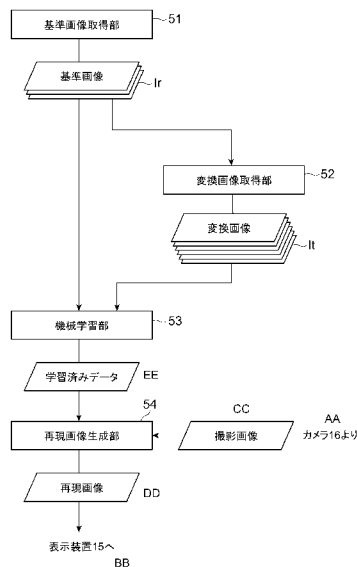
WO 2020/148810 A1

- (51) 国際特許分類:
H04N 7/18 (2006.01) *G06T 7/00* (2017.01)
G06N 99/00 (2019.01) *G06T 19/00* (2011.01)
- (21) 国際出願番号: PCT/JP2019/000964
- (22) 国際出願日: 2019年1月15日(15.01.2019)
- (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: 株式会社 ソニー・インタラクティブエンタテインメント (SONY INTERACTIVE ENTERTAINMENT INC.) [JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: 清水 圭介 (SHIMIZU, Keisuke); 〒1080075 東京都港区港南1丁目7番1号 株式会社ソニー・インタラクティブエンタテインメント内 Tokyo (JP).
- (74) 代理人: 竹 居 信 利 (TAKEI Nobutoshi); 〒1600022 東京都新宿区新宿6丁目7-1 エルプリメント新宿308 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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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,

(54) Title: INFORMATION PROCESSING DEVICE

(54) 発明の名称: 情報処理装置

[図2]



51 Reference image acquisition unit
52 Converted image acquisition unit
53 Machine learning unit
54 Reproduction image generation unit
Lr Reference image
Lt Converted image
AA From camera 16
BB To display device 15
CC Captured image
DD Reproduction image
EE Pre-learned data

(57) Abstract: An information processing device for acquiring a plurality of reference images obtained by imaging an object that is to be reproduced, acquiring a plurality of converted images obtained by enlarging or shrinking each of the plurality of reference images, executing machine learning using a plurality of images to be learned, as teaching data, that include the plurality of converted images, and generating pre-learned data that is used for generating a reproduction image that represents the appearance of the object.

(57) 要約: 再現対象となる対象物を撮影して得られる複数の基準画像を取得し、当該複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を取得し、当該複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行し、対象物の外観を表す再現画像を生成するために用いられる学習済みデータを生成する情報処理装置である。

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： 情報処理装置

技術分野

[0001] 本発明は、対象物を表す画像を生成するための情報処理装置、情報処理方法、プログラム、及び学習済みデータに関する。

背景技術

[0002] 機械学習によって画像を自動生成する技術が研究されている。このような試みの一つとして、人などの対象物を撮影して得られる複数の撮影画像を教師データとして用いて、同じ対象物を再現する画像を生成する手法が検討されている。

発明の概要

発明が解決しようとする課題

[0003] 上記技術においては、限られた種類の画像を教師データとして利用して機械学習を行ったとしても、その限られた教師データに比較的近いパターンの再現画像しか生成することができず、教師データの内容から大きく逸脱したパターンの再現画像を精度よく生成することは難しい。一方で、生成可能な再現画像のバリエーションを増やすためには多様なパターンの画像を撮影しなければならず、手間がかかってしまう。

[0004] 本発明は上記実情を考慮してなされたものであって、その目的の一つは、比較的少ない手間に対象物を再現する再現画像のバリエーションを増やすことのできる情報処理装置、情報処理方法、プログラム、及び学習済みデータを提供することにある。

課題を解決するための手段

[0005] 本発明の一態様に係る情報処理装置は、再現対象となる対象物を撮影して得られる複数の基準画像を取得する基準画像取得部と、前記複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を取得する変換画像取得部と、前記複数の変換画像を含む複数の学習対象画像を教師データとし

て用いて機械学習を実行し、前記対象物の外観を表す再現画像を生成するために用いられる学習済みデータを生成する機械学習部と、を含むことを特徴とする。

[0006] 本発明の一態様に係る情報処理方法は、再現対象となる対象物を撮影して得られる複数の基準画像を取得するステップと、前記複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を取得するステップと、前記複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行し、前記対象物の外観を表す再現画像を生成するために用いられる学習済みデータを生成するステップと、を含むことを特徴とする。

[0007] 本発明の一態様に係るプログラムは、再現対象となる対象物を撮影して得られる複数の基準画像を取得するステップと、前記複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を取得するステップと、前記複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行し、前記対象物の外観を表す再現画像を生成するために用いられる学習済みデータを生成するステップと、をコンピュータに実行させるためのプログラムである。このプログラムは、コンピュータ読み取り可能で非一時的な情報記憶媒体に格納されて提供されてよい。

[0008] 本発明の一態様に係る学習済みデータは、再現対象となる対象物を撮影して得られる複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行した結果得られる、前記対象物の外観を表す再現画像を生成するために用いられる学習済みデータである。

[0009] 本発明の一態様に係る情報処理装置は、再現対象となる対象物の姿勢を表す姿勢データを取得する取得部と、対象物を撮影して得られる複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行した結果得られる学習済みデータを用いて、前記姿勢データが表す姿勢を前記対象物が取っている様子を表す再現画像を生成する再現画像生成部と、を含むことを特徴とする。

図面の簡単な説明

[0010] [図1]本発明の実施の形態に係る情報処理装置の構成ブロック図である。

[図2]情報処理装置の機能ブロック図である。

[図3]基準画像の一例を示す図である。

[図4]変換画像の一例を示す図である。

[図5]基準画像に対する姿勢特定処理の結果の一例を示す図である。

[図6]利用者を撮影する様子を示す図である。

[図7]再現画像の一例を示す図である。

発明を実施するための形態

[0011] 以下、本発明の実施の形態について、図面に基づき詳細に説明する。

[0012] 図1は、本発明の一実施形態に係る情報処理装置10の構成ブロック図である。情報処理装置10は、パーソナルコンピュータ等の情報処理装置であって、図1に示されるように、制御部11と、記憶部12と、インタフェース部13と、を含んで構成される。また、情報処理装置10は、操作デバイス14、表示装置15、及びカメラ16と接続されている。

[0013] 制御部11は、CPU等のプロセッサを少なくとも一つ含み、記憶部12に記憶されているプログラムを実行して各種の情報処理を実行する。なお、本実施形態において制御部11が実行する処理の具体例については、後述する。記憶部12は、RAM等のメモリデバイスを少なくとも一つ含み、制御部11が実行するプログラム、及び当該プログラムによって処理されるデータを格納する。

[0014] インタフェース部13は、操作デバイス14、表示装置15及びカメラ16との間のデータ通信のためのインタフェースである。情報処理装置10は、インタフェース部13を介して有線又は無線のいずれかで操作デバイス14、表示装置15及びカメラ16のそれぞれと接続される。具体的にインタフェース部13は、情報処理装置10が供給する映像データを表示装置15に送信するために、HDMI（登録商標）（High-Definition Multimedia Interface）等のマルチメディアインタフェースを含むこととする。また、操作

デバイス 14 が受け付けたユーザーの操作内容を示す操作信号や、カメラ 16 によって撮影された撮影画像を受信するために、USB (Universal Serial Bus) 等のデータ通信インタフェースを含んでいる。

[0015] 操作デバイス 14 は、ユーザーからの操作指示を受け付け、その内容を示す操作信号を情報処理装置 10 に送信する。表示装置 15 は、情報処理装置 10 から送信される映像信号に応じた映像を表示して、ユーザーに閲覧させる。

[0016] カメラ 16 は、視野内の様子を撮影して得られる撮影画像を出力する。本実施形態では、後述する再現映像を生成する際に、人の姿勢を表す姿勢データを取得するために用いられる。

[0017] 次に、情報処理装置 10 が実現する機能について、図 2 を用いて説明する。情報処理装置 10 は、再現対象となる物体（対象物）の外観を表す画像を入力として機械学習を行い、その結果を用いて対象物の外観を再現するために用いられる。図 2 に示すように、情報処理装置 10 は、機能的に、基準画像取得部 51 と、変換画像取得部 52 と、機械学習部 53 と、再現画像生成部 54 と、を含んでいる。これらの機能は、制御部 11 が記憶部 12 に記憶されたプログラムを実行することにより実現される。このプログラムは、インターネット等の通信ネットワークを介して情報処理装置 10 に提供されてもよいし、光ディスク等のコンピュータ読み取り可能な情報記憶媒体に格納されて提供されてもよい。

[0018] 基準画像取得部 51 は、後述する機械学習において教師データとして用いられる基準画像 I_r のデータを取得する。本実施形態において基準画像 I_r は、再現対象となる対象物を実際に撮影して得られる撮影画像であるものとする。また、対象物は人であることとし、この人物を対象者 T という。

[0019] 基準画像 I_r は、対象者 T の略全体を含むことが望ましい。ただし、再現対象となる部分が対象者 T の一部分に限られるのであれば、再現対象となる部分のみを含んでいればよい。また、基準画像 I_r は対象者 T 以外の背景などの色情報をできる限り含んでいないことが望ましい。そのため、対象者 T

を撮影して基準画像 I_r を生成する際には、特定色のスクリーンの前で対象者 T を撮影して、グリーンバックなどと呼ばれる手法により、背景の情報を基準画像 I_r から消去してもよい。あるいは、実際に対象者 T を撮影して得られる撮影画像から対象者 T が写っている部分のみを抽出して得られる画像を、基準画像 I_r として使用してもよい。

[0020] また、基準画像 I_r は、対象者 T の再現対象となる部分が略中心に位置する画像であることが望ましい。そのため、撮影時には対象者 T の目の位置が基準画像 I_r の縦方向の中心に略一致するように、カメラの高さを調整する。また、対象者 T の位置が横方向の中心に略一致するようにカメラの向き、及び対象者 T が立つ位置を調整する。あるいは、撮影した画像を後からクリッピングして対象者 T の目の位置が基準画像 I_r の中心に位置するように調整してもよい。図3は、以上説明したような要件を満たす基準画像 I_r の一例を示している。なお、ここでは人の目の位置を中心とすることとしたが、これとは別の部位が基準画像 I_r の中心に位置するよう調整してもよい。

[0021] 基準画像取得部51は、一人の対象者 T を様々な異なる複数の姿勢で撮影して得られる複数の基準画像 I_r のデータを取得する。ただし、撮影時のカメラから対象者 T までの距離は、どの基準画像 I_r についても変化がなく、略一定のままであるものとする。そのため、対象者 T を撮影する際には、対象者 T が同じ場所で様々な姿勢を取りつつ撮影を行えばよく、カメラの位置や対象者 T の位置を変化させて撮影を行う必要はない。

[0022] 変換画像取得部52は、基準画像 I_r を予め定められた倍率で拡大して変換画像 I_t を生成することによって、そのデータを取得する。この変換画像 I_t も、基準画像 I_r とともに教師データとして機械学習に利用される。特に変換画像取得部52は、1枚の基準画像 I_r を互いに異なる倍率で拡大することによって、1枚の基準画像 I_r から複数の変換画像 I_t を生成することとする。なお、以下の説明において倍率の値は、基準画像 I_r の縦横それぞれの向きの長さに対する倍率を意味している。また、変換画像 I_t は、縦横比を維持しつつ基準画像 I_r の一部領域を拡大して得られる、元の基準画

像 I_r と同じサイズの画像であるものとする。

[0023] 1枚の基準画像 I_r に基づいて生成される複数の変換画像 I_t それぞれの基準画像 I_r に対する倍率は、等比数列にしたがって順に増加する値に決定される。具体的に、1枚の基準画像 I_r から X 枚の変換画像 I_t を生成することとし、これら X 枚の変換画像 I_t を倍率が小さい方から順に変換画像 $I_t(1), I_t(2), \dots, I_t(X)$ と表記する。また、最大の倍率（すなわち、変換画像 $I_t(X)$ の倍率）を E_{max} とする。このとき、各変換画像 I_t の倍率は、公比

$$(E_{max})^{1/X}$$

で増加するように決定される。すなわち、 n 番目の変換画像 $I_t(n)$ の倍率 $E(n)$ は、

$$E(n) = (E_{max})^{n/X}$$

で計算される。

[0024] 具体例として、以下では1枚の基準画像 I_r に基づいて9枚の変換画像 $I_t(1) \sim I_t(9)$ が生成されることとする。また、最大の倍率（すなわち、変換画像 $I_t(9)$ の倍率）は4倍とする。このとき、変換画像 $I_t(n)$ の倍率 $E(n)$ は、以下の計算式に従って計算される。

$$E(n) = 4^{n/9}$$

[0025] 各変換画像 I_t は、元の基準画像 I_r と同じ大きさの画像となるように、元の基準画像 I_r の一部分のみを拡大して生成される。例えば倍率4倍の変換画像 $I_t(9)$ については、元の基準画像 I_r に含まれる縦横それぞれ $1/4$ のサイズの領域を拡大することによって生成される。基準画像 I_r が縦横それぞれ 1024 ピクセルの画像であれば、変換画像取得部52は、基準画像 I_r に含まれる縦横それぞれ $256 (= 1024 / 4)$ ピクセルの正方形の領域を拡大して、元の基準画像 I_r と同じ縦横それぞれ 1024 ピクセルの変換画像 $I_t(9)$ を生成することとなる。

[0026] ここで、拡大対象とする領域は、基準画像 I_r 内の中心位置を中心とする領域であることとする。ただし、対象者 T の再現対象としたい部分が基準画

像 I_r の中心に位置しておらず、ずれた位置に含まれている場合には、対象者 T の位置に合わせて中心からずれた領域を拡大対象としてもよい。

[0027] 以上説明したような画像処理によって得られる複数の変換画像 I_t は、対象者 T が実際の撮影時よりもカメラに少しずつ近づいた状態で撮影した撮影画像に近くなると想定される。すなわち、カメラから対象者 T までの距離が様々に変化した場合の撮影画像に相当する画像を、実際に撮影を行わずに生成することができる。このとき、複数の変換画像 I_t の倍率は、順に等比数列にしたがって増加するように決定されているので、これらの変換画像 I_t は、基準画像 I_r と最大倍率の変換画像 I_t (9) との間で等間隔に距離を変化させた際の撮影画像に近い大きさで、対象者 T を含んだ画像となる。図4は、このようにして図3に示す基準画像 I_r から生成された9枚の変換画像 I_t の一例を示している。

[0028] 機械学習部53は、変換画像取得部52によって取得された変換画像 I_t を教師データとして用いて、機械学習を行う。この機械学習は、対象者 T の姿勢と外観との対応関係についての学習であるものとする。以下では、機械学習部53が機械学習に使用する画像を学習対象画像という。学習対象画像は、少なくとも変換画像 I_t を含む。また、学習対象画像は基準画像取得部51によって取得された基準画像 I_r を含んでもよい。以下の説明では、学習対象画像は基準画像 I_r 及び変換画像 I_t の双方を含むものとする。例えば基準画像取得部51が同じ対象者 T を撮影して得られる100枚の基準画像 I_r を取得し、変換画像取得部52が100枚の基準画像 I_r のそれぞれを互いに異なる倍率で拡大することによって、1枚の基準画像 I_r につき9枚の変換画像 I_t を生成したものとする。この場合、100枚の基準画像 I_r と900枚の変換画像 I_t の計1000枚の画像が学習対象画像となる。

[0029] 以下、機械学習部53が実行する機械学習の具体的な内容について説明する。まず機械学習部53は、取得された全ての学習対象画像について、対象者 T の姿勢を特定する処理を行う。この姿勢特定処理は、人の身体を構成する各部位がどのような位置関係にあるかを特定する処理であって、公知の技

術により実現できる。姿勢特定処理の結果得られる姿勢データは、人の身体を構成する関節や特徴点などの画像内における位置を示す座標値の集合によって構成されてよい。図5は、図3の基準画像 I_r に対して姿勢特定処理を実行した結果得られる姿勢データを表す骨格モデルの一例を示している。

[0030] 機械学習部53は、基準画像 I_r だけでなく、その基準画像 I_r を拡大して得られる複数の変換画像 I_t のそれぞれについても、別個独立に姿勢特定処理を実行することとする。変換画像 I_t 内では対象者Tが基準画像 I_r や他の変換画像 I_t とは違う大きさで含まれており、その身体の一部しか含まれないような場合もあり得る。そのため、同じ姿勢の対象者Tを撮影して得られる基準画像 I_r と各変換画像 I_t との間で、姿勢特定処理の結果に相違が生じることもあるからである。

[0031] なお、姿勢特定処理の処理内容や処理の対象は、後述する再現処理においてどの程度対象者Tの外観を詳細に再現したいかによって変化してもよい。例えば対象者Tの手指の動きを再現する必要があるれば、学習対象画像に含まれる手指の姿勢を特定し、機械学習に用いる必要がある。また、対象者Tの表情を再現したい場合には、学習対象画像内の対象者Tの顔に含まれる目や口の位置や状態などを対象者Tの姿勢データの一部として特定してもよい。

[0032] 姿勢特定処理を終えると、機械学習部53は、姿勢特定処理の結果を用いて、実際に機械学習に利用する教師データを選択する。この教師データ選択処理は必須ではなく、取得された全ての学習対象画像を教師データとして機械学習に利用してもよい。しかしながら、学習に寄与しない、又は寄与の度合いが低いと想定される学習対象画像をこの選択処理によって教師データから除外することによって、機械学習に必要な所要時間や計算量等を削減し、処理効率を向上させることができる。

[0033] 教師データ選択処理の一例として、機械学習部53は、対象者Tの特定の部位が含まれない学習対象画像を教師データから除外してもよい。この特定の部位は、例えば対象者Tの頭部である。対象者Tが前屈みになっているときなど、対象者Tの姿勢によっては、対象者Tの頭部が基準画像 I_r の中心

からずれ、その結果基準画像 I_r の中心領域を拡大して変換画像 I_t を生成した際に、変換画像 I_t 内に対象者 T の頭部が含まれなくなってしまう場合がある。このような変換画像 I_t を機械学習の入力に利用しても、対象者 T の姿勢と外観との対応関係を効果的に学習することは難しい。そのため、このように重要な部位が含まれない状態の学習対象画像は、教師データから除外することとする。なお、このような特定の部位が含まれるか否かを判断基準とした選択処理は、変換画像 I_t だけについて実施してもよい。また、そもそも学習対象画像に対する姿勢の特定処理に失敗した場合や、特定された姿勢の精度が低いと判定された場合には、その学習対象画像を教師データから除外することとしてもよい。

[0034] 教師データ選択処理の別の例として、機械学習部 53 は、互いに姿勢が類似すると判定される複数の学習対象画像が発見された場合に、そのうちの一部の学習対象画像だけを教師データとして選択し、それ以外の学習対象画像を教師データから除外してもよい。このような類否判定処理を実現するために、例えば機械学習部 53 は、各学習対象画像から得られる姿勢データ（姿勢特定結果）に対して分類処理を実行して、互いに姿勢が近いと判断される学習対象画像をグルーピングする。この分類処理は、 k 平均法などの公知のクラスタリング技術によって実現されてよい。

[0035] この分類処理は、学習対象画像の全体について実行するのではなく、基準画像 I_r 、及び倍率が互いに等しい変換画像 I_t のグループのそれぞれについて独立に実行することが望ましい。すなわち、機械学習部 53 は、基準画像 I_r だけを含むグループを対象として分類処理を行った結果に基づいて、他の基準画像 I_r と姿勢データが類似すると判定される一部の基準画像 I_r を教師データから除外する。また、倍率 $4^{1/9}$ 倍で拡大された変換画像 I_t (1) のグループ、倍率 $4^{2/9}$ 倍で拡大された変換画像 I_t (2) のグループなど、それぞれ同じ倍率の変換画像 I_t が属する 9 個の変換画像 I_t のグループに対して、それぞれ独立に分類処理を行って、同様に類似する姿勢が他に存在する変換画像 I_t を教師データから除外する。このように倍率ごとに独

立して分類処理を実行するのは、倍率によって学習対象画像に含まれる対象者Tの大きさが異なり、それにより姿勢の類似度に寄与する部位も変化するからである。例えば対象者Tの全身が含まれる基準画像I_rについては、手や足の位置が近ければ姿勢が近いと判断される可能性が高くなる。一方、対象者Tの表情が複数の基準画像I_rの間で互いに異なっていたとしても、基準画像I_r内における目や口の位置は大きく変化しないため、このような顔の中の部位の位置情報は姿勢の類似度に対する寄与の度合いが相対的に小さくなる。これに対して最大倍率の変換画像I_t（9）においては、図4に例示したように対象者Tの顔を中心とした限られた領域のみが画像内に含まれ、手や足などは画像内に含まれなくなる。そのため、手や足の位置の相違は姿勢の類似度に寄与せず、顔に含まれる目や口の位置などが姿勢の類否を判断する上で重要となる。このように倍率のグループ毎に分類処理を実行する結果、ある基準画像I_rは他の基準画像I_rと姿勢が類似すると判断されて教師データから除外される一方で、この基準画像I_rを拡大して得られる変換画像I_tは教師データとして利用されるケースも生じ得る。このような分類処理の実行結果に基づいて実際に機械学習に使用する教師データを絞り込むことで、比較的少ない画像を教師データとして効率的に学習を行うことができる。

[0036] なお、以上説明した複数の選択処理は、互いに組み合わせて実行してもよい。また、以上説明したもの以外にも、各種の判断基準で教師データから除外する学習対象画像を選択してもよい。

[0037] 選択処理によって実際に機械学習に使用する学習対象画像を教師データとして選択すると、機械学習部53は、選択された学習対象画像と、その学習対象画像に基づいて特定された姿勢データの組を入力として、その姿勢を取っている対象者Tの外観がどのような画像になるかを学習する。このような機械学習は、ニューラルネットワークなどを用いた機械学習アルゴリズムによって実現できる。この機械学習によって、対象者Tの姿勢を表す姿勢データから対象者Tの外観を表す画像を推定する推定器を実現するための学習済

みデータが生成される。

[0038] 再現画像生成部54は、機械学習部53が実行する機械学習によって得られる学習済みデータを用いて、対象者Tの外観を再現する画像（再現画像）を生成する。具体的に再現画像生成部54は、対象者Tに取らせた姿勢を表す姿勢データ（骨格モデルデータ）を取得し、機械学習部53が生成した学習済みデータを使用する推定器に入力する。推定器は、その骨格モデルデータが表す姿勢を対象者Tが取った場合にどのような外観になるかを推定し、推定結果の画像を出力する。

[0039] より具体的に、本実施形態では、対象者Tとは別の人物（以下、利用者Uという）の動作を反映して、その動作と対応する動作を対象者Tが取っているかのように見える映像（再現映像）を生成し、出力する。このような制御を実現するため、再現画像生成部54は、カメラ16で利用者Uを撮影して得られる撮影画像を取得する。このとき、カメラ16による利用者Uの撮影条件は、基準画像I_rを用意するために対象者Tを撮影した際の撮影条件に近いことが望ましい。すなわち、カメラ16は、対象者Tを撮影したカメラと同じカメラであっても異なるカメラであってもよいが、解像度などは同じであることが望ましい。また、撮影時のカメラ16の高さも、対象者Tを撮影した際のカメラの高さと近いほうがよい。

[0040] 一方で、カメラ16と利用者Uとの間の距離は一定である必要がなく、利用者Uはカメラ16に近づいたり遠ざかったりする移動を行ってもよい。ただし、その距離範囲は、基準画像I_rを取得するための撮影が行われた際のカメラと対象者Tとの間の距離を超えないことが望ましい。また、以上の説明では変換画像I_tを生成する際の最大倍率を縦横それぞれ4倍としているので、この倍率を超える大きさで写ってしまう距離までカメラ16に近づきすぎることは好ましくない。しかしながら、基準画像I_rに対応する距離と最大倍率の変換画像I_tに対応する距離の間の距離範囲内であれば、任意の位置に移動しながら撮影を行うことができる。図6は、このような利用者Uの撮影の様子を示している。

- [0041] 再現画像生成部54は、利用者Uを含むカメラ16の撮影画像に対して、姿勢特定処理を実行する。この姿勢特定処理は、機械学習部53が学習対象データを生成するために学習対象画像に実行した姿勢特定処理と同じアルゴリズムのものである。これにより、利用者Uの撮影時点における姿勢を表す姿勢データが得られる。
- [0042] 再現画像生成部54は、利用者Uの姿勢データを、機械学習部53が生成した学習済みデータを用いる推定器に入力する。その結果として、対象者Tが利用者Uと同じ姿勢を取っている様子を表す再現画像が生成される。再現画像生成部54は、この再現画像を描画し、表示装置15の画面に表示する。以上説明した手順による利用者Uの撮影画像に基づく再現画像の表示処理を、利用者Uが様々な姿勢を変化させながらリアルタイムで繰り返し実行することにより、利用者Uと同じように対象者Tが動く様子を示す再現映像を生成、表示することができる。図7は、図6に示す利用者Uの姿勢に応じて再現画像生成部54が生成する対象者Tの再現画像の一例を示している。
- [0043] ここで、様々な姿勢を取っている対象者Tを撮影して得られる複数の基準画像 I_r を用いて機械学習を行った結果得られる学習済みデータを利用することで、再現画像生成部54は、対象者Tが現実には取っていないような姿勢を取った際にどのような外観となるかを推定し、再現画像を生成することができる。さらに、基準画像 I_r だけでなく変換画像 I_t を学習対象に含めて機械学習を行った結果の学習済みデータを利用することで、再現画像生成部54は、対象者Tを撮影した際の距離よりも近い位置まで利用者Uがカメラ16に近づいたとしても、対象者Tの外観を比較的精度よく再現することができる。これにより、カメラ16の撮影方向に沿って移動する動きを含め、利用者Uが様々な動作を行った場合に、その動作と同じ動作を取る対象者Tの様子を動画像として表示することができる。このような処理によれば、例えば窓口における受付対応などのように、その場には存在しない対象者Tが現実にかかる状況に反応する動作を行う様子を、リアルタイムで映像として生成し、閲覧者に提示することができる。

[0044] なお、再現画像生成部54が生成する再現画像は、対象者Tが含まれる部分以外の領域を透明な画素に変換した画像であってもよい。このような画素変換は、生成された再現画像内において基準画像I_rの背景色と同じ色の画素を透明色に変換することによって実現できる。このように背景を透明にした再現画像を所与の背景画像と合成することで、対象者Tが実際には行っていない別の場所にいるかのような画像を生成することができる。このような合成を行う際には、対象者Tの明度に合わせて背景画像の明度を調整したり、背景画像と対象者Tとの境界部分にガウシアンノイズを追加したりするなどして境界部分が目立たないようにする処理を実行してもよい。

[0045] 以上説明したように、本実施形態に係る情報処理装置10によれば、固定された位置にいる対象者Tを撮影して得られる画像に基づいて、この対象者Tが前後方向の移動も含めて様々な動作を取る様子を再現する画像を生成することができる。

[0046] なお、本発明の実施の形態は、以上説明したものに限られない。例えば以上の説明では再現対象となる対象物は人であることとしたが、これ以外にも、例えば動物など姿勢が変化する各種の物であってよい。また、以上の説明では一人の対象者Tのみを再現対象とすることとしたが、複数の対象者についてそれぞれ複数の基準画像I_rを取得し、それぞれ独立に機械学習を行って学習済みデータを生成すれば、複数の対象者を再現対象とすることも可能である。この場合、再現画像生成部54は、まず再現対象となる対象者の選択を利用者から受け入れ、選択された対象者に対応する学習済みデータを読み出して、再現画像の生成に利用する。

[0047] また、以上の説明では一つの情報処理装置10が機械学習と再現画像の生成処理の双方を実施することとしたが、これに限らずこれらの処理は互いに別の情報処理装置によって実現されてもよい。この場合、機械学習を行った情報処理装置は、その結果得られる学習済みデータをコンピュータ読み取り可能な情報記憶媒体に格納したり、通信ネットワークを介して配信したりして外部の情報処理装置に提供する。提供を受けた情報処理装置は、提供され

た学習済みデータを利用することで、以上説明したような再現画像生成部54の機能を実現することができる。

[0048] また、以上の説明における変換画像の倍率や1枚の基準画像に基づいて生成する変換画像の数などの各種のパラメータは、例示に過ぎない。また、以上の説明では変換画像は基準画像を拡大して得られる拡大画像であることとしたが、これに限らず、基準画像を所与の倍率で縮小して得られる縮小画像を変換画像として機械学習に利用してもよい。このような縮小画像を機械学習に用いることで、基準画像の撮影時におけるカメラから対象者Tまでの距離よりも離れた位置まで利用者Uがカメラ16から遠ざかった場合にも、精度よく対象者Tを再現する再現画像を生成することができる。

符号の説明

[0049] 10 情報処理装置、11 制御部、12 記憶部、13 インタフェース部、14 操作デバイス、15 表示装置、16 カメラ、51 基準画像取得部、52 変換画像取得部、53 機械学習部、54 再現画像生成部。

請求の範囲

- [請求項1] 再現対象となる対象物を撮影して得られる複数の基準画像を取得する基準画像取得部と、
- 前記複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を取得する変換画像取得部と、
- 前記複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行し、前記対象物の外観を表す再現画像を生成するために用いられる学習済みデータを生成する機械学習部と、
- を含むことを特徴とする情報処理装置。
- [請求項2] 請求項1に記載の情報処理装置において、
- 前記複数の学習対象画像は、前記複数の基準画像を含む
- ことを特徴とする情報処理装置。
- [請求項3] 請求項1又は2に記載の情報処理装置において、
- 前記機械学習部は、前記複数の学習対象画像のそれぞれに対して、当該学習対象画像に含まれる前記対象物の姿勢を特定する処理を実行し、当該処理の結果得られる姿勢データと、当該学習対象画像とを教師データとして用いて前記機械学習を実行し、前記対象物の姿勢データを前記再現画像に変換するために用いられる学習済みデータを生成する
- ことを特徴とする情報処理装置。
- [請求項4] 請求項3に記載の情報処理装置において、
- 前記機械学習部は、前記姿勢を特定する処理を実行した結果、前記対象物の所定の部位を含まないと判定された学習対象画像を前記教師データから除外する
- ことを特徴とする情報処理装置。
- [請求項5] 請求項3又は4に記載の情報処理装置において、
- 前記機械学習部は、前記複数の学習対象画像に対して、それぞれの学習対象画像から得られた姿勢データが互いに類似するか否かを判定す

る類否判定処理を実行し、得られた姿勢データが他の学習対象画像から得られた姿勢データに類似すると判定された一部の学習対象画像を、前記教師データから除外する

ことを特徴とする情報処理装置。

[請求項6]

請求項5に記載の情報処理装置において、

前記変換画像取得部は、1枚の前記基準画像を互いに異なる倍率で拡大又は縮小して複数の前記変換画像を生成し、

前記機械学習部は、前記複数の学習対象画像を、それぞれ互いに同じ倍率で拡大又は縮小された複数の変換画像を含む複数のグループに分類し、当該複数のグループのそれぞれに対して独立に前記類否判定処理を実行して前記教師データから除外する学習対象画像を選択することを特徴とする情報処理装置。

[請求項7]

請求項1から6のいずれか一項に記載の情報処理装置において、

前記変換画像取得部は、1枚の前記基準画像を互いに異なる倍率で拡大又は縮小して複数の前記変換画像を生成し、かつ、前記互いに異なる倍率は、等比数列にしたがって順に変化する倍率である

ことを特徴とする情報処理装置。

[請求項8]

再現対象となる対象物を撮影して得られる複数の基準画像を取得するステップと、

前記複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を取得するステップと、

前記複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行し、前記対象物の外観を表す再現画像を生成するために用いられる学習済みデータを生成するステップと、

を含むことを特徴とする情報処理方法。

[請求項9]

再現対象となる対象物を撮影して得られる複数の基準画像を取得するステップと、

前記複数の基準画像のそれぞれを拡大又は縮小して得られる複数の

変換画像を取得するステップと、

前記複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行し、前記対象物の外観を表す再現画像を生成するために用いられる学習済みデータを生成するステップと、

をコンピュータに実行させるためのプログラム。

[請求項10]

再現対象となる対象物を撮影して得られる複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行した結果得られる、前記対象物の外観を表す再現画像を生成するために用いられる学習済みデータ。

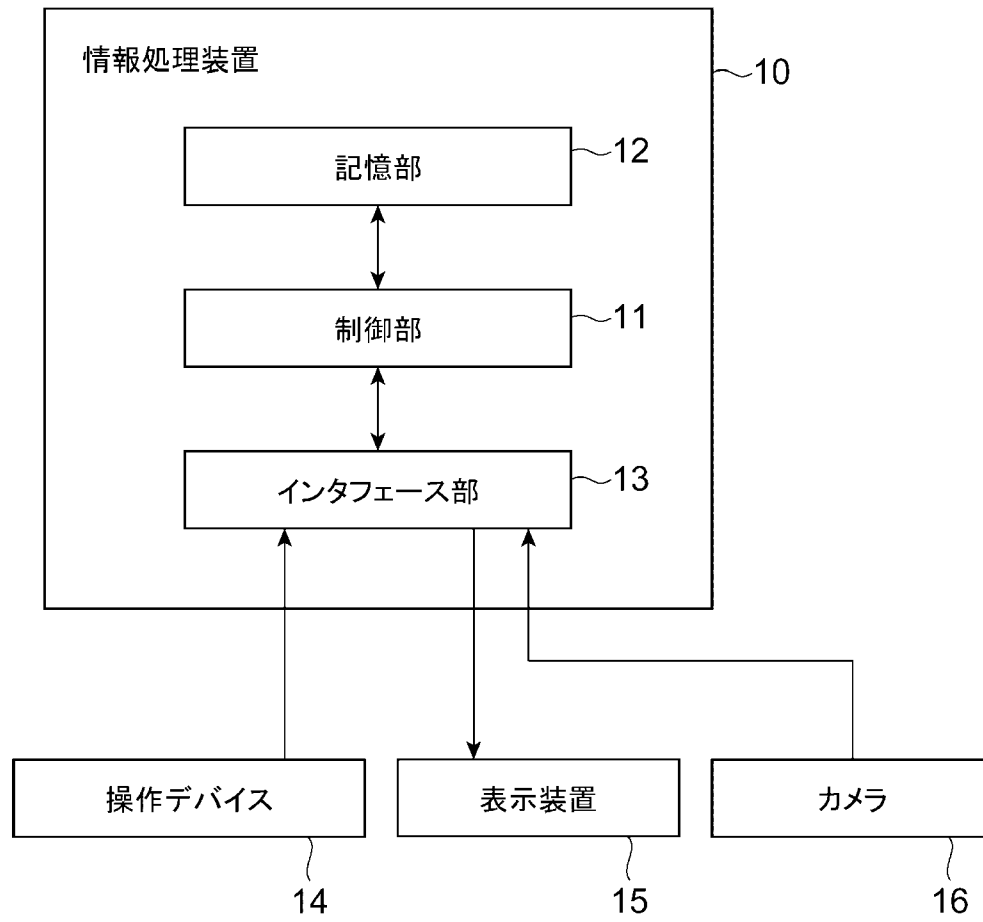
[請求項11]

再現対象となる対象物の姿勢を表す姿勢データを取得する取得部と、

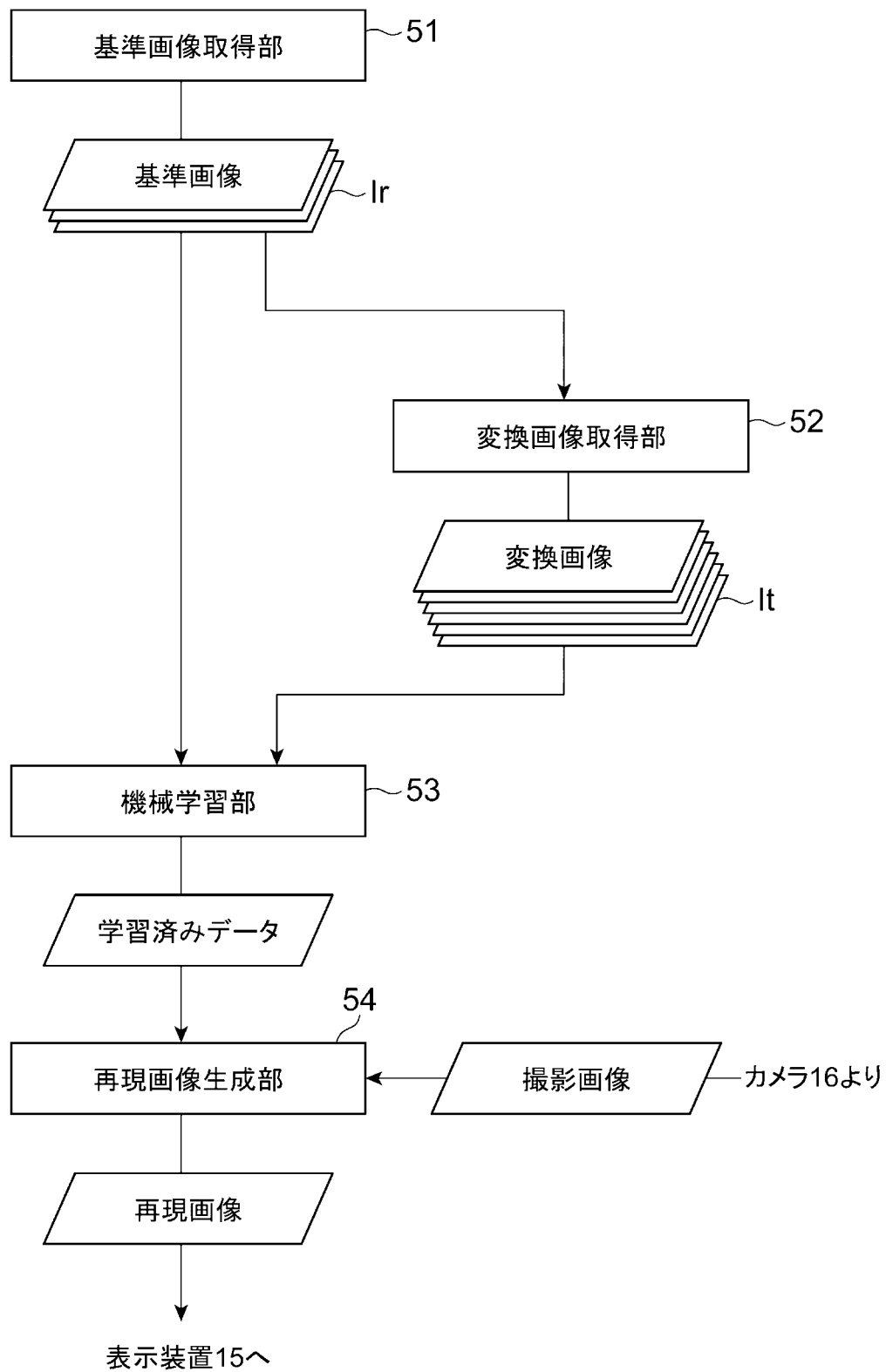
対象物を撮影して得られる複数の基準画像のそれぞれを拡大又は縮小して得られる複数の変換画像を含む複数の学習対象画像を教師データとして用いて機械学習を実行した結果得られる学習済みデータを用いて、前記姿勢データが表す姿勢を前記対象物が取っている様子を表す再現画像を生成する再現画像生成部と、

を含むことを特徴とする情報処理装置。

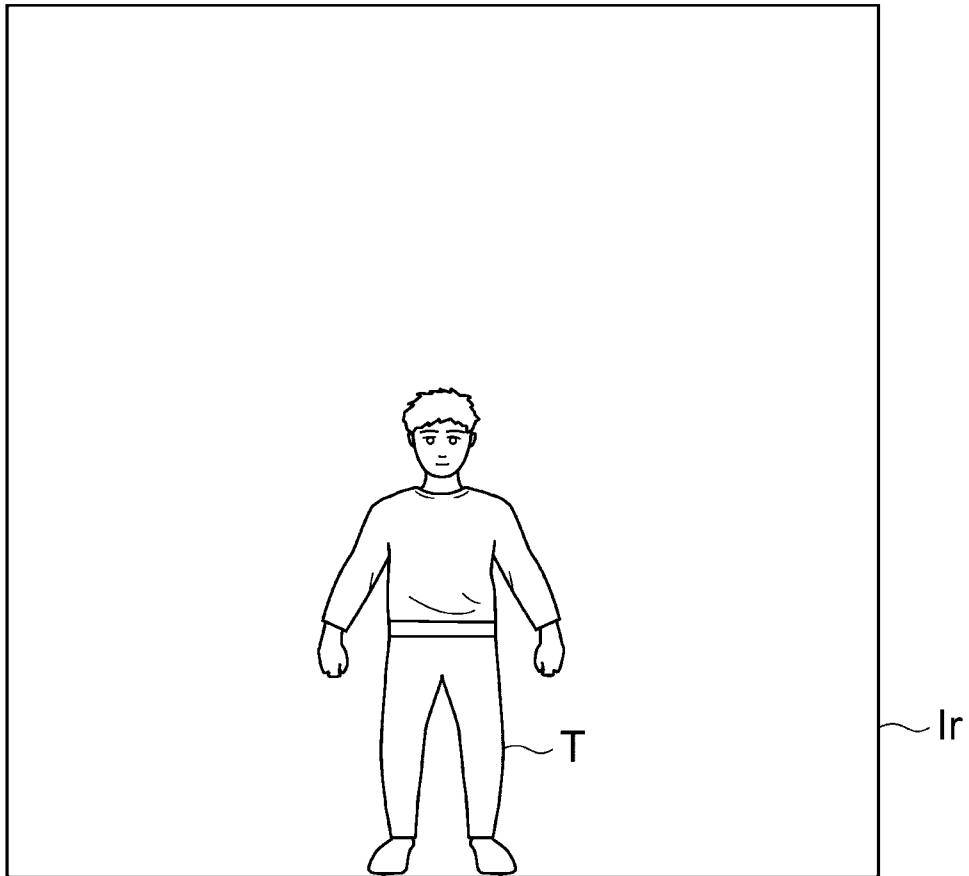
[図1]



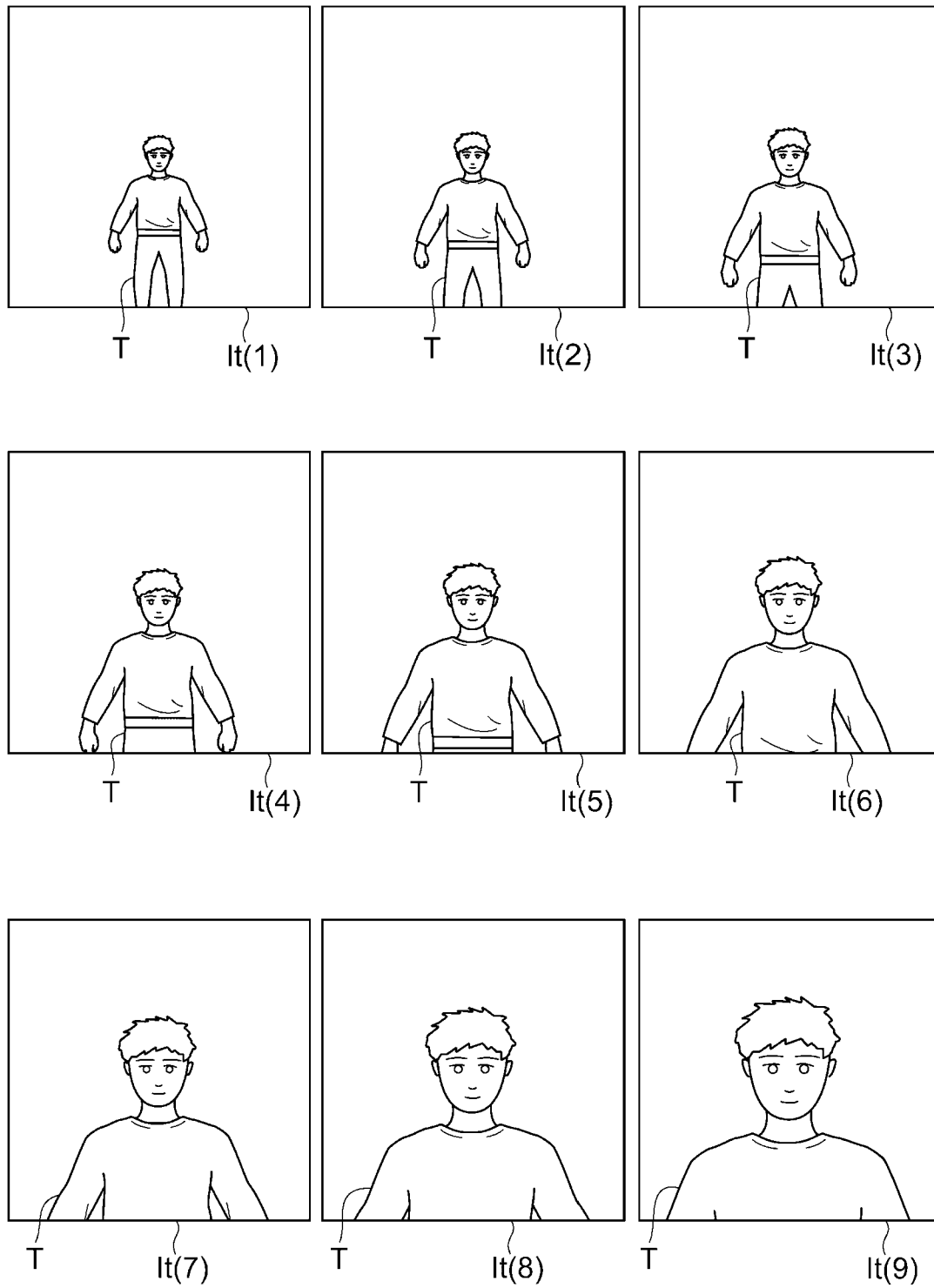
[図2]



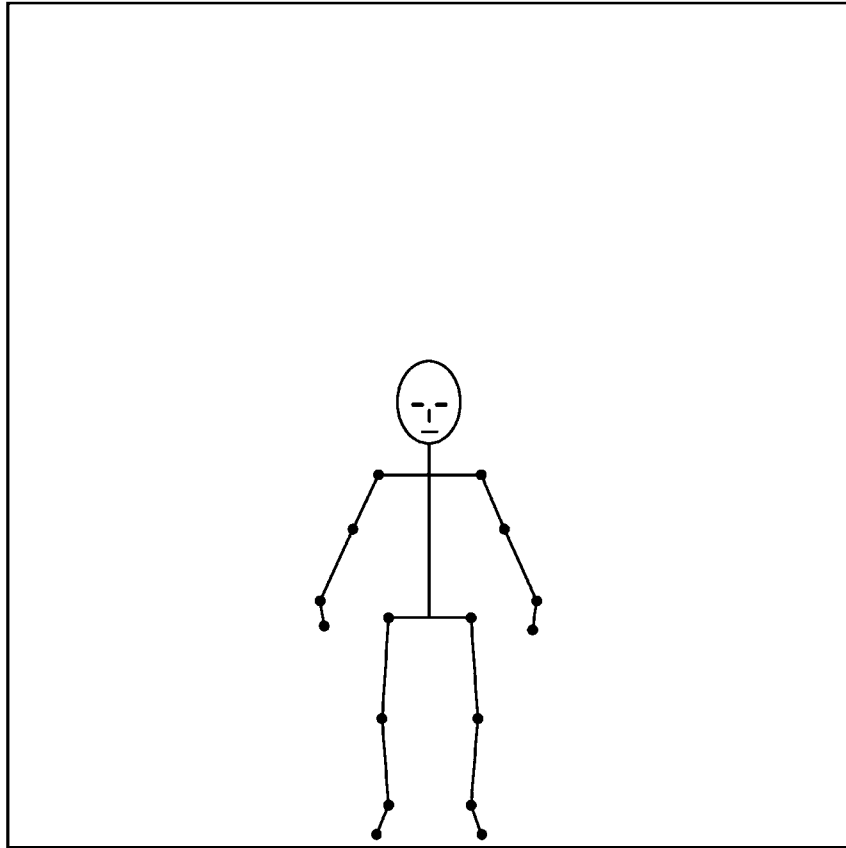
[図3]



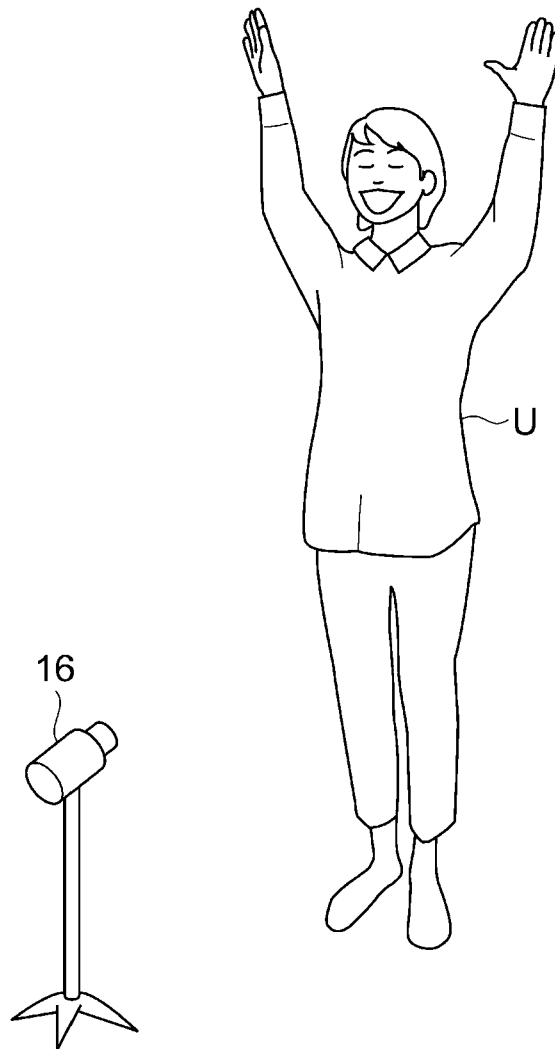
[図4]



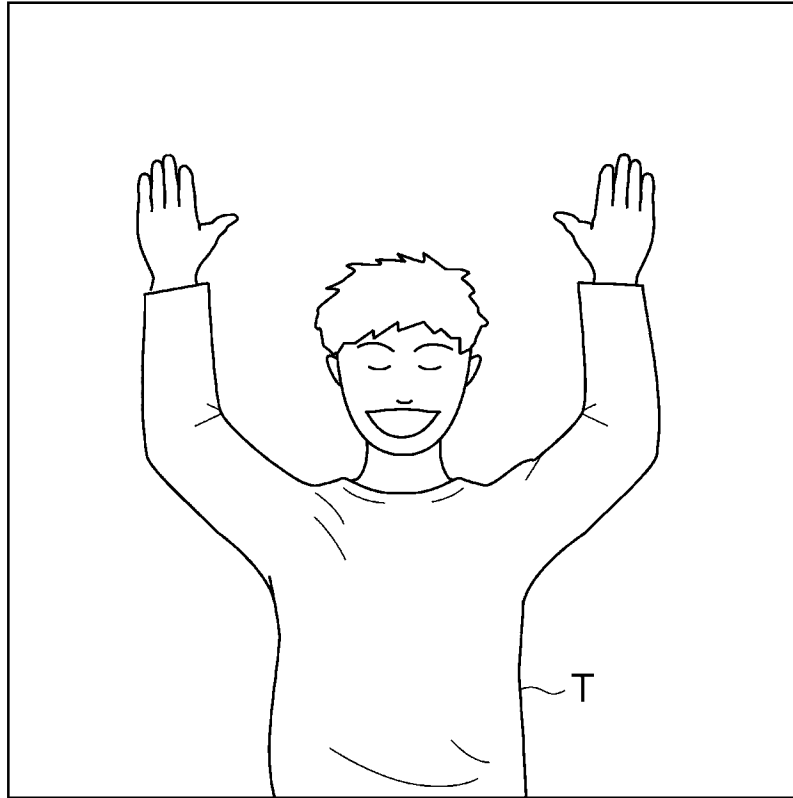
[図5]



[図6]



[図7]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/000964

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. H04N7/18 (2006.01) i, G06N99/00 (2019.01) i, G06T7/00 (2017.01) i, G06T19/00 (2011.01) i

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)

Int.Cl. H04N7/18, G06N99/00, G06T7/00, G06T19/00

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	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

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.
Y A	JP 2018-136632 A (OMRON CORP.) 30 August 2018, paragraphs [0023]-[0051], fig. 1-2 & WO 2018/150901 A1, paragraphs [0023]-[0051], fig. 1-2	1-2, 8-10 3-7, 11
Y A	JP 2018-169672 A (MITSUBISHI ELECTRIC INFORMATION SYSTEMS CORP.) 01 November 2018, paragraphs [0037]-[0038], fig. 4 (Family: none)	1-2, 8-10 3-7, 11
Y A	JP 2018-116599 A (CANON INC.) 26 July 2018, paragraphs [0003], [0052] & US 2018/0211138 A1, paragraphs [0003], [0077]	1-2, 8-10 3-7, 11



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
15 March 2019 (15.03.2019)

Date of mailing of the international search report
26 March 2019 (26.03.2019)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/000964

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2018-129007 A (NIPPON TELEGRAPH AND TELEPHONE CORP.) 16 August 2018, entire text (Family: none)	1-11
A	JP 2012-088787 A (CANON INC.) 10 May 2012, paragraph [0042] (Family: none)	1-11

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. H04N7/18(2006.01)i, G06N99/00(2019.01)i, G06T7/00(2017.01)i, G06T19/00(2011.01)i

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））

Int.Cl. H04N7/18, G06N99/00, G06T7/00, G06T19/00

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019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
Y A	JP 2018-136632 A（オムロン株式会社）2018.08.30, 段落[0023]-[0051], 図 1-2 & WO 2018/150901 A1, 段落[0023]-[0051], 図 1-2	1-2, 8-10 3-7, 11
Y A	JP 2018-169672 A（三菱電機インフォメーションシステムズ株式会 社）2018.11.01, 段落[0037]-[0038], 図 4 (ファミリーなし)	1-2, 8-10 3-7, 11

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矛盾するものではなく、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

15.03.2019

国際調査報告の発送日

26.03.2019

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

鈴木 隆夫

5 P

5891

電話番号 03-3581-1101 内線 3581

C (続き) . 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
Y A	JP 2018-116599 A (キヤノン株式会社) 2018. 07. 26, 段落[0003], [0052] & US 2018/0211138 A1, 段落[0003], [0077]	1-2, 8-10 3-7, 11
A	JP 2018-129007 A (日本電信電話株式会社) 2018. 08. 16, 全文 (ファミリーなし)	1-11
A	JP 2012-088787 A (キヤノン株式会社) 2012. 05. 10, 段落[0042] (ファミリーなし)	1-11