

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2018-132855
(P2018-132855A)

(43) 公開日 平成30年8月23日(2018.8.23)

| | | | | | | |
|----------------------------|-----------|------|-------|-------------|--|--|
| (51) Int.Cl. | F 1 | | | テーマコード (参考) | | |
| G06T 3/00 (2006.01) | G06T 3/00 | 740 | 5B057 | | | |
| G06T 7/00 (2017.01) | G06T 7/00 | 350C | 5L096 | | | |

審査請求 未請求 請求項の数 9 O L (全 13 頁)

(21) 出願番号 特願2017-24688 (P2017-24688)
(22) 出願日 平成29年2月14日 (2017.2.14)

(71) 出願人 504133110
国立大学法人電気通信大学
東京都調布市調布ヶ丘一丁目5番地1

(74) 代理人 100107766
弁理士 伊東 忠重

(74) 代理人 100070150
弁理士 伊東 忠彦

(72) 発明者 柳井 啓司
東京都調布市調布ヶ丘一丁目5番地1 国立大学法人電気通信大学内

Fターム(参考) 5B057 BA23 CA08 CA12 CA16 CB08
CB12 CB16 CD11 DA20 DB02
DB09 DC33 DC40
5L096 AA06 CA22 DA01 HA09 HA11
KA04

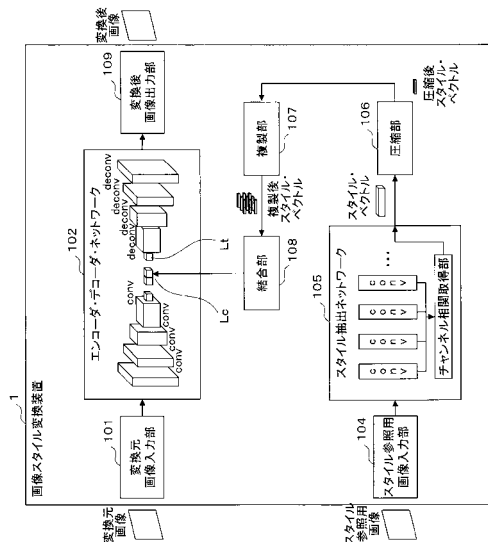
(54) 【発明の名称】 画像スタイル変換装置、画像スタイル変換方法および画像スタイル変換プログラム

(57) 【要約】 (修正有)

【課題】リアルタイム処理を可能としつつ、任意のスタイルに変換することのできる、画像スタイル変換技術を提供する。

【解決手段】画像スタイル変換装置1は、変換元となる画像を入力する変換元画像入力部101と、スタイル参照用の画像を入力するスタイル参照画像入力部104と、変換元となる画像を事前の学習結果に基づいて変換後の画像に変換する、畳み込みニューラルネットワークによるエンコーダ・デコーダ・ネットワーク102と、スタイル参照用の画像のスタイルを示すスタイル・ベクトルを抽出する、畳み込みニューラルネットワークによるスタイル抽出ネットワーク105と、スタイル抽出ネットワークにより抽出されたスタイル・ベクトルを、エンコーダ・デコーダ・ネットワークの中間層に結合する結合部108と、エンコーダ・デコーダ・ネットワークから変換後の画像を取得して出力する変換後画像出力部109とを備える。

【選択図】 図3



【特許請求の範囲】

【請求項 1】

変換元となる画像を入力する第 1 の画像入力部と、
 スタイル参照用の画像を入力する第 2 の画像入力部と、
 前記第 1 の画像入力部により入力された変換元となる画像を、事前の学習結果に基づいて変換後の画像に変換する、畳み込みニューラルネットワークによるエンコーダ・デコーダ・ネットワークと、
 前記第 2 の画像入力部により入力されたスタイル参照用の画像のスタイルを示すスタイル・ベクトルを抽出する、畳み込みニューラルネットワークによるスタイル抽出ネットワークと、
 前記スタイル抽出ネットワークにより抽出されたスタイル・ベクトルを、前記エンコーダ・デコーダ・ネットワークの中間層に結合する結合部と、
 前記エンコーダ・デコーダ・ネットワークから変換後の画像を取得して出力する画像出力部と
 を備えたことを特徴とする画像スタイル変換装置。

10

【請求項 2】

変換元となる画像を入力する画像入力部と、
 前記画像入力部により入力された変換元となる画像を、事前の学習結果に基づいて変換後の画像に変換する、畳み込みニューラルネットワークによるエンコーダ・デコーダ・ネットワークと、
 ユーザ設定に応じてスタイルを示すスタイル・ベクトルを生成するスタイル・ベクトル生成部と、
 前記スタイル・ベクトル生成部により生成されたスタイル・ベクトルを、前記エンコーダ・デコーダ・ネットワークの中間層に結合する結合部と、
 前記エンコーダ・デコーダ・ネットワークから変換後の画像を取得して出力する画像出力部と
 を備えたことを特徴とする画像スタイル変換装置。

20

【請求項 3】

前記結合部に与えられるスタイル・ベクトルを圧縮する圧縮部を備え、
 前記結合部は、圧縮後のスタイル・ベクトルを前記エンコーダ・デコーダ・ネットワークの中間層に結合する
 ことを特徴とする請求項 1 または 2 に記載の画像スタイル変換装置。

30

【請求項 4】

前記圧縮部は、主成分分析、非負行列因子分解法、または、ニューラルネットワークにより、スタイル・ベクトルを圧縮する
 ことを特徴とする請求項 3 に記載の画像スタイル変換装置。

【請求項 5】

前記圧縮後のスタイル・ベクトルを複製する複製部を備え、
 前記結合部は、複製後のスタイル・ベクトルを前記エンコーダ・デコーダ・ネットワークの中間層に結合する
 ことを特徴とする請求項 3 または 4 に記載の画像スタイル変換装置。

40

【請求項 6】

変換元となる画像およびスタイル参照用の画像を入力し、
 入力された前記スタイル参照用の画像のスタイルを示すスタイル・ベクトルを、畳み込みニューラルネットワークのスタイル抽出ネットワークにより抽出し、
 抽出された前記スタイル・ベクトルを、畳み込みニューラルネットワークのエンコーダ・デコーダ・ネットワークの中間層に結合し、
 入力された前記変換元となる画像を、前記エンコーダ・デコーダ・ネットワークにより、事前の学習結果に基づいて変換後の画像に変換し、
 前記エンコーダ・デコーダ・ネットワークから変換後の画像を取得して出力する、

50

処理をコンピュータが実行することを特徴とする画像スタイル変換方法。

【請求項 7】

変換元となる画像を入力し、
 ユーザ設定に応じてスタイルを示すスタイル・ベクトルを生成し、
 生成された前記スタイル・ベクトルを、畳み込みニューラルネットワークのエンコーダ・デコーダ・ネットワークの中間層に結合し、
 入力された前記変換元となる画像を、前記エンコーダ・デコーダ・ネットワークにより、事前の学習結果に基づいて変換後の画像に変換し、
 前記エンコーダ・デコーダ・ネットワークから変換後の画像を取得して出力する、
 処理をコンピュータが実行することを特徴とする画像スタイル変換方法。

10

【請求項 8】

変換元となる画像およびスタイル参照用の画像を入力し、
 入力された前記スタイル参照用の画像のスタイルを示すスタイル・ベクトルを、畳み込みニューラルネットワークのスタイル抽出ネットワークにより抽出し、
 抽出された前記スタイル・ベクトルを、畳み込みニューラルネットワークのエンコーダ・デコーダ・ネットワークの中間層に結合し、
 入力された前記変換元となる画像を、前記エンコーダ・デコーダ・ネットワークにより、事前の学習結果に基づいて変換後の画像に変換し、
 前記エンコーダ・デコーダ・ネットワークから変換後の画像を取得して出力する、
 処理をコンピュータに実行させることを特徴とする画像スタイル変換プログラム。

20

【請求項 9】

変換元となる画像を入力し、
 ユーザ設定に応じてスタイルを示すスタイル・ベクトルを生成し、
 生成された前記スタイル・ベクトルを、畳み込みニューラルネットワークのエンコーダ・デコーダ・ネットワークの中間層に結合し、
 入力された前記変換元となる画像を、前記エンコーダ・デコーダ・ネットワークにより、事前の学習結果に基づいて変換後の画像に変換し、
 前記エンコーダ・デコーダ・ネットワークから変換後の画像を取得して出力する、
 処理をコンピュータに実行させることを特徴とする画像スタイル変換プログラム。

30

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像スタイル変換装置、画像スタイル変換方法および画像スタイル変換プログラムに関する。

【背景技術】

【0002】

絵画等の画像が、その画像に含まれる対象物に対応するコンテンツと、表現の手法に対応するスタイルとに分離可能であることが知られている。そして、畳み込みニューラルネットワーク (Convolutional Neural Networks: CNN) を用いることにより、任意の写真等の画像から、コンテンツを残しつつ、元のスタイルを著名な画家の画風等のスタイルに変換する技術が提案されている (非特許文献 1 ~ 4 等を参照)。

40

【0003】

非特許文献 1 に開示される技術は、反復処理を必要とすることから、画像 1 枚についての処理時間が長く (例: 5 分程度)、リアルタイムの処理は困難であった。

【0004】

非特許文献 2 に開示される技術は、フィードフォワード処理により処理時間を短縮し、リアルタイムの処理を可能としているが、学習された 1 種類のスタイルについてしか適用できないものであった。

【0005】

非特許文献 3 に開示される技術は、リアルタイム処理を可能としつつ、30 種類程度の

50

学習済のスタイルからスタイルを選択可能としているが、任意のスタイルに変換することはできなかった。

【 0 0 0 6 】

非特許文献 4 に開示される技術は、コンテンツ特徴とスタイル特徴から変換後の画像の特徴マップを生成し、事前に学習した逆変換ネットワークによって特徴マップから画像を復元することで、スタイルが入れ替わった画像を得ている。しかし、特徴マップの生成に若干時間がかかり（1 秒程度）、高速ではあるが厳密なリアルタイム処理は難しい。また、本来難しい逆変換が行われているため、生成される画像のクオリティを高めるのは困難である。

【 先行技術文献 】

10

【 非特許文献 】

【 0 0 0 7 】

【 非特許文献 1 】 L. A. Gatys et al.: Image style transfer using convolutional neural networks, CVPR, 2016. (ArXiv:1508.06576, 2015/08)

【 非特許文献 2 】 J. Johnson et al.: Perceptual Losses for Real-Time Style Transfer and Super-Resolution, ECCV, 2016. (ArXiv: 1603.08155)

【 非特許文献 3 】 V. Dumoulin et al.: A Learned Representation for Artistic Style, ArXiv 1610.07629, 2016/10.

【 非特許文献 4 】 Tian Qi Chen et al.: Fast Patch-based Style Transfer of Arbitrary Style, arXiv 1612.04337v1 [cs.CV] 13 Dec 2016.

20

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 8 】

上述したように、従来、リアルタイム処理を可能としつつ、任意のスタイルに変換することは困難であり、その解決が要望されていた。

【 0 0 0 9 】

本発明は上記の従来の問題点に鑑み提案されたものであり、その目的とするところは、リアルタイム処理を可能としつつ、任意のスタイルに変換することのできる、画像スタイル変換技術を提供することにある。

【 課題を解決するための手段 】

30

【 0 0 1 0 】

上記の課題を解決するため、本発明にあつては、変換元となる画像を入力する第 1 の画像入力部と、スタイル参照用の画像を入力する第 2 の画像入力部と、前記第 1 の画像入力部により入力された変換元となる画像を、事前の学習結果に基づいて変換後の画像に変換する、畳み込みニューラルネットワークによるエンコーダ・デコーダ・ネットワークと、前記第 2 の画像入力部により入力されたスタイル参照用の画像のスタイルを示すスタイル・ベクトルを抽出する、畳み込みニューラルネットワークによるスタイル抽出ネットワークと、前記スタイル抽出ネットワークにより抽出されたスタイル・ベクトルを、前記エンコーダ・デコーダ・ネットワークの中間層に結合する結合部と、前記エンコーダ・デコーダ・ネットワークから変換後の画像を取得して出力する画像出力部とを備える。

40

【 発明の効果 】

【 0 0 1 1 】

本発明にあつては、リアルタイム処理を可能としつつ、任意のスタイルに変換することができる。

【 図面の簡単な説明 】

【 0 0 1 2 】

【 図 1 】 本発明の一実施形態にかかる画像スタイル変換装置の学習時における機能構成例を示す図（その 1）である。

【 図 2 】 本発明の一実施形態にかかる画像スタイル変換装置の学習時における機能構成例を示す図（その 2）である。

50

【図3】本発明の一実施形態にかかる画像スタイル変換装置の変換時における機能構成例を示す図(その1)である。

【図4】本発明の一実施形態にかかる画像スタイル変換装置の変換時における機能構成例を示す図(その2)である。

【図5】本発明の一実施形態にかかる画像スタイル変換装置の変換時における機能構成例を示す図(その3)である。

【図6】画像スタイル変換装置のハードウェア構成例を示す図である。

【図7】学習時の処理例を示すフローチャートである。

【図8】変換時の処理例を示すフローチャート(その1)である。

【図9】変換元画像、スタイル参照画像および変換後画像の例を示す図である。

10

【図10】変換時の処理例を示すフローチャート(その2)である。

【発明を実施するための形態】

【0013】

以下、本発明の好適な実施形態につき説明する。

【0014】

<構成>

図1は本発明の一実施形態にかかる画像スタイル変換装置1の学習時における機能構成例を示す図である。図1において、画像スタイル変換装置1は、変換元画像入力部101とエンコーダ・デコーダ・ネットワーク102とロス・ネットワーク103とを備えている。また、画像スタイル変換装置1は、スタイル参照用画像入力部104とスタイル抽出ネットワーク105と圧縮部106と複製部107と結合部108とを備えている。

20

【0015】

変換元画像入力部101は、学習用の変換元画像(例えば、8万枚程度)を逐次に入力する機能を有している。学習用の変換元画像は、変換時に与えられる可能性のある変換元画像をカバーできるような、偏りのない広範な画像が選ばれる。

【0016】

エンコーダ・デコーダ・ネットワーク102は、変換元画像入力部101により入力された変換元画像に対して変換を行い、変換後画像を出力する機能を有している。エンコーダ・デコーダ・ネットワーク102は、畳み込みニューラルネットワークによる複数のエンコード層(畳み込み層)convと、複数のデコード層deconvとを有している。また、複数のエンコード層とデコード層の中間に、複製後スタイル・ベクトルが結合される結合層Lcと、その直後にデータ(Feature Maps)のサイズが結合前のサイズに戻される調整層Ltとが設けられている。各エンコード層は、1以上の畳み込み処理部と、プーリング処理部とを有し、各層の結果のデータ(Feature Maps)が内部的に保持される。各デコード層は、エンコード層と逆方向の変換を行う。結合層Lcと調整層Ltは、例えば、図示のようにエンコード層とデコード層の繋ぎ目に設けられるものでもよいし、繋ぎ目よりも前のエンコード層の途中に設けられるものでもよい。

30

【0017】

ロス・ネットワーク103は、学習済の畳み込みニューラルネットワークによるものであり、評価対象の画像と、コンテンツ参照用の画像と、スタイル参照用の画像とを入力する。そして、ロス・ネットワーク103は、評価対象の画像のコンテンツとコンテンツ参照用の画像のコンテンツとの差異(コンテンツ・ロス)と、評価対象の画像のスタイルとスタイル参照用の画像のスタイルとの差異(スタイル・ロス)とを出力する機能を有している。コンテンツ・ロスとスタイル・ロスはエンコーダ・デコーダ・ネットワーク102にフィードバックされ、コンテンツ・ロスとスタイル・ロスが最小となるようにバックプロパゲーションにより最適化が行われる。

40

【0018】

ロス・ネットワーク103において、コンテンツについては、例えば、畳み込みニューラルネットワークのいずれかの中間層のデータが比較に用いられる。スタイルについては、例えば、畳み込みニューラルネットワークの複数の中間層のデータのチャンネル相関に

50

より得られるグラム行列 (Gram Matrix) が比較に用いられる。スタイルについては、チャンネル相関に代えて、各チャンネルの平均、分散等の、特徴マップ (Feature Maps) のチャンネルに関する分布統計量が用いられてもよい。評価対象の画像にはエンコーダ・デコーダ・ネットワーク 102 の出力が用いられる。コンテンツ参照用の画像は変換元画像入力部 101 により入力された変換元画像が用いられる。スタイル参照用の画像はスタイル参照用画像入力部 104 により入力されたスタイル参照用画像が用いられる。

【0019】

スタイル参照用画像入力部 104 は、学習用のスタイル参照用画像 (例えば、5千枚程度) を逐次に入力する機能を有している。学習用のスタイル参照用画像は、変換時に与えられる可能性のあるスタイル参照用画像をカバーできるような、偏りのない広範な画像が選ばれる。

10

【0020】

スタイル抽出ネットワーク 105 は、ロス・ネットワーク 103 と同様に、学習済の畳み込みニューラルネットワークによるものである。スタイル抽出ネットワーク 105 は、入力されたスタイル参照用画像からスタイルを示すデータを抽出し、スタイル・ベクトルとして出力する機能を有している。スタイルを示すデータは、複数の中間層のデータのチャンネル相関により得られるグラム行列のほか、各チャンネルの平均、分散等の、特徴マップ (Feature Maps) のチャンネルに関する分布統計量が用いられる。スタイル・ベクトルについて解説した文献として、「Shin Matsuo and Keiji Yanai: CNN-based Style Vector for Style Image Retrieval, ICMR (2016/06).」がある。分布統計量によるスタイルの表現について言及した文献として、「Yanghao Liy et al.: Demystifying Neural Style Transfer, arXiv 1701.01036v1 [cs.CV] 4 Jan 2017.」がある。

20

【0021】

圧縮部 106 は、スタイル抽出ネットワーク 105 から出力されるスタイル・ベクトルを、スタイルの特徴を損なわずに圧縮する機能を有している。具体的には、主成分分析 (Principal Component Analysis: PCA) や非負行列因子分解法 (Non-negative Matrix Factorization: NMF) が用いられる。学習済のニューラルネットワークが用いられてもよい。PCA と NMF との比較では、実験上、非負の重みが得られる NMF の方が良い結果が得られているが、NMF は反復計算のために若干計算時間が掛るため、PCA の方が高速処理 (変換時にはリアルタイム処理) には向いていると考えられる。

30

【0022】

複製部 107 は、圧縮部 106 により得られた圧縮後スタイル・ベクトルを結合時のサイズ (ピクセル数に対応) に合わせて複製する機能を有している。

【0023】

結合部 108 は、複製部 107 により得られた複製後スタイル・ベクトルをエンコーダ・デコーダ・ネットワーク 102 の中間層である結合層 Lc に結合 (concatenation) する機能を有している。結合は、例えば、結合層 Lc の直前のデコード層のデータ (Feature Maps) の奥行き方向にデータが追加されることで行われる。

【0024】

なお、スタイル・ベクトルがエンコーダ・デコーダ・ネットワーク 102 の中間層のデータに結合されることが重要であり、圧縮部 106 や複製部 107 は副次的なものである。例えば、スタイル抽出ネットワーク 105 において、結合部 108 において直接に結合できる形式のスタイル・ベクトルを出力する場合には、圧縮部 106 や複製部 107 を省略することもできる。

40

【0025】

図 2 は本発明の一実施形態にかかる画像スタイル変換装置 1 の学習時における他の機能構成例を示す図である。図 2 では、スタイル抽出ネットワーク 105 がニューラルネットワークにより構成され、圧縮後と同等のスタイル情報が出力されるようになっており、圧縮部 106 は省略されている。また、スタイル抽出ネットワーク 105 には、エンコーダ・デコーダ・ネットワーク 102 と同様に、ロス・ネットワーク 103 からコンテンツ・

50

ロスとスタイル・ロスがフィードバックされ、ランダムな初期値の状態から同時に一体的に学習（End-to-End学習）が行われる。

【0026】

図3は本発明の一実施形態にかかる画像スタイル変換装置1の変換時における機能構成例を示す図であり、図1に示した構成のうち変換時において用いられる構成である。図3において、図1の構成と比較して、ロス・ネットワーク103がなくなっている点と、エンコーダ・デコーダ・ネットワーク102により変換された画像を変換後画像として外部に出力する変換後画像出力部109が設けられている点異なる。また、変換元画像入力部101により入力される変換元画像は変換の対象となる1枚の画像となり、スタイル参照用画像入力部104により入力されるスタイル参照用画像は流用したいスタイルを示す1枚の画像となる。入力される変換元画像やスタイル参照用画像は学習時に用いられた画像である必要はなく、任意の画像でよい。

10

【0027】

図4は本発明の一実施形態にかかる画像スタイル変換装置1の変換時における機能構成例を示す図であり、図2に示した構成のうち変換時において用いられる構成である。図4において、図2の構成と比較して、ロス・ネットワーク103がなくなっている点と、エンコーダ・デコーダ・ネットワーク102により変換された画像を変換後画像として外部に出力する変換後画像出力部109が設けられている点異なる。また、変換元画像入力部101により入力される変換元画像は変換の対象となる1枚の画像となり、スタイル参照用画像入力部104により入力されるスタイル参照用画像は流用したいスタイルを示す1枚の画像となる。入力される変換元画像やスタイル参照用画像は学習時に用いられた画像である必要はなく、任意の画像でよい。

20

【0028】

図5は本発明の一実施形態にかかる画像スタイル変換装置1の変換時における他の機能構成例を示す図であり、スタイル参照用画像の入力に代え、ユーザ操作によりスタイル設定の入力が行われるようにしたものである。

【0029】

図5において、図3の構成と比較して、スタイル参照用画像入力部104とスタイル抽出ネットワーク105がなくなっている点と、スタイル設定入力部110とスタイル・ベクトル生成部111とが設けられている点異なる。スタイル設定入力部110は、例えば、スタイル・ベクトルの各要素に対応したスライダ等を備えたユーザインタフェースにより、スタイル設定を入力させる。スタイル・ベクトル生成部111は、スタイル設定入力部110により入力されたスタイル設定に基づいてスタイル・ベクトルを生成する。

30

【0030】

また、図4の構成に対応させて、スタイル・ベクトル生成部111によりスタイル抽出ネットワーク105の出力するスタイル情報と同等の情報量のスタイル・ベクトルが生成されるようにすることで、圧縮部106が省略可能である。

【0031】

なお、変換時の実装として、図3～図5における変換元画像入力部101、スタイル参照用画像入力部104、スタイル設定入力部110、変換後画像出力部109をスマートフォン等のクライアント端末に設けることができる。また、画像スタイル変換装置1の他の機能をクラウド上のサーバに配置し、クライアント端末とサーバの間で入出力データの送受信機能を持たせる。これにより、モバイル環境での画像スタイル変換を実現することができる。

40

【0032】

図6は画像スタイル変換装置1のハードウェア構成例を示す図である。図6において、画像スタイル変換装置1は、バス17を介して相互に接続されたCPU（Central Processing Unit）11、ROM（Read Only Memory）12、RAM（Random Access Memory）13、HDD（Hard Disk Drive）/SSD（Solid State Drive）14を備えている。また、画像スタイル変換装置1は、接続I/F（Interface）15、通信I/F16を備え

50

ている。

【0033】

CPU11は、RAM13をワークエリアとしてROM12またはHDD/SSD14等に格納されたプログラムを実行することで、画像スタイル変換装置1の動作を統括的に制御する。図1～図5で説明した画像スタイル変換装置1の機能は、CPU11において所定のプログラムが実行されることで実現される。プログラムは、記録媒体を経由して取得されるものでもよいし、ネットワークを経由して取得されるものでもよいし、ROM組込でもよい。

【0034】

<動作>

図7は学習時の処理例を示すフローチャートであり、図1の構成における動作である。図7において、管理者の指示等に応じて学習が開始されると、変換元画像入力部101とスタイル参照用画像入力部104は、変換元画像とスタイル参照用画像をそれぞれ入力する。そして、変換元画像とスタイル参照用画像の各組み合わせについて以下のループ処理が行われる(ステップS11～S15)。

【0035】

ループ処理において、スタイル抽出ネットワーク105は入力されたスタイル参照用画像からスタイル・ベクトルを抽出する。そして、圧縮部106はスタイル・ベクトルを圧縮し、複製部107は圧縮後スタイル・ベクトルを複製して複製後スタイル・ベクトルを生成する(ステップS12)。

【0036】

次いで、エンコーダ・デコーダ・ネットワーク102は入力された変換元画像を変換するが、その過程で、結合部108はエンコーダ・デコーダ・ネットワーク102の中間層に複製後スタイル・ベクトルを結合する(ステップS13)。

【0037】

次いで、ロス・ネットワーク103はエンコーダ・デコーダ・ネットワーク102の出力画像と変換元画像とスタイル参照用画像とからコンテンツ・ロスとスタイル・ロスを計算する。そして、コンテンツ・ロスとスタイル・ロスはエンコーダ・デコーダ・ネットワーク102にフィードバックされ、バックプロパゲーションが実施される(ステップS14)。

【0038】

その後、変換元画像とスタイル参照用画像の全ての組み合わせについて未処理の場合はループ処理が継続され、全ての組み合わせについて処理が行われた場合はループ処理が終了され(ステップS15)、学習処理が終了される。

【0039】

図2の構成については、ステップS12において明示的な圧縮が行われない点と、ステップS14において、スタイル抽出ネットワーク105についても同時にバックプロパゲーションによる学習が実施される点が異なる。

【0040】

以上の処理により、用意された多数の変換元画像とスタイル参照用画像について、変換元画像のコンテンツを残しつつ、スタイル参照用画像のスタイルが適用された画像へ変換する変換器としてエンコーダ・デコーダ・ネットワーク102が学習される。

【0041】

図8は変換時の処理例を示すフローチャートであり、図3の構成における動作である。図8において、ユーザの指示等に応じて変換処理が開始されると、変換元画像入力部101とスタイル参照用画像入力部104は変換元画像とスタイル参照用画像をそれぞれ入力する(ステップS21)。

【0042】

次いで、スタイル抽出ネットワーク105は入力されたスタイル参照用画像からスタイル・ベクトルを抽出する。そして、圧縮部106はスタイル・ベクトルを圧縮し、複製部

10

20

30

40

50

107は圧縮後スタイル・ベクトルを複製して複製後スタイル・ベクトルを生成する(ステップS22)。

【0043】

次いで、エンコーダ・デコーダ・ネットワーク102は入力された変換元画像を変換するが、その過程で、結合部108はエンコーダ・デコーダ・ネットワーク102の中間層に複製後スタイル・ベクトルを結合する(ステップS23)。

【0044】

次いで、変換後画像出力部109はエンコーダ・デコーダ・ネットワーク102から出力される画像についてフォーマット変換等を行い、変換後画像として外部に出力し(ステップS24)、変換処理が終了される。

【0045】

図4の構成については、ステップS22において明示的な圧縮が行われない点が異なる。

【0046】

以上の変換処理において、エンコーダ・デコーダ・ネットワーク102による画像変換は一方向に流れる処理で完了するため、高速なリアルタイム処理が可能となる。また、任意に与えられるスタイル参照用画像のスタイルを考慮した変換が可能となる。概念的には、学習時における多数のスタイル参照用画像のスタイルが融合された形で、任意に与えられるスタイル参照用画像のスタイルに近似されて適用される。

【0047】

図9は変換の例を示しており、(a)は変換元画像、(b)はスタイル参照用画像(出典：<https://www.wikiart.org/en/vincent-van-gogh/the-starry-night-1889>)、(c)は変換後画像を示している。すなわち、(a)の変換元画像のコンテンツ(山と道の風景)に(b)のスタイル参照用画像のスタイル(画風)が適用された変換後画像が(c)のように得られる。

【0048】

また、スタイルとしては、絵画における画風に限らず、質感等でもよい。例えば、変換元画像のコンテンツを「野球のボール」とし、スタイル参照用画像のスタイルを「岩石の質感を示す画像」とすると、変換元画像の野球のボールの表面が岩石の質感に置き換えられた変換後画像が得られる。

【0049】

図10は変換時の他の処理例を示すフローチャートであり、図5の構成における動作である。図10において、ユーザの指示等に応じて変換処理が開始されると、スタイル設定入力部110はユーザからスタイル設定を入力する(ステップS31)。

【0050】

次いで、変換元画像入力部101は変換元画像を入力する(ステップS32)。なお、ユーザの操作としては、変換元画像の指定がスタイル設定の入力より先でもよい。

【0051】

次いで、スタイル・ベクトル生成部111は、スタイル設定入力部110により入力されたスタイル設定からスタイル・ベクトルを生成する。そして、圧縮部106はスタイル・ベクトルを圧縮し、複製部107は圧縮後スタイル・ベクトルを複製して複製後スタイル・ベクトルを生成する(ステップS33)。

【0052】

次いで、エンコーダ・デコーダ・ネットワーク102は入力された変換元画像を変換するが、その過程で、結合部108はエンコーダ・デコーダ・ネットワーク102の中間層に複製後スタイル・ベクトルを結合する(ステップS34)。

【0053】

次いで、変換後画像出力部109はエンコーダ・デコーダ・ネットワーク102から出力される画像についてフォーマット変換等を行い、変換後画像として外部に出力し(ステップS35)、変換処理が終了される。

10

20

30

40

50

【 0 0 5 4 】

図 4 の構成に対応して、図 5 の構成から圧縮部 1 0 6 が省略される場合、ステップ S 3 3 において明示的な圧縮は行われぬ。

【 0 0 5 5 】

以上の変換処理では、特にスタイル参照用画像を用意しなくても、ユーザがスライダ等のユーザインタフェースを操作することで、任意のスタイルによる所望の変換後画像を行わせることができる。

【 0 0 5 6 】

< 総括 >

以上説明したように、本実施形態によれば、リアルタイム処理を可能としつつ、任意のスタイルに変換することができる。 10

【 0 0 5 7 】

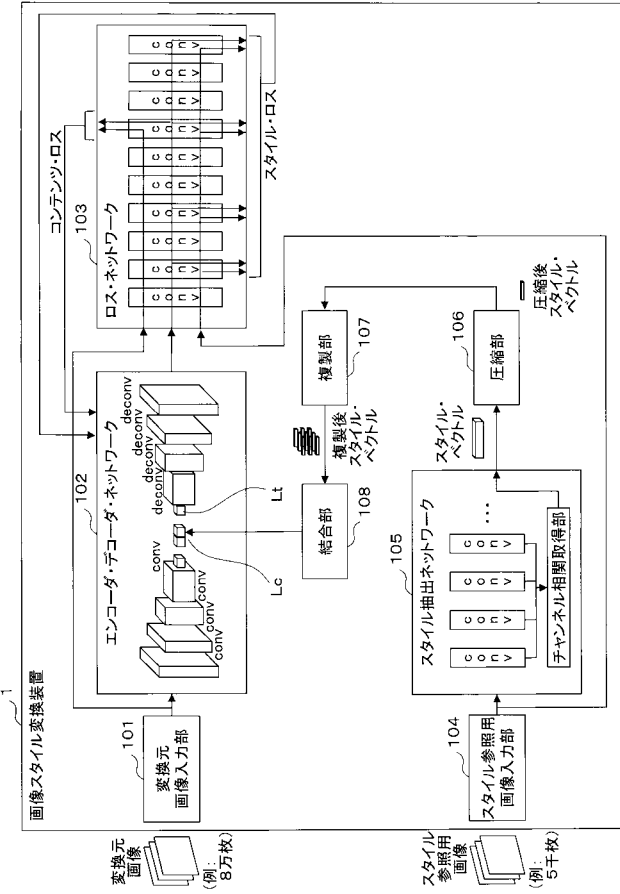
以上、本発明の好適な実施の形態により本発明を説明した。ここでは特定の具体例を示して本発明を説明したが、特許請求の範囲に定義された本発明の広範な趣旨および範囲から逸脱することなく、これら具体例に様々な修正および変更を加えることができることは明らかである。すなわち、具体例の詳細および添付の図面により本発明が限定されるものと解釈してはならない。

【 符号の説明 】

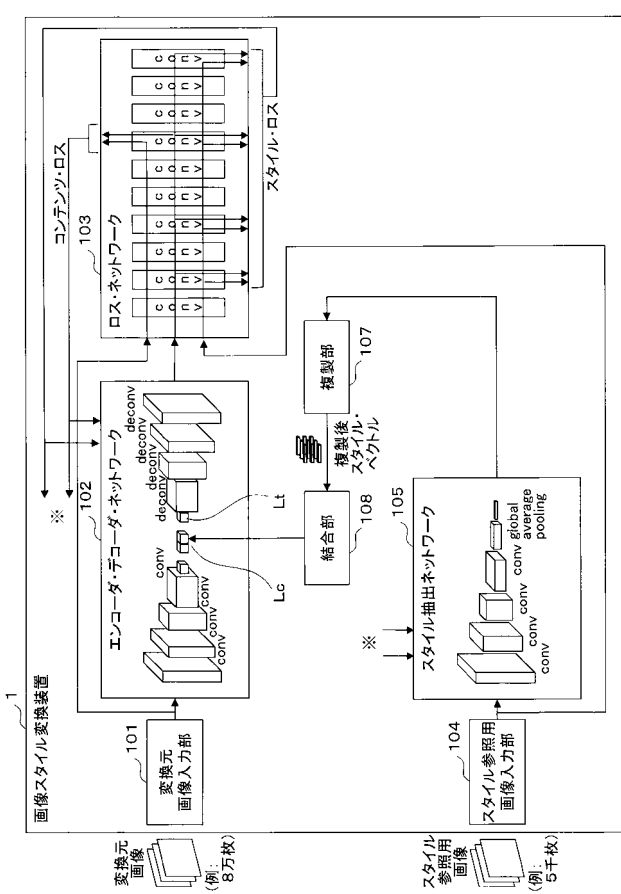
【 0 0 5 8 】

| | | |
|-------|-------------------|----|
| 1 | 画像スタイル変換装置 | 20 |
| 1 0 1 | 変換元画像入力部 | |
| 1 0 2 | エンコーダ・デコーダ・ネットワーク | |
| 1 0 3 | ロス・ネットワーク | |
| 1 0 4 | スタイル参照用画像入力部 | |
| 1 0 5 | スタイル抽出ネットワーク | |
| 1 0 6 | 圧縮部 | |
| 1 0 7 | 複製部 | |
| 1 0 8 | 結合部 | |
| 1 0 9 | 変換後画像出力部 | |
| 1 1 0 | スタイル設定入力部 | 30 |
| 1 1 1 | スタイル・ベクトル生成部 | |

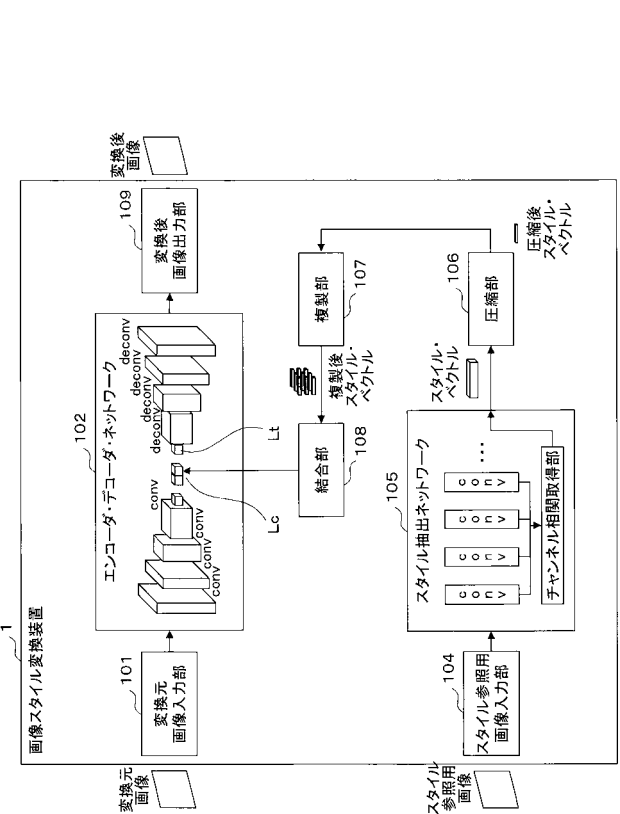
【図 1】



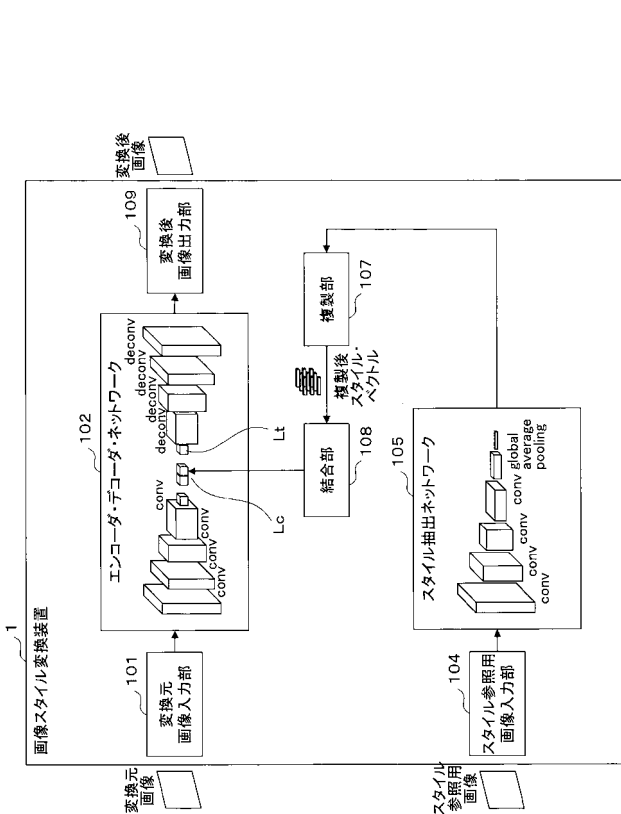
【図 2】



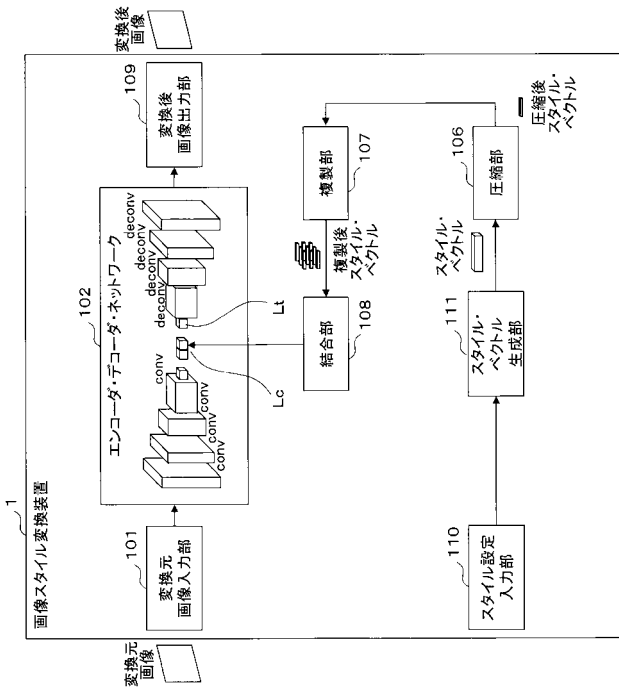
【図 3】



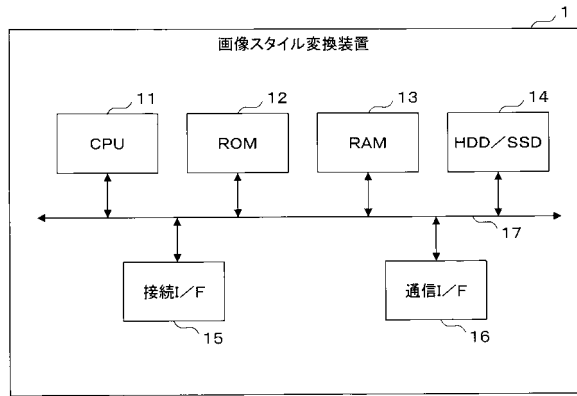
【図 4】



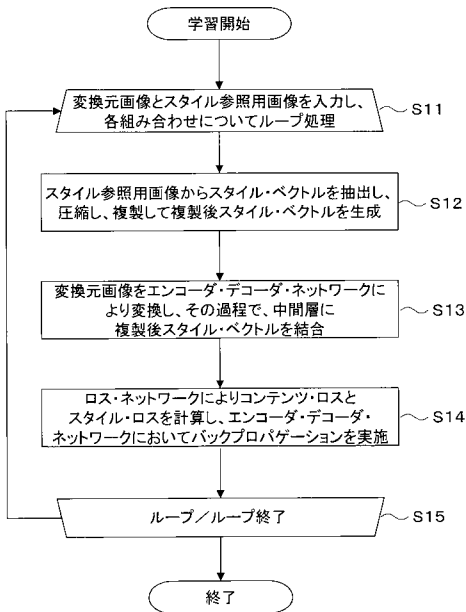
【 図 5 】



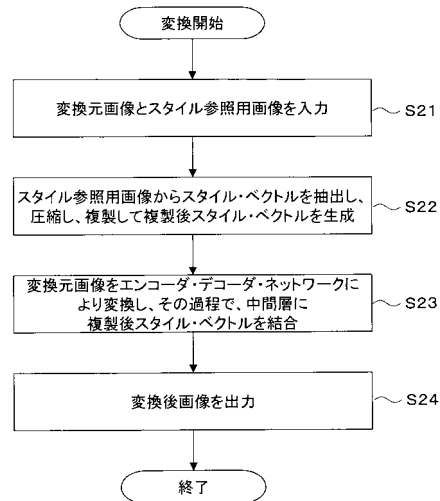
【 図 6 】



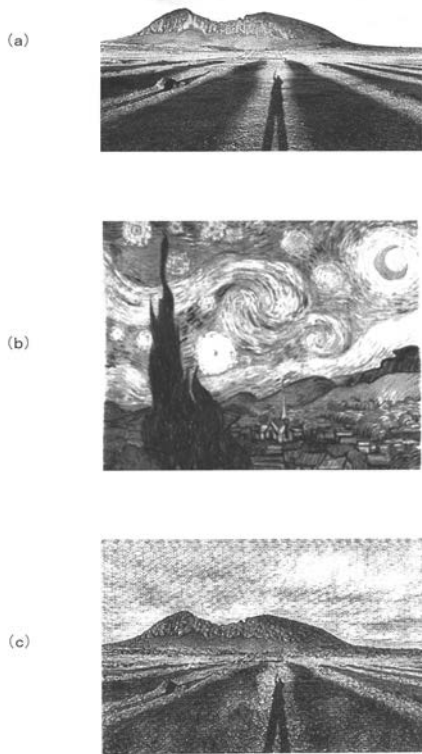
【 図 7 】



【 図 8 】



【 図 9 】



【 図 1 0 】

