

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7417640号  
(P7417640)

(45)発行日 令和6年1月18日(2024.1.18)

(24)登録日 令和6年1月10日(2024.1.10)

(51)国際特許分類 F I  
G 0 6 T 3/4061(2024.01) G 0 6 T 3/40 7 3 5

請求項の数 22 (全32頁)

(21)出願番号	特願2021-575454(P2021-575454)	(73)特許権者	504161984 ホアウェイ・テクノロジーズ・カンパニー・リミテッド 中華人民共和国・518129・グァン ドン・シェンツェン・ロンガン・ディス トリクト・バンティアン・(番地なし) ・ホアウェイ・アドミニストレーション ・ビルディング
(86)(22)出願日	令和1年6月18日(2019.6.18)	(74)代理人	100110364 弁理士 実広 信哉
(65)公表番号	特表2022-536807(P2022-536807 A)	(74)代理人	100133569 弁理士 野村 進
(43)公表日	令和4年8月18日(2022.8.18)	(72)発明者	ウェイ・スー アメリカ合衆国・ワシントン・9805 3・レドモンド・トゥーハンドレッドア 最終頁に続く
(86)国際出願番号	PCT/US2019/037727		
(87)国際公開番号	WO2020/256704		
(87)国際公開日	令和2年12月24日(2020.12.24)		
審査請求日	令和4年1月28日(2022.1.28)		

(54)【発明の名称】 リアルタイム映像超高解像度

(57)【特許請求の範囲】

【請求項1】

デジタル画像の画像解像度を高めるためのコンピュータ実装方法であって、前記方法は、ベース高解像度(HR)画像を生成するために、前記デジタル画像のバイキュービックアップサンプリングを実行するステップと、

前記デジタル画像の輝度を示すグレースケール画像である低解像度(LR)残差画像を生成するために、赤-緑-青(RGB)色空間から輝度(Y)、彩度青差(Cb)、および彩度赤差(Cr)(YCbCr)色空間に前記デジタル画像を変換するステップと、

ニューラルネットワークモデルの複数の畳み込み層を使用して、前記LR残差画像を前記デジタル画像に対応する複数のHR残差サブ画像に変換するステップと、

前記ベースHR画像および前記複数のHR残差サブ画像を使用して、前記デジタル画像に対応するHR画像を生成するステップと、

を含む、コンピュータ実装方法。

【請求項2】

HR残差画像を生成するために、前記複数のHR残差サブ画像を画素シフトするステップをさらに含み、

前記デジタル画像に対応する前記HR画像を生成する前記ステップは、前記HR残差画像と前記ベースHR画像とを合成するステップを含む、

請求項1に記載のコンピュータ実装方法。

【請求項3】

前記ニューラルネットワークモデルは入力層を含み、前記複数の畳み込み層は4つの畳み込み層を含む、請求項1または2に記載のコンピュータ実装方法。

【請求項4】

前記入力層は、前記デジタル画像を受け取るように構成され、前記4つの畳み込み層の出力層は、前記複数のHR残差サブ画像を出力するように構成される、請求項3に記載のコンピュータ実装方法。

【請求項5】

前記複数の畳み込み層のうちの第1の層は、 $3 \times 3$ 画素カーネルおよび8チャンネルで構成され、

前記複数の畳み込み層のうちの第2の層は、 $3 \times 3$ 画素カーネルおよび6チャンネルで構成され、

前記複数の畳み込み層のうちの第3の層は、 $3 \times 3$ 画素カーネルおよび4チャンネルで構成され、

前記複数の畳み込み層のうちの第4の層は、4チャンネルで構成される、

請求項3に記載のコンピュータ実装方法。

【請求項6】

複数の訓練画像対で前記ニューラルネットワークモデルを訓練するステップであって、前記複数の訓練画像対の各訓練画像対は、

訓練画像に対応するLR画像であって、前記LR画像は、劣化した画質を有し、前記ニューラルネットワークモデルへの入力として構成される、LR画像と、

前記訓練画像に対応し、前記ニューラルネットワークモデルの目標出力として構成される複数のHR残差サブ画像と、を含む、ステップ

をさらに含む、請求項1から5のいずれか一項に記載のコンピュータ実装方法。

【請求項7】

前記ニューラルネットワークモデルを訓練するステップは、

フィルタリングされた画像を生成するために、ローパスフィルタを使用して前記訓練画像をフィルタリングするステップと、

ダウンサンプリングされたLR画像を生成するために、前記フィルタリングされた画像をダウンサンプリングするステップと、

前記訓練画像に対応する前記LR画像を生成するために、ノイズおよびアーチファクトを追加することによって、前記ダウンサンプリングされたLR画像の画質を劣化させるステップと、

をさらに含む、請求項6に記載のコンピュータ実装方法。

【請求項8】

前記ニューラルネットワークモデルを訓練するステップは、

コントラスト強調画像を生成するために、前記訓練画像にアンバランスなアンシャープマスクを適用するステップと、

前記訓練画像に対応するHR残差画像を生成するために、前記コントラスト強調画像から前記ダウンサンプリングされたLR画像のアップサンプリングされたバージョンを減算するステップと、

をさらに含む、請求項7に記載のコンピュータ実装方法。

【請求項9】

前記ニューラルネットワークモデルを訓練するステップは、

前記訓練画像に対応する前記複数のHR残差サブ画像を生成するために、前記訓練画像に対応する前記HR残差画像を分割するステップ

をさらに含む、請求項8に記載のコンピュータ実装方法。

【請求項10】

命令を記憶するメモリと、

前記メモリと通信する1つまたは複数のプロセッサであって、前記1つまたは複数のプロセッサは、

10

20

30

40

50

ベース高解像度（HR）画像を生成するために、デジタル画像のバイキュービックアップサンプリングを実行し、

前記デジタル画像の輝度を示すグレースケール画像である低解像度（LR）残差画像を生成するために、赤 - 緑 - 青（RGB）色空間から輝度（Y）、彩度青差（Cb）、および彩度赤差（Cr）（YCbCr）色空間に前記デジタル画像を変換し、

ニューラルネットワークモデルの複数の畳み込み層を使用して、前記LR残差画像を前記デジタル画像に対応する複数のHR残差サブ画像に変換し、

前記ベースHR画像および前記複数のHR残差サブ画像を使用して、前記デジタル画像に対応するHR画像を生成する、前記命令を実行する、1つまたは複数のプロセッサと、  
を備える、システム。

10

【請求項 1 1】

前記1つまたは複数のプロセッサは、

HR残差画像を生成するために、前記複数のHR残差サブ画像を画素シフトし、

前記デジタル画像に対応する前記HR画像を生成するために、前記HR残差画像と前記ベースHR画像とを合成する、

前記命令を実行する、請求項10に記載のシステム。

【請求項 1 2】

前記ニューラルネットワークモデルは入力層を含み、前記複数の畳み込み層は4つの畳み込み層を含み、

前記入力層は、前記デジタル画像を受け取るように構成され、

前記4つの畳み込み層の出力層は、前記複数のHR残差サブ画像を出力するように構成される、

請求項10または11に記載のシステム。

20

【請求項 1 3】

前記複数の畳み込み層のうちの第1の層は、3×3画素カーネルおよび8チャンネルで構成され、

前記複数の畳み込み層のうちの第2の層は、3×3画素カーネルおよび6チャンネルで構成され、

前記複数の畳み込み層のうちの第3の層は、3×3画素カーネルおよび4チャンネルで構成され、

前記複数の畳み込み層のうちの第4の層は、4チャンネルで構成される、

請求項12に記載のシステム。

30

【請求項 1 4】

前記1つまたは複数のプロセッサは、

複数の訓練画像対で前記ニューラルネットワークモデルを訓練する前記命令を実行し、前記複数の訓練画像対の各訓練画像対は、

訓練画像に対応するLR画像であって、前記LR画像は、劣化した画質を有し、前記ニューラルネットワークモデルへの入力として構成される、LR画像と、

前記訓練画像に対応し、前記ニューラルネットワークモデルの目標出力として構成される複数のHR残差サブ画像と、

を含む、請求項10から13のいずれか一項に記載のシステム。

40

【請求項 1 5】

前記ニューラルネットワークモデルを訓練するために、前記1つまたは複数のプロセッサは、

フィルタリングされた画像を生成するために、ローパスフィルタを使用して前記訓練画像をフィルタリングし、

ダウンサンプリングされたLR画像を生成するために、前記フィルタリングされた画像をダウンサンプリングし、

前記訓練画像に対応する前記LR画像を生成するために、ノイズおよびアーチファクトを追加することによって、前記ダウンサンプリングされたLR画像の画質を劣化させる、

50

前記命令を実行する、請求項14に記載のシステム。

【請求項16】

前記ニューラルネットワークモデルを訓練するために、前記1つまたは複数のプロセッサは、

コントラスト強調画像を生成するために、前記訓練画像にアンバランスなアンシャープマスクを適用し、

前記訓練画像に対応するHR残差画像を生成するために、前記コントラスト強調画像から前記ダウンサンプリングされたLR画像のアップサンプリングされたバージョンを減算する、

前記命令を実行する、請求項15に記載のシステム。

10

【請求項17】

前記ニューラルネットワークモデルを訓練するために、前記1つまたは複数のプロセッサは、

前記訓練画像に対応する前記複数のHR残差サブ画像を生成するために、前記訓練画像に対応する前記HR残差画像を分割する

前記命令を実行する、請求項16に記載のシステム。

【請求項18】

デジタル画像の画像解像度を高めるためのコンピュータ命令を記憶するコンピュータ可読媒体であって、前記命令は、1つまたは複数のプロセッサによって実行されると、前記1つまたは複数のプロセッサに、

20

ベース高解像度（HR）画像を生成するために、前記デジタル画像のバイキュービックアップサンプリングを実行するステップと、

前記デジタル画像の輝度を示すグレースケール画像である低解像度（LR）残差画像を生成するために、赤 - 緑 - 青（RGB）色空間から輝度（Y）、彩度青差（Cb）、および彩度赤差（Cr）（YCbCr）色空間に前記デジタル画像を変換するステップと、

ニューラルネットワークモデルの複数の畳み込み層を使用して、前記LR残差画像を前記デジタル画像に対応する複数のHR残差サブ画像に変換するステップと、

前記ベースHR画像および前記複数のHR残差サブ画像を使用して、前記デジタル画像に対応するHR画像を生成するステップと、

を実行させる、コンピュータ可読媒体。

30

【請求項19】

前記命令はさらに、前記1つまたは複数のプロセッサに、

複数の訓練画像対で前記ニューラルネットワークモデルを訓練するステップであって、前記複数の訓練画像対の各訓練画像対は、

訓練画像に対応するLR画像であって、前記LR画像は、劣化した画質を有し、前記ニューラルネットワークモデルへの入力として構成される、LR画像と、

前記訓練画像に対応し、前記ニューラルネットワークモデルの目標出力として構成される複数のHR残差サブ画像と、

を含む、ステップを実行させる、請求項18に記載のコンピュータ可読媒体。

【請求項20】

40

前記命令はさらに、前記1つまたは複数のプロセッサに、

フィルタリングされた画像を生成するために、ローパスフィルタを使用して前記訓練画像をフィルタリングするステップと、

ダウンサンプリングされたLR画像を生成するために、前記フィルタリングされた画像をダウンサンプリングするステップと、

前記訓練画像に対応する前記LR画像を生成するために、ノイズおよびアーチファクトを追加することによって、前記ダウンサンプリングされたLR画像の画質を劣化させるステップと、

を実行させる、請求項19に記載のコンピュータ可読媒体。

【請求項21】

50

前記命令はさらに、前記1つまたは複数のプロセッサに、  
 コントラスト強調画像を生成するために、前記訓練画像にアンバランスなアンシャープマスクを適用するステップと、  
 前記訓練画像に対応するHR残差画像を生成するために、前記コントラスト強調画像から前記ダウンサンプリングされたLR画像のアップサンプリングされたバージョンを減算するステップと、  
 前記訓練画像に対応する前記複数のHR残差サブ画像を生成するために、前記訓練画像に対応する前記HR残差画像を分割するステップと、  
 を実行させる、請求項20に記載のコンピュータ可読媒体。

【請求項22】

デジタル画像の画像解像度を高めるための画像解像度調整システムであって、前記システムは、

ベース高解像度(HR)画像を生成するために、前記デジタル画像のバイキュービックアップサンプリングを実行するアップサンプリング手段と、  
 前記デジタル画像の輝度を示すグレースケール画像である低解像度(LR)残差画像を生成するために、赤-緑-青(RGB)色空間から輝度(Y)、彩度青差(Cb)、および彩度赤差(Cr)(YCbCr)色空間に前記デジタル画像を変換する色空間処理手段と、

前記LR残差画像を前記デジタル画像に対応する複数のHR残差サブ画像に変換する畳み込み手段と、

前記ベースHR画像および前記複数のHR残差サブ画像を使用して、前記デジタル画像に対応するHR画像を生成する追加手段と、  
 を備える、画像解像度調整システム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、コンピューティングデバイス上の映像データ処理に関する。

【背景技術】

【0002】

エレクトロニクス技術の急速な発展に伴い、2kおよび4k高解像度(HD)ディスプレイは、ホームシアターシステム、タブレットコンピュータ、およびモバイルデバイスの標準となっている。モバイルデバイスで再生される映像は、映像キャプチャデバイス、ストレージスペース、ネットワーク帯域幅、および/またはデータフローに関連する制限のために、通常、より低い解像度を有する。したがって、新世代のHDディスプレイを最大限に活用するために、個々のフレームが最新のデバイスに表示される前に映像の効果的な解像度および視覚効果を高めることが重要である。

【0003】

既存の画像解像度向上技術は、主にノイズおよびアーチファクトのない高品質の画像で動作する。入力映像(例えば、モバイルデバイスで視聴される映像)にノイズおよびブロッキングアーチファクトが存在する場合、既存の技術を使用すると、画質がさらに低下することが多い。さらに、計算の複雑さのために、現在の画像解像度向上技術は、従来の補間技術と比較してピーク信号対ノイズ比(PSNR)および構造類似性(SSIM)を改善せず、したがって、HDディスプレイを有するモバイルデバイス上での画像処理における用途が限られている。

【発明の概要】

【0004】

次に、様々な例を示して、以下の詳細な説明でさらに説明される概念のうちの選択したものを簡略化した形で紹介する。この概要は、特許請求される主題の重要な特徴または本質的な特徴を特定するためのものでも、特許請求される主題の範囲の限定に使用するためのものでもない。

【課題を解決するための手段】

10

20

30

40

50

## 【 0 0 0 5 】

本開示の第1の態様によれば、デジタル画像の画像解像度を高めるためのコンピュータ実装方法が提供される。方法は、ベース高解像度（HR）画像を生成するために、デジタル画像のバイキュービックアップサンプリングを実行するステップを含む。デジタル画像は、低解像度（LR）残差画像を生成するために、赤 - 緑 - 青（RGB）色空間から輝度（Y）、彩度青差（Cb）、および彩度赤差（Cr）（YCbCr）色空間に変換される。LR残差画像は、ニューラルネットワークモデルの複数の畳み込み層を使用して、デジタル画像に対応する複数のHR残差サブ画像に変換される。ベースHR画像および複数のHR残差サブ画像を使用して、デジタル画像に対応するHR画像が生成される。

## 【 0 0 0 6 】

このような第1の態様による方法の第1の実装形態において、複数のHR残差サブ画像は、HR残差画像を生成するために画素シフトされる。デジタル画像に対応するHR画像を生成することは、HR残差画像とベースHR画像とを合成することを含む。

## 【 0 0 0 7 】

そのような第1の態様または第1の態様の任意の前述の実装形態による方法の第2の実装形態において、ニューラルネットワークモデルは入力層を含み、複数の畳み込み層は4つの畳み込み層を含む。

## 【 0 0 0 8 】

そのような第1の態様または第1の態様の任意の前述の実装形態による方法の第3の実装形態において、入力層はデジタル画像を受け取るように構成され、4つの畳み込み層の出力層は複数のHR残差サブ画像を出力するように構成される。

## 【 0 0 0 9 】

そのような第1の態様または第1の態様の任意の前述の実装形態による方法の第4の実装形態において、複数の畳み込み層のうちの第1の層は、 $3 \times 3$ 画素カーネルおよび8チャンネルで構成され、複数の畳み込み層のうちの第2の層は、 $3 \times 3$ 画素カーネルおよび6チャンネルで構成され、複数の畳み込み層のうちの第3の層は、 $3 \times 3$ 画素カーネルおよび4チャンネルで構成され、複数の畳み込み層のうちの第4の層は、4チャンネルで構成される。

## 【 0 0 1 0 】

そのような第1の態様または第1の態様の任意の前述の実装形態による方法の第5の実装形態において、ニューラルネットワークモデルは、複数の訓練画像対で訓練される。複数の訓練画像対の各訓練画像対は、訓練画像に対応するLR画像と、訓練画像に対応し、ニューラルネットワークモデルの目標出力として構成される複数のHR残差サブ画像と、を含む。LR画像は、劣化した画質を有し、ニューラルネットワークモデルへの入力として構成される。

## 【 0 0 1 1 】

そのような第1の態様または第1の態様の任意の前述の実装形態による方法の第6の実装形態において、ニューラルネットワークモデルを訓練するステップは、フィルタリングされた画像を生成するために、ローパスフィルタを使用して訓練画像をフィルタリングするステップと、ダウンサンプリングされたLR画像を生成するために、フィルタリングされた画像をダウンサンプリングするステップと、訓練画像に対応するLR画像を生成するために、ノイズおよびアーチファクトを追加することによって、ダウンサンプリングされたLR画像の画質を劣化させるステップと、を含む。

## 【 0 0 1 2 】

そのような第1の態様または第1の態様の任意の前述の実装形態による方法の第7の実装形態において、ニューラルネットワークモデルを訓練するステップは、コントラスト強調画像を生成するために、訓練画像にアンバランスなアンシャープマスクを適用するステップと、訓練画像に対応するHR残差画像を生成するために、コントラスト強調画像からダウンサンプリングされたLR画像のアップサンプリングされたバージョンを減算するステップと、を含む。

## 【 0 0 1 3 】

10

20

30

40

50

そのような第1の態様または第1の態様の任意の前述の実装形態による方法の第8の実装形態において、ニューラルネットワークモデルを訓練するステップは、訓練画像に対応する複数のHR残差サブ画像を生成するために、訓練画像に対応するHR残差画像を分割するステップをさらに含む。

【0014】

本開示の第2の態様によれば、デジタル画像の画像解像度を高めるためのシステムが提供される。システムは、命令を記憶するメモリと、メモリと通信する1つまたは複数のプロセッサと、を備える。1つまたは複数のプロセッサは、ベース高解像度(HR)画像を生成するために、デジタル画像のバイキュービックアップサンプリングを実行する命令を実行する。デジタル画像は、低解像度(LR)残差画像を生成するために、赤-緑-青(RGB)色空間から輝度(Y)、彩度青差(Cb)、および彩度赤差(Cr)(YCbCr)色空間に変換される。LR残差画像をデジタル画像に対応する複数のHR残差サブ画像に変換するために、ニューラルネットワークモデルの複数の畳み込み層はLR残差画像に適用される。ベースHR画像および複数のHR残差サブ画像を使用して、デジタル画像に対応するHR画像が生成される。

10

【0015】

そのような第2の態様によるシステムの第1の実装形態において、1つまたは複数のプロセッサは、HR残差画像を生成するために、複数のHR残差サブ画像を画素シフトし、デジタル画像に対応するHR画像を生成するために、HR残差画像とベースHR画像とを合成する、ようにさらに構成される。

20

【0016】

そのような第2の態様または第2の態様の任意の前述の実装形態によるシステムの第2の実装形態において、ニューラルネットワークモデルは入力層を含み、複数の畳み込み層は4つの畳み込み層を含む。入力層は、デジタル画像を受け取るように構成される。4つの畳み込み層の出力層は、複数のHR残差サブ画像を出力するように構成される。

【0017】

そのような第2の態様または第2の態様の任意の前述の実装形態によるシステムの第3の実装形態において、複数の畳み込み層のうちの第1の層は、 $3 \times 3$ 画素カーネルおよび8チャンネルで構成され、複数の畳み込み層のうちの第2の層は、 $3 \times 3$ 画素カーネルおよび6チャンネルで構成され、複数の畳み込み層のうちの第3の層は、 $3 \times 3$ 画素カーネルおよび4チャンネルで構成され、複数の畳み込み層のうちの第4の層は、4チャンネルで構成される。

30

【0018】

そのような第2の態様または第2の態様の任意の前述の実装形態によるシステムの第4の実装形態において、1つまたは複数のプロセッサは、複数の訓練画像対でニューラルネットワークモデルを訓練するようにさらに構成される。複数の訓練画像対の各訓練画像対は、訓練画像に対応するLR画像と、訓練画像に対応し、ニューラルネットワークモデルの目標出力として構成される複数のHR残差サブ画像と、を含む。LR画像は、劣化した画質を有し、ニューラルネットワークモデルへの入力として構成される。

【0019】

そのような第2の態様または第2の態様の任意の前述の実装形態によるシステムの第5の実装形態において、1つまたは複数のプロセッサは、フィルタリングされた画像を生成するために、ローパスフィルタを使用して訓練画像をフィルタリングし、ダウンサンプリングされたLR画像を生成するために、フィルタリングされた画像をダウンサンプリングし、訓練画像に対応するLR画像を生成するために、ノイズおよびアーチファクトを追加することによって、ダウンサンプリングされたLR画像の画質を劣化させる、ようにさらに構成される。

40

【0020】

そのような第2の態様または第2の態様の任意の前述の実装形態によるシステムの第6の実装形態において、ニューラルネットワークモデルを訓練するために、1つまたは複数のプロセッサは、コントラスト強調画像を生成するために、訓練画像にアンバランスなアン

50

シャープマスクを適用し、訓練画像に対応するHR残差画像を生成するために、コントラスト強調画像からダウンサンプリングされたLR画像のアップサンプリングされたバージョンを減算する、ようにさらに構成される。

【 0 0 2 1 】

そのような第2の態様または第2の態様の任意の前述の実装形態によるシステムの第7の実装形態において、ニューラルネットワークモデルを訓練するために、1つまたは複数のプロセッサは、訓練画像に対応する複数のHR残差サブ画像を生成するために、訓練画像に対応するHR残差画像を分割するようにさらに構成される。

【 0 0 2 2 】

本開示の第3の態様によれば、デジタル画像の画像解像度を高めるための命令を記憶する非一時的コンピュータ可読媒体が提供される。1つまたは複数のプロセッサによって実行されると、命令は、1つまたは複数のプロセッサに、ベース高解像度（HR）画像を生成するために、デジタル画像のバイキュービックアップサンプリングを実行するステップを含む動作を実行させる。デジタル画像は、低解像度（LR）残差画像を生成するために、赤 - 緑 - 青（RGB）色空間から輝度（Y）、彩度青差（Cb）、および彩度赤差（Cr）（YCbCr）色空間に変換される。LR残差画像をデジタル画像に対応する複数のHR残差サブ画像に変換するために、ニューラルネットワークモデルの複数の畳み込み層はLR残差画像に適用される。ベースHR画像および複数のHR残差サブ画像を使用して、デジタル画像に対応するHR画像が生成される。

【 0 0 2 3 】

そのような第3の態様による非一時的コンピュータ可読媒体の第1の実装形態において、動作は、複数の訓練画像対でニューラルネットワークモデルを訓練するステップをさらに含む。複数の訓練画像対の各訓練画像対は、訓練画像に対応するLR画像と、訓練画像に対応し、ニューラルネットワークモデルの目標出力として構成される複数のHR残差サブ画像と、を含む。LR画像は、劣化した画質を有し、ニューラルネットワークモデルへの入力として構成される。

【 0 0 2 4 】

そのような第3の態様または第3の態様の任意の前述の実装形態による非一時的コンピュータ可読媒体の第2の実装形態において、動作は、フィルタリングされた画像を生成するために、ローパスフィルタを使用して訓練画像をフィルタリングするステップと、ダウンサンプリングされたLR画像を生成するために、フィルタリングされた画像をダウンサンプリングするステップと、訓練画像に対応するLR画像を生成するために、ノイズおよびアーチファクトを追加することによって、ダウンサンプリングされたLR画像の画質を劣化させるステップと、をさらに含む。

【 0 0 2 5 】

そのような第3の態様または第3の態様の任意の前述の実装形態による非一時的コンピュータ可読媒体の第3の実装形態において、動作は、コントラスト強調画像を生成するために、訓練画像にアンバランスなアンシャープマスクを適用するステップと、訓練画像に対応するHR残差画像を生成するために、コントラスト強調画像からダウンサンプリングされたLR画像のアップサンプリングされたバージョンを減算するステップと、訓練画像に対応する複数のHR残差サブ画像を生成するために、訓練画像に対応するHR残差画像を分割するステップと、をさらに含む。

【 0 0 2 6 】

本開示の第4の態様によれば、デジタル画像の画像解像度を高めるための画像解像度調整システムが提供される。システムは、ベース高解像度（HR）画像を生成するために、デジタル画像のバイキュービックアップサンプリングを実行するアップサンプリング手段を備える。システムはまた、低解像度（LR）残差画像を生成するために、赤 - 緑 - 青（RGB）色空間から輝度（Y）、彩度青差（Cb）、および彩度赤差（Cr）（YCbCr）色空間にデジタル画像を変換する色空間処理手段を備える。システムは、LR残差画像をデジタル画像に対応する複数のHR残差サブ画像に変換する畳み込み手段をさらに備える。システムはま

10

20

30

40

50

た、ベースHR画像および複数のHR残差サブ画像を使用して、デジタル画像に対応するHR画像を生成する追加手段を備える。

【0027】

前述の例のいずれかを、他の前述の例のいずれか1つ以上と組み合わせて、本開示の範囲内で新しい実施形態を作成することができる。

【0028】

必ずしも縮尺通りに描かれていない図面では、同様の符号は異なる図で同様の構成要素を説明することができる。図面は、一般に、限定ではなく例として、本明細書で説明される様々な実施形態を示す。

【図面の簡単な説明】

【0029】

【図1】いくつかの例示的な実施形態による、深層学習(DL)アーキテクチャ(DLA)を使用したDLモデルの訓練を示すブロック図である。

【図2】いくつかの例示的な実施形態による、DLA内で訓練されたニューラルネットワークモデルを使用した訓練されたDLモデルの生成を示す図である。

【図3】いくつかの例示的な実施形態による、図1のDLAを使用して画像解像度を調整するためのシステムを示す図である。

【図4】いくつかの例示的な実施形態による、図3のシステムによって使用されるDLA内の畳み込み層の構成を示す図である。

【図5】いくつかの例示的な実施形態による、図1のDLAの訓練に使用することができる訓練画像対の生成を示す図である。

【図6】いくつかの例示的な実施形態による、訓練画像対の生成、ならびにDLA訓練および画像解像度調整機能の実行のための対の使用のフローチャートである。

【図7】いくつかの例示的な実施形態による、デジタル画像の画像解像度を高めるための方法のフローチャートである。

【図8】いくつかの例示的な実施形態による、本明細書に記載の様々なデバイスハードウェアとともに使用することができる代表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9】例示的な実施形態による、アルゴリズムを実装し、方法を実行するデバイスのための回路を示すブロック図である。

【発明を実施するための形態】

【0030】

最初に、以下では1つまたは複数の実施形態の例示的な実装形態が提供されるが、図1~図9に関して記載される開示のシステムおよび方法は、現在既知またはまだ既存でないかどうかを問わず、任意の数の技術を使用して実施されうることを理解されたい。本開示は、本明細書において図示または記載される例示的な設計および実装形態を含む、以下に示される例示的な実装形態、図面、および技法に決して限定されるべきでなく、それらの均等物の全範囲とともに添付の特許請求の範囲の範囲内で修正されてよい。

【0031】

以下の説明では、本明細書の一部を形成し、実践され得る特定の実施形態の説明によって示される、添付図面が参照される。これらの実施形態は、当業者が本発明の主題を実施することを可能にするために十分詳細に説明されており、また、他の実施形態が利用され得ること、および本開示の範囲から逸脱することなく構造的、論理的、電気的変更がなされ得ることを理解されたい。したがって、以下の例示的な実施形態の説明は限定的な意味で解釈されるべきではなく、本開示の範囲は添付の特許請求の範囲によって定義される。

【0032】

本開示は、コンピューティングデバイス上の映像データ処理に関する。いくつかの態様は、ニューラルネットワークを使用して画像の画像解像度を変更することに関する。他の態様は、リアルタイム映像超高解像度に関する。

【0033】

10

20

30

40

50

本明細書で使用される場合、画像に関連する「低解像度」（またはLR）および「高解像度」（またはHR）という用語は、画像のサイズ（画素単位）に関連付けられる。例えば、2つの画像が同じシーンを描写し、第1の画像が第2の画像の高さおよび幅と比較して（画素単位で）より大きい高さおよび幅を有する場合、第1の画像は高解像度画像と呼ばれ、第2の画像は低解像度画像と呼ばれる。

【0034】

本明細書で使用される場合、「超解像」（またはSR）という用語は、従来の補間方法と比較して、画素数（例えば、アップスケーリングを介して）を増加させ、ピーク信号対ノイズ比（PSNR）および構造類似性（SSIM）を改善する解像度向上技術を指す。

【0035】

本明細書で使用される場合、「超高解像」（またはUR）という用語は、SR（すなわち、PSNRおよびSSIMとともに映像フレームの解像度を向上させる）の画像解像度向上技術を含むが、ノイズレベルを低減し、映像圧縮によって一般的に引き起こされるブロッキングアーチファクトを除去し、ローカルコントラストを向上させる（SRには機能が存在しない）。

【0036】

本明細書で使用される場合、「順方向計算」および「逆方向計算」という用語は、ニューラルネットワークモデル（または別のタイプのモデル）の訓練に関連してワーカーマシンで実行される計算を指す。順方向および逆方向の計算中に実行される計算は、前の反復からの結果に基づいて（例えば、前の逆方向計算の終了時に生成された勾配に基づいて）重みを修正する。勾配は、ワーカーマシンが計算しているモデルの重みに対する変更ごとにワーカーマシンの出力がどれだけ変化するかを測定値である。勾配は、誤差の変化に関するすべての重みの変化を測定する。勾配値が大きいほど、モデルは速く学習することができる。

【0037】

本明細書で使用される場合、「バイキュービックアップサンプリング」という用語は、バイキュービック補間を使用する画像アップサンプリングを指す。バイキュービック補間は、2次元グリッド上のデータ点を補間するための技術である。画像処理に関連して、バイキュービック補間は16ピクセル（4×4ピクセルマトリックス）を考慮し、補間はラグランジュ多項式、三次スプライン、または三次畳み込みアルゴリズムを介して実行される。

【0038】

本明細書に開示された技術は、モバイルデバイス上の画像解像度をリアルタイムで改善し、超解像度を達成するために使用することができる。URに関連する機能は、限られた数の畳み込み層を介して残差画像情報を抽出し伝播するニューラルネットワークモデルを使用する簡潔な人工ニューラルネットワーク（ANN）によって達成される。より具体的には、残差画像は、LR入力画像に対応するHR画像とLR入力画像のバイキュービックアップサンプリングバージョンとの差であるANNを使用して（例えば、LR入力画像に対応するLR残差画像を使用して）取得される。ニューラルネットワークモデルを使用してLR残差画像を処理すると、畳み込み層を通るデータフローが減少し、モデルが非常に少数のパラメータで動作することが可能になる。「ニューラルネットワーク」（またはNN）および「人工ニューラルネットワーク」（またはANN）という用語は同義語であり、本明細書では互換的に使用される。「畳み込みニューラルネットワーク」（またはCNN）および「深層ニューラルネットワーク」（またはDNN）という用語は同義語であり、複数の畳み込み層を含むニューラルネットワークの種類を指す。

【0039】

本明細書に開示される技術は、訓練画像セットから作成された入出力画像対を用いてニューラルネットワークモデルを訓練するために使用することもできる。訓練セット内の例示的な画像は、モデルがSR処理、ノイズ低減（NR）、およびアーチファクト除去（AR）を実行してモバイルデバイス内で最適なUR処理を達成することを可能にするために、シミュレートされたノイズおよびブロッキングアーチファクトを用いてダウンサンプリングさ

10

20

30

40

50

れ、入力側で劣化される。出力側のHR画像のローカルコントラストを強調して、SR処理中に発見された細部を強調することができる。これに関して、ニューラルネットワークモデルは、UR処理によって許容可能なエネルギー消費で、リアルタイムで達成することができる著しい視覚の改善を伴うUR処理を効率的に実行するように訓練される。

【0040】

従来技術は、画像逆畳み込みアルゴリズムまたは例ベースのスパースコーディングアルゴリズムを使用してSR処理を実行する。例えば、いくつかの従来技術は、パッチ抽出、非線形マッピング、および再構成のためのスパースコーディングの手順をそれぞれシミュレートするために3層畳み込みモデルを使用する。しかしながら、そのような技術は、入力LR画像が畳み込み層に供給される前に高解像度にアップサンプリングされる必要があるため、非効率的である。その結果、すべての畳み込み演算は、かなりの量の冗長計算を含むHR空間で実行される。他の従来技術の技術は、LR入力画像を使用し、畳み込み演算を用いてサブ画素がHR画像に合成されるときに、最後のステップまでLR空間において畳み込みを実行する。しかしながら、そのような技術は、モバイルデバイスで作業するときにも遅い。

10

【0041】

単一画像SR処理における新しい傾向は、敵対的生成ネットワーク（GAN）を使用して、画像に欠けている細部を生成することである。しかしながら、GANは、良好な結果を生み出すために大きなネットワーク容量に依存しており、したがってモバイルデバイスに展開するのに適していない。映像用に設計された他の従来技術のSRアルゴリズムは、フレーム間の動きを補償するために連続するフレーム間のオプティカルフロー評価を含む。オプティカルフロー評価は、GANを使用することに関連するタスクを処理することよりも容易ではない計算集約的なタスクであるため、それらの映像SRアルゴリズムはまた、モバイルデバイス上で動作するのが遅い。

20

【0042】

以下は、画像解像度を改善し、モバイルデバイス上でリアルタイムUR処理を実行するための現在開示されている技術の特徴的な特徴である（これらの特徴は、従来技術の画像処理技術には存在しない）：（1）超解像という唯一のタスクに加えて、全体的な視覚が向上する、モバイルデバイス用途に超解像処理を使用すること、（2）超解像、ノイズ低減、ブロッキングアーチファクト除去、ローカルコントラスト強調の機能を組み合わせ、リアルタイム性能でモバイルデバイスに展開できる簡潔なニューラルネットワークモデル（層およびパラメータの数を減らしたもの）を使用すること（SR画像全体ではなく、残差画像が畳み込み層によって処理されるため、処理リソースが大幅に節約され、UR処理速度が向上する）、（3）畳み込み層の出力で高速画素シフト動作を使用して、従来の手法で使用されていた計算集約型のサブ画素畳み込みを置き換えること、（4）入力側の画質を劣化させると同時に出力側の画質を向上させる新しいニューラルネットワークモデル訓練技術を使用して、より効率的なモデル訓練および優れた視覚の出力映像フレームの生成に貢献すること。

30

【0043】

図1は、いくつかの例示的な実施形態による、深層学習（DL）アーキテクチャ（DLA）を使用して訓練されたDLモデル110を生成するDLモデルの訓練を示すブロック図100である。いくつかの例示的な実施形態では、集合的に機械学習アルゴリズムまたはツールとも呼ばれる深層学習プログラムを含む機械学習プログラム（MLP）は、データまたは他の人工知能（AI）ベースの機能を相関させることに関連する動作を実行するために利用される。

40

【0044】

図1に示すように、深層学習モデル訓練108は、（特徴を含むことができる）訓練データ102に基づいてDLA106内で実行される。深層学習モデル訓練108の間、DLモデルのさらなる訓練の目的のために、訓練データ102からの特徴を評価することができる。DLモデル訓練108は、訓練されたDLモデル110をもたらす。訓練されたDLモデル110は、新デ

50

ータ114に基づいて評価116を提供するために使用することができる1つまたは複数の分類器112を含むことができる。

【0045】

いくつかの態様では、訓練データ102は、低解像度(LR)入力画像104および対応する高解像度(HR)目標出力画像118を含むことができる。LR入力画像104およびHR目標出力画像118は、例えば、図6に関連して説明したように生成され、DLモデル訓練108中に使用され、訓練されたDLモデル110が本明細書で説明したUR関連機能を実行することを可能にする。より具体的には、LR入力画像104およびHR目標出力画像118は、LR映像フレームを含むLR画像に対してリアルタイムUR機能を実行するために、ニューラルネットワークモデル(例えば、図3の畳み込み層306を含む図2のニューラルネットワークモデル204)の畳み込み層を訓練するために使用される。

10

【0046】

深層学習は、機械学習の一部であり、明示的にプログラムされることなく学習する能力をコンピュータに与える研究分野である。機械学習は、既存のデータから学習することができ、データを相関させることができ、新データに関する予測を行うことができる、本明細書ではツールとも呼ばれるアルゴリズムの研究および構築を探求する。そのような機械学習ツールは、出力または評価116として表されるデータ駆動の予測または決定を行うために、例示的な訓練データ(例えば、訓練データ102)からモデルを作ることによって動作する。例示的な実施形態は、いくつかの機械学習ツール(例えば、深層学習アーキテクチャ)に関して提示されているが、本明細書に提示される原理は、他の機械学習ツールに適用されてもよい。

20

【0047】

いくつかの例示的な実施形態では、異なる機械学習ツールを使用することができる。例えば、ロジスティック回帰、単純ベイズ、ランダムフォレスト(RF)、ニューラルネットワーク、行列分解、およびサポートベクターマシン(SVM)ツールを、深層学習モデル訓練108中に(例えば、訓練データ102を相関させるために)使用することができる。

【0048】

機械学習における2つの一般的な種類の問題は、分類問題および回帰問題である。分類問題は、カテゴリ化問題とも呼ばれ、項目をいくつかのカテゴリ値のうちの1つに分類する(例えば、この物体はリンゴかオレンジか?)ことを目的とする。回帰アルゴリズムは、いくつかの項目を(例えば、実数である値を提供することによって)定量化することを目的とする。いくつかの実施形態では、DLA106は、訓練データ102を利用して、結果に影響を及ぼす識別された特徴間の相関を見つける機械学習アルゴリズムを使用するように構成することができる。

30

【0049】

機械学習アルゴリズムは、評価116を生成するために、新データ114を分析するために訓練データ102からの特徴を利用する。特徴は、機械学習モデルを訓練するために観察および使用されている現象の個々の測定可能な特性を含む。特徴の概念は、線形回帰などの統計的手法に用いられる説明変数の概念に関する。情動的、識別的、および独立した特徴の選択は、パターン認識、分類、および回帰におけるMLPの効果的な動作にとって重要である。特徴は、数値特徴、文字列、およびグラフなど、異なるタイプのものであってもよい。いくつかの態様では、訓練データは、異なるタイプのものであり、特徴は、コンピューティングデバイスによって使用するための数値である。

40

【0050】

いくつかの態様では、DLモデル訓練108中に使用される特徴には、次の1つまたは複数を含めることができる: LR画像(例えば、LR入力画像104)、HR画像(例えば、HR目標出力画像118)、複数のセンサ(例えば、オーディオ、モーション、イメージセンサ)からのセンサデータ、複数のアクチュエータ(例えば、ワイヤレススイッチまたは他のアクチュエータ)からのアクチュエータイベントデータ、複数の外部ソースからの外部情報ソース、センサ状態データに関連付けられたタイマーデータ(例えば、時間センサデータが

50

取得される)、アクチュエータイベントデータ、または外部情報ソースデータ、ユーザコミュニケーション情報、ユーザデータ、ユーザ行動データなど。

#### 【0051】

機械学習アルゴリズムは、訓練データ102を利用して、評価116の結果に影響を及ぼす識別された特徴間の相関を見つける。いくつかの例示的な実施形態では、訓練データ102は、ラベル付きデータを含み、ラベル付きデータは、1つまたは複数の識別された特徴および1つまたは複数の結果についての既知のデータである。(識別された特徴を含むことができる)訓練データ102を用いて、DLモデルは、DLA106内のDLモデル訓練108を使用して訓練される。訓練の結果は、訓練されたDLモデル110である。DLモデル110が評価を実行するために使用される場合、新データ114が訓練されたDLモデル110への入力として提供され、DLモデル110は評価116を出力として生成する。例えば、DLA106はモバイルデバイスに展開することができ、新データ114はLR画像(例えば、リアルタイムLR映像フィードなどのLR映像からのフレーム)を含むことができる。DLA106は、LR画像に対してUR機能(例えば、ノイズを低減しながら画像解像度を向上させ、ブロッキングアーチファクトを除去し、画像コントラストを向上させる)を実行して、リアルタイムでHR出力画像を生成する。

10

#### 【0052】

図2は、いくつかの例示的な実施形態による、DLA106内で訓練されたニューラルネットワークモデル204を使用した訓練されたDLモデル206の生成を示す図200である。図2を参照すると、ソースデータ202は、訓練されたDLモデル206(訓練されたDLモデル110と同じであり得る)を生成するために、ニューラルネットワークモデル204(または別のタイプの機械学習アルゴリズムもしくは技術)によって分析することができる。ソースデータ202は、1つまたは複数の特徴によって識別されるデータを含む、102などのデータの訓練セットを含むことができる。本明細書で使用される場合、「ニューラルネットワーク」および「ニューラルネットワークモデル」という用語は交換可能である。

20

#### 【0053】

機械学習技術は、モデルに供給されるデータに対して正確に予測を行うようにモデルを訓練する(例えば、所与の発話においてユーザによって言われたこと;名詞が人、場所、または物であるかどうか;明日の天気はどうか)。学習フェーズの間、モデルは、所与の入力に対する目標出力を正しく予測するようにモデルを最適化するために入力の訓練データセットに対して開発される。一般に、学習フェーズは、教師あり、半教師あり、または教師なしであってもよく、訓練入力に対応して「正しい」出力が提供される減少レベルを示す。教師あり学習フェーズでは、すべての目標出力がモデルに提供され、モデルは、入力を出力にマッピングする一般的な規則またはアルゴリズムを開発するように指示される。対照的に、教師なし学習フェーズでは、モデルが訓練データセット内の関係を発見するためにそれ自体の規則を開発することができるように、所望の出力は入力に提供されない。半教師あり学習フェーズでは、不完全にラベル付けされた訓練セットが提供され、出力の一部は訓練データセットについて知られており、一部は知られていない。

30

#### 【0054】

モデルは、いくつかのエポックの訓練データセットに対して実行することができ、訓練データセットは、その結果を改良するためにモデルに繰り返し供給される(すなわち、データセット全体がエポックの間に処理される)。反復中、モデル(例えば、ニューラルネットワークモデルまたは別のタイプの機械学習モデル)は、データセット全体のミニバッチ(または一部)に対して実行される。教師あり学習フェーズでは、所与の入力セット(例えば、ソースデータ202)に対する目標出力を予測するためのモデルが開発され、訓練データセットの最大数の入力に対する所与の入力に対応すると指定された出力をより確実に提供するために、いくつかのエポックにわたって評価される。別の例では、教師なし学習フェーズの場合、データセットをn個のグループにクラスタ化するためにモデルが開発され、所与の入力を所与のグループにどの程度一貫して配置するか、および各エポックにわたってn個の所望のクラスタをどの程度確実に生成するかに関して、いくつかのエポッ

40

50

クにわたって評価される。

【0055】

エポックが実行されると、モデルが評価され、それらの変数（例えば、重み、バイアス、または他のパラメータ）の値が調整されて、モデルを反復的にさらに改良しようと試みる。本明細書で使用される場合、「重み」という用語は、機械学習モデルによって使用されるパラメータを指すために使用される。逆方向計算中、モデルは勾配を出力することができ、これは順方向計算に関連する重みを更新するために使用することができる。

【0056】

様々な態様では、評価は、偽陰性に対してバイアスされ、偽陽性に対してバイアスされ、またはモデルの全体的な精度に関して均等にバイアスされる。値は、使用される機械学習技術に応じていくつかの方法で調整することができる。例えば、遺伝的または進化的アルゴリズムでは、所望の出力の予測に最も成功したモデルの値は、後続のエポック中に使用するモデルの値を開発するために使用され、追加のデータ点を提供するためのランダムな変動/突然変異を含み得る。当技術分野の通常のスキルの1つは、線形回帰、ランダムフォレスト、決定木学習、ニューラルネットワーク、深層ニューラルネットワークなどを含む、本開示で適用され得るいくつかの他の機械学習アルゴリズムに精通している。

【0057】

各モデルは、入力に影響する1つまたは複数の変数の値を変化させて所望の結果により密接にマッピングすることによって、いくつかのエポックにわたって規則またはアルゴリズムを開発するが、訓練データセットは変化する可能性があり、好ましくは非常に大きく、完全な精度および精密さは達成できない可能性がある。したがって、学習フェーズを構成するエポックの数は、所与の試行数もしくは固定時間/計算予算として設定されてもよく、または所与のモデルの精度が十分に高いか低いか、もしくは精度のプラトーに達したときに、その数/予算に達する前に終了されてもよい。例えば、訓練フェーズがnエポックを実行し、少なくとも95%の精度を有するモデルを生成するように設計され、そのようなモデルがnエポック前に生成される場合、学習フェーズは早期に終了し、最終目標精度閾値を満たす生成されたモデルを使用することができる。同様に、所与のモデルがランダム確率閾値を満たすのに十分に不正確である（例えば、モデルは、所与の入力に対する真/偽の出力を決定する際に55%の精度しかない）場合、そのモデルの学習フェーズは早期に終了することができるが、学習フェーズ内の他のモデルは訓練を継続することができる。同様に、所与のモデルが引き続き同様の精度を提供するか、または複数のエポックにわたって結果が変動する場合（性能のプラトーに達した場合）、所与のモデルの学習フェーズは、エポック数/計算予算に達する前に終了することができる。

【0058】

学習フェーズが完了すると、モデルを最終決定する。いくつかの例示的な実施形態では、最終決定されたモデルは、試験基準に対して評価される。第1の例では、その入力に対する既知の目標出力を含む試験データセットが最終決定モデルに供給されて、訓練されていないデータを処理する際のモデルの精度を判定する。第2の例では、偽陽性率または偽陰性率を使用して、最終決定後のモデルを評価することができる。第3の例では、各モデル内のデータクラス間の描写を使用して、そのデータのクラスターの最も明確な境界を生成するモデルを選択する。

【0059】

いくつかの例示的な実施形態では、DLモデル206は、ネットワークに配置された、ロングショートタームメモリ（LSTM）ノードなどの一連の「ニューロン」を含む、ニューラルネットワークモデル204（例えば、深層学習、深層畳み込み、またはリカレントニューラルネットワーク）によって訓練される。ニューロンは、データ処理および人工知能、特に機械学習で使用されるアーキテクチャ要素であり、所与のニューロンに提供される入力の重みに基づいて、そのメモリに保持されている値をいつ「記憶」し、いつ「忘れる」かを決定することができるメモリを含む。本明細書で使用されるニューロンの各々は、ネットワーク内の他のニューロンから所定数の入力を受け入れて、解析されているフレームの

10

20

30

40

50

コンテンツのための関係出力およびサブ関係出力を提供するように構成される。個々のニューロンは、発話内の各フレームが互いにどのように関連するかについての相互作用および関係学習モデリングを提供するために、ニューラルネットワークの様々な構成において一緒に連鎖されるか、またはツリー構造に編成され得る。

#### 【0060】

例えば、ニューロンとして機能するLSTMは、入力ベクトル（例えば、発話からの音素）、メモリセル、および出力ベクトル（例えば、コンテキスト表現）を処理するためのいくつかのゲートを含む。入力ゲートおよび出力ゲートは、それぞれメモリセルに流入およびメモリセルから流出する情報を制御するが、忘却ゲートは、任意選択的に、ニューラルネットワーク内の先のリンクされたセルからの入力に基づいてメモリセルから情報を除去する。様々なゲートの重みおよびバイアスペクトルは、訓練フェーズの過程で調整され、訓練フェーズが完了すると、それらの重みおよびバイアスは、通常の動作のために最終決定される。当業者は、ニューロンおよびニューラルネットワークが、プログラムによって（例えば、ソフトウェア命令を介して）、または各ニューロンをリンクしてニューラルネットワークを形成する専用ハードウェアを介して構築され得ることを理解するであろう。

10

#### 【0061】

ニューラルネットワークは、データを分析して評価（例えば、音声単位を認識する）を生成するための特徴を利用する。特徴は、観察されている現象の個々の測定可能な特性である。特徴の概念は、線形回帰などの統計的手法に用いられる説明変数の概念に関する。さらに、深い特徴は、深層ニューラルネットワークの隠れ層内のノードの出力を表す。

20

#### 【0062】

ニューラルネットワーク（例えば、ニューラルネットワークモデル204）は、人工ニューラルネットワークまたはニューラルネットワークモデルと呼ばれることもあるが、動物の脳の生物学的ニューラルネットワークを考慮したコンピューティングシステムである。そのようなシステムは、典型的にはタスク固有のプログラミングなしでタスクを実行するために、学習と呼ばれる性能を徐々に改善する。例えば、画像認識では、オブジェクトの名前でタグ付けされた例示的な画像を分析することによってオブジェクトを含む画像を識別するようにニューラルネットワークを教示することができ、オブジェクトおよび名前を学習した後、分析結果を使用してタグ付けされていない画像内のオブジェクトを識別することができる。ニューラルネットワークは、ニューロンと呼ばれる接続されたユニットの集合に基づいており、シナプスと呼ばれるニューロン間の各接続は、接続の強度とともに変化する活性化強度を有する一方向性信号を送信することができる。受信ニューロンは、典型的には、潜在的に多くの送信ニューロンからのものである結合された入力信号が十分な強度であるかどうかに基づいて、それに接続された下流ニューロンに信号をアクティブ化し伝播することができ、強度はパラメータである。

30

#### 【0063】

CNNとも呼ばれるDNNは、複数の畳み込み層で構成されるスタックニューラルネットワークである。層は、計算が行われる場所であるノードから構成され、十分な刺激に遭遇したときに発火する人間の脳内のニューロン上に緩やかにパターン化される。ノードは、データからの入力を、その入力を増幅または減衰させる係数または重みのセットと組み合わせ、アルゴリズムが学習しようとしているタスクの入力に重要性を割り当てる。これらの入力重みの積が合計され、その合計がノードの活性化関数と呼ばれるものに渡され、その信号がネットワークをさらに進んで最終的な結果に影響を与えるかどうか、およびその程度を判定する。DNNは、特徴抽出および変換のために非線形処理ユニットの多くの層のカスケードを使用する。各連続層は、前の層からの出力を入力として使用する。上位レベルの特徴は、下位レベルの特徴から導出されて階層表現を形成する。入力層に続く層は、入力のフィルタリング結果であり、次の畳み込み層によって使用される特徴マップを生成する畳み込み層であってもよい。

40

#### 【0064】

DNNアーキテクチャの訓練において、変数間の関係を推定するための統計的プロセスの

50

セットとして構造化される回帰は、コスト関数の最小化を含むことができる。コスト関数は、出力を補正するためにニューラルネットワークが訓練例をマッピングする際にどれだけ良好に実行されたかを表す数を返す関数として実装されてもよい。訓練において、コスト関数値が所定の範囲内でない場合、既知の訓練画像に基づいて、誤差逆伝播法が使用され、誤差逆伝播法は、確率的勾配降下法（SGD）などの最適化方法とともに使用される人工ニューラルネットワークを訓練する一般的な方法である。

**【 0 0 6 5 】**

誤差逆伝播法の使用は、伝播および重み更新を含むことができる。入力がニューラルネットワークに提示されると、出力層に到達するまで、層ごとにニューラルネットワークを介して順方向に伝播される。次いで、ニューラルネットワークの出力は、コスト関数を使用して所望の目標出力と比較され、出力層内のノードの各々について誤差値が計算される。誤差値は、出力から開始して、各ノードが元の出力への寄与を大まかに表す関連する誤差値を有するまで逆方向に伝播される。誤差逆伝播法は、これらの誤差値を使用して、ニューラルネットワーク内の重みに対するコスト関数の勾配を計算することができる。計算された勾配は、選択された最適化方法に供給され、重みを更新してコスト関数を最小化しようと試みる。

10

**【 0 0 6 6 】**

訓練アーキテクチャ106は、ニューラルネットワークモデルを使用する深層学習アーキテクチャと呼ばれる（訓練されたモデルは、訓練されたDLモデル110および206などの訓練された深層学習モデルと呼ばれる）が、本開示はこれに限定されず、本明細書に開示された技術を使用して、他のタイプの機械学習訓練アーキテクチャもモデル訓練に使用することができる。

20

**【 0 0 6 7 】**

図3は、いくつかの例示的な実施形態による、図1のDLAを使用して画像解像度を調整するためのシステム300を示す図である。図3を参照すると、システム300は、畳み込み層306と、残差生成モジュール322と、バイキュービックアップサンプリングモジュール316と、画素シフトモジュール318と、加算器320とを含む。畳み込み層306は、図2のニューラルネットワークモデル204などのニューラルネットワークモデルとして構成することができる。

**【 0 0 6 8 】**

残差生成モジュール322は、好適な回路、論理、インタフェース、またはコードを備えることができ、入力LR画像302を赤 - 緑 - 青（RGB）色空間から輝度（Y）、彩度青差（Cb）、および彩度赤差（Cr）（またはYCbCr）色空間に変換し、入力LR画像302に対応するYチャンネル（輝度）LR残差画像304（すなわち、輝度を示すグレースケール画像である）を取得するように構成される。

30

**【 0 0 6 9 】**

バイキュービックアップサンプリングモジュール316は、好適な回路、論理、インタフェース、またはコードを備えることができ、ベースHR画像312を生成するために、入力LR画像302に対してバイキュービックアップサンプリングを実行するように構成される。

**【 0 0 7 0 】**

畳み込み層306は、ニューラルネットワークモデル（例えば、ニューラルネットワークモデル204）の一部とすることができ、LR残差画像（例えば、LR残差画像304などのLR入力画像のグレースケールバージョン）に基づいて入力LR画像302に対応する複数のHR残差サブ画像308を生成するように構成される。畳み込み層306のさらなる詳細は、図4に関連して以下に説明される。

40

**【 0 0 7 1 】**

画素シフトモジュール318は、好適な回路、論理、インタフェース、またはコードを備えることができ、HR残差画像310を生成するために、複数のHR残差サブ画像308に対して画素シフトを実行するように構成される。

**【 0 0 7 2 】**

50

動作中、入力LR画像302に対応するベースHR画像312およびLR残差画像304を生成するために、低解像度(LR)入力画像302が(例えば、バイキュービックアップサンプリングモジュール316および残差生成モジュール322によって)処理される。畳み込み層306は、LR残差画像304を含み、入力LR画像302に対応する複数のHR残差サブ画像308を生成するために複数の層を使用する。いくつかの態様では、複数のHR残差サブ画像308は、4つのサブ画像を含む。画素シフトモジュール318は、HR残差画像310を生成するために、複数のHR残差サブ画像308に対して画素シフトを実行する。加算器320は、好適な回路、論理、インタフェース、またはコードを備えることができ、入力LR画像302に対応する出力画像としてHR画像314を生成するために、ベースHR画像312をHR残差画像310と加算するように構成される。

10

【0073】

HR画像314の生成をベースHR画像312およびHR残差画像310の生成に分割することで、畳み込み層306を流れるデータ量が削減され、入力LR画像からHR画像を再構成するためのモデルの有効容量およびその推論効率が向上する。

【0074】

いくつかの態様では、バイキュービックアップサンプリングモジュール316、残差生成モジュール322、画素シフトモジュール318、および加算器320によって実行される1つまたは複数の機能は、畳み込み層306によって実行することができる。

【0075】

図4は、いくつかの例示的な実施形態による、図3のシステム300によって使用されるDLA内の畳み込み層306の構成を示す図である。図4を参照すると、畳み込み層306は、図4に示すように層深さ(またはチャンネル数)420を有する4つの層、すなわち畳み込み層404、406、408、および410を含む。第1の畳み込み層404は8チャンネルの測定値(すなわち、層は8の深さを有する)を生成するように構成され、第2の畳み込み層406は6チャンネルの測定値(すなわち、層は6の深さを有する)を生成するように構成され、第3および第4の畳み込み層(それぞれ408および410)は各々4チャンネルの測定値(すなわち、層は4の深さを有する)を生成するように構成される。畳み込み層404から408は、それぞれサイズが3×3画素の畳み込みカーネル414、416、および418を使用してそれらのチャンネル測定値を生成するように構成される。

20

【0076】

畳み込み層306はまた、畳み込み層404から410への入力として受信された画像に対応する1チャンネル入力層402を含むことができる。例えば、入力層402は、畳み込み層404から410への入力として通信されるLR残差画像304を表すことができる。入力層402は、第1の畳み込み層404への入力として画像データを通信するために使用されるカーネル412(例えば、5×5画素カーネル)を含むことができる。

30

【0077】

動作中、畳み込み層404~410の各々は、複数の畳み込みカーネルを使用して前の層から受信した入力を畳み込む。例えば、第1の畳み込み層404がカーネル412を介して入力層402から入力データを受信した後、第1の畳み込み層404は、入力層402からの入力およびサイズ3×3画素の8つの異なる畳み込みカーネル(畳み込みカーネル414など)を使用して8チャンネルの測定値を生成する。各チャンネルの測定値は、追加の畳み込みのために後続の層に伝達される。第4の畳み込み層410は、複数のHR残差サブ画像308を出力する(例えば、層410の4つのチャンネルの各々は、合計4つのHR残差サブ画像308に対して1つのHR残差サブ画像を出力する)。

40

【0078】

図4は、示された層深さ420を有する4つの別個の畳み込み層404から410を含む畳み込み層306を示しているが、本開示はこれに限定されず、畳み込み層306の異なる構成もまた、UR関連機能を実行するためにシステム300内で使用することができる。

【0079】

図5は、いくつかの例示的な実施形態による、図1のDLAの訓練に使用することができる

50

訓練画像対の生成を示す図500である。図5を参照すると、各訓練画像対は、畳み込み層入力画像532と、1組の畳み込み層目標出力画像536とを含む。畳み込み層入力画像532および畳み込み層目標出力画像536の両方は、訓練フェーズ中に畳み込み層306に入力され、畳み込み層目標出力画像536は、訓練フェーズ中の実際の出力を含まない。実際の畳み込み演算中、532で処理される画像が畳み込み層306に入力され、536で畳み込み層306は処理されたUR画像を出力することを理解されたい。

#### 【0080】

畳み込み層入力画像532を生成するために、ローパスフィルタ(LPF)522が、フィルタリングされた画像504を生成するために例示的な訓練画像1502に適用される。次いで、フィルタリングされた画像504は、ダウンサンプリングされたLR画像 $I_{LR}506$ を生成するためにダウンサンプリングされる(例えば、ダウンサンプリング(DS)モジュール524を使用する)。いくつかの態様では、フィルタリングされた画像504は、以下の式： $I_{LR} = I * N(0, \sigma_{blur})$ によって表されるように0.5倍だけダウンサンプルすることができ、式中、 $*$ は畳み込み演算を表し、 $N(0, \sigma_{blur})$ は、 $\sigma_{blur}$ の平均値0および標準偏差を有するガウス関数を表す。

10

#### 【0081】

いくつかの態様では、LPF522は、高周波信号を除去し、それによってアンダーサンプルアーチファクトを回避するために使用される。いくつかの態様では、DSモジュール524によるダウンサンプリングは、バイキュービック補間によって行われる。ダウンサンプリングされた(例えば、半分のサイズの)LR画像506は、次いで、畳み込み層に供給される前に、ノイズおよびブロッキングアーチファクトによって劣化する。より具体的には、ノイズ付加モジュール526は、ダウンサンプリングされたLR画像506にノイズを導入するために使用される。いくつかの態様では、(1)光子ノイズおよび(2)ガウスノイズの2種類のノイズがシミュレートされる。

20

#### 【0082】

光子ノイズは、光の離散特性をシミュレートするために使用され、以下のようにポアソン確率プロセスでシミュレートされる： $I_{pho} \sim P(I_{LR} \text{光子}) / \text{光子}$ 、式中、 $P(\cdot)$ は、平均 $\lambda$ を有するポアソン分布を表し、光子は、画像が表すことができる最も明るい色を生成する光子の数を表す。

#### 【0083】

ガウスノイズは、画像センサからの確率的ノイズをシミュレートするために使用される。これは、ガウス確率過程を用いて生成されたランダム画像 $I_{gauss}$ を入力画像に加えることによって、以下のように計算される： $I_{noisy} = I_{pho} + I_{gauss}$ 、式中、 $I_{gauss} \sim N(0, \sigma_n)$ であり、 $\sigma_n$ は目標ノイズレベルを表す。

30

#### 【0084】

映像の圧縮および解凍は、通常、映像フレームにブロッキングアーチファクトを引き起こす。アーチファクトは、映像圧縮および解凍によって直接シミュレートされる。入力側で画質が劣化した畳み込み層306を使用してニューラルネットワークモデル(例えば、204)を訓練することにより、ニューラルネットワークモデルノイズの低減およびアーチファクトの除去能力が提供される。

40

#### 【0085】

ノイズ付加モジュール526は、ノイズの多いLR画像508を生成するために、ダウンサンプリングされたLR画像506にノイズを導入する。ノイズの多いLR画像508は、アーチファクトを導入し、訓練画像502に対応するLR画像510を生成するアーチファクト追加モジュール528によってさらに劣化する。LR画像510は、劣化した画質を特徴とし、訓練フェーズのために、畳み込み層306への畳み込み層入力画像532として構成される。畳み込み層入力画像532は、所望の目標出力を生成するために畳み込み層306を訓練するために使用される。ここで、畳み込み層306の所望の出力は、畳み込み層の訓練フェーズ中に畳み込み層目標出力画像536として入力される。

#### 【0086】

50

畳み込みネットワークの畳み込み層目標出力画像536のセットを生成するために、訓練画像1502は、コントラスト強調画像512を生成するローカルコントラスト強調（LCE）モジュール520によって強調される。この処理は、畳み込み層306に高いローカルコントラストを有する画像を生成するように教示することに寄与し、これは超解像によって復元される微細な細部を強化する。いくつかの態様では、LCEモジュール520は、以下のように、アンバランスなアンシャープマスクを適用することによって、ローカルコントラスト強調を実行することができ： $I_{LC} = I + \min(I - I * N(0, \sigma_{um}), \sigma_{max})$ 、式中、訓練画像とそのガウスぼかしバージョンの違いは、正の部分で上限  $\sigma_{max}$  が切り捨てられることである（切り捨ては、人間の目が明るいエッジに敏感であるために表示されるハロアーチファクトを低減するのに有益である）。次いで、HR残差画像516は、減算モジュール540を介してコントラスト強調画像512からダウンサンプルされたLR画像506のアップサンプリングされたLR画像538を減算することによって計算される。アップサンプリングされたLR画像538は、アップサンプリングモジュール530によってダウンサンプリングされたLR画像506をアップサンプリングすることによって生成される。画素分割動作534がHR残差画像516に適用され、複数のHR残差サブ画像518がHR残差画像516から（例えば、 $4 \times 4$ 画素ブロックを4つのサブ画像に分割することによって）生成される。複数のHR残差サブ画像518は、訓練フェーズ中に畳み込み層306に入力される、畳み込み層目標出力画像536として構成される。

【0087】

いくつかの態様では、畳み込み層306を訓練する前に複数の訓練対が生成され、訓練対はデータファイル（例えば、tfrecordフォーマットを使用する）に保存される。ニューラルネットワークモデル訓練（例えば、108）中に、訓練コードは、訓練データを抽出し、そのようなデータを使用してUR処理を実行するための（例えば、畳み込み層306の）モデルパラメータを調整することができる。これにより、複雑な画像処理をその場で回避することでモデル訓練が高速化され、訓練データに関連付けられた様々なパラメータを使用してモデルを繰り返し訓練できるようになる。

【0088】

本明細書で説明するニューラルネットワークモデル訓練技術（例えば、入力側の画質を劣化させ、出力側の画質を向上させる）を使用することにより、畳み込み層306は、LR画像または映像フレームのUR処理中に優れた視覚の画像または映像フレームを生成することができる。さらに、超解像、ノイズ低減、ブロッキングアーチファクトの除去、およびローカルコントラスト強調の機能を組み合わせた簡潔なニューラルネットワークモデル（例えば、4つなどの限られた数の層を有する畳み込み層306）を使用すると、簡潔なニューラルネットワークモデルは、リアルタイムUR処理を実行するためにモバイルデバイス（またはスマートTV、タブレット、ラップトップ、および他のコンピューティングデバイスなどの他の種類の限定リソースデバイス）に展開することができる。

【0089】

動作使用において、処理されるべき実際の画像または映像フレームは、畳み込み層306への入力532を備える。畳み込み層306は、訓練の結果として、出力536においてUR画像または映像フレームを生成するために画像または映像フレームを処理する。出力されたUR画像または映像フレームは、その後、表示、送信、格納、または他の方法で使用することができる。

【0090】

図6は、いくつかの例示的な実施形態による、訓練画像対の生成、ならびにDLA訓練および画像解像度調整機能の実行のための対の使用の方法600のフローチャートである。方法600は、動作602から618を含む。限定ではなく例として、方法600は、超高解像管理モジュール860によって実行されるものとして記載され、超高解像管理モジュール860は、図9に示すデバイス900などのモバイルデバイス内で実行するように構成することができる。

【0091】

10

20

30

40

50

図6を参照すると、訓練画像対の生成は、訓練画像（例えば、502）が受信されたときに動作602で開始することができる。動作604において、フィルタリングされた画像（例えば、504）を生成するために、ローパスフィルタ（例えば、522）を使用して訓練画像がフィルタリングされる。動作606において、ダウンサンプリングされたLR画像（例えば、506）を生成するために、フィルタリングされた画像が（例えば、ダウンサンプリングモジュール524によって）ダウンサンプリングされる。動作608において、訓練画像に対応する劣化した画質（例えば、LR画像510）を有する低解像度画像を生成するために、ダウンサンプリングされたLR画像の画質が、ノイズ（例えば、ノイズ付加モジュール526によって）およびアーチファクト（例えば、アーチファクト追加モジュール528によって）を追加することによって劣化する。動作610において、訓練画像に対応するLR画像が、ニューラルネットワークモデルへの入力として構成される。

10

## 【0092】

訓練対内の出力画像を生成するため、動作612において、コントラスト強調画像を生成するために、アンバランスなアンシャープマスクが訓練画像に適用される（例えば、LCEモジュール520は、コントラスト強調画像512を生成する）。動作614において、訓練画像に対応するHR残差画像を生成するために、ダウンサンプリングされたLR画像のアップサンプリングされたバージョンがコントラスト強調画像から減算される（例えば、HR残差画像516を、アップサンプリングされたLR画像538およびコントラスト強調画像512を使用して減算モジュール540によって生成する）。動作616において、訓練画像に対応する複数のHR残差サブ画像を生成するために、（訓練画像に対応する）HR残差画像が分割される（例えば、複数のHR残差サブ画像518を、HR残差画像516に対して実行される画素分割動作534を介して生成する）。動作618において、訓練画像に対応する複数のHR残差サブ画像は、ニューラルネットワークモデルの出力として構成される（例えば、複数のHR残差サブ画像518は、畳み込み層306を訓練するための畳み込み層目標出力画像536のセットとして構成される）。

20

## 【0093】

図7は、いくつかの例示的な実施形態による、デジタル画像の画像解像度を高めるための方法700のフローチャートである。方法700は、動作702、704、706、および708を含む。限定ではなく例として、方法700は、超高解像管理モジュール860によって実行されるものとして記載され、超高解像管理モジュールは、図9に示すデバイス900などのモバイルデバイス内で実行するように構成することができる。

30

## 【0094】

図7を参照すると、動作702において、ベース高解像度（HR）画像を生成するために、デジタル画像に対してバイキュービックアップサンプリングが実行される。例えば、図3を参照すると、バイキュービックアップサンプリングモジュールは、入力LR画像を使用してベースHR画像を生成する。動作704において、低解像度（LR）残差画像を生成するために、デジタル画像は、赤 - 緑 - 青（RGB）色空間から輝度（Y）、彩度青差（Cb）、および彩度赤差（Cr）（YCbCr）色空間に変換される。いくつかの態様では、RGBからYCrCbへの変換は、LR残差画像を生成するために、ニューラルネットワークの複数の畳み込み層をデジタル画像に適用することによって（または残差生成モジュールを使用して）実行することができる。

40

## 【0095】

動作706において、LR残差画像は、（例えば、複数の畳み込み層を使用することによって）入力LR画像に対応する複数のHR残差サブ画像に変換される。例えば、入力LR画像に対応する複数のHR残差サブ画像を生成するために、畳み込み層は、LR残差画像を使用する。動作708において、ベースHR画像および複数のHR残差サブ画像を使用して、入力LR画像に対応するHR画像が生成される。例えば、複数のHR残差サブ画像を画素シフトすることにより、HR残差画像が生成される。HR画像は、ベースHR画像およびHR残差画像を使用して生成される。

## 【0096】

50

図8は、いくつかの例示的な実施形態による、本明細書に記載の様々なデバイスハードウェアとともに使用することができる代表的なソフトウェアアーキテクチャ800を示すブロック図である。図8は、ソフトウェアアーキテクチャ802の非限定的な例にすぎず、本明細書に記載の機能を容易にするために多くの他のアーキテクチャが実装され得ることが理解されよう。ソフトウェアアーキテクチャ802は、とりわけ、プロセッサ905、メモリ910、ストレージ915および/または920、ならびにI/Oインタフェース925および930を含む図9のデバイス900などのハードウェア上で実行される。

**【0097】**

代表的なハードウェア層804が示されており、例えば、図9のデバイス900を表すことができる。代表的なハードウェア層804は、関連する実行可能命令808を有する1つまたは複数の処理ユニット806を備える。実行可能命令808は、図1～図14の方法、モジュールなどの実装を含む、ソフトウェアアーキテクチャ802の実行可能命令を表す。ハードウェア層804はまた、実行可能命令808も有するメモリまたはストレージモジュール810を含む。ハードウェア層804はまた、デバイス900の一部として示される他のハードウェアなど、ハードウェア層804の任意の他のハードウェアを表す他のハードウェア812を含むことができる。

10

**【0098】**

図8の例示的なアーキテクチャでは、ソフトウェアアーキテクチャ802は、各層が特定の機能を提供する層のスタックとして概念化することができる。例えば、ソフトウェアアーキテクチャ802は、オペレーティングシステム814、ライブラリ816、フレームワーク/ミドルウェア818、アプリケーション820、およびプレゼンテーション層844などの層を含むことができる。動作的には、アプリケーション820または層内の他の構成要素は、ソフトウェアスタックを介してアプリケーションプログラミングインターフェース(API)呼び出し824を呼び出し、API呼び出し824に回答してメッセージ826として示される応答、戻り値などを受信することができる。図8に示す層は本質的に代表的なものであり、すべてのソフトウェアアーキテクチャ802がすべての層を有するわけではない。例えば、一部のモバイルまたは専用オペレーティングシステムはフレームワーク/ミドルウェア818を提供しない場合があるが、他のものはそのような層を提供することができる。他のソフトウェアアーキテクチャは、追加の層または異なる層を含んでもよい。

20

**【0099】**

オペレーティングシステム814は、ハードウェアリソースを管理し、共通サービスを提供することができる。オペレーティングシステム814は、例えば、カーネル828、サービス830、およびドライバ832を含むことができる。カーネル828は、ハードウェアと他のソフトウェア層との間の抽象化層として機能することができる。例えば、カーネル828は、メモリ管理、プロセッサ管理(例えば、スケジューリング)、構成要素管理、ネットワーク、セキュリティ設定などを担当することができる。サービス830は、他のソフトウェア層に他の共通サービスを提供することができる。ドライバ832は、基礎となるハードウェアの制御またはインタフェースを担当することができる。例えば、ドライバ832は、ハードウェア構成に応じて、ディスプレイドライバ、カメラドライバ、Bluetooth(登録商標)ドライバ、フラッシュメモリドライバ、シリアル通信ドライバ(例えば、ユニバーサルシリアルバス(USB)ドライバ)、Wi-Fi(登録商標)ドライバ、オーディオドライバ、電力管理ドライバなどを含むことができる。

30

40

**【0100】**

ライブラリ816は、アプリケーション820または他の構成要素もしくは層によって利用され得る共通のインフラストラクチャを提供することができる。ライブラリ816は、典型的には、他のソフトウェアモジュールが、基礎となるオペレーティングシステム814の機能(例えば、カーネル828、サービス830、またはドライバ832)と直接インタフェースするよりも容易な方法でタスクを実行することを可能にする機能を提供する。ライブラリ816は、メモリ割り当て機能、文字列操作機能、数学的機能などの機能を提供することができるシステムライブラリ834(例えば、C標準ライブラリ)を含むことができる。さら

50

に、ライブラリ816は、メディアライブラリ（例えば、MPEG4、H.264、MP3、AAC、AMR、JPG、PNGなどの様々なメディアフォーマットの提示および操作をサポートするライブラリ）、グラフィックスライブラリ（例えば、ディスプレイ上のグラフィックコンテンツにおいて2Dおよび3Dをレンダリングするために使用され得るOpenGLフレームワーク）、データベースライブラリ（例えば、様々なリレーショナルデータベース機能を提供することができるSQLite）、ウェブライブラリ（例えば、ウェブブラウジング機能を提供することができるWebKit）などのAPIライブラリ836を含むことができる。ライブラリ816はまた、アプリケーション820および他のソフトウェア構成要素/モジュールに多くの他のAPIを提供するための多種多様な他のライブラリ838を含むことができる。

**【0101】**

フレームワーク/ミドルウェア818（ミドルウェアと呼ばれることもある）は、アプリケーション820または他のソフトウェア構成要素/モジュールによって利用され得る高レベルの共通インフラストラクチャを提供することができる。例えば、フレームワーク/ミドルウェア818は、様々なグラフィカルユーザインタフェース（GUI）機能、高レベルリソース管理、高レベルロケーションサービスなどを提供することができる。フレームワーク/ミドルウェア818は、アプリケーション820または他のソフトウェア構成要素/モジュールによって利用され得る広範囲の他のAPIを提供することができ、その一部は特定のオペレーティングシステム814またはプラットフォームに固有であり得る。

**【0102】**

アプリケーション820は、組み込みアプリケーション840、サードパーティアプリケーション842、および超高解像度管理モジュール（URMM）860を含む。いくつかの態様では、URMM860は、適切な回路、論理、インタフェース、またはコードを備えることができ、図1～図7に関連して説明したUR関連機能のうちの1つまたは複数を実行するように構成することができる。

**【0103】**

代表的な組み込みアプリケーション840の例は、連絡先アプリケーション、ブラウザアプリケーション、書籍リーダーアプリケーション、ロケーションアプリケーション、メディアアプリケーション、メッセージングアプリケーション、またはゲームアプリケーションを含むことができるが、これらに限定されない。サードパーティアプリケーション842は、組み込みアプリケーション840のうちのいずれか、ならびに、広範な他のアプリケーションを含み得る。特定の例では、サードパーティアプリケーション842（例えば、特定のプラットフォームのベンダ以外のエンティティによってAndroid（商標）またはiOS（商標）ソフトウェア開発キット（SDK）を使用して開発されたアプリケーション）は、iOS（商標）、Android（商標）、Windows（登録商標）Phone、または他のモバイルオペレーティングシステムなどのモバイルオペレーティングシステム上で実行されるモバイルソフトウェアであってもよい。この例では、サードパーティアプリケーション842は、本明細書で記載される機能を容易にするために、オペレーティングシステム814などのモバイルオペレーティングシステムによって提供されるAPI呼び出し824を呼び出すことができる。

**【0104】**

アプリケーション820は、組み込みオペレーティングシステム機能（例えば、カーネル828、サービス830、およびドライバ832）、ライブラリ（例えば、システムライブラリ834、APIライブラリ836、および他のライブラリ838）、およびフレームワーク/ミドルウェア818を利用して、システムのユーザと対話するためのユーザインタフェースを作成することができる。代替的または追加的に、いくつかのシステムでは、ユーザとの対話は、プレゼンテーション層844などのプレゼンテーション層を介して行われてもよい。これらのシステムでは、アプリケーション/モジュール「論理」は、ユーザと対話するアプリケーション/モジュールの態様から分離することができる。

**【0105】**

一部のソフトウェアアーキテクチャは仮想マシンを利用する。図8の例では、これは仮

10

20

30

40

50

想マシン848によって示されている。仮想マシンは、アプリケーション/モジュールがあたかもハードウェアマシン上で実行されているかのように（例えば、図9のデバイス900のように）実行することができるソフトウェア環境を作成する。仮想マシン848は、ホストオペレーティングシステム（例えば、オペレーティングシステム814）によってホストされ、典型的には、常にではないが、仮想マシン848の動作ならびにホストオペレーティングシステム（すなわち、オペレーティングシステム814）とのインタフェースを管理する仮想マシンモニタ846を有する。ソフトウェアアーキテクチャ802は、オペレーティングシステム850、ライブラリ852、フレームワーク/ミドルウェア854、アプリケーション856、またはプレゼンテーション層858などの仮想マシン848内で実行される。仮想マシン848内で実行されるソフトウェアアーキテクチャのこれらの層は、前述の対応する層と同じであってもよいし、異なってもよい。

10

**【0106】**

図9は、例示的な実施形態による、アルゴリズムを実装し、方法を実行するデバイスのための回路を示すブロック図である。様々な実施形態において、すべての構成要素が使用される必要はない。例えば、クライアント、サーバ、およびクラウドベースのネットワークデバイスは、それぞれ異なる構成要素のセットを使用することができ、またはサーバの場合、例えば、より大きなストレージデバイスを使用することができる。

**【0107】**

コンピュータ900（コンピューティングデバイス900、コンピュータシステム900、またはコンピュータ900とも呼ばれる）の形態のある例示的なコンピューティングデバイスは、プロセッサ905と、メモリ910と、リムーバブルストレージ915と、非リムーバブルストレージ920と、入力インタフェース925と、出力インタフェース930と、通信インタフェース935とを含むことができ、これらはすべてバス940によって接続されている。例示的なコンピューティングデバイスがコンピュータ900として図示および記載されているが、コンピューティングデバイスは、異なる実施形態では異なる形であり得る。

20

**【0108】**

メモリ910は、揮発性メモリ945および不揮発性メモリ950を含み得、プログラム955を記憶し得る。コンピューティングデバイス900は、揮発性メモリ945、不揮発性メモリ950、リムーバブルストレージ915、および非リムーバブルストレージ920などの、様々なコンピュータ可読媒体を含み得るか、またはこれらを含むコンピューティング環境にアクセスし得る。コンピュータストレージには、ランダムアクセスメモリ（RAM）、読出し専用メモリ（ROM）、消去可能プログラマブル読出し専用メモリ（EPROM）、電気的消去可能プログラマブル読出し専用メモリ（EEPROM）、フラッシュメモリもしくは他のメモリ技術、コンパクトディスク読出し専用メモリ（CD ROM）、デジタル多用途ディスク（DVD）、もしくは他の光ディスクストレージ、磁気カセット、磁気テープ、磁気ディスクストレージ、もしくは他の磁気ストレージデバイス、またはコンピュータ可読命令を記憶することが可能な任意の他の媒体が含まれる。

30

**【0109】**

コンピュータ可読媒体に記憶されたコンピュータ可読命令（例えば、メモリ910に記憶されたプログラム955）は、コンピューティングデバイス900のプロセッサ905によって実行可能である。ハードドライブ、CD-ROM、およびRAMは、ストレージデバイスなどの非一時的コンピュータ可読媒体を含む物品のいくつかの例である。コンピュータ可読媒体およびストレージデバイスという用語は、搬送波が一時的すぎると見なされる程度にまで搬送波を含まない。コンピュータ可読非一時的媒体は、磁気記憶媒体、光記憶媒体、フラッシュ媒体、および固体記憶媒体を含むあらゆるタイプのコンピュータ可読媒体を含む。ソフトウェアはコンピュータにインストールされ、コンピュータとともに販売され得ることを理解されたい。あるいは、ソフトウェアは、例えば、ソフトウェア作成者によって所有されているサーバから、またはソフトウェア作成者によって所有されていないが使用されているサーバからを含めて、物理媒体または配布システムを介してソフトウェアを取得することを含めて、取得されてコンピュータにロードされ得る。このソフトウェアは、

40

50

例えば、インターネット上の配布用のサーバに記憶され得る。本明細書で使用される場合、「コンピュータ可読媒体」および「機械可読媒体」という用語は交換可能である。

【0110】

プログラム955は、図8のURMM860と同じであり得る、URMM960などの本明細書で説明されるモジュールを使用する顧客嗜好構造を利用することができる。

【0111】

本明細書に記載されたモジュールのうちの任意の1つまたは複数は、ハードウェア（例えば、機械のプロセッサ、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）、またはそれらの任意の適切な組み合わせ）を使用して実装され得る。さらに、これらのモジュールのうちの任意の2つ以上は、単一のモジュールに組み合わされてもよく、単一のモジュールについて本明細書に記載された機能は、複数のモジュールに細分されてもよい。さらに、様々な例示的な実施形態によれば、単一の機械、データベース、またはデバイス内に実装されるものとして本明細書に記載されたモジュールは、複数の機械、データベース、またはデバイスにわたって分散されてもよい。

10

【0112】

いくつかの態様では、URMM960、ならびにプログラム955の一部である1つまたは複数の他のモジュールは、単一のモジュールとして統合することができ、統合されたモジュールの対応する機能を実行する。

【0113】

以上ではいくつかの実施形態が詳細に説明されているが、他の改変形態も可能である。例えば、図に示されている論理フローは、所望の結果を達成するために、図示されている特定の順序、すなわち順番を必要としない。他のステップが提供されてもよく、または記載されたフローからステップが除去されてもよく、記載されたシステムに他の構成要素が追加されてもよく、または記載されたシステムから除去されてもよい。添付の特許請求の範囲内には他の実施形態があり得る。

20

【0114】

本開示のステップのいずれか1つまたはすべてに関して上述したような処理および動作を容易にする1つまたは複数のコンピュータ実行可能命令を含むソフトウェアは、本開示と一致する1つまたは複数のコンピューティングデバイスにインストールされ、それとともに販売され得ることをさらに理解されたい。代替的に、ソフトウェアは、例えば、ソフトウェア作成者によって所有されているサーバから、またはソフトウェア作成者によって所有されていないが使用されているサーバからを含めて、物理媒体または配布システムを介してソフトウェアを取得することを含めて、取得されて1つまたは複数のコンピューティングデバイスにロードされ得る。このソフトウェアは、例えば、インターネット上の配布用のサーバに記憶され得る。

30

【0115】

また、本開示は、その適用において、説明に記載されたまたは図面に示された構成要素の構造および配置の詳細に限定されないことが当業者によって理解されるであろう。本明細書の実施形態は、他の実施形態が可能であり、様々な方法で実施または実行され得る。また、本明細書で使用される表現および用語は、説明のためのものであり、限定するものと見なされるべきではないことが理解されるであろう。本明細書における「含む」、「備える」または「有する」およびその変形の使用は、その後列挙される項目およびその均等物ならびに追加の項目を包含することを意味する。特に限定されない限り、本明細書における「接続される」、「結合される」、および「取り付けられる」という用語ならびにそれらの変形は、広く使用され、直接的および間接的な接続、結合、および取り付けを包含する。加えて、「接続される」および「結合される」という用語およびそれらの変形は、物理的または機械的な接続または結合に限定されない。さらに、上、下、底部、および上部などの用語は相対的であり、例示を助けるために使用されるが、限定的ではない。

40

【0116】

図示された実施形態により採用される例示的なデバイス、システム、および方法の構成

50

要素は、少なくとも部分的に、デジタル電子回路、アナログ電子回路、またはコンピュータハードウェア、ファームウェア、ソフトウェア、またはそれらの組み合わせで実施され得る。これらの構成要素は、例えば、プログラマブルプロセッサ、コンピュータ、または複数のコンピュータなどのデータ処理装置によって実行される、またはその動作を制御するために、情報キャリア、または機械可読ストレージデバイスに有形に具現化されたコンピュータプログラム、プログラムコードまたはコンピュータ命令などのコンピュータプログラム製品として実装され得る。

【0117】

コンピュータプログラムは、コンパイル言語またはインタプリタ言語を含む任意の形のプログラミング言語で記述され得、スタンドアロンプログラムとして、またはモジュール、コンポーネント、サブルーチン、もしくはコンピューティング環境での使用に適した他のユニットとして含む任意の形で展開され得る。コンピュータプログラムは、1つのコンピュータ上で、または1つのサイトの複数のコンピュータ上で実行されるように、または複数のサイトにわたって分散され、通信ネットワークによって相互接続されるように展開され得る。また、本明細書に記載された技術を達成するための機能プログラム、コード、およびコードセグメントは、本明細書に記載された技術が関連する当業者によって特許請求の範囲内にあると容易に解釈され得る。例示的な実施形態に関連付けられた方法ステップは、機能（例えば、入力データを操作し、または出力を生成することによって）を実行するためのコンピュータプログラム、コード、または命令を実行する1つまたは複数のプログラマブルプロセッサによって実行され得る。方法ステップはまた、例えばFPGA（フィールドプログラマブルゲートアレイ）またはASIC（特定用途向け集積回路）などの専用論理回路によって実行され得、方法を実行するための装置は、専用論理回路として実装され得る。

【0118】

本明細書に開示された実施形態に関連して説明された様々な例示的な論理ブロック、モジュール、および回路は、汎用プロセッサ、デジタル信号プロセッサ（DSP）、ASIC、FPGAもしくは他のプログラマブル論理デバイス、ディスクリートゲートもしくはトランジスタ論理、ディスクリートハードウェア構成要素、または本明細書に記載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により実装または実行され得る。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、代替的に、プロセッサは任意の従来のプロセッサ、コントローラ、マイクロコントローラ、また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コアと連携する1つまたは複数のマイクロプロセッサ、またはその他の任意のこのような構成として実装され得る。

【0119】

コンピュータプログラムの実行に適したプロセッサは、例として、汎用および専用マイクロプロセッサの両方、ならびに任意の種類デジタルコンピュータの任意の1つまたは複数のプロセッサを含む。一般に、プロセッサは、読出し専用メモリまたはランダムアクセスメモリあるいはその両方から命令およびデータを受信する。コンピュータの必要な要素は、命令を実行するためのプロセッサ、ならびに命令およびデータを記憶するための1つまたは複数のメモリデバイスである。一般に、コンピュータはまた、データを記憶するための1つまたは複数の大容量ストレージデバイス、例えば、磁気、光磁気ディスク、または光ディスクを含むか、それらからデータを受信するか、それらにデータを転送するか、それらの両方を行うように動作可能に結合される。コンピュータプログラム命令およびデータを具現化するのに適した情報キャリアは、例として、半導体メモリデバイス、例えば、電氣的プログラマブル読出し専用メモリまたはROM（EPROM）、電氣的消去可能プログラマブルROM（EEPROM）、フラッシュメモリデバイス、またはデータストレージディスク（例えば、磁気ディスク、内蔵ハードディスク、またはリムーバブルディスク、光磁気ディスク、またはCD-ROMおよびDVD-ROMディスク）を含む、あらゆる形の不

10

20

30

40

50

揮発性メモリを含む。プロセッサおよびメモリは、専用論理回路によって補完されうるか、または専用論理回路に組み込まれ得る。

【0120】

当業者は、情報および信号が様々な異なる技術および技法のいずれかを使用して表されうることを理解する。例えば、上記の説明全体で参照されうるデータ、命令、コマンド、情報、信号、ビット、シンボル、およびチップは、電圧、電流、電磁波、磁場もしくは粒子、光場もしくは粒子、またはそれらの任意の組み合わせによって表され得る。

【0121】

本明細書で使用される場合、「機械可読媒体」（または「コンピュータ可読媒体」）は、命令およびデータを一時的または永続的に記憶することができるデバイスを含み、ランダムアクセスメモリ（RAM）、読出し専用メモリ（ROM）、バッファメモリ、フラッシュメモリ、光学媒体、磁気媒体、キャッシュメモリ、他の種類のストレージ（例えば、消去可能プログラマブル読出し専用メモリ（EEPROM））、またはそれらの任意の適切な組み合わせを含みうるが、これらに限定されない。「機械可読媒体」という用語は、プロセッサ命令を記憶することができる単一の媒体または複数の媒体（例えば、集中型もしくは分散型データベース、または関連するキャッシュおよびサーバ）を含むと解釈されるべきである。「機械可読媒体」という用語はまた、命令が、1つまたは複数のプロセッサ905による実行時に、1つまたは複数のプロセッサ905に、本明細書に記載された方法論のうちの任意の1つまたは複数を実行させるように、1つまたは複数のプロセッサ905による実行のための命令を記憶することができる任意の媒体、または複数の媒体の組み合わせを含むと解釈されるべきである。したがって、「機械可読媒体」は、単一のストレージ装置またはデバイス、ならびに複数のストレージ装置またはデバイスを含む「クラウドベースの」ストレージシステムまたはストレージネットワークを指す。本明細書で使用される場合、「機械可読媒体」という用語は、信号自体を除外する。

【0122】

さらに、本開示の範囲から逸脱することなく、様々な実施形態において個別または別個なものとして説明および図示されている技法、システム、サブシステム、および方法は、他のシステム、モジュール、技法、または方法と組み合わせたり一体化したりすることができる。互いに結合もしくは直接結合されるまたは通信するものとして示されまたは論じられる他の品目は、電気的にも、機械的にも、またはその他方法でも、なんらかのインタフェース、装置、または中間構成要素を通じて間接的に結合されるまたは通信することができる。他の変更、置換、および代替の例は、当業者によって確認可能であり、本明細書で開示される範囲から逸脱することなく作成され得る。

【0123】

本開示は当該特定の特徴および実施形態を参照して説明されてきたが、本開示の範囲から逸脱することなく様々な変更および組み合わせをそれに行いうることは明らかである。例えば、他の構成要素は、記載されたシステムに追加されてもよく、または記載されたシステムから除去されてもよい。したがって、明細書および図面は、添付の特許請求の範囲によって定義される開示の単なる例示と見なされるべきであり、本開示の範囲内に含まれるありとあらゆる修正、変形、組み合わせまたは同等物を網羅することが意図されている。添付の特許請求の範囲内には他の態様があり得る。最後に、本明細書で使用される場合、特に明記しない限り、接続詞「または」は、非排他的な「または」を指す。

【符号の説明】

【0124】

- 102 訓練データ
- 104 低解像度（LR）入力画像
- 106 深層学習アーキテクチャ（DLA）
- 108 DLモデル訓練
- 110 訓練されたDLモデル
- 112 分類器

10

20

30

40

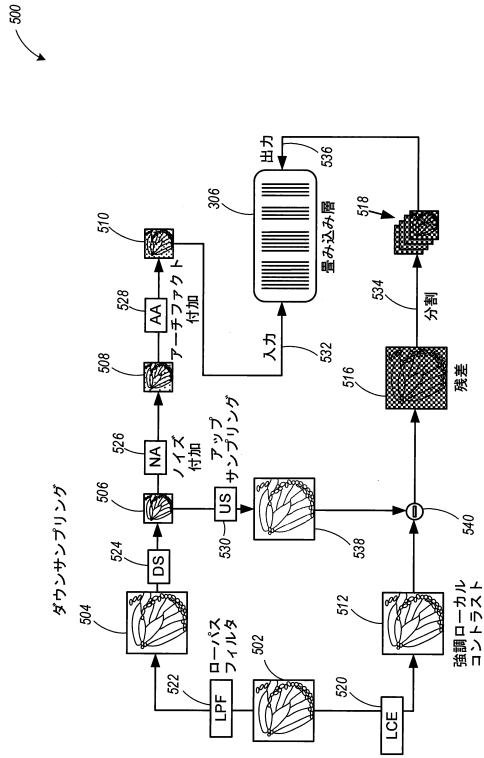
50

114	新データ	
116	評価	
118	高解像度 (HR) 目標出力画像	
202	ソースデータ	
204	ニューラルネットワークモデル	
206	訓練されたDLモデル	
300	システム	
302	入力画像	
302	入力LR画像	
304	LR残差画像	10
306	畳み込み層	
308	HR残差サブ画像	
310	HR残差画像	
312	ベースHR画像	
314	HR画像	
316	バイキュービックアップサンプリングモジュール	
318	画素シフトモジュール	
320	加算器	
322	残差生成モジュール	
402	入力層	20
404	第1の畳み込み層	
406	第2の畳み込み層	
408	第3の畳み込み層	
410	第4の畳み込み層	
412	カーネル	
414	畳み込みカーネル	
416	畳み込みカーネル	
420	層深さ	
502	訓練画像	
504	フィルタリングされた画像	30
506	ダウンサンプリングされたLR画像	
508	ノイズの多いLR画像	
510	LR画像	
512	コントラスト強調画像	
516	HR残差画像	
518	HR残差サブ画像	
520	ローカルコントラスト強調 (LCE) モジュール	
522	ローパスフィルタ (LPF)	
524	ダウンサンプリング (DS) モジュール	
526	ノイズ付加モジュール	40
528	アーチファクト追加モジュール	
530	アップサンプリングモジュール	
532	畳み込み層入力画像	
534	画素分割動作	
536	畳み込み層目標出力画像	
538	アップサンプリングされたLR画像	
540	減算モジュール	
802	ソフトウェアアーキテクチャ	
804	ハードウェア層	
806	処理ユニット	50

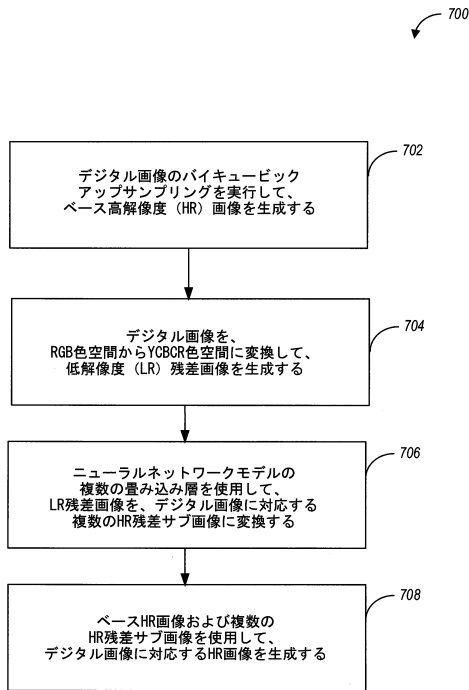
808	実行可能命令	
810	ストレージモジュール	
812	ハードウェア	
814	オペレーティングシステム	
816	ライブラリ	
818	ミドルウェア	
820	アプリケーション	
824	アプリケーションプログラミングインターフェース (API) 呼び出し	
826	メッセージ	
828	カーネル	10
830	サービス	
832	ドライバ	
834	システムライブラリ	
836	APIライブラリ	
840	組み込みアプリケーション	
842	サードパーティアプリケーション	
844	プレゼンテーション層	
846	仮想マシンモニタ	
848	仮想マシン	
850	オペレーティングシステム	20
852	ライブラリ	
854	ミドルウェア	
856	アプリケーション	
858	プレゼンテーション層	
860	超高解像度管理モジュール	
900	コンピューティングデバイス / コンピュータ / コンピュータシステム	
905	プロセッサ	
910	メモリ	
915	ストレージ	
915	リムーバブルストレージ	30
920	非リムーバブルストレージ	
925	入力インタフェース	
930	出力インタフェース	
935	通信インタフェース	
940	バス	
945	揮発性メモリ	
950	不揮発性メモリ	
955	プログラム	



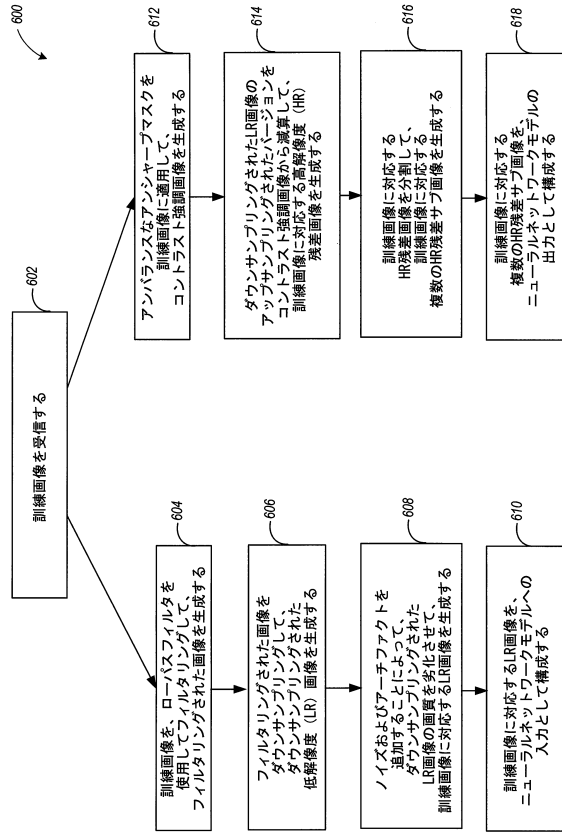
【図5】



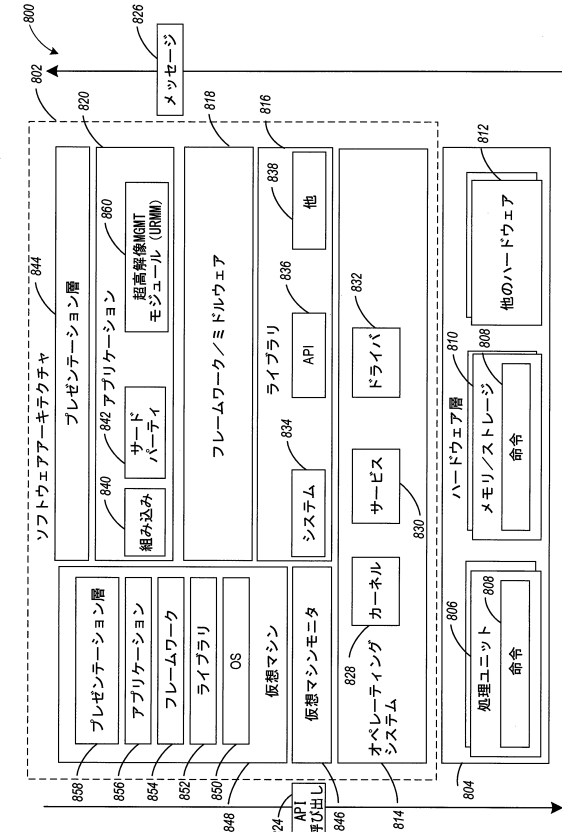
【図7】



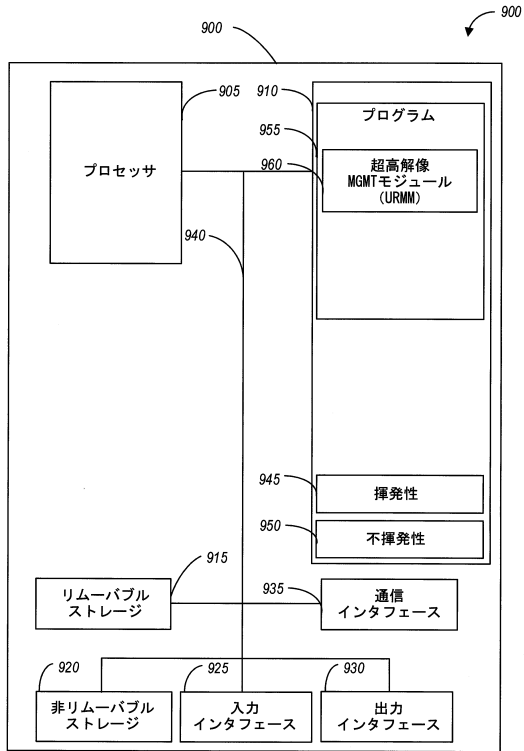
【図6】



【図8】



【図 9】



10

20

30

40

50

## フロントページの続き

- ンドサーティーセブンス・プレイス・ノースイースト・8932  
(72)発明者 シン・リウ  
アメリカ合衆国・ワシントン・98033・カークランド・ノースイースト・シックスティース・ストリート・11208・アパートメント・114
- (72)発明者 ホンユ・スン  
アメリカ合衆国・ワシントン・98074・サマミッシュ・ノースイースト・ナインス・プレイス・21298
- (72)発明者 シャオシン・ジュ  
アメリカ合衆国・ワシントン・98006・ベルビュー・ワンハンドレッドアンドフィフティース・アベニュー・サウスイースト・4412
- (72)発明者 ファン・ジャン  
アメリカ合衆国・ワシントン・98074・サマミッシュ・ノースイースト・フィフティーンス・プレイス・23620
- 審査官 渡部 幸和
- (56)参考文献 特開2019-067078(JP, A)  
米国特許出願公開第2019/0114742(US, A1)  
中国特許出願公開第107358575(CN, A)  
加藤 裕 他, 水平・垂直畳み込みニューラルネットワークを用いた超解像, 電気学会論文誌C, 日本, 一般社団法人電気学会, 2018年07月01日, Vol. 138 No. 7, pp.957-963  
CHEN, Chao et al., Single Image Super-Resolution Using Deep CNN with Dense Skip Connections and Inception-ResNet, 2018 9th International Conference on Information Technology in Medicine and Education (ITME), 米国, 2018年12月27日, pp.999-1003, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8589454>
- (58)調査した分野 (Int.Cl., DB名)  
G06T 3/00