

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7544359号  
(P7544359)

(45)発行日 令和6年9月3日(2024.9.3)

(24)登録日 令和6年8月26日(2024.8.26)

(51)国際特許分類		F I			
G 0 6 T	5/00 (2024.01)	G 0 6 T	5/00		
H 0 4 N	1/387(2006.01)	H 0 4 N	1/387	1 1 0	
G 0 6 T	7/194(2017.01)	G 0 6 T	7/194		
G 0 6 T	7/00 (2017.01)	G 0 6 T	7/00	3 5 0 C	

請求項の数 13 (全38頁)

(21)出願番号	特願2021-83720(P2021-83720)	(73)特許権者	000004226 日本電信電話株式会社 東京都千代田区大手町一丁目5番1号
(22)出願日	令和3年5月18日(2021.5.18)	(73)特許権者	504137912 国立大学法人 東京大学 東京都文京区本郷七丁目3番1号
(65)公開番号	特開2022-177454(P2022-177454 A)	(74)代理人	100121706 弁理士 中尾 直樹
(43)公開日	令和4年12月1日(2022.12.1)	(74)代理人	100128705 弁理士 中村 幸雄
審査請求日	令和5年9月22日(2023.9.22)	(74)代理人	100147773 弁理士 義村 宗洋
		(72)発明者	吹上 大樹 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

最終頁に続く

(54)【発明の名称】 最適化装置、訓練装置、合成装置、それらの方法、およびプログラム

(57)【特許請求の範囲】

【請求項1】

前景マスク画像と、前記前景マスク画像を適用したアルファマップ と、前景画像と、背景画像を用いて、前記アルファマップ を用いてブレンドした画像において前記前景画像の視認性を計算する対象となる画像特徴成分である第一対象成分の知覚コントラストマップと、前記前景マスク画像を用いて不透明にブレンドした画像において前記前景画像の視認性を計算する対象となる画像特徴成分である第二対象成分の知覚コントラストマップを計算する知覚コントラスト計算部と、

前記第二対象成分の知覚コントラストマップを用いて、画像特徴に対する重みベクトルを計算する重み計算部と、

前記第一対象成分の知覚コントラストマップと前記重みベクトルを用いて前記前景画像の視認性を示す視認性マップ $v$ を得る特徴集約部と、

前記視認性マップ $v$ と所望の視認性マップ $^hv$ とから視認性損失を計算し、前記前景画像と前記アルファマップ とから前記アルファマップ が空間的に滑らかに変化するような制約を与える平滑度損失を計算し、前記視認性損失と前記平滑度損失を用いて損失 $L_F$ を計算する損失計算部と、

前記損失 $L_F$ を用いて、前記アルファマップ を更新するアルファ値更新部と、を含む、最適化装置。

【請求項2】

請求項1の最適化装置であって、

前記損失計算部は、前記アルファマップ でブレンドした画像中に含まれている背景以外の前記第一対象成分と、前記前景画像を不透明にブレンドした際の背景以外の前記第二対象成分の相関が高くなる効果を与えるような画像忠実度損失を計算し、前記視認性損失と前記平滑度損失と前記画像忠実度損失を用いて前記損失 $L_F$ を計算する、

最適化装置。

【請求項 3】

請求項 1 または請求項 2 の最適化装置であって、

前記背景画像の視認性を示す背景視認性マップと背景画像の視認性の下限を示す所望の視認性下限マップ $v_B$ とから背景についての背景視認性損失を計算し、前記背景画像と前記アルファマップ とから前記アルファマップ が空間的に滑らかに変化するような制約

10

を与える背景平滑度損失を計算し、前記背景視認性損失と前記背景平滑度損失を用いて損失 $L_B$ を計算する背景損失計算部と、

前記損失 $L_F$ と前記損失 $L_B$ を用いて損失 $L$ を求める損失加算部とを含み、

前記アルファ値更新部は、前記損失 $L$ を用いて、前記アルファマップ を更新する、

最適化装置。

【請求項 4】

請求項 3 の最適化装置であって、

前記知覚コントラスト計算部は、 $M$ 枚の前景マスク画像と、 $M$ 枚の前記前景マスク画像を適用した $M$ 枚のアルファマップ  $(m)$ と、 $M$ 枚の前景画像と、前記背景画像を用いて、前記アルファマップ  $(m)$ を用いてブレンドした画像において $M$ 枚の前記前景画像の視認性を

20

計算する対象となる画像特徴成分である $M$ 個の第一対象成分の知覚コントラストマップと、 $M$ 枚の前記前景マスク画像を用いて不透明にブレンドした画像において $M$ 枚の前記前景画像の視認性を計算する対象となる画像特徴成分である $M$ 個の第二対象成分の知覚コントラストマップを計算し、

前記重み計算部は、 $M$ 個の前記第二対象成分の知覚コントラストマップを用いて、画像特徴に対する $M$ 個の重みベクトルを計算し、

前記特徴集約部は、 $M$ 個の前記第一対象成分の知覚コントラストマップと $M$ 個の前記重みベクトルを用いて $M$ 枚の前記前景画像の視認性を示す $M$ 個の視認性マップ $v$ を得、

前記損失計算部は、 $M$ 個の前記視認性マップ $v$ と $M$ 個の所望の視認性マップ $v_B$ とから $M$ 個の視認性損失を計算し、 $M$ 枚の前記前景画像と $M$ 枚の前記アルファマップ とから $M$ 枚の

30

前記アルファマップ が空間的に滑らかに変化するような制約を与える $M$ 個の平滑度損失を計算し、 $M$ 個の前記視認性損失と $M$ 個の前記平滑度損失を用いて $M$ 個の損失 $L_F$ を計算し、

前記損失加算部は、 $M$ 個の前記損失 $L_F$ と前記損失 $L_B$ を用いて損失 $L$ を求める、

最適化装置。

【請求項 5】

前景画像と背景画像を深層ニューラルネットワークから成る共通のエンコーダによってエンコードしたのち、エンコードした特徴と所望の視認性マップから他の深層ニューラルネットワークから成るデコーダを用いてデコードすることによってアルファマップを生成するアルファ値生成部と、

前景マスク画像と、前記前景マスク画像を適用した前記アルファマップ と、前記前景画像と、前記背景画像を用いて、前記アルファマップ を用いてブレンドした画像において前記前景画像の視認性を計算する対象となる画像特徴成分である第一対象成分の知覚コントラストマップと、前記前景マスク画像を用いて不透明にブレンドした画像において前記前景画像の視認性を計算する対象となる画像特徴成分である第二対象成分の知覚コントラストマップを計算する知覚コントラスト計算部と、

40

前記第二対象成分の知覚コントラストマップを用いて、画像特徴に対する重みベクトルを計算する重み計算部と、

前記第一対象成分の知覚コントラストマップと前記重みベクトルを用いて前記前景画像の視認性を示す視認性マップ $v$ を得る特徴集約部と、

前記視認性マップ $v$ と所望の視認性マップ $v_B$ とから視認性損失を計算し、前記前景画像

50

と前記アルファマップ とから前記アルファマップ が空間的に滑らかに変化するような制約を与える平滑度損失を計算し、前記視認性損失と前記平滑度損失を用いて損失 $L_F$ を計算する損失計算部と、

前記損失 $L_F$ を用いて、前記エンコーダおよび前記デコーダのパラメータを更新するパラメータ更新部と、を含む、

訓練装置。

【請求項 6】

請求項 5 の訓練装置で訓練したパラメータを用いる合成装置であって、

合成対象の前景画像と合成対象の背景画像を、前記パラメータを用いる前記エンコーダによってエンコードしたのち、エンコードした特徴と所望の視認性マップから前記パラメータを用いる前記デコーダを用いてデコードすることによってアルファマップを生成するアルファ値生成部と、

前記合成対象の前景画像と前記合成対象の背景画像を前記アルファマップを用いて合成し、合成画像を生成するブレンド部とを含む、

合成装置。

【請求項 7】

最適化装置が、前景マスク画像と、前記前景マスク画像を適用したアルファマップ と、前景画像と、背景画像を用いて、前記アルファマップ を用いてブレンドした画像において前記前景画像の視認性を計算する対象となる画像特徴成分である第一対象成分の知覚コントラストマップと、前記前景マスク画像を用いて不透明にブレンドした画像において前記前景画像の視認性を計算する対象となる画像特徴成分である第二対象成分の知覚コントラストマップを計算する知覚コントラスト計算ステップと、

前記最適化装置が、前記第二対象成分の知覚コントラストマップを用いて、画像特徴に対する重みベクトルを計算する重み計算ステップと、

前記最適化装置が、前記第一対象成分の知覚コントラストマップと前記重みベクトルを用いて前記前景画像の視認性を示す視認性マップ $v$ を得る特徴集約ステップと、

前記最適化装置が、前記視認性マップ $v$ と所望の視認性マップ $^hv$ とから視認性損失を計算し、前記前景画像と前記アルファマップ とから前記アルファマップ が空間的に滑らかに変化するような制約を与える平滑度損失を計算し、前記視認性損失と前記平滑度損失を用いて損失 $L_F$ を計算する損失計算ステップと、

前記最適化装置が、前記損失 $L_F$ を用いて、前記アルファマップ を更新するアルファ値更新ステップと、を含む、

最適化方法。

【請求項 8】

請求項 7 の最適化方法であって、

前記損失計算ステップは、前記アルファマップ でブレンドした画像中に含まれている背景以外の前記第一対象成分と、前記前景画像を不透明にブレンドした際の背景以外の前記第二対象成分の相関が高くなる効果を与えるような画像忠実度損失を計算し、前記視認性損失と前記平滑度損失と前記画像忠実度損失を用いて前記損失 $L_F$ を計算する、

最適化方法。

【請求項 9】

請求項 7 または請求項 8 の最適化方法であって、

前記最適化装置が、前記背景画像の視認性を示す背景視認性マップと背景画像の視認性の下限を示す所望の視認性下限マップ $^hv_B$ とから背景についての背景視認性損失を計算し、前記背景画像と前記アルファマップ とから前記アルファマップ が空間的に滑らかに変化するような制約を与える背景平滑度損失を計算し、前記背景視認性損失と前記背景平滑度損失を用いて損失 $L_B$ を計算する背景損失計算ステップと、

前記最適化装置が、前記損失 $L_F$ と前記損失 $L_B$ を用いて損失 $L$ を求める損失加算ステップとを含み、

前記アルファ値更新ステップは、前記損失 $L$ を用いて、前記アルファマップ を更新す

10

20

30

40

50

る、

最適化方法。

【請求項 10】

請求項 9 の最適化方法であって、

前記知覚コントラスト計算ステップは、M枚の前景マスク画像と、M枚の前記前景マスク画像を適用したM枚のアルファマップ<sup>(m)</sup>と、M枚の前景画像と、前記背景画像を用いて、前記アルファマップ<sup>(m)</sup>を用いてブレンドした画像においてM枚の前記前景画像の視認性を計算する対象となる画像特徴成分であるM個の第一対象成分の知覚コントラストマップと、M枚の前記前景マスク画像を用いて不透明にブレンドした画像においてM枚の前記前景画像の視認性を計算する対象となる画像特徴成分であるM個の第二対象成分の知覚

10

コントラストマップを用いて、画像特徴に対するM個の重みベクトルを計算し、

前記特徴集約ステップは、M個の前記第一対象成分の知覚コントラストマップとM個の前記重みベクトルを用いてM枚の前記前景画像の視認性を示すM個の視認性マップ $v$ を得、

前記損失計算ステップは、M個の前記視認性マップ $v$ とM個の所望の視認性マップ $^hv$ とからM個の視認性損失を計算し、M枚の前記前景画像とM枚の前記アルファマップ<sup>(m)</sup>とからM枚の前記アルファマップ<sup>(m)</sup>が空間的に滑らかに変化するような制約を与えるM個の平滑度損失を計算し、M個の前記視認性損失とM個の前記平滑度損失を用いてM個の損失 $L_F$ を計算し、

20

前記損失加算ステップは、M個の前記損失 $L_F$ と前記損失 $L_B$ を用いて損失 $L$ を求める、最適化方法。

【請求項 11】

訓練装置が、前景画像と背景画像を深層ニューラルネットワークから成る共通のエンコーダによってエンコードしたのち、エンコードした特徴と所望の視認性マップから他の深層ニューラルネットワークから成るデコーダを用いてデコードすることによってアルファマップを生成するアルファ値生成ステップと、

前記訓練装置が、前景マスク画像と、前記前景マスク画像を適用した前記アルファマップと、前記前景画像と、前記背景画像を用いて、前記アルファマップ<sup>(m)</sup>を用いてブレンドした画像において前記前景画像の視認性を計算する対象となる画像特徴成分である第一対象成分の知覚コントラストマップと、前記前景マスク画像を用いて不透明にブレンドした画像において前記前景画像の視認性を計算する対象となる画像特徴成分である第二対象成分の知覚コントラストマップを計算する知覚コントラスト計算ステップと、

30

前記訓練装置が、前記第二対象成分の知覚コントラストマップを用いて、画像特徴に対する重みベクトルを計算する重み計算ステップと、

前記訓練装置が、前記第一対象成分の知覚コントラストマップと前記重みベクトルを用いて前記前景画像の視認性を示す視認性マップ $v$ を得る特徴集約ステップと、

前記訓練装置が、前記視認性マップ $v$ と所望の視認性マップ $^hv$ とから視認性損失を計算し、前記前景画像と前記アルファマップ<sup>(m)</sup>とから前記アルファマップ<sup>(m)</sup>が空間的に滑らかに変化するような制約を与える平滑度損失を計算し、前記視認性損失と前記平滑度損失を用いて損失 $L_F$ を計算する損失計算ステップと、

40

前記訓練装置が、前記損失 $L_F$ を用いて、前記エンコーダおよび前記デコーダのパラメータを更新するパラメータ更新ステップと、を含む、

訓練方法。

【請求項 12】

請求項 11 の訓練方法で訓練したパラメータを用いる合成方法であって、

合成装置が、合成対象の前景画像と合成対象の背景画像を、前記パラメータを用いる前記エンコーダによってエンコードしたのち、エンコードした特徴と所望の視認性マップから前記パラメータを用いる前記デコーダを用いてデコードすることによってアルファマップを生成するアルファ値生成ステップと、

50

前記合成装置が、前記合成対象の前景画像と前記合成対象の背景画像を前記アルファマップを用いて合成し、合成画像を生成するブレンドステップとを含む、

合成方法。

【請求項 13】

請求項 1 から請求項 4 の何れかの最適化装置、または、請求項 5 の訓練装置、または、請求項 6 の合成装置としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ある画像を別の画像に半透明に合成する場面において視認性をコントロールする技術に関する。

10

【背景技術】

【0002】

視覚情報を別の背景に重畳する際、2つの画像の画素値を単一のパラメータ（アルファ値）で重み付き線形加算する手法（アルファブレンド）がデファクトスタンダードとしてあらゆる場面で用いられている。しかし、このような単純なアルファブレンドを用いた場合、ブレンドした前景画像の視認性が背景画像に大きく依存してしまう。一般にユーザは視認性を元にパラメータを決定するわけであるから、これはユーザにとって不親切かつ非直感的なインターフェースであると言える。

【0003】

20

この問題を解決する有効なアプローチとして、視認性に関わる人間の知覚モデルを利用し、モデルに基づいてブレンディングパラメータを最適化するという方法が考えられる。視認性の背景への依存は、コントラスト知覚のメカニズムによって説明することができる。例えば、よく知られているコントラストマスクングという現象では、同じコントラストのパターンであっても、背景に高いコントラストを持つ別のパターンがあると見えづらくなる。これは、物理的なコントラストが大きくなるほど応答の増分が小さくなるという視覚系の非線形なコントラスト応答が原因であると考えられている。こうした視覚系の振る舞いを再現可能な計算モデルは、ノイズの視認性を予測する画質評価の手法にも組み込まれ活用されてきた。こうした研究の流れを受け、非特許文献 1 は、画質評価のために開発されたモデルを利用し、アルファ値を最適化する手法を提案している。

30

【先行技術文献】

【非特許文献】

【0004】

【文献】Taiki Fukiage, Takeshi Oishi, and Katsushi Ikeuchi, "Visibility-based blending for real-time applications", IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2014.

【発明の概要】

【発明が解決しようとする課題】

【0005】

非特許文献 1 は、ブレンド画像の視認性に対するモデルの予測精度が十分ではなく、またアルファ値の最適化方法にも以下の課題が残されている。

40

【0006】

(1) 非特許文献 1 では、画像に含まれるノイズの視認性を予測するための画質評価手法を、ブレンド画像中に含まれる前景画像成分の視認性を予測するために用いている。この画質評価手法では、ノイズの足された画像と元画像をそれぞれ初期視覚野の計算モデルにより知覚的な画像表現に変換し、この知覚的な画像表現において互いの距離（＝知覚的な距離）を計算することでノイズの視認性を評価する。非特許文献 1 ではこれに倣い、ブレンド画像とブレンド前の背景画像の知覚的な距離を計算することで、ブレンド画像内に含まれる前景画像の視認性を評価する。しかし、アルファブレンドにおいては、ブレンディングによって前景画像が足されるだけでなく、背景画像のコントラストが減衰する効果も

50

生じる。よって、ブレンド画像と背景画像の間の知覚的な距離は、前景画像成分だけでなく背景画像成分の減衰分も含まれていることになる。よって、こうした画質評価モデルを素直にアルファブレンド画像に適用するだけでは、ブレンド画像中の前景画像の視認性を正しく推定することはできない。

【0007】

(2) 上記問題に加えて、一般的な画像中に含まれる画像特徴はバラエティに富むため、ノイズの視認性予測のために開発された画質評価手法は任意の画像の視認性を高精度に予測できない。非特許文献1で用いられたものも含む多くの画質評価手法では、入力画像をまず複数の異なる帯域の空間周波数・色、方位等の特徴に分解し、それぞれの特徴に対する知覚的なコントラストを計算する。その後、これら特徴毎のコントラストから計算された特徴毎の視認性を1つの値に集約することで視認性の予測値とする。この際、特徴毎に重み付けを行ってから集約することもしばしば行われる。しかし、特徴毎の重みは固定されているため、色・空間周波数成分等の画像特徴を多く含む画像ほど視認性が過大に評価されやすく、逆にぼけた画像や滑らかな画像のように特徴が少ない画像ほど視認性が過小評価されやすいといった問題が生じる。

10

【0008】

本発明は、背景成分の減衰の効果が混入することなく正確に前景画像の視認性が推定可能となり、多様な前景画像に対する視認性の予測精度を大幅に改善することができる最適化装置、訓練装置、合成装置、それらの方法、およびプログラムを提供することを目的とする。

20

【課題を解決するための手段】

【0009】

上記の課題を解決するために、本発明の一態様によれば、最適化装置は、前景マスク画像と、前景マスク画像を適用したアルファマップと、前景画像と、背景画像を用いて、アルファマップを用いてブレンドした画像において前景画像の視認性を計算する対象となる画像特徴成分である第一対象成分の知覚コントラストマップと、前景マスク画像を用いて不透明にブレンドした画像において前景画像の視認性を計算する対象となる画像特徴成分である第二対象成分の知覚コントラストマップを計算する知覚コントラスト計算部と、第二対象成分の知覚コントラストマップを用いて、画像特徴に対する重みベクトルを計算する重み計算部と、第一対象成分の知覚コントラストマップと重みベクトルを用いて前景画像の視認性を示す視認性マップ $v$ を得る特徴集約部と、視認性マップ $v$ と所望の視認性マップ $^hv$ とから視認性損失を計算し、前景画像とアルファマップとからアルファマップが空間的に滑らかに変化するような制約を与える平滑度損失を計算し、視認性損失と平滑度損失を用いて損失 $L_F$ を計算する損失計算部と、損失 $L_F$ を用いて、アルファマップを更新するアルファ値更新部と、を含む。

30

上記の課題を解決するために、本発明の他の態様によれば、訓練装置は、前景画像と背景画像を深層ニューラルネットワークから成る共通のエンコーダによってエンコードしたのち、エンコードした特徴と所望の視認性マップから他の深層ニューラルネットワークから成るデコーダを用いてデコードすることによってアルファマップを生成するアルファ値生成部と、前景マスク画像と、前景マスク画像を適用したアルファマップと、前景画像と、背景画像を用いて、アルファマップを用いてブレンドした画像において前景画像の視認性を計算する対象となる画像特徴成分である第一対象成分の知覚コントラストマップと、前景マスク画像を用いて不透明にブレンドした画像において前景画像の視認性を計算する対象となる画像特徴成分である第二対象成分の知覚コントラストマップを計算する知覚コントラスト計算部と、第二対象成分の知覚コントラストマップを用いて、画像特徴に対する重みベクトルを計算する重み計算部と、第一対象成分の知覚コントラストマップと重みベクトルを用いて前景画像の視認性を示す視認性マップ $v$ を得る特徴集約部と、視認性マップ $v$ と所望の視認性マップ $^hv$ とから視認性損失を計算し、前景画像とアルファマップとからアルファマップが空間的に滑らかに変化するような制約を与える平滑度損失を計算し、視認性損失と平滑度損失を用いて損失 $L_F$ を計算する損失計算部と、損失 $L_F$ を用い

40

50

て、エンコーダおよびデコーダのパラメータを更新するアルファ値更新部と、を含む。

【発明の効果】

【0010】

本発明によれば、背景成分の減衰の効果が混入することなく正確に前景画像の視認性が推定可能となり、多様な前景画像に対する視認性の予測精度を大幅に改善することができるという効果を奏する。

【図面の簡単な説明】

【0011】

【図1】第1～第3実施形態に係る合成装置の機能ブロック図。

【図2】第1～第5実施形態に係る合成装置の処理フローの例を示す図。

10

【図3】第1実施形態に係るアルファ値生成部の機能ブロック図。

【図4】第1実施形態に係るアルファ値生成部の処理フローの例を示す図。

【図5】第1実施形態に係る視認性予測部の機能ブロック図。

【図6】第1実施形態に係る視認性予測部の処理フローの例を示す図。

【図7】第1実施形態に係る知覚コントラスト計算部の機能ブロック図。

【図8】第1実施形態に係る知覚コントラスト計算部の処理フローの例を示す図。

【図9】第1実施形態に係る損失計算部の機能ブロック図。

【図10】第1実施形態に係る損失計算部の処理フローの例を示す図。

【図11】所望の視認性マップ、最適化済のアルファマップの例を示す図。

【図12】第2実施形態に係るアルファ値生成部の機能ブロック図。

20

【図13】第2実施形態に係るアルファ値生成部の処理フローの例を示す図。

【図14】第2実施形態に係る視認性予測部の機能ブロック図。

【図15】第2実施形態に係る視認性予測部の処理フローの例を示す図。

【図16】第2実施形態に係る知覚コントラスト計算部の機能ブロック図。

【図17】第2実施形態に係る知覚コントラスト計算部の処理フローの例を示す図。

【図18】第2実施形態に係る損失計算部の機能ブロック図。

【図19】第2実施形態に係る損失計算部の処理フローの例を示す図。

【図20】第3実施形態に係るアルファ値生成部の機能ブロック図。

【図21】第3実施形態に係るアルファ値生成部の処理フローの例を示す図。

【図22】第4実施形態に係る合成装置の機能ブロック図

30

【図23】第4実施形態に係るアルファ値生成部の機能ブロック図。

【図24】第4実施形態に係るアルファ値生成部の処理フローの例を示す図。

【図25】第4実施形態に係る前景視認性予測部の機能ブロック図。

【図26】第4実施形態に係る前景視認性予測部の処理フローの例を示す図。

【図27】第4実施形態に係る知覚コントラスト計算部の機能ブロック図。

【図28】第4実施形態に係る知覚コントラスト計算部の処理フローの例を示す図。

【図29】第5実施形態に係る訓練装置の機能ブロック図。

【図30】第5実施形態に係る訓練装置の処理フローの例を示す図。

【図31】第5実施形態に係るアルファ値生成部の機能ブロック図。

【図32】第5実施形態に係るアルファ値生成部の処理フローの例を示す図。

40

【図33】第5実施形態に係る合成装置の機能ブロック図。

【図34】本手法を適用するコンピュータの構成例を示す図。

【発明を実施するための形態】

【0012】

以下、本発明の実施形態について、説明する。なお、以下の説明に用いる図面では、同じ機能を持つ構成部や同じ処理を行うステップには同一の符号を記し、重複説明を省略する。以下の説明において、テキスト中で使用する記号「^」「\_」等は、本来直後の文字の真上に記載されるべきものであるが、テキスト記法の制限により、当該文字の直前に記載する。式中においてはこれらの記号は本来の位置に記述している。また、ベクトルや行列の各要素単位で行われる処理は、特に断りが無い限り、そのベクトルやその行列の全ての

50

要素に対して適用されるものとする。

【 0 0 1 3 】

< 第 1 実施形態 >

図 1 は第 1 実施形態に係る合成装置 1 0 0 の機能ブロック図を、図 2 はその処理フローを示す。

【 0 0 1 4 】

第 1 実施形態に係る合成装置 1 0 0 は、アルファ値生成部 1 1 0 とブレンド部 1 2 0 とを含む。

【 0 0 1 5 】

合成装置 1 0 0 は、前景画像  $I^fg$   $R^3 \times H \times W$  と背景画像  $I^bg$   $R^3 \times H \times W$ 、前景マスク画像  $M$   $R^H \times W$ 、アルファマップの初期値  $i^n$   $R^H \times W$ 、所望の視認性マップ  $\wedge v$   $R^H \times W$  を入力として、前景画像  $I^fg$  と背景画像  $I^bg$  とを合成し、合成画像  $I^{bl}$   $R^3 \times H \times W$  を出力する。ここで、 $R$  は実数全体の集合、 $H$ 、 $W$  はそれぞれ画像の高さと幅を表す。各画像の各画素値は、3 個の色成分からなる。なお、所望の視認性マップは、前景画像の所望の視認性を示すものである。

【 0 0 1 6 】

アルファ値生成部 1 1 0 は、アルファマップの初期値  $i^n$  と前景画像  $I^fg$  と背景画像  $I^bg$  と前景マスク画像  $M$  と所望の視認性マップ  $\wedge v$  を入力として、アルファ値  $out_i$   $[0,1]$  を生成し (S 1 1 0)、アルファマップ  $out$   $R^H \times W$  を出力する。 $i$  は画像の各画素のインデックスを表し、 $out_i$  はアルファマップ  $out$  の  $i$  番目の画素のアルファ値を表す。

【 0 0 1 7 】

ブレンド部 1 2 0 は、前景画像  $I^fg$  と背景画像  $I^bg$  とアルファ値生成部 1 1 0 から出力されるアルファマップ  $out$  を入力として、前景画像  $I^fg$  と背景画像  $I^bg$  を次式によりブレンドし (S 1 2 0)、合成画像  $I^{bl}$   $R^3 \times H \times W$  を出力する。

【 0 0 1 8 】

$$I^{bl}_i = out_i I^{fg}_i + (1 - out_i) I^{bg}_i \quad (1)$$

ただし、 $I^{bl}_i$  は合成画像  $I^{bl}$  の  $i$  番目の画素値、 $I^{fg}_i$  は前景画像  $I^{fg}$  の  $i$  番目の画素値、 $I^{bg}_i$  は背景画像  $I^{bg}$  の  $i$  番目の画素値をそれぞれ表す。

【 0 0 1 9 】

合成装置および後述する、最適化装置、訓練装置は、例えば、中央演算処理装置 (CPU: Central Processing Unit)、主記憶装置 (RAM: Random Access Memory) などを有する公知又は専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。合成装置、最適化装置および訓練装置は、例えば、中央演算処理装置の制御のもとで各処理を実行する。合成装置、最適化装置および訓練装置に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて中央演算処理装置へ読み出されて他の処理に利用される。合成装置、最適化装置および訓練装置の各処理部は、少なくとも一部が集積回路等のハードウェアによって構成されていてもよい。合成装置、最適化装置および訓練装置が備える各記憶部は、例えば、RAM (Random Access Memory) などの主記憶装置、またはリレーショナルデータベースやキーバリューストアなどのミドルウェアにより構成することができる。ただし、各記憶部は、必ずしも合成装置、最適化装置および訓練装置がその内部に備える必要はなく、ハードディスクや光ディスクもしくはフラッシュメモリ (Flash Memory) のような半導体メモリ素子により構成される補助記憶装置により構成し、合成装置、最適化装置および訓練装置の外部に備える構成としてもよい。

【 0 0 2 0 】

以下、アルファ値生成部 1 1 0 について説明する。

【 0 0 2 1 】

< アルファ値生成部 1 1 0 >

図 3 はアルファ値生成部 1 1 0 の機能ブロック図を、図 4 はその処理フローの例を示す。

【 0 0 2 2 】

10

20

30

40

50

アルファ値生成部 1 1 0 は、視認性予測部 1 1 1、損失計算部 1 1 3、アルファ値更新部 1 1 5 を含む。

【 0 0 2 3 】

アルファ値生成部 1 1 0 は、前景画像  $I^f$ 、背景画像  $I^b$ 、前景マスク画像  $M$ 、アルファマップの初期値  $\alpha^{in}$ 、所望の視認性マップ  $v$  を入力として、最適化済のアルファマップ  $\alpha^{out}$   $R^H \times W$  を生成し、出力とする。前景マスク画像  $M$  は、全画像領域中で前景画像  $I^f$  が存在する画素で 1、それ以外で 0 をとるマスク画像を表す。

【 0 0 2 4 】

処理の順序としては、まず、視認性予測部 1 1 1 が、前景画像  $I^f$ 、背景画像  $I^b$ 、前景マスク画像  $M$ 、アルファマップの初期値  $\alpha^{in}$  を入力として、視認性マップ  $v$   $R^H \times W$  を求め ( S 1 1 1 )、出力する。視認性マップ  $v$  は、前景画像の視認性を示すものである。

10

【 0 0 2 5 】

次に、損失計算部 1 1 3 が、視認性マップ  $v$  と前景画像  $I^f$ 、アルファマップ  $\alpha^{in}$   $R^H \times W$  を入力として、損失  $L$  を計算し ( S 1 1 3 )、出力する。

【 0 0 2 6 】

次いで、アルファ値更新部 1 1 5 が、損失  $L$  を入力として、損失  $L$  を元にアルファマップを更新し ( S 1 1 5 )、更新したアルファマップ  $\alpha^{out}$  を出力する。

【 0 0 2 7 】

その後は、初期値  $\alpha^{in}$  の代わりに更新したアルファマップ  $\alpha^{out}$  を入力として視認性予測部 1 1 1 が視認性マップ  $v$  を出力する。このループを繰り返し、一定条件を満たしたら ( S 1 1 5 - 2 の yes )、アルファ値更新部 1 1 5 が最適化済のアルファマップ  $\alpha^{out}$  を出力して処理を終える。ループの制御はアルファ値更新部 1 1 5 に含まれる。

20

【 0 0 2 8 】

< 視認性予測部 1 1 1 >

図 5 は視認性予測部 1 1 1 の機能ブロック図を、図 6 はその処理フローの例を示す。

【 0 0 2 9 】

視認性予測部 1 1 1 は、マスク適用部 1 1 1 1、知覚コントラスト計算部 1 1 1 3、重み計算部 1 1 1 5、特徴集約部 1 1 1 7 を含む。

【 0 0 3 0 】

入力：前景画像  $I^f$ 、背景画像  $I^b$ 、前景マスク画像  $M$ 、アルファマップ  $\alpha^{in}$  または初期値  $\alpha^{in}$  (以下、単に「アルファマップ  $\alpha^{in}$ 」ともいう)

30

出力：視認性マップ  $v$

処理：

マスク適用部 1 1 1 1 は、アルファマップ  $\alpha^{in}$  と前景マスク画像  $M$  とを入力として、アルファマップ  $\alpha^{in}$  と前景マスク画像  $M$  とを画素毎に乗算することで、アルファマップ  $\alpha^{in}$  にマスクを適用し ( S 1 1 1 1 )、マスク適用済のアルファマップ  $\alpha'$  を出力する

知覚コントラスト計算部 1 1 1 3 は、マスク適用済のアルファマップ  $\alpha'$ 、前景マスク画像  $M$ 、前景画像  $I^f$ 、背景画像  $I^b$  を入力として、第一対象成分  $I^f$  の知覚コントラストマップ  $P$   $R^{3N \times H \times W}$ 、第二対象成分  $I^b$  の知覚コントラストマップ  $P'$   $R^{3N \times H \times W}$  を計算し ( S 1 1 1 3 )、出力する。ここで、第一対象成分  $I^f$  は入力されたマスク適用済のアルファマップ  $\alpha'$  を用いてブレンドした画像において、視認性を計算する対象となる画像特徴成分 (前景画像成分とマスク適用済のアルファマップ  $\alpha'$  の構造を合わせたもの) を表し、第二対象成分  $I^b$  は前景マスク画像  $M$  を用いて不透明にブレンドした画像において、視認性を計算する対象となる画像特徴成分を表す。画像特徴成分は画像を  $N$  個の空間周波数帯域と 3 個の色成分に分解して得られる  $3 \times N$  次元の特徴として表現される。 $N$  は入力画像サイズによって上限が決まり、画像サイズが大きいほど  $N$  を大きく設定できる。 $N$  が大きいほど画像特徴数が増加するため、一般にモデルの予測力も向上すると考えられる。本実施形態ではモデルパラメータのフィッティングのために実施した実験で用いた刺激サイズに合わせ  $N=6$  とした。また、画像特徴として空間周波数と色に加えて方位成分を考慮しても構わない。計算コストや精度を考慮して画像特徴を設定すればよい。知覚コントラスト計

40

50

算部 1 1 1 3 における処理の詳細は後述する。

【 0 0 3 1 】

重み計算部 1 1 1 5 は、第二対象成分  $I^{fg}$  の知覚コントラストマップ  $P'$  を入力として、各画素の画像特徴に対する重みベクトル  $w \in R^{3N}$  を計算し ( S 1 1 1 5 )、出力する。重み  $w$  は、第二対象成分  $I^{fg}$  の知覚コントラストマップ  $P'$  のある画素の特徴ベクトル  $r' \in R^{3N}$  に特徴間のインタラクションを表現する行列  $X \in R^{3N \times 3N}$  を掛け合わせ、バイアス  $b \in R^{3N}$  を加算したものを sigmoid 関数に通すことで得られる。つまり、重み  $w$  は、次式により、得られる。

【 0 0 3 2 】

$$w = \text{sigmoid}(Xr' + b) \quad (2)$$

ここで、 $X$  および  $b$  は、後述する < 視認性予測部 1 1 1 のパラメータ > の手順により、データへのフィッティングによって求められる。

【 0 0 3 3 】

特徴集約部 1 1 1 7 は、第一対象成分  $I^{tg}$  の知覚コントラストマップ  $P$  と重み計算部 1 1 1 5 で得られた各画素の重みベクトル  $w \in R^{3N}$  とを入力として、重みベクトル  $w$  を用いて、第一対象成分  $I^{tg}$  の各画素における知覚コントラストマップ  $P$  のある画素の特徴ベクトル  $r \in R^{3N}$  を重み付けしつつ特徴集約を行い、次式により、各画素の視認性  $v_i$  を得る ( S 1 1 1 7 )。

【 数 1 】

$$v_i = \left( \sum_{n=1}^{3N} w_n r_n^p \right)^{\frac{1}{p}} \quad (3)$$

ただし、 $p$  はモデルキャリブレーションで最適化されたパラメータであり、 $w_n$  および  $r_n^p$  はそれぞれ重みベクトル  $w$  および特徴ベクトル  $r$  の  $n$  番目の成分を示す。

【 0 0 3 4 】

最後に、特徴集約部 1 1 1 7 は、各画素の視認性を表す視認性マップ  $v \in R^H \times W$  を出力する。

【 0 0 3 5 】

次に、知覚コントラスト計算部 1 1 1 3 の詳細について説明する。

【 0 0 3 6 】

< 知覚コントラスト計算部 1 1 1 3 >

図 7 は知覚コントラスト計算部 1 1 1 3 の機能ブロック図を、図 8 はその処理フローの例を示す。

【 0 0 3 7 】

知覚コントラスト計算部 1 1 1 3 は、色変換部 1 1 1 3 A、アルファブレンド部 1 1 1 3 B、対象成分抽出部 1 1 1 3 C、バンド分解部 1 1 1 3 D、バンドリミテッドコントラスト計算部 1 1 1 3 E、第一正規化部 1 1 1 3 F、第二正規化部 1 1 1 3 G、連結部 1 1 1 3 H を含む。

【 0 0 3 8 】

知覚コントラスト計算部 1 1 1 3 は、マスク適用済のアルファマップ  $\alpha$ 、前景マスク画像  $M$ 、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$  を入力として、第一対象成分  $I^{tg}$  の知覚コントラストマップ  $P$ 、第二対象成分  $I^{fg}$  の知覚コントラストマップ  $P'$  を計算し ( S 1 1 1 3 )、出力する。しかし、実際には ( マスク適用済のアルファマップ  $\alpha$ 、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$  ) を入力として第一対象成分  $I^{tg}$  の知覚コントラストマップ  $P$  を計算する処理と ( 前景マスク画像  $M$ 、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$  ) を入力として第二対象成分  $I^{fg}$  の知覚コントラストマップ  $P'$  を計算する処理に分けることができ、それぞれの処理は入力が異なるだけで中身は完全に同一である。このため、ここではマスク適用済のアルファマップ  $\alpha$  と前景マス

10

20

30

40

50

ク画像Mを区別せず「アルファマップ」と呼び、また第一対象成分  $I^{tg}$  と第二対象成分  $I^{fg}$  を区別せず単に「対象成分」と呼ぶこととし、一方の処理についてのみ記述する。

【0039】

入力：アルファマップ、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$

出力：対象成分の知覚コントラストマップ  $P$ 、 $P'$

処理：

まず、色変換部 1113A は、前景画像  $I^{fg}$  と背景画像  $I^{bg}$  を入力として、RGB色空間をYUV色空間に変換し (S1113A)、出力する。

【0040】

対象成分抽出部 1113C は、色変換後の前景画像  $I^{fg}$  と背景画像  $I^{bg}$ 、およびアルファマップを入力として、対象成分  $I^{tg}$   $R^{3(N-1) \times H \times W}$  を抽出し (S1113C)、出力する。対象成分抽出部 1113Cの詳細は後述する。

10

【0041】

アルファブレンド部 1113B は、色変換後の前景画像  $I^{fg}$  と背景画像  $I^{bg}$ 、およびアルファマップを入力として、式(1)にしたがってアルファブレンド画像  $I^{bl}$   $R^{3 \times H \times W}$  を得 (S1113B)、出力する。

【0042】

$$I^{bl}_i = i I^{fg}_i + (1 - i) I^{bg}_i \quad (1)$$

バンド分解部 1113D は、アルファブレンド画像  $I^{bl}$  と色変換後の背景画像  $I^{bg}$  を入力として、アルファブレンド画像  $I^{bl}$  のバンドパス成分  $I^{bl}$ 、アルファブレンド画像  $I^{bl}$  のガウシアンピラミッド画像  $I^{bl}$ 、アルファブレンド画像  $I^{bl}$  のローパス残余成分  $L^{bl}$ 、背景画像  $I^{bg}$  のローパス残余成分  $L^{bg}$  を得 (S1113D)、出力する。

20

【0043】

アルファブレンド画像  $I^{bl}$  のバンドパス分解には、ラプラシアンピラミッドを用い、N-1個の空間周波数帯と1個のローパス残余成分に分解する。

【0044】

アルファブレンド画像  $I^{bl}$  のガウシアンピラミッド画像  $I^{bl}$  については、N+1個分の異なる解像度 (スケール) の画像を生成する。なお、ガウシアンピラミッドのスケール数をN+1としたのは、後述のバンドリミテッドコントラスト計算時に、N-1個それぞれのバンドパス画像を、各周波数帯の2つ下のスケールのガウシアンピラミッド画像で割る処理を行うためである。アルファブレンド画像  $I^{bl}$  および背景画像  $I^{bg}$  のローパス残余成分  $L^{bl}$  および  $L^{bg}$  は、それぞれのラプラシアンピラミッドの最も粗いスケールのローパス残余画像に相当する。アルファブレンド画像  $I^{bl}$  のバンドパス成分  $I^{bl}$  とガウシアンピラミッド画像  $I^{bl}$  はバンドリミテッドコントラスト計算部 1113Eへの入力として用いられ、アルファブレンド画像  $I^{bl}$  および背景画像  $I^{bg}$  のローパス残余成分  $L^{bl}$  および  $L^{bg}$  は第二正規化部 1113Gへの入力として用いられる。

30

【0045】

バンドリミテッドコントラスト計算部 1113E は、対象成分  $I^{tg}$  とアルファブレンド画像  $I^{bl}$  のバンドパス成分  $I^{bl}$ 、アルファブレンド画像  $I^{bl}$  のガウシアンピラミッド画像  $I^{bl}$  を入力として、次式により、対象成分  $I^{tg}$  とアルファブレンド画像  $I^{bl}$  のバンドパス成分  $I^{bl}$  それぞれについてバンドリミテッドコントラスト  $C^{tg}$ 、 $C^{bl}$  を計算し (S1113E)、第一正規化部 1113Fに向けて出力する。

40

【数2】

$$C_{k,i}^{tg} = \frac{\Delta I_{k,i}^{tg}}{\sigma + \sqrt{I_{k,i}^{bl}}}, \quad C_{k,i}^{bl} = \frac{\Delta I_{k,i}^{bl}}{\sigma + \sqrt{I_{k,i}^{bl}}} \quad (4)$$

ここで、 $I_{k,i}^{tg}$  および  $I_{k,i}^{bl}$  はそれぞれ対象成分  $I^{tg}$  およびバンドパス成分  $I^{bl}$  のk番目のスケールのi番目の画素であり、 $I_{k,i}^{bl}$  はガウシアンピラミッド画像  $I^{bl}$  のk番目のスケ

50

ールの*i*番目の画素である。は、ゼロによる除算を避けるために必要なスカラー定数であり、例えば、 $\beta = 0.1$ である。式(4)の計算は、色チャンネルごとに独立して行われる。式(4)では、*N*-1個のバンドパス成分画像それぞれについて、2つ下のスケールのガウシアンピラミッド画像を同一スケールにアップサンプルした画像で割る処理を行っている。これは、人間が知覚するコントラストが、付近の平均輝度が小さいほど大きく、大きいほど小さく感じられるという特性を再現する処理に相当する。

【0046】

第一正規化部1113Fは、対象成分 $I^{tg}$ とアルファブレンド画像 $I^{bl}$ のバンドパス成分 $I^{bl}$ のバンドリミテッドコントラスト $C^{tg}$ 、 $C^{bl}$ を入力として、次式により、対象成分 $I^{tg}$ に対する知覚コントラスト $P^{band}$ を計算し(S1113F)、連結部1113Hに向けて出力する。

10

【数3】

$$P_{c,k,i}^{band} = \frac{|C_{c,k,i}^{tg}|}{\left(\beta_{c,k}^2 + \frac{1}{|N|} \sum_{j \in N(i)} |C_{c,k,j}^{bl}|\right)^{\gamma}} \quad (5)$$

ここで、 $P_{c,k,i}^{band}$ は、色チャンネル $c$  ( $c \in \{Y, U, V\}$ )の*k*番目のスケールの*i*番目の画素における対象成分 $I^{tg}$ の正規化されたコントラスト応答を表す。*N*(*i*)は、*i*番目のピクセルに隣接するピクセルのグループを表す。例えば、周辺の5×5ピクセルを近隣の定義として使用した場合、 $|N| = 25$ である。 $\beta_{c,k}$ は応答の飽和が始まるポイントを定義するパラメータであり、 $\gamma$ は非線形応答関数の傾きを変調するパラメータである。 $\beta_{c,k}$ の画像特徴間(色チャンネルおよびスケール間)の相対的な大きさは、各画像特徴のバンドリミテッドコントラストの平均二乗画素値の比と同じにすることで正規化する。 $\beta_{c,k}$ を得るために、各画像特徴のバンドリミテッドコントラストの平均二乗値を校正画像パッチ間で平均化し、パラメータ $\beta_{c,k}$ でスケールリングする。 $\gamma$ は、モデルキャリブレーションで最適化されるパラメータである。

20

【0047】

第二正規化部1113Gは、アルファブレンド画像 $I^{bl}$ および背景画像 $I^{bg}$ のローパス残余成分 $L^{bl}$ および $L^{bg}$ を入力として、次式により、前景画像 $I^{fg}$ のローパス成分に対する知覚コントラスト $P^{low}$ を計算し(S1113G)、連結部1113Hに向けて出力する。

30

【数4】

$$P_{c,N,i}^{low} = \frac{|L_{c,i}^{bl} - L_{c,i}^{bg}|}{\left(\hat{\beta}_c^2 + |L_{c,i}^{bl}|^2\right)^{\hat{\gamma}}} \quad (6)$$

ここで、 $\hat{\beta}_c$ 、 $\hat{\gamma}_c$ は式(5)の $\beta_{c,k}$ 、 $\gamma$ と同じ役割を持つパラメータであるが、ローパス成分はバンドパス成分とは質的に異なる応答関数を持ちうるため、異なる値を取りうる。 $\hat{\beta}_c$ 、 $\hat{\gamma}_c$ と同様に、色チャンネル間の相対的な大きさを、ローパス残差成分の各色チャンネルの平均二乗画素値の比と同じにすることで $\hat{\beta}_c$ を正規化する。 $\hat{\beta}_c$ を得るために、各色チャンネルの平均二乗値を校正画像パッチ間で平均化し、パラメータ $\hat{\beta}_c$ でスケールリングする。 $\hat{\gamma}_c$ 、 $\hat{\gamma}$ は、モデルキャリブレーションで最適化されるパラメータである。

40

【0048】

連結部1113Hは、対象成分 $I^{tg}$ に対する知覚コントラスト $P^{band}$ と前景画像 $I^{fg}$ のローパス成分に対する知覚コントラスト $P^{low}$ を入力として、これらを合わせて対象成分 $I^{tg}$ に対する知覚コントラストのラプラシアンピラミッド(スケール数=*N*個)を構成する。その後、ラプラシアンピラミッドの各スケールの解像度を元画像の解像度と一致するようにアップスケールリングを行い、1つのテンソル $P \in \mathbb{R}^{3N \times H \times W}$ に連結し(S1113H)、対象成分 $I^{tg}$ の知覚コントラストとして知覚コントラスト計算部1113から出力する。

50

$P=[r_1 | g_1 | \dots | H \times W]$   $R^{3N \times H \times W}$ であり、各ピラミッドレベル $k \in \{1, 2, \dots, N\}$ の各色チャンネル $c \in \{Y, U, V\}$ の応答値が連結され、1つのベクトル $r \in R^{3N}$ となる。

【0049】

<対象成分抽出部 1 1 1 3 C >

対象成分抽出部 1 1 1 3 Cでは、視認性を計算する対象となる画像特徴成分を抽出する。ブレンド画像中の前景画像 $I^f$ の視認性を計算する場合、アルファ値が空間的に一様であれば前景画像 $I^f$ とアルファマップ  $A$  を掛け合わせたものを対象成分  $I^t$ とすればよい。しかし、アルファ値が空間的に変化する条件においては、アルファマップ  $A$  の構造によって生じるコントラストも対象成分  $I^t$ に含める必要がある。例えば、アルファマップ  $A$  が0から1に変化するエッジ付近では、前景画像 $I^f$ の色と背景画像 $I^b$ の色によってコントラストが定義されるので、前景画像 $I^f$ だけでなく背景画像 $I^b$ の色も考慮しつつ対象成分  $I^t$ を抽出する必要がある。ただし、背景画像 $I^b$ 中に含まれる構造は対象成分  $I^t$ に含まれてはならないという制約があるため、前景画像 $I^f$ とローパスフィルタリングした背景画像 $I^b$ をブレンドしてコントラストを計算する。しかし、この方法は、フィルタリングしすぎると背景画像 $I^b$ の局所的な明るさの違いを考慮できず、フィルタリングが足りなければ背景の構造が対象成分  $I^t$ に混入してしまうというトレードオフの問題に直面する。この問題を解決するため、本実施形態では、ラプラシアンピラミッドによるバンドパス分解を利用する。ラプラシアンピラミッドでは、スケール毎に含まれる空間周波数の帯域に限られる。このため、各スケールにおいて、そこに含まれる空間周波数以上の帯域をカットするように背景画像 $I^b$ のローパスフィルタリングを行うことで、背景画像 $I^b$ の構造が対象成分  $I^t$ に含まれないようにしつつ、背景画像 $I^b$ の局所的な色の変動を最大限考慮することが可能となる。具体的な処理は以下に記載する。

10

20

【0050】

入力：アルファマップ  $A$ 、色変換後の前景画像 $I^f$ 、色変換後の背景画像 $I^b$

出力：対象成分  $I^t \in R^{3(N-1) \times H \times W}$

処理：

まず、対象成分抽出部 1 1 1 3 Cは、色変換後の背景画像 $I^b$ からガウシアンピラミッド ( $N+2$ スケール)を生成する。 $N+2$ としたのは、後の処理において、 $k$ 番目のバンドの対象成分を抽出する際、 $k+3$ スケール目のガウシアンピラミッド画像を使用することになるため、最大で( $N-1$ 番目のバンドの対象成分抽出のために) $N+2$ 番目のスケールのガウシアンピラミッド画像が必要となるためである。

30

【0051】

以降は、対象成分のラプラシアンピラミッドの各スケール( $k \in \{1, 2, \dots, N-1\}$ )を抽出する処理となる。冗長性を省くため、全部で $N-1$ 回繰り返す処理のうち、任意の $k$ 番目の処理についてのみ記述する。また、ラプラシアンピラミッドの最も粗いスケール( $k=N$ )に相当するローパス残余成分については、後の処理で使用しないため生成しない。

【0052】

対象成分抽出部 1 1 1 3 Cは、色変換後の背景画像 $I^b$ のガウシアンピラミッドから、現在のスケール( $k$ )より3つ粗いスケール( $k+3$ )の画像を取り出し、元画像の解像度までアップスケーリングを行う。

40

【0053】

対象成分抽出部 1 1 1 3 Cは、アップスケーリングした色変換後の背景画像 $I^b$ と、色変換後の前景画像 $I^f$ とをアルファマップ  $A$  を用いてブレンドする。なお、本実施形態で用いたRGB色空間からYUV色空間への変換は線形変換のため、アルファブレンド後に色変換を行うのと同じ結果が得られる。

【0054】

対象成分抽出部 1 1 1 3 Cは、ブレンドした画像からラプラシアンピラミッドを生成し、そのうちの $k$ 番目のスケールを対象成分の $k$ 番目のバンドパス成分として出力する。なお、 $k$ 番目以外のスケールは使用しないため、これ以外のスケールのバンドパス成分の生成は省略してもよい。

50

【 0 0 5 5 】

< 視認性予測部 1 1 1 のパラメータ >

視認性予測部 1 1 1 で用いた知覚モデルは複数のパラメータ  $\{\alpha, \beta, \gamma, \delta, \rho, \lambda, b\}$  をもつため、これらのパラメータを事前にキャリブレーションし、決定しておかねばならない。このために、本実施形態では視認性マッチング課題を用いて様々な画像に対する視認性評価データを集め、データを最もよく説明できるように最尤推定によりパラメータをフィッティングする。視認性マッチング課題では、参照画像として提示したアルファブレンド画像の視認性とマッチするように、別の画像パッチ同士のブレンド画像（テスト画像）のアルファ値を調整する。モデルのパラメータ  $\theta$  が与えられたとき、参照画像に使用した{前景画像  $I^f$ 、背景画像  $I^b$ 、前景マスク画像  $M$ 、アルファ値}の組  $R_x$  を入力として得られる視認性予測値を  $v(R_x | \theta)$ 、その参照画像と合わせて提示されたテスト画像の{前景画像  $I^f$ 、背景画像  $I^b$ 、前景マスク画像  $M$ 、アルファ値（観察者の回答）}の組  $T_x$  を入力として得られる視認性予測値を  $v(T_x | \theta)$  とすると、観察者があるアルファ値  $\alpha_x$  を回答として選択する確率は次式のようにモデル化できる。

10

【数 5】

$$P(\alpha_x | R_x, T_x, \theta) = \frac{\exp(-s|v(R_x|\theta) - v(T_x(\alpha_x)|\theta)|^2)}{\sum_{\alpha'} \exp(-s|v(R_x|\theta) - v(T_x(\alpha')|\theta)|^2)} \quad (7)$$

ここで、 $T_x(\alpha_x)$  は、アルファ値  $\alpha_x$  でブレンドされたテスト画像である。モデルパラメータ  $\theta$  を用いて  $s$  を最適化する。 $\alpha'$  は 0.01 間隔で離散的にサンプリングした。  $x$  は実験条件を表すインデックスとする。  $s$  は回答の不確かさ（ばらつき）を表現するスケール変数で、モデルパラメータ  $\theta$  と一緒にフィッティングする。ここで、視認性予測値  $v$   $\{v(R_x | \theta), v(T_x | \theta)\}$  は単一のスカラー値として与えられる必要があるため、以下の式でモデルが出力した視認性マップ  $v_i$  を空間的に集約する。

20

【数 6】

$$\bar{v} = \left( \frac{1}{|N^{tg}|} \sum_{i \in N^{tg}} v_i^q \right)^{\frac{1}{q}} \quad (8)$$

30

ここで  $N^{tg}$  は視認性を計算すべき対象成分が存在する画素の集合を表し、  $|N^g|$  は  $N^g$  に含まれる画素数を表す。パラメータ  $q$  はモデルパラメータ  $\theta$  と一緒にフィッティングする。ここから、全ての実験条件  $x$  について得られた回答値から negative log likelihood を次式のように計算できる。

【数 7】

$$L(x|\theta) = - \sum_{x \in \mathcal{X}} P(\alpha_x | R_x, T_x, \theta) \quad (9)$$

40

パラメータのフィッティングは、式(9)を最小化することで得られる。最小化は Adam optimizer などの確率的勾配降下法で行う。

【 0 0 5 6 】

< 損失計算部 1 1 3 >

図 9 は損失計算部 1 1 3 の機能ブロック図を、図 1 0 はその処理フローの例を示す。

【 0 0 5 7 】

損失計算部 1 1 3 は、視認性損失計算部 1 1 3 1、平滑度損失計算部 1 1 3 3、加算部 1 1 3 5 を含む。

【 0 0 5 8 】

損失計算部 1 1 3 は、所望の視認性マップ  $\hat{v}$  と視認性マップ  $v$  と前景画像  $I^f$ 、アルファ

50

マップ を入力として、損失Lを計算し ( S 2 1 3 )、出力する。

【 0 0 5 9 】

視認性損失計算部 1 1 3 1 は、現在のアルファマップ から予測された視認性マップ $v$ と  
 所望の視認性マップ $\hat{v}$ とを入力とし、視認性マップ $v$ と $\hat{v}$ との距離を計算し、損失 $L_v$ とし  
 、出力する。この視認性損失 $L_v$ だけでは前景画像 $I^{fg}$ の元の構造が保たれる保証がないため  
 、平滑度損失 $L_s$ でアルファマップ が空間的に滑らかに変化するような制約を与える。加  
 算部 1 1 3 5 はこれらの損失の重み付き和として最終的な損失Lを計算する。

【 0 0 6 0 】

< 視認性損失計算部 1 1 3 1 >

入力：所望の視認性マップ $\hat{v}$   $R^H \times W$ 、視認性マップ $v$   $R^H \times W$

10

出力：視認性損失 $L_v$

処理：視認性損失計算部 1 1 3 1 は、次式によって視認性損失 $L_v$ を計算する ( S 1 1 3 1 )。

【 数 8 】

$$L_v = \frac{1}{|N^{tg}|} \sum_{i \in N^{tg}} (\hat{v}_i - v_i)^2 \quad (10)$$

< 平滑度損失計算部 1 1 3 3 >

エッジ保存型平滑化の効果を与えるため、前景画像 $I^{fg}$ の勾配の高いところではアルファ  
 マップ が空間的に大きく変化することを許容しつつ、前景画像 $I^{fg}$ の勾配の低いところでは  
 アルファマップ の空間的な変化を抑制するような損失関数を用いる。前景画像 $I^{fg}$ の勾  
 配の計算には、輝度情報のみを用いる。

20

【 0 0 6 1 】

入力：前景画像 $I^{fg}$ 、アルファマップ

出力：平滑度損失 $L_s$

処理：

まず、平滑度損失計算部 1 1 3 3 は、前景画像 $I^{fg}$ をYUV色空間に変換し、輝度成分である  
 Yチャンネルだけを取り出す。例えば、平滑度損失計算部 1 1 3 3 は、色変換部 1 1 1  
 3 Aと同様の処理を行い、前景画像 $I^{fg}$ をYUV色空間に変換してもよいし、色変換部 1 1 1  
 3 Aの出力値を利用してよい。

30

【 0 0 6 2 】

次に、平滑度損失計算部 1 1 3 3 は、次式により平滑度損失 $L_s$ を計算する ( S 1 1 3 3 )。

【 数 9 】

$$L_s = \frac{1}{HW} \sum_{x=1}^W \sum_{y=1}^H \left( \left| \frac{\partial}{\partial x} \alpha_{x,y} \right| e^{-\left| \frac{\partial}{\partial x} I^{fg}_{x,y} \right|} + \left| \frac{\partial}{\partial y} \alpha_{x,y} \right| e^{-\left| \frac{\partial}{\partial y} I^{fg}_{x,y} \right|} \right) \quad (11)$$

40

ここで、 $\alpha_{x,y}$ と $I^{fg}_{x,y}$ は、それぞれ、 $y$ 番目の行の $x$ 番目の列にあるアルファマップと前景  
 画像(輝度成分であるYチャンネル)のピクセル値を表す。空間微分にはソーベルフィルタの  
 畳み込みを用いたが、これ以外の微分フィルタを用いてもよい。また、エッジ保存型平滑  
 化ではない通常の平滑化を用いても良い。この場合、損失関数は前景画像 $I^{fg}$ を用いない以  
 下の式で定義される。

【 数 1 0 】

$$L_s = \frac{1}{HW} \sum_{x=1}^W \sum_{y=1}^H \left( \left| \frac{\partial}{\partial x} \alpha_{x,y} \right| + \left| \frac{\partial}{\partial y} \alpha_{x,y} \right| \right) \quad (12)$$

50

<加算部 1 1 3 5 >

入力：視認性損失 $L_v$ 、平滑度損失 $L_s$

出力：損失 $L$

処理：加算部 1 1 3 5 は、次式により、損失 $L$ を計算する ( S 1 1 3 5 )。例えば、2つの損失の重みは ( $v_s, s_s$ )=(0.1,1)とする。

【 0 0 6 3 】

$$L = v_s L_v + s_s L_s \quad (13)$$

<アルファ値更新部 1 1 5 >

アルファ値更新部 1 1 5 は、損失 $L$ とアルファマップ  $\alpha$  を入力として、損失 $L$ のアルファマップ  $\alpha$  中の各画素に関する勾配に基づいて、アルファマップ  $\alpha$  を更新する ( S 1 1 5 )。本実施形態ではAdam optimizerを用いて更新を行う。アルファ値更新部 1 1 5 は、所定の条件を満たしたところで更新を終了し ( S 1 1 5 - 2 のyes)、その時点でのアルファマップ  $\alpha$  を最適化済のアルファマップ  $\alpha'$  として出力する。所定の条件とは、損失 $L$ が所定の閾値(例えば0.2)未満となるか、損失 $L$ の減少が一定回数(例えば50回)以上観察されないとき、あるいは一定回数(例えば500回)の更新を終えた時点等である。

【 0 0 6 4 】

<効果>

このような構成により、背景成分の減衰の効果が混入することなく正確に前景画像の視認性が推定可能となり、多様な前景画像に対する視認性の予測精度を大幅に改善することができるという効果を奏する。

【 0 0 6 5 】

より詳しく説明すると、本実施形態ではまず視認性を計算する対象となる成分(対象成分)を抽出し、これに対して視覚系のコントラスト知覚メカニズムを模した非線形処理を与えることで対象成分の知覚的なコントラストの推定値を計算する。なお、対象成分は、前景画像成分とアルファマップに含まれる構造成分を両方とも考慮して計算される。対象成分を明示的に抽出することで、背景成分の減衰の効果が混入することなく正確に前景画像の視認性が推定可能となる。

【 0 0 6 6 】

また、本実施形態ではコンテンツ適応的な特徴集約機構をモデルに付け加える。具体的には、色や空間周波数といった画像特徴毎に計算された視認性を集約する際に、前景画像のブレンド前の見た目に応じて、各画像特徴に対する重みを適応的に調節する。例えば元の前景画像がきめ細かいテクスチャであれば、高空間周波数成分に対する重みを大きくし、元の画像がぼけた模様であれば、低空間周波数成分に対する重みを大きくする。これにより、多様な前景画像に対する視認性の予測精度を大幅に改善することができる。

【 0 0 6 7 】

<変形例>

アルファ値生成部を最適化装置として機能させてもよい。この場合、最適化装置は、前景画像 $I^a$ 、背景画像 $I^b$ 、前景マスク画像 $M$ 、アルファマップの初期値  $\alpha^{in}$ 、所望の視認性マップ $\alpha^v$ を入力として、最適化済のアルファマップ  $\alpha^{out}$   $R^H \times W$ を生成し、出力する。以降の実施形態についても同様である。

【 0 0 6 8 】

<第2実施形態>

第1実施形態と異なる部分を中心に説明する。

【 0 0 6 9 】

画像の全ての領域で所望の視認性を得られるようにするため、非特許文献1では画素毎にアルファ値を独立に最適化するアルゴリズムを採用している。したがって、得られるアルファ値は隣接する画素間で不連続なものとなり、ブレンドの後で前景画像の元の構造や特徴が保たれる保証がない。これを防ぐため、非特許文献1ではブレンド前にアルファ値を空間的に平滑化する。しかし、この方法では、平滑化を行う空間的な範囲が広過ぎれば所望の視認性から逸脱してしまうし、狭過ぎれば元の画像構造が保たれなくなるというト

10

20

30

40

50

レードオフが生じ、人手によるチューニングが必要となる。

【0070】

本実施形態では、この課題を解決するため、アルファ値を最適化する際、所望の視認性に近くなるようにするだけでなく、同時に前景画像の元の構造をできるだけ保つような制約を与える。具体的には、ブレンド画像中の視認性計算の対象成分と元の前景画像成分の相関が高いほどコストが下がるような損失関数を用いた。これにより、前景画像の構造を保持しつつ所望の視認性が得られるようなブレンド結果を自動的に得ることができる。さらに、前景画像の構造を保つ最適化の副次的な効果として、ユーザが画像領域毎に異なる「所望の視認性」を与えた場合に、その視認性のマップが画像中のオブジェクトを正確に塗り分けていなくても、アルファ値がオブジェクトの輪郭に沿って自然に変化するよう

10

【0071】

第2実施形態の損失計算部213は、第1実施形態の損失計算部113に画像忠実度損失計算部2134を加えたものとなる。画像忠実度損失 $L_f$ は、現在のアルファマップでブレンドした画像中に含まれている背景以外の成分（第一対象成分 $I^{t^g}$ ）と、前景画像 $I^{f^g}$ を不透明にブレンドした際の背景以外の成分（第二対象成分 $I^{f^g}$ ）の相関が高くなる効果を与えるような損失である。前景画像 $I^{f^g}$ の構造を保持しつつ所望の視認性が得られるようなアルファマップを自動的に得ることができる。さらに、前景画像 $I^{f^g}$ の構造を保つ最適化の副次的な効果として、ユーザが画像領域毎に異なる「所望の視認性」を与えた場合に、その視認性マップが画像中のオブジェクトを正確に塗り分けていなくても、アルファ値

20

【0072】

図1は第1実施形態に係る合成装置200の機能ブロック図を、図2はその処理フローを示す。

【0073】

第2実施形態に係る合成装置100は、アルファ値生成部210とブレンド部120を含む。ブレンド部120の処理については、第1実施形態と同様なので、説明を省略する。

30

【0074】

<アルファ値生成部210>

図12はアルファ値生成部210の機能ブロック図を、図13はその処理フローの例を示す。

【0075】

アルファ値生成部210は、視認性予測部211、損失計算部213、アルファ値更新部115を含む。

【0076】

アルファ値生成部210は、前景画像 $I^{f^g}$ 、背景画像 $I^{b^g}$ 、前景マスク画像 $M$ 、アルファマップの初期値 $I^a$ 、所望の視認性マップ $v$ を入力として、最適化済のアルファマップ $I^a$ を生成し（S210）、出力とする。

40

【0077】

処理の順序としては、まず、視認性予測部211が、前景画像 $I^{f^g}$ 、背景画像 $I^{b^g}$ 、前景マスク画像 $M$ 、アルファマップの初期値 $I^a$ を入力として、視認性マップ $v$ と第一対象成分 $I^{t^g}$ 、第二対象成分 $I^{f^g}$ を求め（S211）、出力する。第一対象成分 $I^{t^g}$ 、第二対象成分 $I^{f^g}$ は共に、視認性予測部211での視認性マップ $v$ の計算と、損失計算部213での画像忠実度損失 $L_f$ の計算に用いるが、いずれの用途でも、 $N$ 個の空間周波数帯のうち、最も低い空間周波数帯であるローパス残余成分は用いない。このため、対象成分の次元数は $3 \times (N-1)$ 次元となる。 $N$ については第1実施形態で説明した通りである。

【0078】

50

次に、損失計算部 2 1 3 が、視認性マップ $v$ 、第一対象成分  $I^{tg}$ 、第二対象成分  $I^{fg}$ 、前景画像  $I^{fg}$ 、アルファマップ を入力として、損失 $L$ を計算し ( S 2 1 3 )、出力する。

【 0 0 7 9 】

アルファ値更新部 1 1 5 の処理 S 1 1 5 は第 1 実施形態と同様であり説明を省略する。

【 0 0 8 0 】

その後は、初期値  $i^{in}$ の代わりに更新したアルファマップ を入力として視認性予測部 2 1 1 が視認性マップ $v$ と第一対象成分  $I^{tg}$ 、第二対象成分  $I^{fg}$ を出力する。なお、第二対象成分  $I^{fg}$ は更新されたアルファマップ に依存しないため、初回に 1 度だけ計算したものを保持しておき、2回目以降はこれを再利用してもよい。このループを繰り返し、一定条件を満たしたら ( S 1 1 5 - 2 の yes )、アルファ値更新部 1 1 5 が最適化済のアルファマップ  $^{out}$ を出力して処理を終える。ループの制御はアルファ値更新部 1 1 5 に含まれる。

10

【 0 0 8 1 】

< 視認性予測部 2 1 1 >

図 1 4 は視認性予測部 2 1 1 の機能ブロック図を、図 1 5 はその処理フローの例を示す。

【 0 0 8 2 】

視認性予測部 2 1 1 は、マスク適用部 1 1 1 1、知覚コントラスト計算部 2 1 1 3、重み計算部 1 1 1 5、特徴集約部 1 1 1 7を含む。

【 0 0 8 3 】

入力：前景画像  $I^{fg}$ 、背景画像  $I^{bg}$ 、前景マスク画像  $M$ 、アルファマップ または  $i^{in}$

20

出力：第一対象成分  $I^{tg}$ 、第二対象成分  $I^{fg}$ 、視認性マップ $v$

処理：

マスク適用部 1 1 1 1、重み計算部 1 1 1 5、特徴集約部 1 1 1 7における処理 S 1 1 1 1, S 1 1 1 5, S 1 1 1 7 は第 1 実施形態と同様であり、説明を省略する。

【 0 0 8 4 】

知覚コントラスト計算部 2 1 1 3 は、マスク適用済のアルファマップ  $'$ 、前景マスク画像  $M$ 、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$ を入力として、第一対象成分  $I^{tg}$ とその知覚コントラストマップ  $P$ 、第二対象成分  $I^{fg}$ とその知覚コントラストマップ  $P'$ を得 ( S 2 1 1 3 )、出力する。知覚コントラストマップ  $P$ を得る過程で得られる第一対象成分  $I^{tg}$ 、第二対象成分  $I^{fg}$ は、そのまま視認性予測部 2 1 1 の出力として出力する。

30

【 0 0 8 5 】

< 知覚コントラスト計算部 2 1 1 3 >

図 1 6 は知覚コントラスト計算部 2 1 1 3 の機能ブロック図を、図 1 7 はその処理フローの例を示す。

【 0 0 8 6 】

知覚コントラスト計算部 2 1 1 3 は、色変換部 1 1 1 3 A、アルファブレンド部 1 1 1 3 B、対象成分抽出部 2 1 1 3 C、バンド分解部 1 1 1 3 D、バンドリミテッドコントラスト計算部 1 1 1 3 E、第一正規化部 1 1 1 3 F、第二正規化部 1 1 1 3 G、連結部 1 1 1 3 Hを含む。

【 0 0 8 7 】

40

知覚コントラスト計算部 2 1 1 3 はマスク適用済のアルファマップ  $'$ 、前景マスク画像  $M$ 、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$ を入力として、第一対象成分  $I^{tg}$ 、第二対象成分  $I^{fg}$ 、第一対象成分  $I^{tg}$ の知覚コントラストマップ  $P$ 、第二対象成分  $I^{fg}$ の知覚コントラストマップ  $P'$ を計算し ( S 2 1 1 3 )、出力する。しかし、実際には ( マスク適用済のアルファマップ  $'$ 、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$  ) を入力として ( 第一対象成分  $I^{tg}$ 、第一対象成分  $I^{tg}$ の知覚コントラストマップ  $P$  ) を計算する処理と ( 前景マスク画像  $M$ 、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$  ) を入力として ( 第二対象成分  $I^{fg}$ 、第二対象成分  $I^{fg}$ の知覚コントラストマップ  $P'$  ) を計算する処理に分けることができ、それぞれの処理は入力が異なるだけで中身は完全に同一である。このため、ここではマスク適用済のアルファマップ  $'$ と前景マスク画像  $M$ を区別せず「アルファマップ  $'$ 」と呼び、また第一対象成分  $I^{tg}$ と第二対象成分  $I$

50

$f^g$ を区別せず単に「対象成分」と呼ぶこととし、一方の処理についてのみ記述する。

【0088】

入力：アルファマップ、前景画像 $f^g$ 、背景画像 $b^g$

出力：対象成分  $I^g$ 、 $I^g$ 、対象成分の知覚コントラストマップ $P$ 、 $P'$

処理：

色変換部 1 1 1 3 A、アルファブレンド部 1 1 1 3 B、バンド分解部 1 1 1 3 D、バンドリミテッドコントラスト計算部 1 1 1 3 E、第一正規化部 1 1 1 3 F、第二正規化部 1 1 1 3 G、連結部 1 1 1 3 Hにおける処理 S 1 1 1 3 A、S 1 1 1 3 B、S 1 1 1 3 D、S 1 1 1 3 E、S 1 1 1 3 F、S 1 1 1 3 G、S 1 1 1 3 Hは第1実施形態と同様であり、説明を省略する。

10

【0089】

対象成分抽出部 2 1 1 3 Cは、色変換後の前景画像 $f^g$ と背景画像 $b^g$ 、およびアルファマップを入力として、対象成分  $I^g$ 、 $I^g$ を抽出し(S 2 1 1 3 C)、出力する。第1実施形態との違いは、対象成分抽出部 2 1 1 3 Cが、対象成分  $I^g$ 、 $I^g$ をバンドリミテッドコントラスト計算部 1 1 1 3 Eに向けて出力するとともに、視認性予測部 2 1 1 の出力として損失計算部 2 1 3 に向けて出力する点である。

【0090】

< 損失計算部 2 1 3 >

図 1 8 は損失計算部 2 1 3 の機能ブロック図を、図 1 9 はその処理フローの例を示す。

【0091】

損失計算部 2 1 3 は、視認性損失計算部 1 1 3 1、画像忠実度損失計算部 2 1 3 4、平滑度損失計算部 1 1 3 3、加算部 2 1 3 5を含む。

20

【0092】

損失計算部 2 1 3 は、所望の視認性マップ $\hat{v}$ と視認性マップ $v$ と前景画像 $f^g$ 、アルファマップ、第一対象成分  $I^g$ と第二対象成分  $I^g$ を入力として、損失 $L$ を計算し(S 2 1 3)、出力する。

【0093】

視認性損失計算部 1 1 3 1では、現在のアルファマップ から予測された視認性マップ $v$ と所望の視認性マップ $\hat{v}$ との距離を計算し、視認性損失 $L_v$ とする。この視認性損失 $L_v$ だけでは前景画像 $f^g$ の元の構造が保たれる保証がないため、何らかの正規化が必要となる。本実施形態では、このために画像忠実度損失 $L_f$ と平滑度損失 $L_s$ を用いる。画像忠実度損失 $L_f$ は、現在のアルファマップ でブレンドした画像中に含まれている背景以外の成分(第一対象成分  $I^g$ )と、前景画像 $f^g$ を不透明にブレンドした際の背景以外の成分(第二対象成分  $I^g$ )の相関が高くなる効果を与えるような損失である。視認性損失 $L_v$ と画像忠実度損失 $L_f$ だけでも、多くの場合十分良い結果を与えるが、稀に最適化結果がローカルミニマムに陥ったり、バンディングアーティファクトを生じたりするケースが見られる。これを防ぎ、さらに安定してよりよい結果を生成できるようにするため、アルファマップ に対してエッジ保存型平滑化の効果を与える平滑度損失 $L_s$ を追加で加える。加算部 1 1 3 5 はこれら全ての損失の重み付き和として最終的な損失 $L$ を計算する。

30

【0094】

視認性損失計算部 1 1 3 1、平滑度損失計算部 1 1 3 3における処理 S 1 1 3 1、S 1 1 3 3は第1実施形態と同様であり、説明を省略する。

40

【0095】

< 画像忠実度損失計算部 2 1 3 4 >

入力：第一対象成分  $I^g$   $R^{3(N-1)} \times H \times W$ 、第二対象成分  $I^g$   $R^{3(N-1)} \times H \times W$

出力：画像忠実度損失 $L_f$

処理：画像忠実度損失計算部 2 1 3 4 は、第一対象成分  $I^g$ 、第二対象成分  $I^g$ を用いて、次式によって画像忠実度損失 $L_f$ を計算する(S 2 1 3 4)。

【数 1 1】

50

$$L_f = \frac{1}{N-1} \sum_{k=1}^{N-1} (1 - \rho(\Delta I_k^{fg}, \Delta I_k^{tg})) \quad (14)$$

ただし、 $\rho(a,b)$ は、画像a,b間の相関を評価する関数である。

【 0 0 9 6 】

< 加算部 2 1 3 5 >

入力：視認性損失 $L_v$ 、画像忠実度損失 $L_f$ 、平滑度損失 $L_s$

出力：損失L

処理：加算部 2 1 3 5 は、次式により、損失Lを計算する ( S 2 1 3 5 )。例えば、3つの損失の重みは $(v, f, s)=(0.1, 1, 1)$ とする。

【 0 0 9 7 】

$$L = vL_v + fL_f + sL_s \quad (15)$$

< 効果 >

上述の構成により、第1実施形態と同様の効果を得ることができる。さらに、前景画像の構造を保持しつつ所望の視認性が得られるようなブレンド結果を自動的に得ることができる。さらに、ユーザが画像領域毎に異なる「所望の視認性」を与えた場合に、アルファ値がオブジェクトの輪郭に沿って自然に変化するような結果が得られる。

【 0 0 9 8 】

< 第3実施形態 >

第1実施形態および第2実施形態と異なる部分を中心に説明する。

【 0 0 9 9 】

第1実施形態および第2実施形態では、前景画像 $I^{fg}$ の視認性をユーザが所望の値に設定できたが、前景画像 $I^{fg}$ に加えて、背景画像 $I^{bg}$ の視認性をユーザが設定できるようにすることもできる。ただし、上述の式(1)によるアルファブレンディングでは、前景の透過度 $\alpha$ が決まると自動的に背景の透過度 $(1-\alpha)$ も決定される。したがって、前景の視認性と背景の視認性が共に任意のレベルになるようアルファマップを最適化することは一般に不可能である。このため、第3実施形態では、前景画像 $I^{fg}$ については所望の視認性を、背景画像 $I^{bg}$ については視認性の下限をユーザが指定する。すなわち、背景画像 $I^{bg}$ の視認性が下限値を下回らない範囲で、前景画像 $I^{fg}$ の視認性が所望の値に最も近づくようにアルファマップを最適化する。これにより、ユーザは背景が適度に視認できる状態を保ちつつ、前景の視認性を望んだレベルで表示できるようになる。

【 0 1 0 0 】

図1は第1実施形態に係る合成装置300の機能ブロック図を、図2はその処理フローを示す。

【 0 1 0 1 】

第3実施形態に係る合成装置300は、アルファ値生成部310とブレンド部120を含む。ブレンド部120の処理は第1実施形態および第2実施形態と同様であり、説明を省略する。

【 0 1 0 2 】

< アルファ値生成部 3 1 0 >

図20はアルファ値生成部310の機能ブロック図を、図21はその処理フローの例を示す。

【 0 1 0 3 】

アルファ値生成部310は、前景視認性予測部311F、前景損失計算部313F、背景視認性予測部311B、背景損失計算部313B、損失加算部314、アルファ値更新部115を含む。このうち前景視認性予測部311Fと前景損失計算部313Fは、それぞれ第2実施形態における視認性予測部211と損失計算部213と同様の処理 ( S 3 1 1 F , S 3 1 3 F ) を行う。

10

20

30

40

50

## 【0104】

アルファ値生成部310は、前景画像 $I^f$ 、背景画像 $I^b$ 、前景マスク画像 $M$ 、アルファマップの初期値 $\alpha^i$ 、所望の視認性マップ $\hat{v}$ 、視認性下限マップ $\hat{v}_B$   $R^H \times W$ を入力として、最適化済のアルファマップ $\alpha^{out}$   $R^H \times W$ を生成し(S310)、出力する。なお、視認性下限マップ $\hat{v}_B$ は背景画像の視認性の下限を示すものである。

## 【0105】

<背景視認性予測部311B>

背景視認性予測部311Bは、前景視認性予測部311Fと共通の処理S311Bを行う。ただし、前景画像 $I^f$ と背景画像 $I^b$ の役割が逆転する。具体的には、全ての処理について前景画像 $I^f$ を背景画像 $I^b$ に置き換え、背景画像 $I^b$ を前景画像 $I^f$ に置き換える。

10

## 【0106】

また、背景視認性予測部311Bは、背景視認性予測部311B内のマスク適用部から出力されるマスク適用済のアルファマップ $\alpha'$ の各画素値 $\alpha'_i$ を1から差し引くことで、背景用のアルファマップを生成する。つまり、 $\alpha'_i - 1$ とし、これをマスク適用済のアルファマップ $\alpha'$ に置き換えて利用する。

## 【0107】

さらに、背景視認性予測部311B内の知覚コントラスト計算部において、背景の第二対象成分 $I^{fg,B}$ を計算する際の入力として、前景マスク画像 $M$ のかわりに「全ての画素が1であるような画像」をアルファマップとして与える。

## 【0108】

上述の処理を行った上で、背景視認性予測部311Bは、前景画像 $I^f$ 、背景画像 $I^b$ 、前景マスク画像 $M$ 、アルファマップの初期値 $\alpha^i$ を入力として、視認性マップ $v^B$ と第一対象成分 $I^{tg,B}$ 、第二対象成分 $I^{fg,B}$ を求め(S311B)、出力する。なお、視認性マップ $v^B$ は、背景画像の視認性を示すものである。

20

## 【0109】

<背景損失計算部313B>

背景損失計算部313Bは、所望の視認性下限マップ $\hat{v}_B$ と視認性マップ $v^B$ と前景画像 $I^f$ 、アルファマップ $\alpha$ 、第一対象成分 $I^{tg,B}$ と第二対象成分 $I^{fg,B}$ を入力として、損失 $L_B$ を計算し(S313B)、出力する。

## 【0110】

例えば、背景損失計算部313B内の視認性損失計算部は、背景画像 $I^b$ の各画素 $i$ における視認性の下限値 $\hat{v}_{i,B}$ と、現在のアルファマップ $\alpha$ で計算した背景の各画素 $i$ における視認性 $v_{i,B}$ を用いて、背景についての視認性損失 $L_{v,B}$ を次のように計算する。

## 【数12】

$$L_{v,B} = \frac{1}{|N|} \sum_{i \in N} |\max(\hat{v}_{i,B} - v_{i,B}, 0)|^2 \quad (16)$$

ここで $N$ は「ブレンドを行う画像領域に属する画素の集合」を表す。上の式は前景損失計算部313Fにおける視認性損失 $L_v$ と基本的に同様であるが、現在の視認性が視認性の下限値を上回る場合には損失としないよう変更が加えられている。なお、背景画像 $I^b$ に対する画像忠実度損失 $L_f$ と平滑度損失 $L_s$ は、前景損失計算部313Fと同様に計算して損失に加えてもよいし、省略してもよい。つまり、背景損失計算部313B内の加算部は、視認性損失 $L_{v,B}$ をそのまま損失 $L_B$ として出力してもよいし、視認性損失 $L_{v,B}$ に画像忠実度損失 $L_f$ と平滑度損失 $L_s$ の少なくとも何れかを加えて損失 $L_B$ として出力してもよい。

40

## 【0111】

<損失加算部314>

最後に、損失加算部314は、前景損失計算部313Fで得られた損失 $L_F$ と、背景損失計算部313Bで得られた損失 $L_B$ とを入力とし、これらの値を加算し(S314)、和を損失 $L(=L_F+L_B)$ としてアルファ値更新部115に出力する。

50

【 0 1 1 2 】

アルファ値更新部 1 1 5 における処理は基本的に第 1 実施形態、第 2 実施形態と同様であるが、更新したアルファマップ を前景視認性予測部 3 1 1 F に出力すると同時に、同じアルファマップ を背景視認性予測部 3 1 1 B にも出力する点のみ異なる。

【 0 1 1 3 】

< 第 4 実施形態 >

これまでの実施形態では 2 枚の画像をブレンドする状況のみ考えてきたが、第 4 実施形態ではこれを任意の枚数の画像同士のブレディングに拡張する。1 枚の背景画像  $I_i^{bg}$  上に M 枚の前景画像  $I_i^{fg}$  をブレンドする式は以下のように定義する。

【 0 1 1 4 】

$$I_i^{bl} = \alpha_i^{(1)} I_i^{fg1} + \alpha_i^{(2)} I_i^{fg2} + \dots + \alpha_i^{(M)} I_i^{fgM} + (1 - \alpha_i^{(1)} - \alpha_i^{(2)} - \dots - \alpha_i^{(M)}) I_i^{bg} \quad (4-1)$$

ここで、 $I_i^{fgm}$  は m 番目の前景画像  $I_i^{fgm}$  の i 番目の画素値、 $\alpha_i^{(m)}$  は m 番目の前景画像  $I_i^{fgm}$  のアルファマップ  $\alpha_i^{(m)}$  の i 番目の画素のアルファ値をそれぞれ表す。各アルファ値は、

$$\alpha_i^{(1)} + \alpha_i^{(2)} + \dots + \alpha_i^{(M)} = 1, \quad 0 \leq \alpha_i^{(m)} \leq 1, \quad m=1, 2, \dots, M \quad (4-2)$$

を満たすものとする。

【 0 1 1 5 】

図 2 2 は第 4 実施形態に係る合成装置 4 0 0 の機能ブロック図を、図 2 はその処理フローを示す。

【 0 1 1 6 】

第 4 実施形態に係る合成装置 4 0 0 は、アルファ値生成部 4 1 0 とブレンド部 4 2 0 を含む。

【 0 1 1 7 】

合成装置 4 0 0 は、M 枚の前景画像  $I_i^{fgm}$   $R^3 \times H \times W$  と 1 枚の背景画像  $I_i^{bg}$   $R^3 \times H \times W$ 、M 枚の前景マスク画像  $M^m$   $R^H \times W$ 、M 個のアルファマップの初期値  $\alpha_{in}^m$   $R^H \times W$ 、M 枚の所望の視認性マップ  $\alpha_v^m$   $R^H \times W$  を入力として、M 枚の前景画像  $I_i^{fgm}$  と 1 枚の背景画像  $I_i^{bg}$  とを合成し、合成画像  $I_i^{bl}$   $R^3 \times H \times W$  を出力する。

【 0 1 1 8 】

アルファ値生成部 4 1 0 は、M 枚のアルファマップの初期値  $\alpha_{in}^m$   $R^H \times W$  と M 枚の前景画像  $I_i^{fgm}$  と 1 枚の背景画像  $I_i^{bg}$  と M 枚の前景マスク画像  $M^m$  と M 枚の所望の視認性マップ  $\alpha_v^m$  を入力として、アルファ値  $\alpha_{out}^{(m)} \in [0, 1]$  を生成し ( S 4 1 0 )、M 枚のアルファマップ  $\alpha_{out}^{(m)}$   $R^H \times W$  を出力する。 $\alpha_{out}^{(m)} \in [0, 1]$  は m 番目のアルファマップ  $\alpha_{out}^{(m)}$  の i 番目の画素のアルファ値を表す。

【 0 1 1 9 】

ブレンド部 4 2 0 は、M 枚の前景画像  $I_i^{fgm}$  と 1 枚の背景画像  $I_i^{bg}$  とアルファ値生成部 4 1 0 から出力される M 枚のアルファマップ  $\alpha_{out}^{(m)}$  を入力として、M 枚の前景画像  $I_i^{fgm}$  と背景画像  $I_i^{bg}$  を次式によりブレンドし ( S 4 2 0 )、合成画像  $I_i^{bl}$   $R^3 \times H \times W$  を出力する。

【 0 1 2 0 】

$$I_i^{bl} = \alpha_i^{out(1)} I_i^{fg1} + \alpha_i^{out(2)} I_i^{fg2} + \dots + \alpha_i^{out(M)} I_i^{fgM} + (1 - \alpha_i^{out(1)} - \alpha_i^{out(2)} - \dots - \alpha_i^{out(M)}) I_i^{bg} \quad (4-1)$$

以下、アルファ値生成部 4 1 0 の詳細について説明する。

【 0 1 2 1 】

< アルファ値生成部 4 1 0 >

図 2 3 はアルファ値生成部 4 1 0 の機能ブロック図を、図 2 4 はその処理フローの例を示す。

【 0 1 2 2 】

アルファ値生成部 4 1 0 は、マスク適用部 4 1 1 1、M 個の前景視認性予測部 4 1 1 F - m、背景視認性予測部 4 1 1 B、M 個の前景損失計算部 4 1 3 F - m、背景損失計算部 4 1 3 B、損失加算部 4 1 4、アルファ値更新部 4 1 5 を含む。図 2 3 中、前景視認性予測部 4 1 1 F 2 ~ 前景視認性予測部 4 1 1 F - ( M - 1 )、前景損失計算部 4 1 3 F 2 ~ 前景損失計算部 4 1 3 F - ( M - 1 ) は省略している。

10

20

30

40

50

## 【 0 1 2 3 】

アルファ値生成部 4 1 0 は、M枚の前景画像 $I^{fgm}$ 、背景画像 $I^{bg}$ 、M枚の前景マスク画像 $M^m$ 、M枚のアルファマップの初期値 $in^{(m)}$ 、背景用のアルファマップの初期値 $in^{(bg)}$ 、M枚の前景画像 $I^{fgm}$ に対応するM枚の所望の視認性マップ $\wedge^{vm}$ 、背景画像 $I^{bg}$ に対する所望の視認性マップ $\wedge^{vb}$ を入力として、最適化済のアルファマップ $out^{(m)}$   $R^{H \times W}$ 、 $out^{(bg)}$   $R^{H \times W}$ を生成し ( S 4 1 0 )、出力とする。

## 【 0 1 2 4 】

< マスク適用部 4 1 1 1 >

本実施形態では、視認性予測部に含まれていたマスク適用部 ( 図 5 参照 ) を、視認性予測部から独立させることで、M+1個ある視認性予測部にて同一の処理を行わなくて済むようにしている。

10

## 【 0 1 2 5 】

M枚のアルファマップ $\alpha^{(m)}$ をそれぞれ独立に最適化すると、式(4-2)の制約を満たすことは難しい。そこで、背景用のアルファマップ $\alpha^{(bg)}$ も最適化対象に加え、次の式により常に制約(4-2)を満たすようなM枚のアルファマップ $\alpha^{(m)}$ を生成する。

## 【 数 1 3 】

$$\alpha_i^{(m)} = \frac{\exp(\alpha_i^{(m)})}{\exp(\alpha_i^{(bg)}) + \sum_{m'=1}^M \exp(\alpha_i^{(m')})} \quad (4-3)$$

20

ここで、 $\alpha^{(m)}$ 、 $\alpha^{(bg)}$ は正規化前の前景アルファマップと正規化前の背景アルファマップを表すものとする。

## 【 0 1 2 6 】

マスク適用部 4 1 1 1 は、M個の正規化前の前景アルファマップ $\alpha^{(m)}$ と正規化前の背景アルファマップ $\alpha^{(bg)}$ を入力として、以下の処理によりマスク適用済のアルファマップ $\alpha^{(m)}$ を生成する。なお、正規化前の前景アルファマップ $\alpha^{(m)}$ は初期値 $in^{(m)}$ またはアルファ値更新部 4 1 5 の出力する前景アルファマップ $\alpha^{(m)}$ であり、正規化前の背景アルファマップ $\alpha^{(bg)}$ は初期値 $in^{(bg)}$ またはアルファ値更新部 4 1 5 の出力する背景アルファマップ $\alpha^{(bg)}$ である。

30

## 【 0 1 2 7 】

入力：M個の正規化前の前景アルファマップ $\alpha^{(m)}$ 、正規化前の背景アルファマップ $\alpha^{(bg)}$ 、M個の前景マスク画像 $M^m$

出力：M個のマスク適用済のアルファマップ $\alpha^{(m)}$

処理：

まず、マスク適用部 4 1 1 1 は、式(4-3)によりM個の正規化後のアルファマップ $\alpha^{(m)}$ を得る。

## 【 0 1 2 8 】

マスク適用部 4 1 1 1 は、M枚の正規化後の前景アルファマップ $\alpha^{(m)}$ それぞれに対応するM枚の前景マスク画像 $M^m$ を画素毎に掛け合わせ、M枚のマスク適用済のアルファマップ $\alpha^{(m)}$ を生成し ( S 4 1 1 1 )、出力する。

40

## 【 0 1 2 9 】

< 前景視認性予測部 4 1 1 F - m >

前景視認性予測部 4 1 1 F 1 ~ 4 1 1 F M は処理が共通するため、ここではm番目の前景視認性予測部 4 1 1 F - m の処理のみ記述する。第 2 実施形態における視認性予測部 2 1 1 と基本的に同一の処理を行うが、マスク適用部 4 1 1 1 が前景視認性予測部 4 1 1 F - m の外側に移動したことで、マスク適用済のアルファマップ $\alpha^{(m)}$ を入力にとる点と、M枚のマスク適用済のアルファマップ $\alpha^{(m)}$ とm番目の前景画像 $I^{fgm}$ を入力にとる点が異なる。その他の相違点は知覚コントラスト計算部 1 1 1 3 にて詳述する。

## 【 0 1 3 0 】

50

図 2 5 は前景視認性予測部 4 1 1 F - m の機能ブロック図を、図 2 6 はその処理フローの例を示す。

【 0 1 3 1 】

前景視認性予測部 4 1 1 F - m は、知覚コントラスト計算部 4 1 1 3、重み計算部 1 1 1 5、特徴集約部 1 1 1 7 を含む。

【 0 1 3 2 】

入力：M 枚の前景画像  $I^{fg,m}$ 、m 番目の前景マスク画像  $M^m$ 、背景画像  $I^{bg}$ 、M 枚のマスク適用済のアルファマップ  $\alpha^{(m)}$

出力：前景視認性マップ  $v^m$

処理：

知覚コントラスト計算部 4 1 1 3 は、M 枚のマスク適用済のアルファマップ  $\alpha^{(m)}$ 、m 番目の前景マスク画像  $M^m$ 、M 枚の前景画像  $I^{fg,m}$ 、背景画像  $I^{bg}$  を入力として、前景画像  $I^{fg,m}$  に対する第一対象成分  $I^{tg,m}$  とその知覚コントラストマップ  $P^m$ 、前景画像  $I^{fg,m}$  に対する第二対象成分  $I^{fg,m}$  とその知覚コントラストマップ  $P'^m$  を得 ( S 4 1 1 3 )、出力する。第一対象成分  $I^{tg,m}$ 、第二対象成分  $I^{fg,m}$  は、そのまま前景視認性予測部 4 1 1 F - m の出力として出力する。

【 0 1 3 3 】

重み計算部 1 1 1 5 は、第二対象成分  $I^{fg,m}$  の知覚コントラストマップ  $P'^m$  を入力として、各画素の画像特徴に対する重みベクトル  $w^m \in R^{3N}$  を計算し ( S 1 1 1 5 )、出力する。

【 0 1 3 4 】

特徴集約部 1 1 1 7 は、第一対象成分  $I^{tg,m}$  の知覚コントラストマップ  $P^m$  と重み計算部 1 1 1 5 で得られた各画素の重みベクトル  $w^m \in R^{3N}$  とを入力として、重みベクトル  $w^m$  を用いて、第一対象成分  $I^{tg,m}$  の各画素における知覚コントラスト  $r^m \in R^{3N}$  を重み付けしつつ特徴集約を行い、各画素の視認性  $v_i^m$  を得る ( S 1 1 1 7 )。

【 0 1 3 5 】

最後に、特徴集約部 1 1 1 7 は、各画素の視認性を表す視認性マップ  $v^m \in R^{H \times W}$  を出力する。

【 0 1 3 6 】

< 知覚コントラスト計算部 4 1 1 3 >

図 2 7 は知覚コントラスト計算部 4 1 1 3 の機能ブロック図を、図 2 8 はその処理フローの例を示す。

【 0 1 3 7 】

第 1、第 2 実施形態と同様に、前景画像  $I^{fg,m}$  をマスク適用済のアルファマップ  $\alpha^{(m)}$  でブレンドした際の前景画像  $I^{fg,m}$  ( 前景画像  $I^{fg,m}$  の第一対象成分  $I^{tg,m}$  ) に対する知覚コントラストマップ  $P^m$  を計算する処理と、前景画像  $I^{fg,m}$  を前景マスク画像  $M^m$  で不透明にブレンドした際の前景画像  $I^{fg,m}$  ( 前景画像  $I^{fg,m}$  の第二対象成分  $I^{fg,m}$  ) に対する知覚コントラストマップ  $P'^m$  を計算する処理とを行う。2 つの処理はブレンドに用いるアルファマップ  $\alpha^{(m)}$  ( マスク適用済のアルファマップまたは前景マスク画像 ) が異なる以外は同一である。また、アルファマップ  $\alpha^{(m)}$  によるブレンドを行うアルファブレンド部 4 1 1 3 B と対象成分抽出部 4 1 1 3 C 以外の処理は第 2 実施形態と同一であるため、以下ではアルファブレンド部 4 1 1 3 B と対象成分抽出部 4 1 1 3 C の処理についてのみ記述する。

【 0 1 3 8 】

< アルファブレンド部 4 1 1 3 B >

入力：M 枚のマスク適用済のアルファマップ  $\alpha^{(m)}$ 、m 番目の前景画像  $I^{fg,m}$  に対応する 1 枚の前景マスク画像  $M^m$ 、M 枚の色変換後の前景画像  $I^{fg,m}$ 、色変換後の背景画像  $I^{bg}$

出力：m 番目のブレンド画像  $I^{bl,tg,m}$

アルファブレンド部 4 1 1 3 B は、前景画像  $I^{fg,m}$  の第一対象成分  $I^{tg,m}$  を計算するために、式(4-1)にしたがってアルファブレンド画像  $I^{bl,tg,m} \in R^{3 \times H \times W}$  を得 ( S 4 1 1 3 B )、出力する。

10

20

30

40

50

【 0 1 3 9 】

$$I_i^{bl,tgm} = I_i^{(1)fg1} + I_i^{(2)fg2} + \dots + I_i^{(M)fgM} + (1 - I_i^{(1)} - I_i^{(2)} - \dots - I_i^{(M)}) I_i^{bg} \quad (4-1)$$

また、アルファブレンド部 4 1 1 3 B は、前景画像  $I_i^{fgm}$  の第二対象成分  $I_i^{fgm}$  を計算するために、次式にしたがってアルファブレンド画像  $I_i^{bl,fgm}$   $R^{3 \times H \times W}$  を得 ( S 4 1 1 3 B )、出力する。

【 0 1 4 0 】

$$I_i^{bl,fgm} = M_i^m I_i^{fgm} + (1 - M_i^m) \{ I_i^{(1)fg1} + \dots + I_i^{(m-1)fgm-1} + I_i^{(m+1)fgm+1} + \dots + I_i^{(M)fgM} + (1 - I_i^{(1)} - \dots - I_i^{(m-1)} - I_i^{(m+1)} - \dots - I_i^{(M)}) I_i^{bg} \} \quad (4-4)$$

ただし、 $M_i^m$  は前景マスク画像  $M^m$  の  $i$  番目の画素のアルファ値とする。

【 0 1 4 1 】

< 対象成分抽出部 4 1 1 3 C >

入力：  $M$  枚のアルファマップ  $(^m)$ 、 $m$  番目の前景画像  $I_i^{fgm}$  に対応する 1 枚の前景マスク画像  $M^m$ 、 $M$  枚の色変換後の前景画像  $I_i^{fgm}$ 、色変換後の背景画像  $I_i^{bg}$

出力：  $m$  番目の前景画像  $I_i^{fgm}$  に対応する 1 個の第一対象成分  $I_i^{tgm}$ 、 $m$  番目の前景画像  $I_i^{fgm}$  に対応する 1 個の第二対象成分  $I_i^{fgm}$

処理：第一対象成分  $I_i^{tgm}$  と第二対象成分  $I_i^{fgm}$  の計算で基本的に共通の処理 S 4 1 1 3 C となるが、以下の(4)のブレンド処理のみ両者で異なる計算方法を用いる。

【 0 1 4 2 】

(1)まず、対象成分抽出部 4 1 1 3 C は、色変換後の  $m$  番目の前景画像  $I_i^{fgm}$  を除く全て ( $M-1$  個) の前景画像  $I_i^{fgm}$  と背景画像  $I_i^{bg}$  のガウシアンピラミッド ( $N+2$  スケール) を生成する。

【 0 1 4 3 】

(2)以降は、対象成分のラプラシアンピラミッドの各スケール ( $k \mid k=1, 2, \dots, N-1$ ) を抽出する処理となる。冗長性を省くため、全部で  $N-1$  回繰り返す処理のうち、任意の  $k$  番目の処理についてのみ記述する。また、ラプラシアンピラミッドの最も粗いスケール ( $k=N$ ) に相当するローパス残余成分については、後の処理で使用しないため生成しない。

【 0 1 4 4 】

(3)対象成分抽出部 4 1 1 3 C は、色変換後の  $m$  番目の前景画像  $I_i^{fgm}$  を除く全て ( $M-1$  個) の前景画像  $I_i^{fgm}$  と背景画像  $I_i^{bg}$  のガウシアンピラミッドから、現在のスケール ( $k$ ) より 3 つ粗いスケール ( $k+3$ ) の画像を取り出し、元画像の解像度までアップスケーリングを行う。

【 0 1 4 5 】

(4)対象成分抽出部 4 1 1 3 C は、アップスケーリングした画像と、色変換後の前景画像  $I_i^{fgm}$  とをアルファマップ  $(^m)$  を用いてブレンドする。対象成分抽出部 4 1 1 3 C は、前景画像  $I_i^{fgm}$  の第一対象成分  $I_i^{tgm}$  を計算する際には、上述の式(4-1)によりブレンド画像を生成する。対象成分抽出部 4 1 1 3 C は、前景画像  $I_i^{fgm}$  の第二対象成分  $I_i^{fgm}$  を計算する際には、上述の式(4-4)を用いる。

【 0 1 4 6 】

(5)対象成分抽出部 4 1 1 3 C は、ブレンドした画像からラプラシアンピラミッドを生成し、そのうちの  $k$  番目のスケールを対象成分の  $k$  番目のバンドパス成分として出力する。なお、 $k$  番目以外のスケールは使用しないため、これ以外のスケールのバンドパス成分の生成は省略してもよい。

【 0 1 4 7 】

< 背景視認性予測部 4 1 1 B >

図 2 5 は前景視認性予測部 4 1 1 B の機能ブロック図を、図 2 6 はその処理フローの例を示す。

【 0 1 4 8 】

前景視認性予測部 4 1 1 B は、知覚コントラスト計算部 4 1 1 B 3、重み計算部 1 1 1 5、特徴集約部 1 1 1 7 を含む。

【 0 1 4 9 】

前景視認性予測部 4 1 1 B の処理 S 4 1 1 B は、基本的に前景視認性予測部 4 1 1 F -  $m$  と同一の処理 S 4 1 1 F -  $m$  となるが、背景画像  $I_i^{bg}$  は全ての画像領域に存在するため、

10

20

30

40

50

画像の存在領域を指定するためのマスク画像を入力にとらない。それ以外の前景視認性予測部 4 1 1 F - m からの相違点は知覚コントラスト計算部 4 1 1 B 3 のみに存在するため、以下では背景視認性予測部 4 1 1 B 内の知覚コントラスト計算部 4 1 1 B 3 についてのみ記述する。

【 0 1 5 0 】

< 知覚コントラスト計算部 4 1 1 B 3 ( 背景用 ) >

知覚コントラスト計算部 4 1 1 B 3 ( 背景用 ) の処理 S 4 1 1 B 3 は、前景用の知覚コントラスト計算部 4 1 1 3 の処理 S 4 1 1 3 と基本的に同一の処理となるが、画像のブレンドの計算が異なる。ここでも、背景画像  $I^{bg}$  をマスク適用済のアルファマップ  $(m)$  でブレンドした際の背景画像  $I^{bg}$  ( 背景画像  $I^{bg}$  の第一対象成分  $I^{tg,B}$  ) に対する知覚コントラストマップ  $P^{bg}$  を計算する処理と、不透明な背景画像  $I^{bg}$  ( 背景画像  $I^{bg}$  の第二対象成分  $I^{fg,B}$  ) に対する知覚コントラストマップ  $P^{bg}$  を計算する処理とを行う。なお、図 2 5、図 2 7 では第一対象成分  $I^{tg,B}$ 、第二対象成分  $I^{fg,B}$ 、知覚コントラストマップ  $P^{bg}$ 、知覚コントラストマップ  $P^{bg}$  を破線で囲んで示す。

10

【 0 1 5 1 】

2 つの処理はブレンドに用いるアルファマップ  $(m)$  が異なる以外は同一である。また、アルファマップ  $(m)$  によるブレンドを行うアルファブレンド部 4 1 1 3 B と対象成分抽出部 4 1 1 3 C 以外の処理は第 2 実施形態と同一であるため、以下ではアルファブレンド部 4 1 1 3 B と対象成分抽出部 4 1 1 3 C の処理についてのみ記述する。

【 0 1 5 2 】

< アルファブレンド部 4 1 1 3 B ( 背景用 ) >

入力：M 枚のマスク適用済のアルファマップ  $(m)$ 、色変換後の M 枚の前景画像  $I^{fgm}$ 、色変換後の背景画像  $I^{bg}$

出力：ブレンド画像  $I^{bl}$

アルファブレンド部 4 1 1 3 B は、背景第一対象成分  $I^{tg,B}$  を計算する際には、上述の式(4-1)によりブレンド画像  $I^{bl}$  を生成する。

20

【 0 1 5 3 】

アルファブレンド部 4 1 1 3 B は、背景第二対象成分  $I^{fg,B}$  を計算する際には、色変換後の背景画像  $I^{bg}$  をそのままブレンド画像  $I^{bl}$  として出力する。

【 0 1 5 4 】

< 対象成分抽出部 4 1 1 3 C >

入力：M 枚のマスク適用済のアルファマップ  $(m)$ 、色変換後の M 枚の前景画像  $I^{fgm}$ 、色変換後の背景画像  $I^{bg}$

出力：背景第一対象成分  $I^{tg,B}$ 、背景第二対象成分  $I^{fg,B}$

処理：

( 第一対象成分  $I^{tg,B}$  の生成処理 )

(1) まず、対象成分抽出部 4 1 1 3 C は、色変換後の M 枚の前景画像  $I^{fgm}$  のガウシアンピラミッド (  $N+2$  スケール ) を生成する。

【 0 1 5 5 】

(2) 以降は、対象成分のラプラシアンピラミッドの各スケール (  $k \mid k=1, 2, \dots, N-1$  ) を抽出する処理となる。冗長性を省くため、全部で  $N-1$  回繰り返す処理のうち、任意の  $k$  番目の処理についてのみ記述する。また、ラプラシアンピラミッドの最も粗いスケール (  $k=N$  ) に相当するローパス残余成分については、後の処理で使用しないため生成しない。

30

40

【 0 1 5 6 】

(3) 対象成分抽出部 4 1 1 3 C は、色変換後の M 枚の前景画像  $I^{fgm}$  のガウシアンピラミッドから、現在のスケール (  $k$  ) より 3 つ粗いスケール (  $k+3$  ) の画像を取り出し、元画像の解像度までアップスケーリングを行う。

【 0 1 5 7 】

(4) 対象成分抽出部 4 1 1 3 C は、アップスケーリングした画像と、色変換後の背景画像  $I^{bg}$  とをアルファマップ  $(m)$  を用いてブレンドする。ブレンドには、上述の式(4-1)を用い

50

る。

【 0 1 5 8 】

(5)対象成分抽出部 4 1 1 3 C は、ブレンドした画像からラプラシアンピラミッドを生成し、そのうちのk番目のスケールを対象成分のk番目のバンドパス成分として出力する。なお、k番目以外のスケールは使用しないため、これ以外のスケールのバンドパス成分の生成は省略してもよい。

【 0 1 5 9 】

(第二対象成分  $I^{fg,B}$  の生成処理)

対象成分抽出部 4 1 1 3 C は、色変換後の背景画像  $I^{bg}$  から直接ラプラシアンピラミッドを生成し、そのうちローパス残余成分を除いたバンドパス成分を第二対象成分  $I^{fg,B}$  とする。

10

【 0 1 6 0 】

< 前景損失計算部 4 1 3 F - m、背景損失計算部 4 1 3 B >

前景損失計算部 4 1 3 F - m は、第 2 実施形態の損失計算部 2 1 3 と同様の処理 S 4 1 3 F - m を行う。背景損失計算部 4 1 3 B は、第 3 実施形態の背景損失計算部 3 1 3 B と同様の処理 S 4 1 3 B を行う

< 損失加算部 4 1 4 >

損失加算部 4 1 4 は、M個の損失  $L_{F,m}$  と損失  $L_B$  とを入力とし、これらの値を加算し ( S 4 1 4 )、和を損失  $L (=L_{F,1}+...+L_{F,M}+L_B)$  としてアルファ値更新部 4 1 5 に出力する。M個の前景画像  $I^{fg,m}$ 、背景画像  $I^{bg}$  それぞれについて、どの画像の視認性を最も所望の値に近づけたいかによって、それぞれの損失を重み付けしてから加算してもよい。例えば、前景画像  $I^{fg,m}$  の視認性を優先して所望の値に近づけたい場合、前景損失計算部 4 1 3 F - m の損失をより大きく重み付けする。

20

【 0 1 6 1 】

< アルファ値更新部 4 1 5 >

アルファ値更新部 4 1 5 は、損失  $L$  と M 個の正規化前のアルファマップ  $\alpha_i^{(m)}$  と正規化前のアルファマップ  $\alpha_i^{(bg)}$  を入力として、M個の正規化前のアルファマップ  $\alpha_i^{(m)}$  および正規化前のアルファマップ  $\alpha_i^{(bg)}$  の各画素に関する勾配

【 数 1 4 】

$$\frac{\partial L}{\partial \alpha_i^{(m)}} (m = 1, 2, \dots, M), \frac{\partial L}{\partial \alpha_i^{(bg)}}$$

30

に基づいて、正規化前の前景アルファマップ  $\alpha_i^{(m)}$  および正規化前の背景アルファマップ  $\alpha_i^{(bg)}$  を更新し ( S 4 1 5 )、出力する。更新方法は第 1、第 2 実施形態と同様の方法で行う。

【 0 1 6 2 】

< 第 5 実施形態 >

これまでの実施形態のアルファ値生成部では、アルファマップ は損失計算部で計算された損失  $L$  に基づいて少しずつ更新して最適化することにより得る。しかし、この方法は最適なアルファマップ を得るまでに時間がかかるため、実時間動作するアプリケーションには不向きである。第 5 実施形態では、最適なアルファマップ を生成するアルファ値生成器を損失計算部による損失  $L$  に基づいて訓練する方法について説明する。この方法で事前に訓練した生成器を用いることで、1回のパスで高速にアルファマップ を生成することが可能となる。

40

【 0 1 6 3 】

< 訓練装置 5 0 0 >

図 2 9 は訓練装置 5 0 0 の機能ブロック図を、図 3 0 はその処理フローの例を示す。

【 0 1 6 4 】

訓練装置 5 0 0 は、アルファ値生成部 5 1 0 と視認性予測部 2 1 1 と損失計算部 2 1 3

50

とパラメータ更新部 5 1 5 とを含む。視認性予測部 2 1 1 と損失計算部 2 1 3 は第 2 実施形態のもの共通のものを用いることができる。

【 0 1 6 5 】

アルファ値生成部 5 1 0 は、前景画像  $I^{fg}$ 、背景画像  $I^{bg}$ 、所望の視認性マップ  $\wedge v$  を入力として、アルファマップ を生成し ( S 5 1 0 )、出力する。

【 0 1 6 6 】

アルファ値生成部 5 1 0 は、前景画像  $I^{fg}$  と背景画像  $I^{bg}$  を深層ニューラルネットワーク (DNN) から成る共通のエンコーダによってエンコードしたのち、エンコードした特徴と所望の視認性マップ  $\wedge v$  から別の DNN から成るデコーダを用いてデコードすることによってアルファマップを生成し ( S 5 1 5 )、出力する。

10

【 0 1 6 7 】

パラメータ更新部 5 1 5 は、損失計算部 2 1 3 から得られた損失  $L$  を入力とし、損失  $L$  の勾配に基づき、エンコーダ、デコーダのパラメータを更新し ( S 5 1 5 )、更新したパラメータを出力する。

【 0 1 6 8 】

< アルファ値生成部 5 1 0 >

図 3 1 はアルファ値生成部 5 1 0 の機能ブロック図を、図 3 2 はその処理フローの例を示す。

【 0 1 6 9 】

アルファ値生成部 5 1 0 は、エンコーダ 5 1 1 F、5 1 1 B、ガウシアンピラミッド生成部 5 1 3、アフィンパラメータ推定部 5 1 6、デコーダ 5 1 7 を含む。

20

【 0 1 7 0 】

入力：前景画像  $I^{fg}$ 、背景画像  $I^{bg}$ 、所望の視認性  $\wedge v$

出力：アルファマップ

処理：

エンコーダ 5 1 1 F、5 1 1 B は、それぞれ前景画像  $I^{fg}$ 、背景画像  $I^{bg}$  を入力とし、エンコードし ( S 5 1 1 F、S 5 1 1 B )、特徴量を抽出し、出力する。例えば、エンコーダ 5 1 1 F、5 1 1 B は、それぞれ前景画像  $I^{fg}$ 、背景画像  $I^{bg}$  それぞれに対して畳み込みとダウンサンプリングを交互に繰り返すことで階層的な特徴抽出を行い、出力する。

【 0 1 7 1 】

30

前景画像用のエンコーダ 5 1 1 F と背景画像用のエンコーダ 5 1 1 B は、重み ( パラメータ ) を共有する。エンコーダには畳み込み-ダウンサンプリング-非線形アクティベーションから成る処理を複数層重ねた畳み込みニューラルネットワークを用いる。非線形アクティベーションには、例えば ReLU 関数を用いる。ReLU 関数の前に Batch normalization を挟んでも良い。層の最大数は入力画像の解像度に依存し、例えば  $256 \times 256$  の画像解像度では最大 8 層のものを用いることができる。特徴数は層を通過するごとに増やしていき、例えば第 1 層では 64、第 2 層では 128、第 3 層では 256...、と増加させる。ただし、計算量削減のため最大特徴数を 512 程度に留めてもよい。

【 0 1 7 2 】

デコーダ 5 1 7 は、前景画像  $I^{fg}$  のエンコーダ 5 1 1 F の最後の出力と背景画像  $I^{bg}$  のエンコーダ 5 1 1 B の最後の出力を連結させたものを入力とし、それに対し転置畳み込みとアップサンプリングを交互に繰り返し適用することで徐々に解像度を元のサイズまで増やしていく。各層における具体的な処理は、転置畳み込み-アップサンプリング-(Batch normalization)- Feature-wise Linear Modulation (後述)-非線形アクティベーションの一連の処理で構成される。非線形アクティベーションには、例えば ReLU 関数を用いる。ただし、デコーダ 5 1 7 は、最終層では Tanh 関数または Sigmoid 関数を用いることで、出力の値域に制限をかける。このとき、参考文献 1 の U-Net の構造に倣い、前景画像  $I^{fg}$  と背景画像  $I^{bg}$  のエンコーダ 5 1 1 F、5 1 1 B で抽出された各階層の特徴をデコーダの対応する階層に連結して加えることで、高次で低解像度の特徴から低次で高解像度の特徴まで考慮しながらアルファマップ を生成することができる。

40

50

## 【 0 1 7 3 】

(参考文献 1) Olaf Ronneberger, Philipp Fischer, Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation", arXiv:1505.04597, 2015.

デコーダ 5 1 7 の各階層の特徴数は、入力時点では前景画像  $I^g$  のエンコーダ 5 1 1 F の出力と背景画像  $I^b$  のエンコーダ 5 1 1 B の出力を連結させるためエンコーダ 5 1 1 F , 5 1 1 B の出力の特徴数の 2 倍となる。デコーダ 5 1 7 は、第 1 層の転置畳み込み処理において特徴数を半分に減らしつつ解像度を 2 倍に増やす。デコーダ 5 1 7 は、第 2 層では、第 1 層の出力に対して、エンコーダ 5 1 1 F , 5 1 1 B から同じ解像度をもつ階層の出力を連結させたものを入力とするため、第 2 層の入力の特徴数は、「第 1 層の出力特徴数 + 2 × エンコーダ中間層出力の特徴数」となる。その後、デコーダ 5 1 7 は、第 2 層の畳み込み処理において、この特徴数を第 1 層の出力の特徴数の半分に減らす。デコーダ 5 1 7 は、このようにして特徴数を徐々に減らしつつ解像度を上げていき、最終層の畳み込み処理で特徴数を 1 に減らし、出力を Tanh 関数または Sigmoid 関数にかけてアルファマップ  $\alpha$  を得 ( S 5 1 7 )、出力する。Tanh 関数を用いる場合は、出力の値域が  $-1 \sim 1$  となるので、これが  $0 \sim 1$  となるよう正規化する。

10

## 【 0 1 7 4 】

デコーダ 5 1 7 では、参考文献 2 の Feature-wise Linear Modulation という手法を用い、入力として与えられる所望の視認性マップ  $\hat{v}$  に応じたアルファマップ  $\alpha$  を生成できるようにする。

## 【 0 1 7 5 】

(参考文献 2) Dumoulin, et al., "Feature-wise transformations", Distill, 2018.

このために、ガウシアンピラミッド生成部 5 1 3 は、所望の視認性マップ  $\hat{v}$  を入力とし、所望の視認性マップ  $\hat{v}$  からガウシアンピラミッドを生成し ( S 5 1 3 )、デコーダ 5 1 7 の各階層の空間サイズと一致する視認性マップ  $V_l$  を得、出力する。

20

## 【 0 1 7 6 】

さらに、アフィンパラメータ推定部 5 1 6 は、視認性マップ  $V_l$  を入力とし、各階層の視認性マップ  $V_l$  をそれぞれ 2 種類の畳み込みニューラルネットワーク ( 解像度に変更は加えない ) に入力し、画素毎、特徴毎に 2 つのアフィンパラメータ (  $\mu_{l,f}$ ,  $\sigma_{l,f}$  ) を得 ( S 5 1 6 )、出力する。このときの特徴数は、デコーダ 5 1 7 の対応する階層の出力の特徴数と一致させる。このときの畳み込みニューラルネットワークの層の数は、例えば 3 層程度とする。3 層とした場合、中間層の特徴数は、例えば出力特徴数の半分程度とする。

30

## 【 0 1 7 7 】

最後に、デコーダ 5 1 7 は、アフィンパラメータ (  $\mu_{l,f}$ ,  $\sigma_{l,f}$  ) を入力とし、各階層  $l$  の畳み込み処理後 ( Batch normalization を行う場合はその直後 ) の各特徴  $f$  のマップ  $X_{l,f}$  を、 (  $\mu_{l,f}$ ,  $\sigma_{l,f}$  ) を用いてアフィン変換 (  $Y_{l,f} = \sigma_{l,f} X_{l,f} + \mu_{l,f}$  ) により変調する。

## 【 0 1 7 8 】

< パラメータ更新部 5 1 5 >

パラメータ更新部 5 1 5 では、損失計算部 2 1 3 から与えられる損失  $L$  の、アルファ値生成部 5 1 0 の各重みについての勾配に基づいて、アルファ値生成部 5 1 0 の各重みを更新し ( S 5 1 0 )、出力する。本実施形態では Adam optimizer を用いて更新を行う。更新は所定の条件を満たしたところで終了し ( S 5 1 5 - 2 の yes )、その時点でのパラメータを保存して固定させる。所定の条件とは、損失  $L$  が所定の閾値 ( 例えば 0.2 ) 未満となるか、損失  $L$  の減少が一定回数 ( 例えば 50 回 ) 以上観察されないとき、あるいは一定回数 ( 例えば 500 回 ) の更新を終えた時点とする。

40

## 【 0 1 7 9 】

< 合成装置 6 0 0 >

図 3 3 は第 5 実施形態に係る合成装置 6 0 0 の機能ブロック図を、図 2 はその処理フローを示す。

## 【 0 1 8 0 】

アルファ値生成部 5 1 0 の訓練終了後は、合成装置 6 0 0 においてブレンド画像  $I^b$  を生

50

成する。合成装置 600 は、アルファ値生成部 510 とブレンド部 520 を含む。

【0181】

入力：前景画像 $I^fg$ 、背景画像 $I^bg$ 、所望の視認性マップ $\wedge v$ 、前景マスク画像 M

出力：ブレンド画像 $I^bl$

処理：

アルファ値生成部 510 は、前景画像 $I^fg$ 、背景画像 $I^bg$ 、所望の視認性マップ $\wedge v$ を入力とし、これらの値を用いて、アルファマップ を生成し ( S 5 1 0 )、出力する。

【0182】

ブレンド部 520 は、アルファマップ と前景マスク画像 M と前景画像 $I^fg$ 、背景画像 $I^bg$  を入力とし、まず、アルファマップ に前景マスク画像 M を画素毎に掛け合わせ、マスク適用済のアルファマップ ' を得る。

【0183】

次に、ブレンド部 520 は、マスク適用済のアルファマップ '、前景画像 $I^fg$ 、背景画像 $I^bg$ を用いて上述の式(1)によりブレンド画像 $I^bl$ を生成し ( S 5 2 0 )、出力する。

【0184】

<効果>

このような構成により、第1実施形態と同様の効果を得ることができる。さらに、高速にアルファマップ を生成することができる。

【0185】

<その他の変形例>

本発明は上記の実施形態及び変形例に限定されるものではない。例えば、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能である。

【0186】

<プログラム及び記録媒体>

上述の各種の処理は、図34に示すコンピュータの記憶部2020に、上記方法の各ステップを実行させるプログラムを読み込ませ、制御部2010、入力部2030、出力部2040などに動作させることで実施できる。

【0187】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。

【0188】

また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

【0189】

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記録媒体に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサ

10

20

30

40

50

ービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの（コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等）を含むものとする。

【0190】

また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

10

20

30

40

50

【図面】

【図 1】

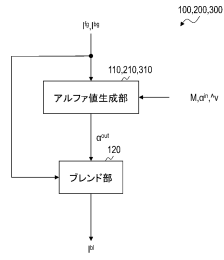


図1

【図 2】

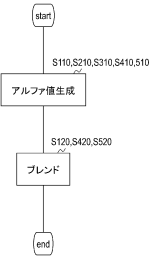


図2

10

【図 3】

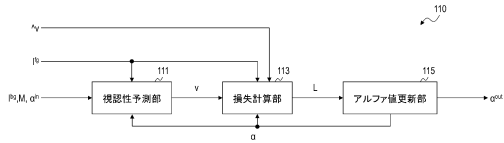


図3

【図 4】

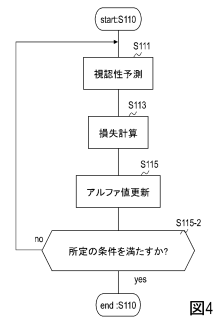


図4

20

30

40

50

【 図 5 】

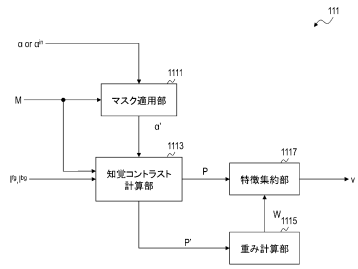


図5

【 図 6 】

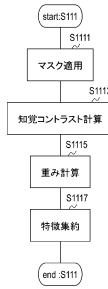
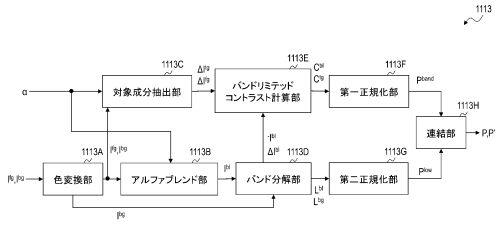


図6

10

【 図 7 】



1113

図7

【 図 8 】

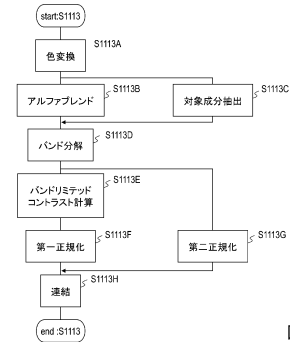
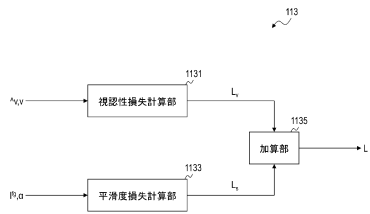


図8

20

【 図 9 】



113

図9

【 図 10 】

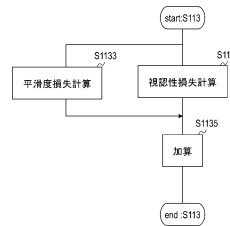


図10

30

40

50

【 図 1 1 】

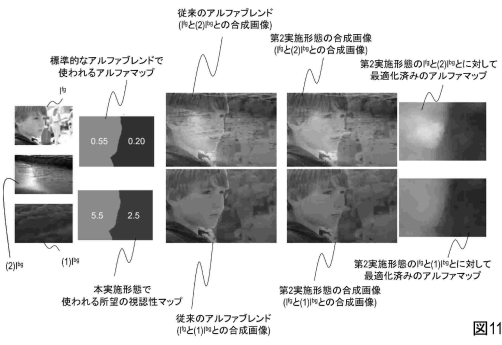


図11

【 図 1 2 】

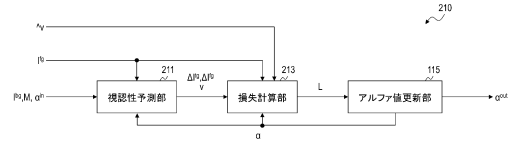


図12

【 図 1 3 】

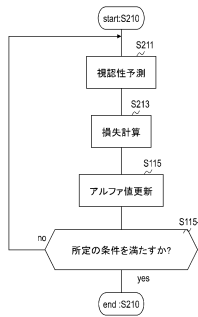


図13

【 図 1 4 】

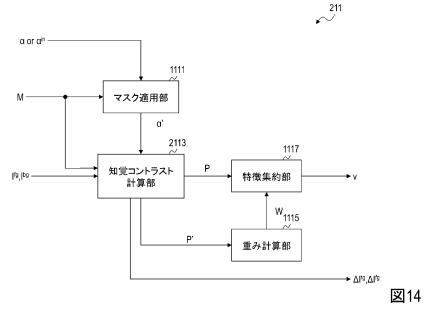


図14

【 図 1 5 】

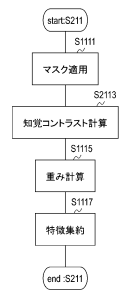


図15

【 図 1 6 】

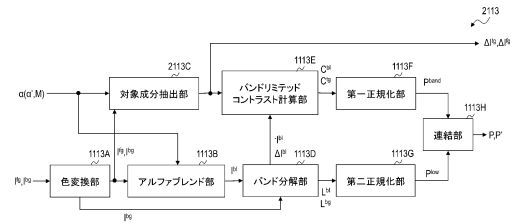


図16

10

20

30

40

50

【 図 17 】

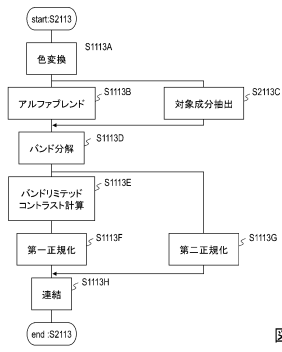


図17

【 図 18 】



図18

10

【 図 19 】

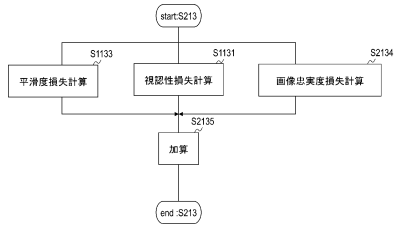


図19

【 図 20 】

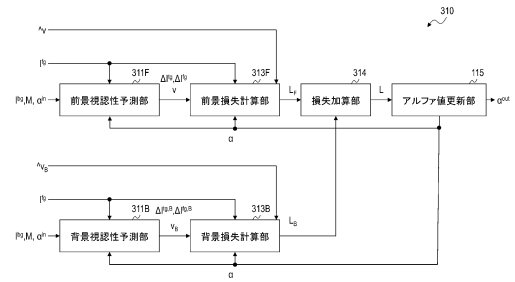


図20

20

【 図 21 】

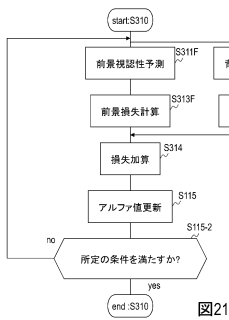


図21

【 図 22 】

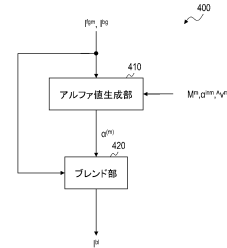


図22

30

40

50

【図 23】

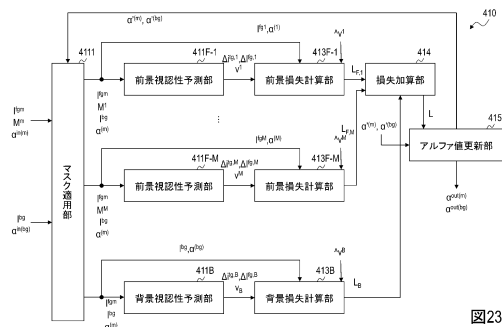


図23

【図 24】

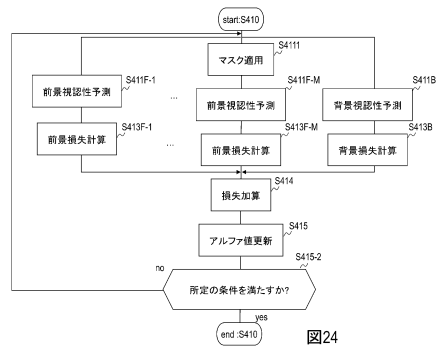


図24

10

【図 25】

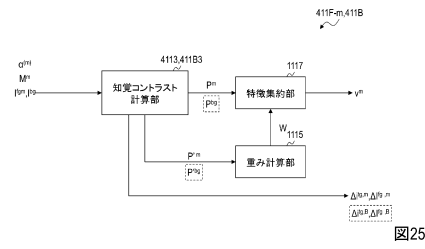


図25

【図 26】

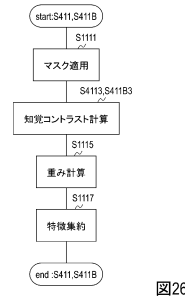


図26

20

【図 27】

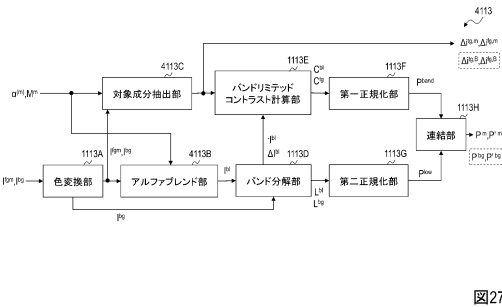


図27

【図 28】

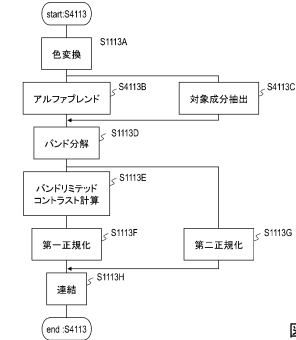


図28

30

40

50



## フロントページの続き

(72)発明者 大石 岳史

東京都文京区本郷七丁目3番1号 国立大学法人東京大学内

審査官 佐田 宏史

(56)参考文献 国際公開第2016/039301(WO, A1)

特開2012-059158(JP, A)

吹上 大樹、外2名, “透明視を利用した複合現実感環境下における遮蔽矛盾解消手法の提案”, 情報処理学会研究報告, 日本, 一般社団法人情報処理学会, 2013年02月15日, Vol.2012-CVIM-184, No.1, pp.1-8

(58)調査した分野 (Int.Cl., DB名)

G06T 1/00, 3/00, 5/00, 7/00 - 7/90

G06V 10/00 - 10/98

H04N 1/387