

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7277342号

(P7277342)

(45)発行日 令和5年5月18日(2023.5.18)

(24)登録日 令和5年5月10日(2023.5.10)

(51)国際特許分類

F I

G 0 6 T 7/60 (2017.01)

G 0 6 T

7/60

2 0 0 Z

請求項の数 22 (全18頁)

(21)出願番号	特願2019-212351(P2019-212351)	(73)特許権者	504407000
(22)出願日	令和1年11月25日(2019.11.25)		パロ アルト リサーチ センター インコ
(65)公開番号	特開2020-98589(P2020-98589A)		ーポレイテッド
(43)公開日	令和2年6月25日(2020.6.25)		アメリカ合衆国 カリフォルニア州 9 4
審査請求日	令和4年11月25日(2022.11.25)		3 0 4 パロ アルト カイオーテ ヒル
(31)優先権主張番号	16/223,519		ロード 3 3 3 3
(32)優先日	平成30年12月18日(2018.12.18)	(74)代理人	100094569
(33)優先権主張国・地域又は機関	米国(US)		弁理士 田中 伸一郎
早期審査対象出願		(74)代理人	100109070
			弁理士 須田 洋之
		(74)代理人	100067013
			弁理士 大塚 文昭
		(74)代理人	100086771
			弁理士 西島 孝喜
		(74)代理人	100109335

最終頁に続く

(54)【発明の名称】 幾何学的先行物を用いる曲線オブジェクトセグメンテーション

(57)【特許請求の範囲】

【請求項1】

曲線オブジェクトセグメンテーションのための方法であって、
 曲線特徴を含む少なくとも1つの入力画像を受信することと、
 プロセッサを使用して、前記少なくとも1つの画像を、パラメータ Θ_R を用いるパラメータを有する表現モジュール $f(\cdot)$ 、及びパラメータ Θ_T を有するタスクモジュール $F(\cdot)$ を有するディープネットワークを使用して、前記曲線特徴のセグメンテーションマップにマッピングすることと、を含み、前記マッピングが、
グラウンドトゥルースセグメンテーションマップ Y_g と、前記表現モジュール $f(\cdot)$ の出力を入力とする前記タスクモジュール $F(\cdot)$ の出力とのネットワーク誤差を用いた損失関数

10

$$E = L(Y_g, F(f(X, \Theta_R), \Theta_T))$$

に基づいて、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成された学習可能フィルタを使用して、前記表現モジュールにおいて前記入力画像を変換することと、
 前記タスクモジュールにおいて前記変換された入力画像を使用して前記セグメンテーションマップを生成することと、を含む、方法。

【請求項2】

20

前記入力画像が、道路、血管、機械亀裂、及び皮膚のしわのうちの１つ以上を含む、請求項１に記載の方法。

【請求項３】

前記マッピングが、損失関数を最小化することを含む訓練プロセスによって導出され、前記損失関数が、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成されている、請求項１に記載の方法。

【請求項４】

前記損失関数が、訓練誤差と曲線セグメンテーションのタスクに関する事前知識を組み込んだ幾何学的正則化項との合計を含む、請求項３に記載の方法。

【請求項５】

前記学習可能フィルタのうちの少なくとも１つが、ガボールフィルタとして初期化される、請求項３に記載の方法。

【請求項６】

前記ガボールフィルタが、訓練プロセス中に変化して、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成されている、請求項５に記載の方法。

【請求項７】

前記少なくとも１つの学習可能フィルタが、少なくとも１つの学習可能なパラメータを有する、請求項３に記載の方法。

【請求項８】

前記少なくとも１つの学習可能パラメータが、搬送周波数、偏心、向き、及びスケールのうちの１つ以上を含む、請求項７に記載の方法。

【請求項９】

少なくとも１つの訓練幾何学的パターン画像を用いて前記学習可能フィルタを訓練することを更に含む、請求項３に記載の方法。

【請求項１０】

前記少なくとも１つの訓練幾何学的パターン画像が、第１の訓練幾何学的パターン画像及び第２の訓練幾何学的パターン画像のうちの少なくとも１つを含み、前記第１の訓練幾何学的パターン画像が第１の向きのパターンを有し、前記第２の訓練幾何学的パターン画像が第２の向きのパターンを有し、前記第１の向きが前記第２の向きとは異なる、請求項９に記載の方法。

【請求項１１】

前記第１の向きが、前記第２の向きに直交する、請求項１０に記載の方法。

【請求項１２】

前記少なくとも１つの学習可能フィルタが、前記第１の訓練幾何学的パターン画像に対する高応答と、前記第２の訓練幾何学的パターン画像に対する低応答とを生成するように構成されている、請求項１０に記載の方法。

【請求項１３】

前記学習可能フィルタが、複数の空間スケールで導出される、請求項３に記載の方法。

【請求項１４】

前記タスクモジュールが、U-Netディープニューラルネットワークである、請求項１に記載の方法。

【請求項１５】

曲線オブジェクトセグメンテーションのためのシステムであって、
プロセッサと、
コンピュータプログラム命令を記憶するメモリと、を含み、前記コンピュータプログラム命令が、前記プロセッサによって実行されると、前記プロセッサに、
曲線特徴を含む少なくとも１つの入力画像を受信することと、
パラメータ Θ_R を用いるパラメータを有する表現モジュール $f(\cdot)$ 、及びパラメータ Θ_I を有するタスクモジュール $F(\cdot)$ を有するディープネットワークを使用して、前記曲線特徴のセグメンテーションマップを生成するために、プロセッサを使用して前記少なくとも１

10

20

30

40

50

つの画像をマッピングすることと、を含む動作を行わせ、前記マッピングが、
グラウンドトゥルスセグメンテーションマップ Y_g と、前記表現モジュール $f(\cdot)$ の出力
 を入力とする前記タスクモジュール $F(\cdot)$ の出力とのネットワーク誤差を用いた損失関数

$$L(Y_g, F(f(X, \theta_R), \theta_T))$$

に基づいて、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成された学習
 可能フィルタを使用して、前記表現モジュールにおいて前記入力画像を変換することと、
 前記タスクモジュールにおいて前記変換された入力画像を使用して前記セグメンテ
 ーションマップを生成することと、を含む、システム。

10

【請求項 16】

訓練画像の第 1 のセット及び訓練画像の第 2 のセットを用いて前記学習可能フィルタを
 訓練することを更に含み、前記訓練画像の第 1 のセットが、第 1 の向きのパターンを有し
 、前記訓練画像の第 2 のセットが、第 2 の向きのパターンを有し、前記第 1 の向きが、前
 記第 2 の向きとは異なる、請求項 15 に記載のシステム。

【請求項 17】

前記少なくとも 1 つの学習可能フィルタが、前記訓練画像の第 1 のセットに対する高応
 答と、前記訓練画像の第 2 のセットに対する低応答とを生成するように構成されている、
 請求項 16 に記載の方法。

20

【請求項 18】

前記プロセッサが、損失関数を最小化するように構成され、前記損失関数が、曲線形状
 の認識と訓練誤差の減少とのバランスをとるように構成されている、請求項 15 に記載の
 システム。

【請求項 19】

前記損失関数が、訓練誤差と曲線セグメンテーションのタスクに関する事前知識を組み
 込んだ幾何学的正則化項との合計を含む、請求項 18 に記載のシステム。

【請求項 20】

前記学習可能フィルタのうちの少なくとも 1 つが、ガボールフィルタとして初期化され
 る、請求項 18 に記載の方法。

30

【請求項 21】

前記少なくとも 1 つの学習可能フィルタが、少なくとも 1 つの学習可能なパラメータを
 有する、請求項 18 に記載の方法。

【請求項 22】

曲線オブジェクトセグメンテーションのためのコンピュータプログラム命令を記憶する
 非一時的コンピュータ可読媒体であって、前記コンピュータプログラム命令が、プロセッ
 サによって実行されると、前記プロセッサに、

曲線特徴を含む少なくとも 1 つの入力画像を受信することと、
パラメータ θ_R を用いるパラメータを有する表現モジュール $f(\cdot)$ 、及びパラメータ θ_T を
 有するタスクモジュール $F(\cdot)$ を有するディープネットワークを使用して、前記曲線特徴
 のセグメンテーションマップを生成するために、プロセッサを使用して前記少なくとも 1
 つの画像をマッピングすることと、を含む動作を行わせ、前記マッピングが、
グラウンドトゥルスセグメンテーションマップ Y_g と、前記表現モジュール $f(\cdot)$ の出力
 を入力とする前記タスクモジュール $F(\cdot)$ の出力とのネットワーク誤差を用いた損失関数

40

$$E = L(Y_g, F(f(X, \theta_R), \theta_T))$$

に基づいて、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成された学習可能フィルタを使用して、前記表現モジュールにおいて前記入力画像を変換することと、

前記タスクモジュールにおいて前記変換された入力画像を使用して前記セグメンテーションマップを生成することと、を含む、非一時的コンピュータ可読媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、画像処理、特に画像セグメンテーションを対象とする。

【0002】

曲線オブジェクトは、自然界の多くの場面において見られ、一般的な例としては、人間の顔のしわ、道路の亀裂、網膜内の血管、及び衛星画像の道路を含む。したがって、曲線オブジェクトの画素単位の自動検出とセグメンテーションは、様々なアプリケーションに影響を与える重要なコンピュータビジョントaskである。

10

【0003】

本明細書に記載される実施形態は、曲線オブジェクトセグメンテーションのための方法を含む。曲線特徴を含む少なくとも1つの入力画像が受信される。少なくとも1つの入力画像は、表現モジュール及びタスクモジュールを有するディープネットワークを使用して、曲線特徴のセグメンテーションマップにマッピングされる。マッピングには、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成された学習可能フィルタを使用して、表現モジュールにおいて入力画像を変換することが含まれる。セグメンテーションマップは、タスクモジュールにおいて変換された入力画像を使用して生成される。

20

【0004】

実施形態は、曲線オブジェクトセグメンテーションのためのシステムを含む。システムは、プロセッサと、プロセッサによって実行されると、プロセッサに動作を行わせるコンピュータプログラム命令を記憶するメモリとを含む。動作は、曲線特徴を含む少なくとも1つの入力画像を受信することを含む。少なくとも1つの画像は、表現モジュール及びタスクモジュールを有するディープネットワークを使用して曲線特徴のセグメンテーションマップを生成するために、プロセッサを使用してマッピングされる。マッピングには、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成された学習可能フィルタを使用して、表現モジュールにおいて入力画像を変換することが含まれる。セグメンテーションマップは、タスクモジュールにおいて変換された入力画像を使用して生成される。

30

【0005】

本明細書に記載される実施形態は、曲線オブジェクトセグメンテーションのためのコンピュータプログラム命令を記憶する非一時的コンピュータ可読媒体を含む。コンピュータプログラム命令は、プロセッサによって実行されると、プロセッサに動作を行わせる。動作は、曲線特徴を含む少なくとも1つの入力画像を受信することを含む。少なくとも1つの画像は、表現モジュール及びタスクモジュールを有するディープネットワークを使用して曲線特徴のセグメンテーションマップを生成するために、プロセッサを使用してマッピングされる。マッピングには、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成された学習可能フィルタを使用して、表現モジュールにおいて入力画像を変換することが含まれる。セグメンテーションマップは、タスクモジュールにおいて変換された入力画像を使用して生成される。

40

【図面の簡単な説明】

【0006】

【図1A】本明細書に記載される様々な実施形態による、幾何学的先行物を使用する曲線画像セグメンテーションのための方法を示す。

【図1B】本明細書に記載される様々な実施形態による、幾何学的先行物を使用する曲線画像セグメンテーションのための方法を示す。

【図2A】本明細書に記載される様々な実施形態による、ノイズ先行物を使用する曲線画像セグメンテーションのための方法を示す。

50

【図 2 B】本明細書に記載される様々な実施形態による、ノイズ先行物を使用する曲線画像セグメンテーションのための方法を示す。

【図 3】本明細書に記載される実施形態による、曲線セグメンテーションのためのディープラーニングアーキテクチャを示す。

【図 4 A】本明細書に記載される実施形態による、例示的な第 1 の幾何学的訓練画像を示す。

【図 4 B】本明細書に記載される実施形態による、例示的な第 2 の幾何学的訓練画像を示す。

【図 5 A】本明細書に記載される実施形態による、学習可能フィルタなしのデータ駆動型訓練を使用したフィルタの可視化を示す。

10

【図 5 B】本明細書に記載される実施形態による、学習可能ではないフィルタの可視化を示す。

【図 5 C】本明細書に記載される実施形態による、学習可能フィルタの可視化を示す。

【図 6 A】本明細書に記載される実施形態による、実施例 1 及び 2 の元の画像を示す。

【図 6 B】本明細書に記載される実施形態による、実施例 1 及び 2 の元の画像のためのグラウンドトゥールースセグメンテーションマップを示す。

【図 7 A】本明細書に記載される実施形態による、異なる技術を使用して、実施例 1 の結果として得られるセグメンテーションマップを示す。

【図 7 B】本明細書に記載される実施形態による、異なる技術を使用して、実施例 1 の結果として得られるセグメンテーションマップを示す。

20

【図 8 A】本明細書に記載される実施形態による、異なる技術を使用して、実施例 8 の結果として得られるセグメンテーションマップを示す。

【図 8 B】本明細書に記載される実施形態による、異なる技術を使用して、実施例 8 の結果として得られるセグメンテーションマップを示す。

【図 9】本明細書に記載される実施形態を実装することができるシステムのブロック図を示す。

【発明を実施するための形態】

【0007】

画像セグメンテーションとは、類似の画像特性を有する画像の一部をグループ化することを指す。例えば、本明細書に記載される実施形態は、曲線オブジェクトセグメンテーションを伴う。ディープラーニングアーキテクチャは、画像の曲線セグメンテーションに使用することができる。ディープラーニングに伴う重大な課題は、大量の訓練サンプルを取得する必要があることである。画素レベルのオブジェクトセグメンテーションなどのタスクの場合、訓練画像の注釈付けは、根気のいる、時間と費用のかかるタスクである。この問題に取り組む 1 つの方法は、ドメイン及び/又はタスクに関する事前知識を組み込むことである。この事前情報により、ネットワーク最適化の訓練サンプルへの依存性を効果的に低下させる。本明細書に記載される実施形態は、ニューラルネットワークの初期表現層に導入される正則化のセットを含む。正則化は、多数のスケール及び配向で曲線画像特徴に強く応答するようにフィルタを訓練するために使用され得る。様々な実施形態によれば、正則化は、出力セグメンテーションマップにおけるノイズを抑制するようにフィルタを訓練するために使用される。正則化を介して表される事前情報は、ドメイン及び/又はタスクの制約を捕捉するネットワーク損失関数に追加の項を導入することによって、ネットワーク訓練に直接組み込まれる。

30

40

【0008】

図 1 A は、本明細書に記載される実施形態による、曲線オブジェクトセグメンテーションのための方法を示す。図 1 A によれば、少なくとも 1 つの入力画像が受信される (110)。様々な構成によれば、入力画像は、少なくとも 1 つの曲線特徴を含む。例えば、入力画像には、道路、血管、機械亀裂、皮膚のしわのうちの 1 つ以上が含まれる場合がある。画像は、プロセッサを使用して曲線特徴のセグメンテーションマップにマッピングされる。様々な実施形態によれば、少なくとも 1 つの画像は曲線特徴の認識と訓練誤差の減少

50

とのバランスをとるように構成された損失関数の出力を最小化することによってマッピングされる。入力画像は、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成された学習可能フィルタを使用して変換される(130)。変換された画像を使用して、セグメンテーションマップを生成する(140)。

【0009】

様々な実施形態によれば、ディープネットワークを使用して、セグメンテーションマップを生成する。ディープネットワークは、例えば、表現モジュール及びタスクモジュールを含んでもよい。入力画像は、目的の曲線形状を認識するように訓練された学習可能フィルタを使用して、表現モジュールにおいて変換される(130)。タスクモジュールは、表現モジュールから受信した変換された入力画像を使用してセグメンテーションマップを生成するように構成される(140)。様々な構成によれば、表現モジュール及びタスクモジュールは、入力モジュールがタスクモジュールに入る前に表現モジュールにおいて変換されるように、直列に構成される。様々な構成によれば、タスクモジュールは、U-Net ディープニューラルネットワークである。

10

【0010】

様々な実施形態によれば、学習可能フィルタのうちの少なくとも1つは、ガボールフィルタとして初期化される。ガボールフィルタは、訓練プロセス中に変化して、曲線形状の認識と訓練誤差の減少とのバランスをとるように構成されてもよい。場合によっては、学習可能フィルタのうちの少なくとも1つは、少なくとも1つの学習可能パラメータを有する。例えば、少なくとも1つの学習可能パラメータは、搬送周波数、偏心、向き、及びスケールのうちの1つ以上を含んでもよい。

20

【0011】

図1Bは、本明細書に記載される実施形態による、幾何学的制約を有する曲線オブジェクトセグメンテーションネットワークを訓練するためのプロセスを示す。訓練は、以下のように反復処理として進む。ディープ曲線セグメンテーションネットワークは、フィルタ重みのセットを用いて初期化される。曲線特徴を含む入力訓練画像が受信される(150)。入力画像は、推定ネットワークセグメンテーションマップを生成するために、表現ネットワーク及びタスクネットワーク(160)の学習可能な表現フィルタ(155)を介したマッピングを含むディープ曲線セグメンテーションネットワークを介して送信される。

【0012】

30

並行して、フィルタリングされたパターンを生成する(157)のために、少なくとも2つの訓練幾何学的パターン画像が生成され(152)、学習可能フィルタを介してマッピングされる(155)。様々な実施形態によれば、第1の訓練幾何学的パターン画像は、第1の向きであり、第2の訓練幾何学的パターン画像は、第1の向きとは異なる第2の向きである。例えば、第1の向きは、第2の向きに直交してもよい。

【0013】

(160)からの推定セグメンテーションマップ及び訓練入力画像に対応する真のセグメンテーションマップ(162)は、損失計算モジュールに送信され(165)、これにより、真のセグメンテーションマップと推定セグメンテーションマップ間の誤差の測定値が計算される。更に、(157)からのフィルタリングされた幾何学的パターン画像は、向きの保存を確実にするために損失計算モジュールに入力される(165)。損失関数は、セグメンテーション誤差を最小化することと、幾何学的な向きの制約を保持することというネットワークの2つの目標のバランスをとる。幾何学的な制約は、第1の訓練幾何学的パターン画像からの高応答と、第2の訓練幾何学的パターン画像からの低応答とを選ぶことによって保持される。様々な構成によれば、学習可能フィルタは、異なるサイズ及び厚さの曲線構造が検出され得るように、複数の空間スケールで導出される。

40

【0014】

損失が収束しているかどうかを判定する(170)。いくつかの実施形態では、損失が前の反復から実質的に低下しなかった場合に、損失が収束したと判定され得る(170)。場合によっては、前の反復からの損失の変化が所定の閾値未満である場合、損失が収束

50

したと判定される。収束が生じたと判定された場合（１７０）、訓練プロセスは終了する（１８５）。収束が生じていないと判定された場合（１７０）、ネットワークパラメータは標準的な逆伝搬技術を介して更新され、次の反復が開始される。少なくとも１つの入力画像が受信される（１６０）。少なくとも１つの画像は、プロセッサを使用して曲線特徴のセグメンテーションマップにマッピングされる（１７０）。入力画像は、学習可能フィルタを使用して変換される（１８０）。変換された画像を使用して、セグメンテーションマップを生成する（１９０）。

【００１５】

図２Ａは、本明細書に記載される実施形態による、曲線オブジェクトセグメンテーションにおけるノイズ抑制のためのプロセスを示す。少なくとも１つの入力画像が受信される（２１０）。様々な構成によれば、入力画像は、少なくとも１つの曲線特徴を含む。少なくとも１つの画像は、プロセッサを使用して曲線特徴のセグメンテーションマップにマッピングされる。様々な実施形態によれば、少なくとも１つの画像は、訓練画像におけるノイズ抑制と曲線セグメンテーション誤差とのバランスをとるように構成された損失関数の出力を最小化することにより、セグメンテーションマップにマッピングされる。様々な構成によれば、損失関数は、訓練画像及びノイズ正則化項に対応する推定されたセグメンテーションマップと真のセグメンテーションマップとの間の訓練誤差の合計を含む。入力画像は、少なくとも１つの入力画像のドメイン及びタスクのうちの１つ以上におけるノイズを抑制するように構成された学習可能フィルタを使用して、変換される（２３０）。様々な実施形態によれば、学習可能フィルタのうちの少なくとも１つは、ガボールフィルタとして初期化される。変換された画像を使用して、セグメンテーションマップを生成する（２４０）。様々な構成によれば、タスクモジュールは、U-Net ディープニューラルネットワークである。

【００１６】

様々な実施形態によれば、ディープネットワークを使用して、セグメンテーションマップを生成する。上述のように、ディープネットワークは、表現モジュール及びタスクモジュールを含み得る。入力画像は、ノイズ抑制と曲線セグメンテーションとのバランスをとるように構成された学習可能フィルタを使用して、表現モジュールにおいて変換される（２３０）。タスクモジュールは、表現モジュールから受信した変換された入力画像を使用してセグメンテーションマップを生成するように構成される（２４０）。

【００１７】

図２Ｂは、本明細書に記載される実施形態による、ノイズ制約を有する曲線オブジェクトセグメンテーションネットワークを訓練するためのプロセスを示す。訓練は、以下のように反復処理として進む。ディープ曲線セグメンテーションネットワークは、フィルタ重みのセットを用いて初期化される。曲線特徴を含む訓練画像が受信される（２５０）。入力画像は、推定ネットワークセグメンテーションマップを生成するために、表現ネットワーク及びタスクネットワーク（２６０）の学習可能な表現フィルタ（２５５）を介したマッピングを含むディープ曲線セグメンテーションネットワークを介して送信される。

【００１８】

並行して、フィルタリングされたパターンを生成する（２５７）ために、少なくとも１つの訓練幾何学的パターン画像が生成され（２５２）、学習可能フィルタを介してマッピングされる（２５５）。様々な構成によれば、少なくとも１つの訓練ノイズ画像は、入力画像のドメイン及びタスクに固有のノイズパターンを含む。目的のラベル付き曲線オブジェクトを含まない訓練画像からバックグラウンドパッチを抽出し、曲線構造に応答するフィルタを介して抽出されたバックグラウンドパッチを処理し、及び最も高いフィルタ応答を有する所定数のバックグラウンドパッチを訓練ノイズ画像として選択することにより、１つ以上のノイズ画像が作られる。いくつかの実施形態では、曲線構造に応答するフィルタは、ガボールフィルタとして選択される。

【００１９】

２６０からの推定セグメンテーションマップ及び訓練入力画像に対応する真のセグメン

10

20

30

40

50

テーションマップ (2 6 2) は、損失計算モジュールに送信され (2 6 5)、これにより、真のセグメンテーションマップと推定セグメンテーションマップ間の誤差の測定値が計算される。更に、2 5 7 からのフィルタリングされたノイズパターン画像は、損失計算モジュールに入力され (2 6 5)、推定されたセグメンテーションマップにおけるノイズ測定値を提供する。損失関数は、セグメンテーション誤差を最小化すること、及び推定セグメンテーションマップのノイズを最小化することというネットワークの2つの目標のバランスをとる。

【 0 0 2 0 】

損失が収束しているかどうかを判定する (2 7 0)。いくつかの実施形態では、損失が収束したことは、損失が前の反復から実質的に低下しなかった場合であると判定されてもよい (2 7 0)。場合によっては、前の反復からの損失の変化が所定の閾値未満である場合、損失が収束していると判定される。収束が生じたと判定された場合 (2 7 0)、訓練プロセスは終了する (2 8 5)。収束が生じていないと判定された場合 (2 7 0)、ネットワークパラメータは標準的な逆伝搬技術を介して更新され、次の反復が開始される。

【 0 0 2 1 】

曲線特徴を含む少なくとも1つの入力画像が受信される (2 6 0)。画像は、プロセッサを使用して曲線特徴のセグメンテーションマップにマッピングされる (2 7 0)。入力画像は、ノイズを抑制するように構成された学習可能フィルタを使用して変換される (2 8 0)。変換された画像を使用して、セグメンテーションマップを生成する (2 9 0)。セグメンテーションマップを生成するために、訓練誤差の減少と曲線形状の認識とのバランスをとるように構成されたフィルタ、及びノイズを抑制するように構成されたフィルタの任意の組み合わせを使用できることを理解されたい。

【 0 0 2 2 】

図 3 は、本明細書に記載される実施形態による、曲線セグメンテーションのためのディープラーニングアーキテクチャを示す。システムは、表現モジュール 3 1 0 及びタスクモジュール 3 2 0 を含む。画像 3 1 0 は、表現モジュールに入力される。表現モジュールは、1つ以上の幾何学的訓練画像 3 4 0 及び/又は1つ以上のノイズ訓練画像 3 4 5 で訓練された学習可能フィルタ 3 3 0 を使用する。学習可能フィルタ 3 5 0 は、変換された入力画像 3 6 0 を出力する。変換された入力画像 3 6 0 は、タスクモジュール 3 2 0 に入力され、変換された画像は、セグメンテーションマップを生成するように訓練されたタスクモジュールの1つ以上のフィルタ 3 7 0 に入力される。出力セグメンテーションマップ 3 8 0 は、タスクモジュールから出力される。

【 0 0 2 3 】

本明細書に記載されている様々な実施形態によれば、幾何学的先行物が表現ネットワークに導入されて、学習可能フィルタが曲線画像の幾何学的形状に強く応答するように促す。例えば、学習可能フィルタは、曲線特徴と一般的に関連付けられている微細な線に応答するように訓練されてもよい。全体的なシステムは、式 1 の損失関数を使用して曲線形状に関する訓練誤差のバランスをとるために最適化され得る。

【 0 0 2 4 】

【数 1】

$$E = L(Y_g, F(f(X, \theta_R), \theta_T)) + L_R(\theta_R) \quad (1)$$

【 0 0 2 5 】

ここで、 $f()$ はパラメータ θ_R を用いるパラメータを有する表現モジュールを表し、 $F()$ は、パラメータ θ_T を有するタスクモジュールを表す。 Y_g は、グラウンドトゥルースセグメンテーションマップであり、 $L()$ は訓練サンプル上のネットワーク誤差を表し、 $L_R()$ は、曲線セグメンテーションのタスクに関する事前知識を組み込んだ幾何学的正則化項であり、式 2 に示される。

【 0 0 2 6 】

【数 2】

$$L_{R1} = \sum_{k=1}^M \sum_{i=1}^N \|W_{ik} \circledast I_{O_{ik}}\|_F^2 - \|W_{ik} \circledast I_{S_{ik}}\|_F^2, W_{ik} \in \Theta_R \quad (2)$$

【0027】

ここで、 $I_{S_{ik}}$ は、 i 番目の向き及び k 番目のスケールに沿った強い曲線パターンを含む第1の訓練幾何学的パターン画像を表す（ S_{ik} と表記される）。 $I_{O_{ik}}$ は、 $I_{S_{ik}}$ に直交する方向のパターンを有する第2の訓練画像を表す。これらの2つの画像の例を図4A及び図4Bに示す。図4Aは、例示的な第1の幾何学的訓練画像を示し、図4Bは、第2の幾何学的訓練画像の例を示す。これらの2つの画像は共に、 k 番目のスケールでの表現フィルタが S_{ik} に沿って配向するように促す。様々な実施形態によれば、画像パターンは、所与のサイズ（例えば 22×22 ）の長方形グリッドを特定の角度 i だけ回転させて得られた座標に適用される、所与のスケール k の2次元ガウス関数によって形成される。例示の長方形グリッドは正方形として表されているが、長方形グリッドは正方形ではない場合があることを理解されたい。

10

【0028】

式1及び2を使用して、 Θ_R パラメータは、ガボールフィルタとして初期化され、ネットワーク訓練プロセス中に進化して、曲線形状 L_{R1} を尊重しながら、訓練誤差 L を最小化するバランスのとれた目標を達成する。いくつかの事例において、学習可能フィルタは、訓練プロセス前のガボールフィルタであり、学習プロセス中に変化するように学習可能フィルタが構成され、曲線特徴のターゲットセグメンテーションが提供される。図5Aは、初期段階におけるドメイン及び/又はタスクに関する事前知識を組み込んでいない訓練から学習されたフィルタの可視化を示す。図5Bは、静的ガボールフィルタを示す。図5Cは、本明細書に記載される実施形態による、曲線特徴の認識の訓練後の学習可能フィルタの可視化を示す。本明細書に記載されるような学習可能なガボールフィルタを使用することは、画像内の曲線形状に応答するガボールのような構造を保持する一方で、訓練標本に基づいてガボール構造を調節及び適応させる能力を保持する。

20

【0029】

様々な構成によれば、学習可能フィルタは、少なくとも1つの学習可能なパラメータを持つガボールフィルタである。式3は、本明細書に記載される実施形態によるガボールフィルタを表す式を示す。

30

【0030】

【数 3】

$$W_R[i, j] = K_{R1} e^{\frac{-(i^2 + \gamma j^2)}{2\sigma^2}} \cos(2\pi f(i \cos \theta + j \sin \theta)) \quad (3)$$

【0031】

式3において、 $[i, j]$ は画像の画素座標であり、 γ は離心率、 f は搬送周波数であり、 θ は方向、 σ はスケールである。様々な実施形態によれば、 γ 、 f 、 θ 、及び σ のうちの1つ以上は、システムの訓練中に訓練データを使用して学習可能な値である。この実施形態では、損失関数は、訓練サンプル項 $L()$ 上のデータネットワーク誤差のみを含み、式1における第2の正則化項（ L_{R1} ）は、式4に示されるように使用されない。

40

【0032】

【数 4】

$$E = L(Y_g, F(f(X, \theta_R), \theta_T)) \quad (4)$$

【0033】

様々な実施形態によれば、本明細書に記載されるような学習可能なパラメータを有するフィルタを使用することは、全てのフィルタ重みを学習しなければならない従来の畳み込

50

みニューラルネットワークと比較した場合、訓練可能なパラメータの数が大幅に減少する。訓練パラメータの減少により、従来の方法よりも複雑さが軽減され、訓練要件が減少される。訓練標本に基づいていくつかの重要なパラメータを学習する機能は、静的フィルタ表現よりも優れた適応性と表現力をネットワークに与える。式 4 は、全ての 4 つのガボールパラメータに関して平滑であり、微分可能であるため、パラメータは、逆伝播を介して学習され得る。

【 0 0 3 4 】

様々な実施形態によれば、画像セグメンテーションに使用されるネットワークは、スパースな訓練の場合に、ノイズに対してより敏感になり得る。本明細書に記載される実施形態は、データ内のノイズに対するディープセグメンテーションネットワークの堅牢性を改善するように構成される。具体的には、バックグラウンドにフォアグラウンドオブジェクトと同様の特徴を含む場合に発生し得る誤検知の問題に対処する。ドメイン及びノ又はタスクに関する事前知識をディープラーニングタスクに組み込むことで、ノイズがある訓練サンプルへのネットワーク最適化の依存性を低下させることができる。本明細書に記載される実施形態は、ノイズに対して堅牢である方法でセグメンテーションを実行するためのシステム及び方法を説明する。

10

【 0 0 3 5 】

様々な構成によれば、学習可能フィルタを使用して、曲線画像セグメンテーションにおけるノイズを抑制することができる。式 5 は、曲線画像セグメンテーションにおけるノイズ抑制と訓練誤差とをバランスさせるように構成された損失関数を示す。

20

【 0 0 3 6 】

【数 5】

$$E = L(Y_g, F(f(X, \Theta_R), \Theta_T)) + L_{R2}(\Theta_R) \quad (5)$$

【 0 0 3 7 】

ここで、 $f()$ はパラメータ Θ_R を有するパラメータを有する表現モジュールを表し、 $F()$ は、パラメータ Θ_T を有するタスクモジュールを表す。 Y_g は、グラウンドトゥールースセグメンテーションマップであり、 $L()$ は訓練サンプル上のネットワーク誤差を表し、 $L_{R2}()$ は、データ内のノイズを抑制しようと試みるノイズ正則化項である。様々な実施形態によれば、 L_{R2} は、曲線データセグメンテーションにおけるノイズをモデル化するように構成される。 L_{R2} の計算は、式 6 に示される。

30

【 0 0 3 8 】

【数 6】

$$L_{R2} = \sum_{k=1}^M \sum_{i=1}^N \sum_{j=1}^K \|W_{ik} \odot I_j\|_F^2, W_{ik} \in \Theta_R \quad (6)$$

【 0 0 3 9 】

ここで、 I_j は、ドメイン及びタスクに固有のノイズパターンを含む補助訓練画像のセットであり、 W_{ik} は Θ_R におけるスケール k のフィルタ重みである。式 5 の $L()$ 項と組み合わされると、表現モジュールは、タスク性能（すなわち、曲線セグメンテーション）とノイズ抑制とのバランスをとる。

40

【 0 0 4 0 】

本明細書に記載される様々な実施形態によれば、ノイズパターンは、標識された曲線オブジェクトを含まない訓練画像からバックグラウンドパッチを抽出することによって作られる。これらのパッチは、例えばガボールフィルタなどのフィルタを介して送信される。最大の応答を与える所定数のパッチが潜在的なノイズのあるパッチ（つまり、曲線構造と混同される可能性が最も高いパッチ）として選択される。様々な実施形態によれば、ノイズがあるパッチのパッチサイズは 64×64 に選択され、最大応答が 100 パッチとなるように選択される。

50

【実施例 1】

【0041】

本明細書に記載される実施形態は、網膜血管セグメンテーションの適用例について実証される。これは、医療専門家がヒト網膜の血管経路における病理を迅速に診断することを可能にする重要な機能である。DRIVE [11] (J. Staal, M. D. Abramoff, M. Niemeijer and M. Van Ginneken, 「Ridge-Based Vessel Segmentation of the Retina」, IEEE Transactions on Medical Imaging, vol. 23, no. 4, pp. 501 - 509, 2004を参照)、及び STARE (A. Hoover, V. Kouznetsova and M. Goldbaum, 「Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response」, IEEE Transactions on Medical Imaging, vol. 19, no. 3, pp. 203 - 210, 2000を参照) データセットを使用して、本手法を訓練及び評価した。ネットワークアーキテクチャは、以下のとおりである。表現層には、サイズが 11×11 で、 $0 \sim 180$ 度の方向で均一にサンプリングした 12 個のガボールフィルタを使用した。タスクネットワークのために、U-Net アーキテクチャを使用した。表現ネットワークは、Adam オプティマイザを活用した標準の確率的勾配降下法を使用して訓練し、学習率は .0005 に設定した。サイズ 128×128 のパッチを訓練画像から抽出し、20 個の訓練画像から約 7000 個のパッチを得た。限定訓練シナリオも評価し、2 つの画像から 700 個のパッチを抽出した。ネットワークの訓練用に 64 のパッチサイズを使用した。初期検証では、単一スケールのガボール正規化のみをテストした。推論では、ネットワークは、パッチではなく画像全体で使用された。以下の方法は、ガボール正規化、ガボールパラメータ化、及び深層網膜画像理解 (DRIU) で評価された。

【0042】

表 1 及び表 2 は、それぞれ、完全訓練及び限定訓練シナリオの結果を示す。結果は、2 つのバイナリマップ間のオーバーラップ係数の正規化された尺度である、Dice 係数に関して表される。ガボール正規化及びガボールパラメータ化の方法は、両方のシナリオにおいて競合的であり、限定訓練シナリオで大幅なマージンを確保し、最先端の DRIU よりも優れている。限定データセットで訓練された提案されたアプローチは、完全データセットで訓練された DRIU よりも優れている。

【0043】

【表 1】

表 1

方法	Diceオーバーラップスコア
DRIU	0.7389
ガボール正規化	0.8031
ガボールパラメータ化	0.7956

【0044】

【表 2】

表 2

方法	Diceオーバーラップスコア
DRIU	0.6318
ガボール正規化	0.7430

【0045】

図 6 A は、実施例 1 の元の入力を示す。望ましい曲線特徴のグラウンドトゥルスマップを図 6 B に示す。図 7 A は、D R I U を使用した結果セグメンテーションマップを示しており、灰色部分が誤った結果を表し、白色部分が正しい結果を表す。図 7 B は、本明細書に記載のガボール正則化システムを使用した結果を示す。図 7 A 及び図 7 B からわかるように、D R I U の結果と比較すると、ガボール正則化されたセグメンテーション結果には、より多くの白色とより少ない灰色が存在する。

【実施例 2】

【0046】

実施例 2 は、網膜画像のセグメンテーションにおけるノイズ抑制の結果を示す。正確かつ安定したセグメンテーションは、疾患の正確かつ安定した診断に有用である。ここでも、D R I V E 及び S T A R E 画像データセットを様々な手法の訓練及びテストの両方に使用した。表現層には、サイズが 11×11 で、 $0 \sim 180$ 度の方向で均一にサンプリングした 12 個のガボールフィルタを使用した。タスクネットワークでは、セグメンテーションに U n e t アーキテクチャを使用した。ネットワークは、A d a m オプティマイザを活用した標準の確率的勾配降下法を使用して訓練し、学習率は . 0 0 0 5 に設定した。サイズ 128×128 のパッチを訓練画像から抽出した。20 個の訓練画像から合計約 7 0 0 0 個のパッチを抽出した。限定訓練シナリオでは、2 つの画像からのパッチを抽出し、合計で約 7 0 0 個のパッチを抽出した。64 個のパッチのパッチサイズを訓練用に使用した。推論では、ネットワークは、パッチではなく画像全体で使用された。

【0047】

以下の方法は、本明細書及び D R I U に記載されているように、ノイズ正則化曲線セグメンテーションネット (N R C S N) で評価した。表 3 は、推定セグメンテーションマップとグラウンドトゥルスセグメンテーションマップとの間のオーバーラップの尺度である、D i c e オーバーラップ係数を用いた D R I U 技術との比較である。提案されたアプローチは、完全訓練のケースの明確なマージンと、低訓練シナリオにおける、より広いマージンによって D R I U よりも優れている。興味深いことに、スパースな訓練を伴う提案されたアプローチは、完全な訓練を伴う D R I U よりも優れている。

【0048】

【表 3】

表 3

方法	Dice オーバーラップスコア
DRIU	0. 6318
NRCSN	0. 7459

【0049】

図 6 A は、実施例 2 の元の入力を示す。望ましい曲線特徴のグラウンドトゥルスマップを図 6 B に示す。図 8 A は、D R I U を使用した結果セグメンテーションマップを示しており、灰色部分が誤った結果を表し、白色部分が正しい結果を表す。図 8 B は、本明細書に記載の N R C S N システムを使用した結果を示す。図 8 A 及び図 8 B からわかるように、D R I U の結果と比較すると、N R C S N セグメンテーション結果には、より多くの白色とより少ない灰色が存在する。

【0050】

上述の方法は、周知のコンピュータプロセッサ、メモリユニット、記憶デバイス、コンピュータソフトウェア、及び他の構成要素を使用してコンピュータ上に実施することができる。このようなコンピュータの高レベルのブロック図を図 9 に示す。コンピュータ 9 0 0 は、プロセッサ 9 1 0 を含み、プロセッサ 9 1 0 は、そのような動作を定義するコンピュータプログラム命令を実行することにより、コンピュータ 9 0 0 の全体的な動作を制御する。コンピュータプログラム命令は、記憶デバイス 9 2 0 (例えば、磁気ディスク) 内に記憶され、コンピュータプログラム命令の実行が所望されるときにメモリ 9 3 0 にロー

ドされてもよい。したがって、本明細書に記載の方法のステップは、メモリ 930 に記憶されたコンピュータプログラム命令によって定義され、コンピュータプログラム命令を実行するプロセッサ 910 によって制御されてもよい。画像取得デバイス 940 は、コンピュータ 900 に画像データを入力するためにコンピュータ 900 に接続されてもよい。様々な構成によれば、画像取得デバイス 940 は、コンピュータ 900 上に配置される。場合によっては、画像取得デバイス 940 及びコンピュータ 900 は、ネットワークを通じて無線で通信する。様々な実施態様によれば、コンピュータ 900 は、画像取得デバイス 940 から遠隔に位置してもよく、コンピュータ 900 は、サーバ又はクラウドベースのサービスの一部として方法ステップを実行してもよい。コンピュータ 900 は、ネットワークを介して他のデバイスと通信するための 1 つ以上のネットワークインターフェース 950 を含んでもよい。コンピュータ 900 はまた、コンピュータ 900 とのユーザインタラクションを可能にする他の入力/出力デバイス 960 (例えば、ディスプレイ、キーボード、マウス、スピーカ、ボタンなど)を含む。このような入力/出力デバイス 960 は、画像取得デバイス 900 から受信したボリウムに注釈を付ける注釈ツールとしてコンピュータプログラムのセットと組み合わせて使用されてもよい。様々な実施形態によれば、図 9 は、例示目的のためのコンピュータの可能な構成要素の高レベル表現であり、コンピュータは他の構成要素を含むことができる。

【0051】

別途記載のない限り、本明細書及び特許請求の範囲で使用される特徴のサイズ、量、及び物理的特性を表す全ての数字は、用語「約」によって修飾されるものとして理解されるべきである。したがって、反対に指示されない限り、前述の明細書及び添付の特許請求の範囲に記載される数値パラメータは、本明細書に開示される教示を利用して当業者が得ようとする所望の特性に応じて変化し得る近似値である。端点による数値範囲の使用は、その範囲内の全ての数を含む(例えば、1~5 は、1、1.5、2、2.75、3、3.80、4 及び 5 を含む)、及びその範囲内の任意の範囲を含む。

【0052】

上述の様々な実施形態は、相互作用して特定の結果を提供する回路及び/又はソフトウェアモジュールを使用して実装されてもよい。コンピューティング技術の当業者は、当該技術分野において一般的に知られている知識を使用して、モジュールレベル又は全体でのいずれかで、記載された機能を容易に実施することができる。例えば、本明細書に示されるフローチャートは、プロセッサによる実行のためのコンピュータ可読命令/コードを作成するために使用されてもよい。このような命令は、当技術分野で知られているように、コンピュータ可読媒体に記憶され、実行のためにプロセッサに転送されてもよい。上記の構造及び手順は、上記のインクジェット排出器の診断を容易にするために使用することができる実施形態の代表的な例に過ぎない。

10

20

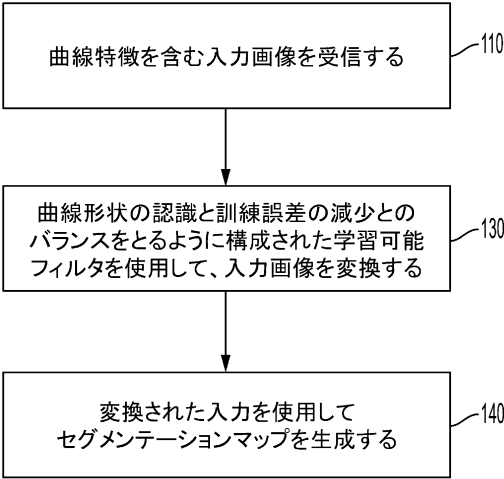
30

40

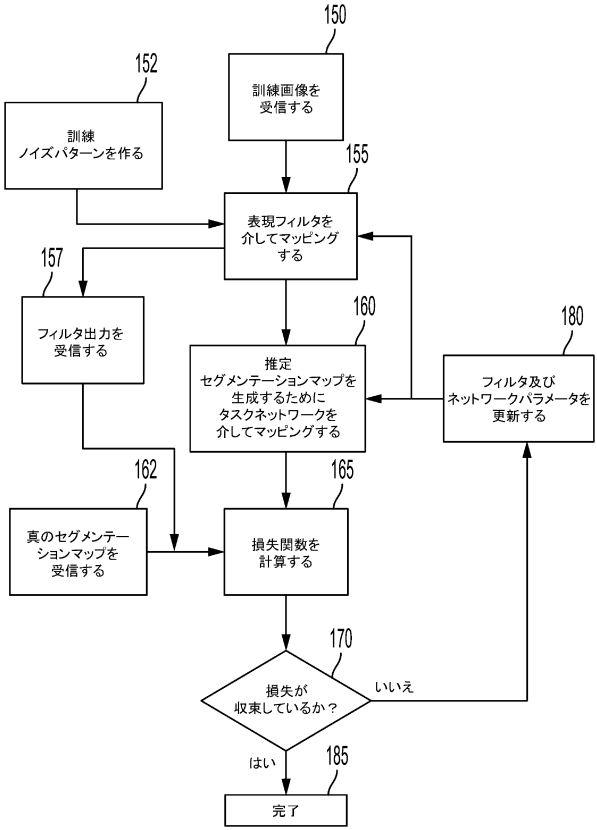
50

【図面】

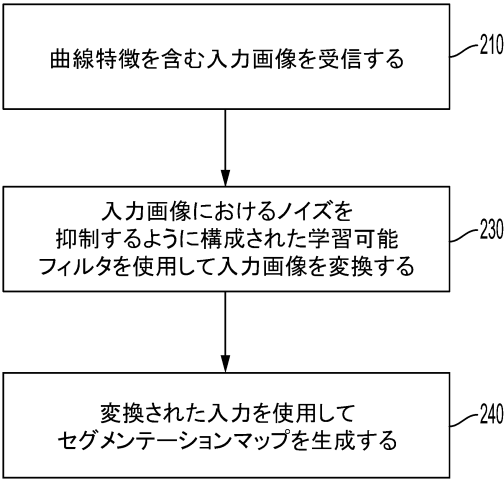
【図 1 A】



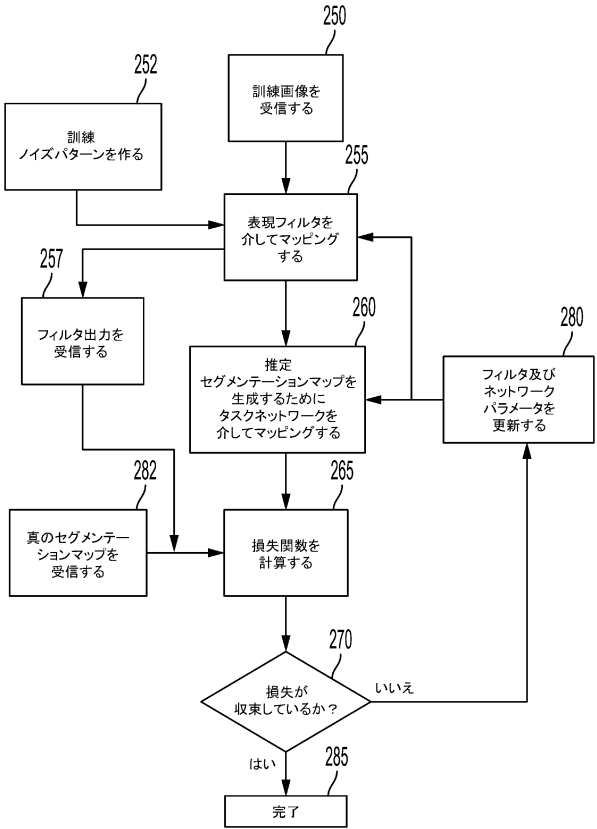
【図 1 B】



【図 2 A】



【図 2 B】



10

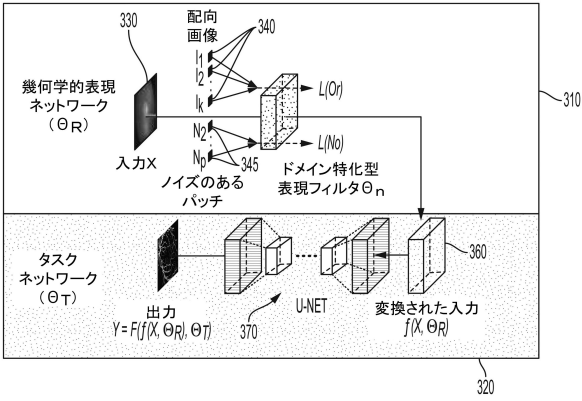
20

30

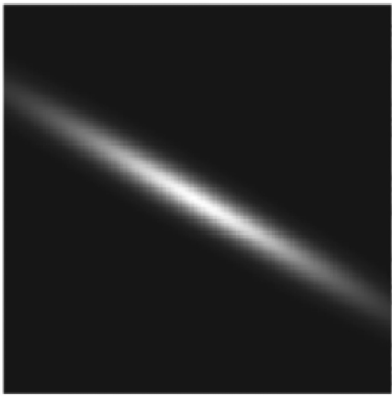
40

50

【図 3】

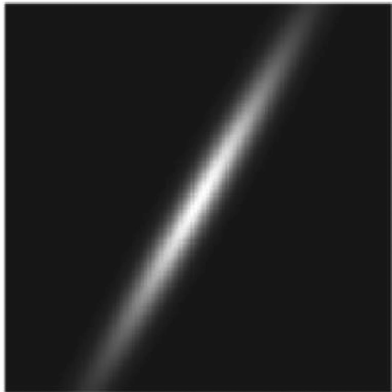


【図 4 A】

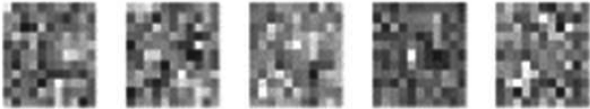


10

【図 4 B】



【図 5 A】

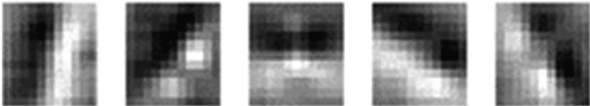


20

【図 5 B】



【図 5 C】



30

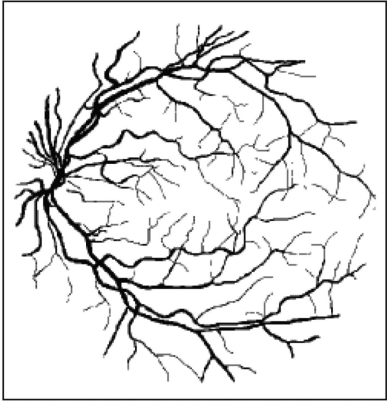
40

50

【図 6 A】

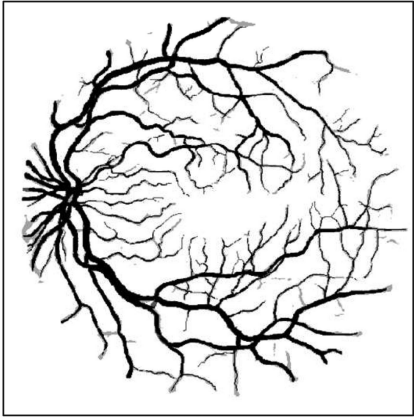


【図 6 B】

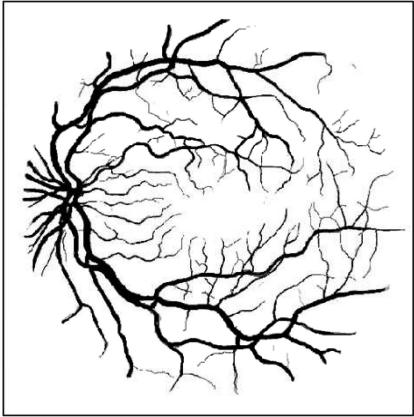


10

【図 7 A】

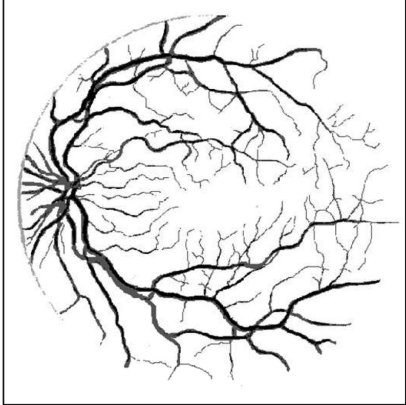


【図 7 B】

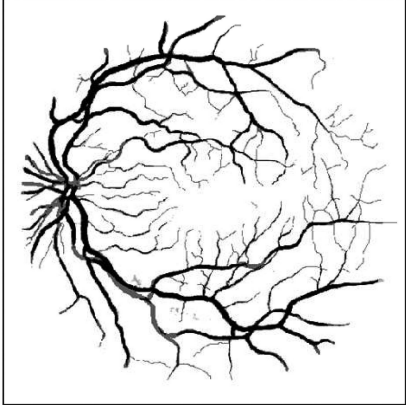


20

【図 8 A】



【図 8 B】

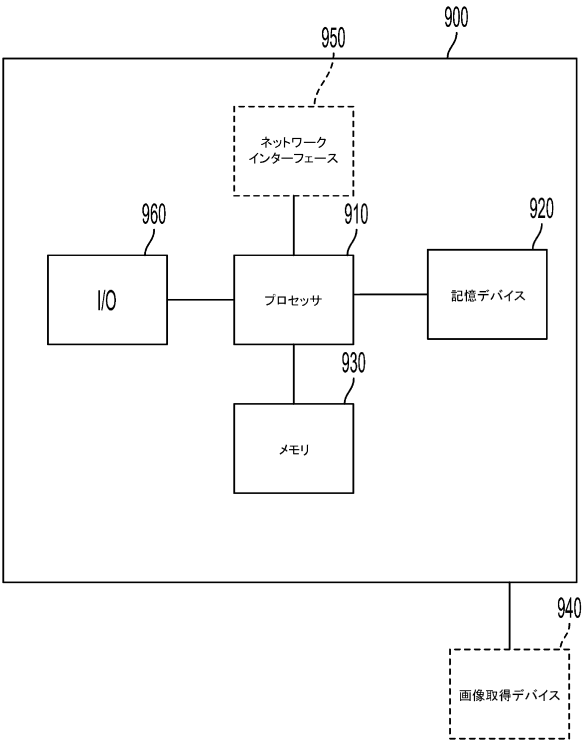


30

40

50

【図 9】



10

20

30

40

50

フロントページの続き

- 弁理士 上杉 浩
(74)代理人 100120525
弁理士 近藤 直樹
(74)代理人 100139712
弁理士 那須 威夫
(72)発明者 ラジャ・バーラ
アメリカ合衆国 ニューヨーク州 1 4 5 3 4 ピッツフォード ウッドグリーン・ドライブ 6 5
(72)発明者 ヴェンカテシュワラオ・チェルクリ
アメリカ合衆国 カリフォルニア州 9 4 4 0 4 フォスター・シティ カタマラン・ストリート 6
1 3 アpartment 3
(72)発明者 ヴィジャイ・クマール・ビー・ジー
アメリカ合衆国 カリフォルニア州 9 5 1 1 8 サンノゼ ブランハム・レーン 1 5 2 0 アパー
トメント 8 1
審査官 山田 辰美
(56)参考文献 Nicola Strisciuglio et al. , Supervised vessel delineation in retinal fundus images with the a
utomatic selection of B-COSFIRE filters , Machine Vision and Applications , Vol.27, no.8 ,
米国 , Springer , 2016年06月24日 , p.1137 - p.1149 , <https://link.springer.com/article/10.1007/s00138-016-0781-7>
(58)調査した分野 (Int.Cl. , D B 名)
G 0 6 T 7 / 0 0 - 7 / 9 0