

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-208073

(P2017-208073A)

(43) 公開日 平成29年11月24日(2017.11.24)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G06F 3/01 (2006.01)</b>	G06F 3/01 510	5B057
<b>G06T 1/00 (2006.01)</b>	G06T 1/00 500A	5B087
<b>G09G 5/00 (2006.01)</b>	G09G 5/00 550C	5C058
<b>G09G 5/10 (2006.01)</b>	G09G 5/10 B	5C122
<b>G09G 5/36 (2006.01)</b>	G09G 5/00 530H	5C182

審査請求 未請求 請求項の数 48 O L 外国語出願 (全 49 頁) 最終頁に続く

(21) 出願番号 特願2017-48499 (P2017-48499)  
 (22) 出願日 平成29年3月14日 (2017.3.14)  
 (31) 優先権主張番号 15/069,829  
 (32) 優先日 平成28年3月14日 (2016.3.14)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 517090495  
 リッチモンド、 ロバート エル  
 アメリカ合衆国 98199 ワシントン  
 州 シアトル パーキンズ レーン ウェ  
 スト 3019  
 (74) 代理人 100083806  
 弁理士 三好 秀和  
 (74) 代理人 100095500  
 弁理士 伊藤 正和  
 (74) 代理人 100111235  
 弁理士 原 裕子  
 (72) 発明者 リッチモンド、 ロバート エル  
 アメリカ合衆国 98199 ワシントン  
 州 シアトル パーキンズ レーン ウェ  
 スト 3019

最終頁に続く

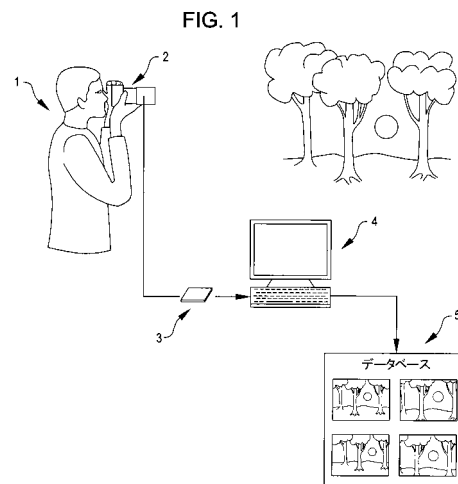
(54) 【発明の名称】 デジタル媒体と観察者の相互作用の構成及び実現

(57) 【要約】 (修正有)

【課題】 画像観察体験を向上させる。

【解決手段】 観察者が見ている場所を追跡することによって、画像処理及び視線検出の概念を結合する。その結果、オリジナルのシーンを見ているような観察体験をもたらす、又は、自動的に、又は写真家により撮られた1つ又は複数の画像の特定部分を観察者が見ると何が起こるべきかについて写真家が前もって指定した意図によって、オリジナルの体験を超えた新しいやり方で観察体験を向上させる。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

1 つ以上のカメラが 1 つ以上のディスプレイスクリーンの前の 1 人以上の観察者の顔の画像をキャプチャすることを介して収集されるデータに応じて動作するアルゴリズムに従って 1 つ以上のディスプレイスクリーン上に表示される 1 つ以上のオリジナルの画像を修正するための方法。

**【請求項 2】**

収集されるデータは、1 人以上の観察者の視線の位置が前記オリジナルの画像内のどこに向けられているか、又は 1 人以上の観察者の顔貌の変化、又は 1 人以上の観察者の識別の 1 つ以上である、請求項 1 に記載の方法。

10

**【請求項 3】**

収集されるデータは、周辺光、又はディスプレイスクリーンからの観察者の距離、又は観察者の数の 1 つ以上である、請求項 1 に記載の方法。

**【請求項 4】**

修正は、前記オリジナルの画像の一部又は全ての輝度を変更することである、請求項 1 に記載の方法。

**【請求項 5】**

修正は、前記オリジナルの画像を代替画像と置換することを含み、前記代替画像のコンテンツは前記オリジナルの画像と同じであるが一部又は全てのピクセルが修正されている、請求項 4 に記載の方法。

20

**【請求項 6】**

修正は、前記オリジナルの画像の全て又は一部の焦点を変更することである、請求項 1 に記載の方法。

**【請求項 7】**

修正は、前記オリジナルの画像を代替画像と置換することであり、前記代替画像のコンテンツは前記オリジナルの画像と同じであるが一部又は全てのピクセルが修正されている、請求項 6 に記載の方法。

**【請求項 8】**

修正は、前記画像の一部又は全てにおいて接近して見えるオブジェクトに対する焦点を変更することである、請求項 6 に記載の方法。

30

**【請求項 9】**

修正は、前記画像の一部又は全てにおいて離れて見えるオブジェクトに対する焦点を変更することである、請求項 6 に記載の方法。

**【請求項 10】**

修正は、前記画像の一部又は全ての見掛けの被写界深度を変更することである、請求項 6 に記載の方法。

**【請求項 11】**

観察者が前記 1 つ以上の画像を凝視しているとき、前記 1 つ以上のカメラから集められたデータの処理は、前記観察者があたかも離れているオブジェクトを見ているかのように目を開散させているかどうかを決定し、そうであれば、修正は、前記画像において近くに見えるオブジェクトが焦点をぼかされ且つ遠くに見えるオブジェクトが鋭く焦点を合わされるように前記画像の部分の焦点を変更することである、請求項 6 に記載の方法。

40

**【請求項 12】**

観察者が前記 1 つ以上の画像を凝視しているとき、修正は、焦点を合わせられたエリアの周囲の前記画像の部分の焦点をぼかし、且つ前記画像を横断して観察者の視線を引くために前記画像を横断して焦点を合わせられたエリアを移動させる、請求項 6 に記載の方法。

**【請求項 13】**

修正は、前記オリジナルの画像の一部を中心としてズームすることによって拡大又は縮小することである、請求項 1 に記載の方法。

50

**【請求項 14】**

修正は、前記オリジナルの画像を代替画像と置換することを含む、請求項 13 に記載の方法。

**【請求項 15】**

修正は、ディスプレイスクリーンから観察者への測定距離の変化と同期して拡大又は縮小することである、請求項 13 に記載の方法。

**【請求項 16】**

修正は、観察者の視線が固定されている場合に、前記オリジナルの画像における場所を中心としてズームすることによって拡大することである、請求項 13 に記載の方法。

**【請求項 17】**

修正は、前記画像の一部又は全ての色バランス又は彩度又は色濃度又はコントラストを変更することである、請求項 1 に記載の方法。

**【請求項 18】**

修正は、前記オリジナルの画像を代替画像と置換することであり、前記代替画像のコンテンツは前記オリジナルの画像と同じであるが一部又は全てのピクセルが修正されている、請求項 17 に記載の方法。

**【請求項 19】**

修正は、前記オリジナルの画像に伴う音声の再生を引き起こすこと又は変更することである、請求項 1 に記載の方法。

**【請求項 20】**

修正は、前記オリジナルの画像内のスプライトの移動又は移動中止を引き起こすことである、請求項 1 に記載の方法。

**【請求項 21】**

修正は、多分岐動画の分岐を選択することである、請求項 1 に記載の方法。

**【請求項 22】**

前記 1 つ以上のオリジナルの画像は静止画である、請求項 1 に記載の方法。

**【請求項 23】**

前記 1 つ以上のオリジナルの画像は動画である、請求項 1 に記載の方法。

**【請求項 24】**

前記 1 つ以上のオリジナルの画像は移動するスプライトを有する静止画である、請求項 1 に記載の方法。

**【請求項 25】**

前記 1 つ以上のオリジナルの画像は 2 次元である、請求項 1 に記載の方法。

**【請求項 26】**

前記 1 つ以上のオリジナルの画像は 3 次元である、請求項 1 に記載の方法。

**【請求項 27】**

前記アルゴリズムは、前記 1 つ以上の画像を選択したアーティストにより決定されるカスタムである、請求項 1 に記載の方法。

**【請求項 28】**

前記アルゴリズムは、前記 1 つ以上のオリジナルの画像を選択したアーティスト又はコンテンツ作成者により選択及び使用されるソフトウェアを供給した会社により事前決定される、請求項 1 に記載の方法。

**【請求項 29】**

前記アルゴリズムは、ソフトウェアを供給した会社により事前決定され、且つ前記アルゴリズムは、任意の画像又は画像の組と協働する、請求項 1 に記載の方法。

**【請求項 30】**

前記アルゴリズムは、段階的变化の修正が行われる速さをアーティスト又はコンテンツ作成者が指定することを可能にする、請求項 1 に記載の方法。

**【請求項 31】**

観察者が画像内のスクリーンの一部にタッチすることによってタッチから収集されるデ

10

20

30

40

50

ータに応じて動作するアルゴリズムに従って１つ以上のディスプレイスクリーン上に表示される１つ以上のオリジナルの画像を修正する方法であって、

修正は、１つ以上の代替画像と前記オリジナルの画像の置換であり、前記代替画像のコンテンツは前記オリジナルの画像と同じであるが一部又は全てのピクセルが修正されている、方法。

【請求項３２】

タッチから収集される前記データは、１本以上の指によるタッチ、タッチの力、タッチの持続時間、及びタッチの移動の１つ以上を含む、請求項３１に記載の方法。

【請求項３３】

修正は、前記オリジナルの画像の全て又は一部の焦点を変更することである、請求項３１に記載の方法。

【請求項３４】

修正は、

前記画像の一部又は全てにおいて接近して見えるオブジェクトに対する焦点を変更すること、又は

前記画像の一部又は全てにおいて離れて見えるオブジェクトに対する焦点を変更すること、又は

前記画像の一部又は全ての見掛けの被写界深度を変更すること  
の１つ以上を含む、請求項３３に記載の方法。

【請求項３５】

修正は、

前記オリジナルの画像の一部又は全ての輝度を変更すること、又は

前記オリジナルの画像の一部から拡大又は縮小すること、又は

前記オリジナルの画像の一部又は全ての色バランス又は彩度又は色濃度又はコントラストを変更すること、又は

前記オリジナルの画像内のスプライトの移動又は移動中止を引き起こすこと、又は  
多分岐動画の分岐を選択すること

の１つ以上を含む、請求項３１に記載の方法。

【請求項３６】

前記１つ以上のオリジナルの画像は、

静止画、又は

動画、又は

移動するスプライトを有する静止画、又は

２次元、又は

３次元

である、請求項３１に記載の方法。

【請求項３７】

アルゴリズムは、

前記１つ以上の画像を選択したアーティストにより決定されるカスタム、又は

前記１つ以上のオリジナルの画像を選択したアーティスト又はコンテンツ作成者により  
選択及び使用されるソフトウェアを供給した会社により事前決定されること、又は

ソフトウェアを供給した会社により事前決定され、前記アルゴリズムは任意の画像又は  
画像の組と協働すること、又は

段階的变化の修正が行われる速さをアーティスト又はコンテンツ作成者が指定すること  
を可能にすること

の１つ以上を含む、請求項３１に記載の方法。

【請求項３８】

声音から収集されるデータに応じて動作するアルゴリズムに従って１つ以上のディスプレイスクリーン上に表示される１つ以上のオリジナルの画像を修正する方法であって、修正  
は１つ以上の代替画像と前記オリジナルの画像の置換であり、前記代替画像のコンテン

10

20

30

40

50

ツは前記オリジナルの画像と同じであるが一部又は全てのピクセルが修正されている、方法。

【請求項 39】

修正は、

前記オリジナルの画像の一部又は全ての輝度を変更すること、

前記オリジナルの画像の一部又は全ての焦点を変更すること、

前記オリジナルの画像の一部又は全てから拡大又は縮小すること、

前記オリジナルの画像の一部又は全ての色バランス又は彩度又は色濃度又はコントラストを変更すること、

前記オリジナルの画像の一部又は全ての鮮明さを変更すること、

前記オリジナルの画像内のスプライトの移動又は移動中止を引き起こすこと、多分岐動画の分岐を選択すること

の 1 つ以上を含む、請求項 38 に記載の方法。

10

【請求項 40】

声音から収集されるデータから、話している人物の推定の識別が決定され、修正は前記推定の識別によって少なくとも部分的に決定される、請求項 38 に記載の方法。

【請求項 41】

ディスプレイスクリーンの筐体に埋め込まれた 1 つ以上の加速度計から収集されるデータに応じて動作するアルゴリズムに従ってハンドヘルドディスプレイスクリーン上に表示されるオリジナルの画像を修正する方法であって、修正は、1 つ以上の左境界、右境界、上部境界又は下部境界の各々を拡大又は縮小することによってオリジナルの画像の視野を変更することである、方法。

20

【請求項 42】

前記 1 つ以上の加速度計から収集される前記データは、観察者から離れているディスプレイの第 1 の端部の反対の端部に対する傾斜であり、傾斜は前記第 1 の端部に近い前記画像のより多くを視界に入れさせる、請求項 41 に記載の方法。

【請求項 43】

修正は、前記オリジナルの画像の一部を中心としてズームすることによって拡大又は縮小することである、請求項 41 に記載の方法。

【請求項 44】

サーバコンピュータ及びクライアントコンピュータを有するシステムにおいて、著者から指導を受信し、前記指導に基づいて、観察者のアクションに基づいて変化する画像を示すデータセットを生成するための方法であって、

30

(a) 前記サーバコンピュータにおいて、一連の複数の画像の仕様をクライアントコンピュータから受信するステップと、

(b) 前記サーバコンピュータにおいて、前記一連の画像におけるオリジナルの画像から前記一連の画像における第 2 の画像への遷移をトリガする収集されるデータの仕様を前記クライアントコンピュータから受信するステップと、

(c) 前記一連の画像における前記オリジナルの画像から前記一連の画像における前記第 2 の画像への遷移の速さの仕様を前記クライアントコンピュータから受信するステップと

40

(d) 観察者のコンピュータに送信され得るデータセットを前記サーバコンピュータにおいてアセンブルするステップであって、前記データセットは前記観察者のコンピュータ上で観察されることが可能であり、前記観察者のアクションから受信されるデータ入力は、一連のオリジナルの画像を指定された遷移の速さで前記第 2 の画像と置換することによって観察される画像の修正を引き起こすステップと

を含む、方法。

【請求項 45】

前記データ入力は、

前記オリジナルの画像内の 1 人以上の観察者の視線の位置、又は

50

1人以上の観察者の顔貌の変化、又は  
周辺光、又は  
観察者のディスプレイスクリーンからの距離、又は  
観察者の数、又は  
音声入力、又は  
タッチ入力  
の1つ以上を含む、請求項44に記載の方法。

【請求項46】

修正は、

前記オリジナルの画像の一部又は全ての輝度を変更すること、

前記オリジナルの画像の一部の焦点を変更すること、

前記オリジナルの画像の一部から拡大又は縮小すること、

前記オリジナルの画像の一部又は全ての色バランス又は彩度又は色濃度又はコントラストを変更すること、

前記オリジナルの画像内のスプライトの移動又は移動中止を引き起こすこと、

多分岐動画の分岐を選択すること

の1つ以上を含む、請求項44に記載の方法。

【請求項47】

前記オリジナルの画像は、

静止画、又は

動画、又は

移動するスプライトを有する静止画、又は

2次元、又は

3次元

である、請求項44に記載の方法。

【請求項48】

前記クライアントコンピュータ及び前記サーバコンピュータは、単一のコンピュータ筐体の中に含まれ、且つ単一ユーザによりアクティブ化される、請求項44に記載の方法。

【発明の詳細な説明】

【背景技術】

【0001】

今日のデジタルカメラ及びスマートフォンはコンピュータパワーを使用して、他の装置の画面上で即座に又は後で見るために画像を向上させる。一例はHDR又はハイダイナミックレンジであり、カメラは異なる露出で複数の写真を素早く撮り、全ての部分が、最も明るい及び最も暗い部分でさえも露出されて全ての細部を引き出す画像を作成する。

【0002】

また、目に対してより魅力的にするために、又はどのくらいの時間（例えば、平均時間、合計時間）を何人の観察者がそれに費やしているかに関して広告の価値を測るために、観察者がウェブページ上の所定の広告をどのくらいの長さ見ているかを決定するために視線検出を使用する既存のシステムも有る。また、動画の再生を停止してバッテリー電源を節約するか又は観察者が動画の何れの部分も見逃さないように、観察者がもはやスマートフォンのディスプレイ等のディスプレイを見ていないかどうかを確認するために視線検出を使用するシステムも有る。

【発明の概要】

【0003】

実施形態は、観察者が見ている場所を追跡することによって画像観察体験を向上させるためにこうした2つの概念を結合する。その結果、オリジナルのシーンを見ているような体験をもたらす、又は、自動的に、又は写真家により撮られた1つ又は複数の画像の特定の部分を観察者が見ると何が起こるべきかについて写真家が予め指定した意図によって、

10

20

30

40

50

オリジナルの体験を超えた新しいやり方でそれを向上させる。

【 0 0 0 4 】

実施形態は、アーティストがディスプレイスクリーン用のオリジナルの画像又は複数のディスプレイスクリーン用の複数のオリジナルの画像を作成することを可能にし、それは観察者が誰であるか又は観察者が何をするかに応じて、観察者が画像を見るときに変化する。画像を変化させる観察者のアクションは、観察者がどこに視線を向けるか又はスクリーンをタッチするか、又は観察者が何を言うか又は観察者がディスプレイからどれくらい離れているか、又は何人の観察者が居るか、又は特定の期待観察者の識別等であってもよい。

【 0 0 0 5 】

実施形態は、1つ以上のカメラが1つ以上のディスプレイスクリーンの前の1人以上の観察者の顔の画像をキャプチャすることを介して収集されるデータに応じて1つ以上のディスプレイスクリーン上に表示される1つ以上のオリジナルの画像を修正するための方法である。収集されるデータは、1人以上の観察者の視線の位置がオリジナルの画像内のどこに向けられているか、又は1人以上の観察者の顔貌の変化、又は周辺光の量、又は観察者のディスプレイスクリーンからの距離、又は観察者の数、又は観察者の識別の1つ以上であってもよい。

【 0 0 0 6 】

[ 輝度 ]

修正は、オリジナルの画像の一部又は全ての輝度を変更することであってもよい。輝度の変更は、オリジナルの画像を代替画像と置換することによってもたらされることができ、代替画像のコンテンツはオリジナルの画像と同じであるが代替画像の少なくとも一部の輝度がオリジナルの画像の少なくとも対応する部分の輝度とは異なるように一部又は全てのピクセルが修正されている。例えば、一对の写真画像がカメラで撮られてもよく、画像は異なる露出を除いて基本的に同じである。1枚の画像の一部又は全てが、輝度を変更するために他の画像の一部又は全てを置換し得る。代替的に、アルゴリズムは、1人以上の観察者によるアクションに応じて、全ての画像ピクセルではないが一部の画像ピクセルの数値輝度を調節することができる。

【 0 0 0 7 】

[ 焦点 ]

修正は、オリジナルの画像の一部又は全ての焦点を変更することであってもよい。変更は、オリジナルの画像と代替画像との置換によってもたらされることができ、代替画像のコンテンツはオリジナルの画像と同じであるが代替画像の少なくとも一部の焦点がオリジナルの画像の少なくとも対応する部分の焦点とは異なるように一部又は全てのピクセルが修正されている。修正は、画像の一部又は全てにおいて接近して見えるオブジェクトに対する焦点を変更することであってもよい。修正は、画像の一部又は全てにおいて離れて見えるオブジェクトに対する焦点を変更することであってもよい。修正は、画像の一部又は全ての見掛けの被写界深度を変更することであってもよい。修正は、オリジナルの画像の一部を中心としてズームすることによって拡大又は縮小することであってもよい。修正は、代替画像とオリジナルの画像の置換を含んでもよい。修正は、ディスプレイスクリーンから観察者への測定距離の変化と同期して拡大又は縮小することであってもよい。

【 0 0 0 8 】

[ 色 ]

修正は、画像の一部又は全ての色バランス又は彩度又は色濃度又はコントラストを変更することであってもよい。これは、オリジナルの画像と代替画像との置換によってもたらされることができ、代替画像のコンテンツはオリジナルの画像と同じであるが代替画像の少なくとも一部の色（例えば、色バランス又は彩度）がオリジナルの画像の少なくとも対応する部分の色とは異なるように一部又は全てのピクセルが修正されている。

【 0 0 0 9 】

[ 鮮明さ ]

修正は、画像の一部又は全ての鮮明さを変更することであってもよい。これは、オリジナルの画像と代替画像との置換によってもたらされることができ、代替画像のコンテンツはオリジナルの画像と同じであるが代替画像の少なくとも一部の鮮明さがオリジナルの画像の対応する部分の鮮明さとは異なるように一部又は全てのピクセルが修正されている。

【 0 0 1 0 】

[ 音声出力 ]

修正は、オリジナルの画像に伴う音声の再生を引き起こすこと又は変更することであってもよい。

【 0 0 1 1 】

[ スプライト ]

修正は、オリジナルの画像内のスプライトの移動又は移動中止を引き起こすことであってもよい。

【 0 0 1 2 】

[ アニメーション G I F ( G r a p h i c s I n t e r c h a n g e F o r m a t ) ]

修正は、アニメーション G I F を共に形成する少数の画像と画像の全て又は一部を置換することであってもよい。

【 0 0 1 3 】

[ 動画分岐 ]

修正は、多分岐動画の分岐を選択することであってもよい。

【 0 0 1 4 】

上記の実施形態の何れかにおいて、1つ以上のオリジナルの画像は、静止画、又は動画、又は移動するスプライトを有する静止画であってもよい。オリジナルの画像は、2次元又は3次元であってもよい。

【 0 0 1 5 】

上記の実施形態の何れかにおいて、1つ又は複数のオリジナルの画像の修正のためのアルゴリズムは、1つ以上の画像を選択したアーティスト又は構成者又はコンテンツ作成者により決定されるカスタムであってもよい。修正のためのアルゴリズムは、1つ以上のオリジナルの画像を選択したアーティスト又はコンテンツ作成者により選択及び使用されるソフトウェアを供給した会社により事前決定されてもよい。アルゴリズムは、ソフトウェアを供給した会社により事前決定されてもよく、アルゴリズムが任意の画像又は画像の組と協働するようなものであってもよい。アルゴリズムは、段階的变化の修正（例えば、輝度の段階的变化、焦点の段階的变化、色の段階的变化、微小（又は部分）画像から完全画像へと新しい画像を拡大すること、又は他の遷移方法）が行われる速さをアーティスト又はコンテンツ作成者が指定することを可能にしてもよい。

【 0 0 1 6 】

[ タッチ又はマウス入力 ]

別の実施形態は、例えば、観察者が画像内のスクリーンの一部をタッチ又はポインティングすることによってタッチ、マウス又は他の入力装置から収集されるデータに応じて動作するアルゴリズムに従って1つ以上のディスプレイスクリーン上に表示される1つ以上のオリジナルの画像を修正する方法であって、修正は1つ以上の代替画像とオリジナルの画像の置換であり、代替画像のコンテンツはオリジナルの画像と同じであるが一部又は全てのピクセルが修正されている。マウスにより収集されるデータは、クリック又はマウスオーバ又はクリック及びドラッグ及び移動の1つ以上により識別されるデータを含んでもよい。タッチから収集されるデータは、1本以上の指によるタッチ、タッチの力、タッチの持続時間、タッチの移動、及びユーザの手／指が画像ディスプレイスクリーンに接触しないが接近するジェスチャ（例えば、ピンチ操作、ウェービング、ポインティング）の1つ以上によって識別されるデータを含んでもよい。修正は、上記の修正の何れかであってもよい。1つ以上のオリジナルの画像は、静止画又は動画又は移動するスプライトを有する静止画、又は2次元又は3次元であってもよい。アルゴリズムは、上記の要素の何れか

10

20

30

40

50



を含んでもよい。

【 0 0 1 7 】

[ 音声入力 ]

別の実施形態は、声音から収集されるデータに応じて動作するアルゴリズムに従って 1 つ以上のディスプレイスクリーン上に表示される 1 つ以上のオリジナルの画像を修正する方法であって、修正は 1 つ以上の代替画像とオリジナルの画像の置換であり、代替画像のコンテンツはオリジナルの画像と同じであるが一部又は全てのピクセルが修正されている。修正は、上記の修正の何れかであってもよい。1 つ以上のオリジナルの画像は、静止画又は動画又は移動するスプライトを有する動画、又は 2 次元又は 3 次元であってもよい。アルゴリズムは、上記の要素の何れかを含んでもよい。

10

【 0 0 1 8 】

[ 加速度計入力 ]

別の実施形態は、ディスプレイスクリーンの筐体に埋め込まれた 1 つ以上の加速度計から収集されるデータに応じて動作するアルゴリズムに従ってハンドヘルドディスプレイスクリーン上に表示されるオリジナルの画像を修正する方法であって、修正は、1 つ以上の左境界、右境界、上部境界又は下部境界の各々を拡大又は縮小することによってオリジナルの画像の視野を変更することである。1 つ以上の加速度計から収集されるデータは、観察者から離れているディスプレイの第 1 の端部の反対の端部に対する傾斜であり、この傾斜は第 1 の端部に近い画像のより多くを視界に入れさせる。修正は、オリジナルの画像の（必ずしも中心ではない）部分を中心としてズームすることによって拡大又は縮小することであってもよい。

20

【 0 0 1 9 】

[ オーサリングツール ]

別の実施形態は、著者から指導を受信し、こうした指導に基づいて、観察者のアクションに基づいて変化する画像を示すデータセットを生成するためのサーバコンピュータ及びクライアントコンピュータを有するシステムにおける方法である。方法は、( a ) サーバコンピュータにおいて、一連の複数の画像の仕様をクライアントコンピュータから受信するステップと、( b ) サーバコンピュータにおいて、一連の画像におけるオリジナルの画像から一連の画像における第 2 の画像への遷移をトリガするべく収集されるデータの仕様をクライアントコンピュータから受信するステップと、( c ) 一連の画像におけるオリジナルの画像から一連の画像における第 2 の画像への遷移の速さの仕様をクライアントコンピュータから受信するステップと、( d ) 観察者のコンピュータに送信され得るデータセットをサーバコンピュータにおいてアセンブルするステップとを含み、データセットは観察者のコンピュータ上で観察されることが可能であり、観察者のアクションから受信されたデータ入力は、一連のオリジナルの画像を指定された遷移の速さで第 2 の画像と置換することによって観察画像の修正を引き起こす。

30

【 0 0 2 0 】

データ入力は、限定されないが、オリジナルの画像内の 1 人以上の観察者の視線の位置、又は 1 人以上の観察者の顔貌の変化、又は周辺光、又は観察者のディスプレイスクリーンからの距離、又は観察者の数、又は観察者の識別、又は音声入力、又はタッチ入力の 1 つ以上を含んでもよい。

40

【 0 0 2 1 】

修正は、限定されないが、オリジナルの画像の一部又は全ての輝度を変更すること、オリジナルの画像の一部の焦点を変更すること、オリジナルの画像の一部から拡大又は縮小すること、オリジナルの画像の一部又は全ての色バランス又は彩度又は色濃度又はコントラストを変更すること、オリジナルの画像内のスプライトの移動又は移動の中止を引き起こすこと、多分岐動画の分岐を選択することを含んでもよい。

【 0 0 2 2 】

オリジナルの画像は、限定されないが、静止画又は動画又は移動するスプライトを有する動画、又は 2 次元又は 3 次元であってもよい。

50

## 【 0 0 2 3 】

クライアントコンピュータ及びサーバコンピュータは、各々の単一のコンピュータ筐体の中に含まれる各ソフトウェアプログラムをそれぞれ含んでもよく、単一ユーザによりアクティブ化されてもよい。

## 【 図面の簡単な説明 】

## 【 0 0 2 4 】

【 図 1 】一実施形態に従って、写真家が単一対象の複数画像を撮影することを示す。

【 図 2 】一実施形態に従って、一組の画像に対する露出マップを作成するためのフローチャートである。

【 図 3 】一実施形態に従って、一組の画像に対する焦点マップを作成するためのフローチャートである。

【 図 4 】一実施形態に従って、一組の画像に対する露出マップがどのように記憶されるかを示す。

【 図 5 】一実施形態に従って、観察者が画像の異なる部分を見るとどのように画像が変化するかを構成者がどのように決定するかを示す。

【 図 6 】一実施形態に従って、観察者が画像の異なる部分を見ると画像を変化させるシステムを示す。

## 【 発明を実施するための形態 】

## 【 0 0 2 5 】

本発明の 1 つ以上の実施形態は、自動的に、又は、写真家 / 構成者により前もって決定され及び、例えば、観察者が画像内のどこを見ているかによって後で駆動されるように、写真を向上させ、例えば、輝度、焦点、被写界深度及び他の品質の次元で観察者と静止画及び動画をインタラクティブにすることである。

## 【 0 0 2 6 】

最も単純且つ限定的な形態では、実施形態は、オリジナルのシーンが露出又は深度について非常に大きな変化を有している写真の中でキャプチャされたシーンの主題を誰かがどのように見るかを再現する。以下の例はこの概念を説明する。

## 【 0 0 2 7 】

## 1 . 輝度

背後で夕日が沈んでいる木立を想像する。観察者が太陽の方を見ると、虹彩が収縮し、それによって瞳孔はサイズが減少し、より少ない光を入らせ且つ輝度を低減させ、観察者は、赤い太陽の残りの半円（木により遮られていない太陽の部分）を見るが、木が黒いシルエットになるのを見る。しかしながら、単一の特定の木をみると、瞳孔は再び開いて、木の皮の特徴及びテクスチャが見えるようになる。

## 【 0 0 2 8 】

このようなシーンの通常の従来の写真は、沈む夕日の形状を見るのに十分に暗く、この場合、木は黒いか、又は木の皮を見るのに十分に明るく、この場合、夕日の背景は太陽が無く白い「ブローンアウト」エリアになるかの 1 つの露出に限定されるであろう。

## 【 0 0 2 9 】

この問題を補償することを意図した既存の写真法があり、この方法は“ H D R ”又はハイダイナミックレンジと呼ばれている。H D R では、動きを回避するために複数の露出が間断なく撮られ、画像の全てのエリアが正しい露出を有するように自動的に混合される。しかしながら、H D R 補償した画像は「フェイク」又は「サッカリン」又は「非現実的」に見えることが多く、従って、観察者にとって不満足となることが多い。

## 【 0 0 3 0 】

その代わり、実施形態は、観察者自身の瞳孔露出システムのアクションを再現する。H D R と同様に、それは、撮られたときに使用される露出のみが異なる一群の写真を使用する。しかしながら、写真の何れか 1 つが表示されると、（H D R 技術と比較して）全く異なることが起こる。

## 【 0 0 3 1 】

10

20

30

40

50

例えば、太陽が正しく露出された写真がモニタ上で最初に明るく表示される。次に、観察者がこの例の木のように最初に露出不足だった画像のより暗い部分を見ると、システムの視線検出部が画像のより暗い部分への観察者の視線の移行を検出して、システムは木のエリアがより露出されている組から別の写真を探して表示する。

#### 【0032】

このようにして、観察者が非常に暗いエリアを見ているので観察者の瞳孔が拡大するとき、観察者はより多くの画像の細部が発展するという報酬を与えられる。太陽を再び見ることは反対を行う。即ち、エリアは非常に明るくてブローンアウトされているが、シルエットになった木の間で赤く沈む太陽へと迅速に落ち着く。このようにして、オリジナルのシーンを見る「感覚」が取り戻され且つ再現され、観察の体験が向上する。

10

#### 【0033】

実施形態は、静止画の写真を撮るときに決定する単なる輝度を超える広範なカメラ及び表示の変化を記録するやり方を写真家／構成者に与えると共に、例えば、画像ディスプレイスクリーン上の視線検出又はタッチ操作又はマウス操作によって観察者が自然にこうした変化をトリガし且つ探索することを可能にする。トリガは、手動、自動又は事前設定した意図に従ったものであり得る。

#### 【0034】

代替的に、スワップに利用可能な複数の画像と画像スワップする代わりに、アルゴリズムは、1人以上の観察者によるアクションに応じて画像の全ての画像ピクセルではないが一部の画像ピクセルの輝度を（例えば、数値的に）調節し得る。

20

#### 【0035】

##### 2. [焦点]

同じ方法が焦点に使用されてもよい。例えば、シーンが様々な距離に有るオブジェクトを有する場合、通常の写真は、全てのオブジェクトに完全に焦点を合わせることができず、少なくとも有限の被写界深度及び有限の焦点距離を有する写真を撮ったカメラを与えられる。従って、実施形態は、写真内のどのオブジェクトを観察者の目が見ているのかを追跡して、それぞれ異なる焦点で取られた多くの異なる写真の組から、そのオブジェクトに対して正確に焦点が合わされた写真を選択することができる。

#### 【0036】

##### 3. 「窓への歩み寄り」

30

観察者がディスプレイ全体を占めない写真又は他の画像の表示に向かって歩くと、観察者とディスプレイとの間の距離の変化から目に対する角度の増加が計算され、ディスプレイに設置されたカメラにより距離の変化が測定され、写真のより多くを同時に表示しながら写真の外フレームが増加され得る。観察者が実際の窓に接近するとより多くのコンテンツが現れるが、観察者が追加のコンテンツを見るためには窓の左から右まで「見渡す」必要があるという点で、この技術は実際の窓に歩み寄ることをシミュレートしている。

#### 【0037】

##### 4. 後処理効果：クロッピング／スクローリング

カメラ設定が変更され、写真が撮られた後で、写真家は、画像がどのように表示されるべきかを決定する。例えば、クロッピングにおいて、写真家は、観察者に提示するために画像のより芸術的なサブセットを選ぶ。しかしながら、写真家は、以下のやり方で、観察者が興味を有する場合に画像のより多くが見られることを意図してもよい。しかしながら、観察者の目が見下ろすか又は見上げる場合、画像は上下にスクロールされ得る。正しく行われると、これは、画像のコンテンツを覗き込む体験を提供し得る。例えば、遠い山の写真において、観察者は見下ろすと、視点に至る小道を見るであろう。言い換えれば、観察者の前の小道の一部がフレームの大部分を占めるように、水平線が遠い山から小道の数フィート前にシフトするであろう。即ち、観察者が見下ろすと画像は上にスクロールされ、観察者が見上げると画像は下にスクロールされ得る。同様に、観察者が左又は右を見る場合、画像は右又は左にスクロールされ得る。

40

#### 【0038】

50

ディスプレイが上下又は横方向に動かされる場合、画像をスクロールさせるために加速度計データを使用することによってハンドヘルドディスプレイにより同じ効果が達成され得る。即ち、ハンドヘルドディスプレイは、画像が窓よりも大きい場合に、観察者が画像の所望の部分を観察するために動かすことができる携帯窓を模倣するように作られることが可能である。

#### 【0039】

##### 5. シャッター時間及び移動

複数のカエデの翼果が下降している静止画写真を想像する。一実施形態では、短い露出時間の複数の写真が、より長い時間を有する写真と共に撮られ得る。観察者は、より長い露出時間を有する写真における個々の種子を最初に見るとき、中心の周囲がぼやけた円としてその翼弁を見る。しかしながら、観察者は、それを見つめるとき、間断なく短い露出時間を有する写真（又は種子を含む写真の部分）によって形成される複数のフレームを見せられ、種子が実際に素早く回転し且つ僅かに落ちてからその場所で動かなくなるのを見る（例えば、短い露出時間を有する写真によって形成される最後のフレームが表示されるとき）。観察者が種子を含む写真の部分から目を逸らしている間、種子はその初期位置及びぼやけた外観に戻るであろう（例えば、より長い露出時間を有する写真又はその一部が再び表示されるであろう）。次に、上記のサイクルは、観察者が種子を再び見るときに繰り返されるであろう。

#### 【0040】

代替的に、観察者が、上記のカエデの種子等の移動しているかもしれないように見えるオブジェクトを見るとき、静止画像の一部がオブジェクトが移動していることを示す動画又はスプライト又はアニメーションGIFと置換され得る。

#### 【0041】

また、これは、体操選手又はスキーヤー又はジャンパーが単一のバックグラウンドに対して複数の位置でキャプチャされる「連続写真術 (Sequential Photography)」にも適用される。一実施形態では、1つの位置だけが示されるであろう。観察者がディスプレイを見るとき、他のショットが示され、フレームを横断する各位置を通して、そして最後のものから逆行して目を案内するであろう。単一の静止画像は移動を描く一連の画像よりも典型的には少なく表されるので、これは、キャプチャ対象が実際にどのように移動しているか見たい人物（例えば、コーチ又は医者）にとっては価値があるであろう。

#### 【0042】

##### 6. 顔貌

人物が微笑んでいるか又は眉をひそめているか又は他の特定の感情を示しているかを決定するために画像認識が使用され得る。このような顔の変化は、画像の一部又は画像の全ての変化を引き起こし得る。画像認識は観察者の識別を決定するために使用可能であり、決定された識別は画像を変化させ得る。

#### 【0043】

##### 7. ディスプレイから観察者の距離

ディスプレイから観察者の距離を計算するためにディスプレイ上に設置されたカメラが使用されてもよく、距離は画像の一部又は全てへの変化を引き起こすために使用され得る。（距離を測定する方法は、焦点の詳細に赤外線パルスを使用する光学方法又は音波パルスを使用する超音波方法等、オートフォーカスを有するカメラにより使用される任意の複数の方法であり得る。）即ち、オートフォーカスのためにカメラが使用し得る距離測定機構は、カメラが設置される又は他のやり方で関連付けられるディスプレイスクリーンから観察者の距離を決定するために利用され得る。

#### 【0044】

##### 8. 観察者によるタッチ

観察者からタッチ入力を受信するためにディスプレイの表面のタッチセンサ式入力レイヤが使用されてもよく、このタッチ入力は画像の一部又は全てへの変化を引き起こすため

10

20

30

40

50

に使用され得る。

【 0 0 4 5 】

9 . 観察者からの声音

観察者から音声入力を受信するためにディスプレイに設置される又は他のやり方で関係付けられるマイクロホンが使用されてもよく、この音声入力は画像の一部又は全てへの変化を引き起こすために使用され得る。声音は観察者を識別するために使用されてもよく、観察者の識別に基づいて画像が修正され得る。

【 0 0 4 6 】

1 0 . ディスプレイの加速度

ディスプレイがハンドヘルドである場合、ディスプレイを動かしている観察者から入力を受信するために加速度計からの入力を使用されてもよく、入力は画像の一部又は全てへの変化を引き起こすために使用され得る。

【 0 0 4 7 】

1 1 . 周辺光

周辺光の量を決定するためにディスプレイ上に設置されるカメラ又は他の光センサが使用されてもよく、これは画像の輝度等、画像の一部又は全てへの変化を引き起こすために使用され得る。

【 0 0 4 8 】

1 2 . 更なる機能

先の例は、既存の撮像技術よりも高い忠実性で、オリジナルのシーンがどのように見えたかを再現しようとすることで、オリジナルのシーンを再現する観察体験を向上させる。しかしながら、こうした処理技術を正確なシーン再現に限定する理由はない。また、こうした技術は、作品に何を意図したかを写真家 / 構成者が観察者に案内するための芸術ツールとしての役割を果たし得る。例えば、構成者は観察者が写真の一部を始めてみる時に、2 回目とは異なる輝度及び焦点が有るべきこと、又は任意の後続の回で、観察者が写真の一部を見ることを意図してもよい。又は、構成者は、観察者が集中している画像の部分に暗示される概念又は意味を強調するために観察者が観ている場所の近くの画像の部分に意図的にぼやけさせ又は焦点をぼかすか、又は観察者の集中を意図的に逸らしてもよい。

【 0 0 4 9 】

[ 詳細な例 ]

この節では、実施形態が画像を観察する体験を向上させること、及び他人による画像の観察に関する意図をより完全に表現するためのツールを写真家に与えることという全体的な目的をどのように達成するかが示される。また、関連するコンポーネント及びそれらがどのように協働するかも示される。この説明は、写真の可能な変化のサブセットを詳細に示しているが、同じ基本処理を使用して上記のように他の変化に容易に拡張され得る。従って、説明は実施形態の非限定的な例であることが意図されている。

【 0 0 5 0 】

図 1 は、単純な実施形態における画像及び前処理段階を示す。手動で又はカメラ内プログラムを使用してシーンの一組の画像を取るカメラ 2 を使用する写真 1 が示されており、輝度、焦点及びシャッター時間等、様々な写真パラメータが各画像ごとに变化されている。

【 0 0 5 1 】

写真は、セキュア・デジタル ( S D ) カード 3 又は任意の他の転送機構等を介してコンピュータ 4 に転送され ( 又はキャプチャ機構及び回路は既にコンピュータであり又は既にコンピュータを含んでおり ) 、ラベルされてデータベース 5 に記憶されてもよい。次に、画像は自動的に解析されて、各々が観察者の期待されるアクションに関係付けられる。例えば、1 つの画像において露出不足又は露出過度である及び別の画像において標準的に露出された各画像の部分が自動的に注記される。「露出不足」は、写真を撮られたシーンの細部が黒又は明度ゼロと区別できないことを意味する。「露出過度」は、オリジナルの細部が白又は明度 1 0 0 % と区別できないことを意味する。先の例の夕日の正面の木の皮を

参照すると、他の画像の同じ場所において標準的に露出されたエリアと比較して、木の皮のエリアにおける暗部を有する画像が注記されて、その情報は保存される。後になって観察者が画像の露出不足の場所を見ると、標準的に露出された同じ場所を有する写真（又はその部分）にスワップされるであろう。従って、区別可能な細部が無い露出不足の暗く且つ黒いセクションを見ている観察者は、その代わりに、裂け目及び凹凸の有る外観を有する木の皮を見るであろう。逆に、観察者は、空等の非常に白く且つ写真用語で「ブローンアウト」している写真の部分を見ているときは、その代わりに、雲の全ての形状及び細部が可視であって、僅かに明るさが少ない青に対して明るい白い雲を見るであろう。

#### 【0052】

図2は、一実施形態に従って、露出不足又は露出過度の画像のエリアを見つけるために、非圧縮画像ピクセル情報がどのように解析され得るかの単純な例を示す。露出不足のエリアを見つけるために、そのエリアにおけるピクセルの平均暗さが測定されて、それらが既定の量だけ異なるかどうかを確認するためにそのエリア内のより小さな領域内の平均と比較される。例えば、全エリアに対する平均輝度  $B_{avg}$  が最大輝度  $B_{max}$  の 5% である場合であって、エリア内の 10 正画素の領域の少なくとも 75% がエリアの平均輝度  $B_{avg}$  から  $B_{max}$  の僅か  $\pm 2\%$  しか変化しない場合、全エリアが露出不足と見なされ且つラベルされる。露出過度に対して同様の決定が行われ得る。例えば、全エリアに対する平均輝度  $B_{avg}$  が最大輝度  $B_{max}$  の 95% である場合であって、エリア内の 10 正画素の領域の少なくとも 75% がエリアの平均輝度  $B_{avg}$  から  $B_{max}$  の僅か  $\pm 2\%$  しか変化しない場合、全エリアが露出過度と見なされ且つラベルされる。画像のこうした部分の  $x-y$  座標が、その画像に関係付けられるデータベース5に保存される。観察者の視線がそれらに向けられている場合、これらは後で同じ部分が先に定義されたように露出不足又は露出過度ではない別の画像（又はその一部）と置換される部分である。「露出不足」及び「露出過度」の例示の定義が先に定義されているが、写真家若しくは別の人物又はソフトウェアアプリケーションが任意の適切なやり方で「露出不足」及び「露出過度」を定義できると考えられることに留意するべきである。例えば、エリアの平均輝度  $B_{avg}$  がそのエリアが有し得る最大輝度  $B_{max}$  の  $x\%$  より少ない場合であって、エリア内の  $m \cdot n$  ピクセルの領域の  $s\%$  の少なくとも  $t\%$  がエリアの平均輝度  $B_{avg}$  から  $B_{max}$  の  $\pm v\%$  しか変化しない場合に、画像のエリアは「露出不足」であってもよく、ここで  $x$  は 5 - 15 等の任意の範囲を有してもよく、 $t$  は 50 - 80 等の任意の範囲を有してもよく、 $s$  は 50 - 75 等の任意の範囲を有してもよく、 $m \cdot n$  は 4 - 500 ピクセル<sup>2</sup> 等の任意の範囲を有してもよく、 $v$  は 1 - 10 等の任意の範囲を有してもよい。同様に、エリアの平均輝度  $B_{avg}$  がそのエリアが有し得る最大輝度  $B_{max}$  の  $y\%$  より大きい場合であって、エリア内の  $m \cdot n$  ピクセルの領域の  $s\%$  の少なくとも  $t\%$  がエリアの平均輝度  $B_{avg}$  から  $B_{max}$  の  $\pm v\%$  しか変化しない場合に、画像のエリアは「露出過度」であってもよく、ここで  $y$  は 85 - 95 等の任意の範囲を有してもよく、 $t$  は 50 - 80 等の任意の範囲を有してもよく、 $s$  は 50 - 75 等の任意の範囲を有してもよく、 $m \cdot n$  は 4 - 500 ピクセル<sup>2</sup> 等の任意の範囲を有してもよく、 $v$  は 1 - 10 等の任意の範囲を有してもよい。そして、 $B_{max}$  はピクセルが有し得る最大デジタル輝度値（例えば、8ビットシステムでは 255）であってもよく、 $B_{min}$  はピクセルが有し得る最小デジタル輝度値（例えば、8ビットシステムでは 0）であってもよい。

#### 【0053】

図3は、（焦点の合っていない）低空間変動を有するものから（明瞭に焦点の合っている）高空間変動を有するエリアを区別するために、画像のサブ部分に対してどのように空間的高速フーリエ変換（FFT）が行われるかを示す。（焦点が外れた、露出不足の（例えば、黒い）、又は露出過度の（例えば、白い））高周波数ではない、即ち、低空間変動のエリアは、同じエリアにおいて高周波数を有する表示時間における別の画像と置換されるようにマーキングされる。「高空間変動H」は事前決定されてもよく、例えば、 $m \cdot n$  ピクセルブロック内の最も明るいピクセルの輝度に対して  $m \cdot n$  ピクセルブロックの垂直寸法又は水平寸法における輝度の  $\pm a\%$  の変化が有ることを意味してもよく、ここで  $a$

は70 - 95等の任意の適切な範囲を有してもよく、 $m \cdot n$ は4 - 500ピクセル<sup>2</sup>等の任意の適切な範囲を有してもよい。同様に、「低空間変動L」は事前決定されてもよく、例えば、 $m \cdot n$ ピクセルブロック内の最も明るいピクセルの輝度に対して $m \cdot n$ ピクセルブロックの垂直寸法又は水平寸法における輝度の $\pm c\%$ の変化が有ることを意味してもよく、ここで $c$ は0 - 30等の任意の適切な範囲を有してもよく、 $m \cdot n$ は4 - 500ピクセル<sup>2</sup>等の任意の適切な範囲を有してもよい。「高空間変動」及び「低空間変動」の例示の定義が先に定義されているが、写真家若しくは別の人物又はソフトウェアアプリケーションが任意の適切なやり方で「高空間変動」及び「低空間変動」を定義できると考えられることに留意するべきである。

#### 【0054】

図4は、木の背後の太陽の写真がどのようにデータベースにおける異なる $x$ 、 $y$ 座標の矩形に分解され得るかを例示する。矩形は、木の各部分を含む各矩形が木の各部分を覆い且つ木と木の間の空間を覆わないか又はその僅かな部分のみを覆うように、十分に小さくてもよい。図4において、写真内の場所を参照するために矩形を番号付ける任意のシステムが使用される。この場合、矩形の番号付けは左下から始まる。各矩形は、多くのピクセル（例えば、100 - 500ピクセル）から成る。 $x$ が1であり且つ $y$ が3である矩形（1, 3）は木の一部を含むので大部分が露出不足であり、矩形（4, 4）は太陽の一部を含むので大部分が露出過度である。しかしながら、矩形（3, 4）は、木の一部及び太陽の一部の両方を含むので、効果的に半分露出不足及び半分露出過度であり、従って、観察者がこの矩形を見るときにアルゴリズムはより多く露出された又はより少なく露出された写真を代わりに使うべきかどうか分からないであろう。しかし、より小さな矩形が使用される場合、人間の目にとって興味深い写真は既定のサイズの矩形よりも多くのピクセル、例えば、縦10ピクセル横10ピクセルより多くを占めるオブジェクトを含むので（写真が、可能性が無視できる程に小さい矩形と全く同じサイズである白黒の正方形の格子縞を有するのでない限り）、この問題が発生する可能性は低い。例えば、上記の状況において、コンピュータは、各矩形が主として露出過度又は露出不足になるまで矩形（3, 4）をより小さな矩形に分解することによってアルゴリズムの解像度を増加し得る。より少ないピクセルを有するより小さな矩形を解析するオーバーヘッドはマイクロプロセッサ又はマイクロコントローラ等の今日のコンピュータ集積回路（IC）の多くにとっては取るに足らないものなので、視線検出器又はマウス又は観察者がどこを見ているかを示し得る任意の他のやり方の解像度よりも少なくなり得ることを除いては、より小さな矩形を使用しない理由はない。この解像度は、使用されるハードウェアの最大解像度及びディスプレイから観察者の距離によって定義される。従って、写真家はハードウェア解像度及び距離を前もって指定するか又はデフォルト値を使用することができる。

#### 【0055】

他の写真パラメータが変化して、観察者が後で観る $x$  -  $y$ エリアに関係付けられ得る。更に、特定の潜在的観察者の識別が前もって指定され且つ表示時間に識別される場合、システムは、どの個々の観察者が $x$  -  $y$ 座標を見ているかを入力として受け入れて、観察者の識別に基づいてその応答を変化させることができる。

#### 【0056】

##### [画像構成者により撮られるステップ]

図5に示されるように、コンピュータディスプレイ12上の単純なユーザインターフェースを使用して、写真家又は他の画像構成者11が、観察者が観察するときにデータベース内の画像のどの部分が（もし有れば）どのアクションを、又はどのような順番でトリガするべきかを以下のように迅速に示し且つ指定することができる。

#### 【0057】

構成者11には、観察されるとき（以下、「表示時間」とも呼ばれる）それらに関係付けられるアクションを有する可能性が高い画像のエリアのリストが提示される。例えば、リストは、露出不足のエリア、露出過度のエリア、高空間解像度のエリア、又は低空間解像度のエリアを含んでもよい。次に、構成者は、リストされたエリアごとにアクションを

10

20

30

40

50

指定する。限定されないが、アクションは、以下の一部であるように指定され得る。即ち、画像全体又は画像のその部分だけを、露出不足又は露出過度でもない異なる露出値を有する、又は例えば、オリジナルのカメラの解像度に対して最大の詳細である又は「焦点が合っている」空間周波数の異なる変化を有する別の画像と置換すること、画像番号  $x$  と置換すること、 $n$  番目の置換であれば画像番号  $y$  と置換すること、エリア  $b$  のみが見られていた場合に画像  $z$  と置換すること等である。異なるアクションがデータベース内の各画像に関係付けられてもよく、又は一組のアクションが一次画像（以下を参照）に関係付けられてもよい。アクションは、インターネット又は他のシステムから引き出された画像を観察者が見ているときにリアルタイムで利用可能な画像又は動画を選択することにまで拡張され得る。

10

**【0058】**

単一の写真が、観察される画像であり且つその中の  $x - y$  矩形を見るときのための全ての方向を有する一次画像として指定可能であり、又は任意の写真がその写真が見られるときのためのこうした方向を有し得る。この例示では、構成者がアクションを指定しない場合、下記のようにデフォルトは「最良自動的」である。

**【0059】**

観察者が取るかもしれない全てのアクション及び結果として生ずる画像の全て又は一部の全ての置換を構成者が終わると、この全ての情報は、1つの「ショー (show)」又は「Eyetrinary<sup>TM</sup>」として保存される。このショーは、全てがクラウドにオンラインで、又は任意の他の適切な電子記憶装置又は場所に記憶されてもよい。

20

**【0060】**

図6に示されるように、観察者13が画像を観るとき、ディスプレイ14は、コンピュータ17に記憶され且つ読み込まれるデータベース5から写真を表示する。例えば、ディスプレイ14に設置される1つ以上のカメラから成る視線検出ハードウェア装置16は、写真内のどこを観察者の目が見ているかの  $x - y$  座標を与える。システムはこうした座標を使用して、データベース5において構成者が指定したアクションを検索し、システムは、その  $x - y$  矩形に対する一次写真又は現在の写真のエントリの何れかを参照することによってアクションを行う。

**【0061】**

アクションがデフォルトの「最良自動的」である場合、システムは、例えば、観察者が見ている写真を、観察者が現在観察している  $x - y$  矩形に関して最良に露出された写真と置換するであろう。既に述べたように、ここでの「最良に露出された」は、一組の写真家が同意するものであり、最も多くの細部を示し、明度が最も「ブローンアウト」されていないか又は標準的な定義が利用可能な黒ではないものである。この置換は迅速なディゾルブであってもよく、その性質及び時間は観察者が観ているときに使用されるように前もって指定されてもよく、又はデフォルト値が使用されるであろう。

30

**【0062】**

先に言及されたデータベースは初期画像及び一組のリンク又はその画像の部分に関係付けられるジャバスクリプト呼び出しから成る `html` のページとしてクラウドにオンラインで記憶されてもよいことに留意することが重要である。視線検出は、あたかも観察者の目がマウスであるかのように、その目が見ている場所をブラウザに教え、ブラウザはリンクの目標である新しい画像を表示することによって、又は、例えば、画像の一部のアスペクトを変更するためにジャバスクリプト関数を呼び出すことによって、アクションを自動的に行うであろう。このようにして、構成者は、全てがサーバのディレクトリ内に有る `html` ページ及び一組の画像を指し示すリンクを観察者に送りさえすれば十分で有り得る。

40

**【0063】**

構成者は、ディスプレイ及びポインタを有するユーザインターフェースを使用して前もって指定する代わりに、視線検出ハードウェアを有するディスプレイを使用して、特定の  $x - y$  エリアを見ている間に、音声、又はマウス、又は任意の他の方法を介するシステム

50



へのコマンドによって他のどの写真（又はその一部）を使用すべきかを指定することができる。他の写真（又はその部分）とこのようにした他の相関は、データベースに記憶されるであろう。例えば、写真家は、図4における矩形（1, 3）を見ている場合に、音声認識システムに向かって、実際に、「私が今見ているエリアを観察者が見ているときに、示されている写真を“tree-exposed.jpg”というファイル名の写真と置換しなさい」と言うことができる。

【0064】

〔付加的使用〕

実施形態は、アーティストの意図を表現することより多くのために使用され得る。例えば、視力療法士は、スクリーンの前後に注意を引くことによって両目を調整する患者の能力を増加させたい場合がある。パイロットは、既存の組の画像又はスプライトに基づいて、パイロットが見ていなかったどのような場所にも危険な航空機を出現させることによって、このような危険な航空交通を迅速に見つけるためのディスプレイを有するフライトシミュレータシステムで訓練され得る。観察者の目が離れて行くこと（水平線に向かって、更に遠くを見ること）を許される場合、実施形態は、更に遠くの焦点で撮られた写真を代用することができる。従って、観察者は、3Dの形態に焦点を合わせる及び外す並びに内外に「移動（travel）」するように自身を訓練することができる。

【0065】

〔視線検出画像ツール〕

実施形態は、画像ツール用の視線検出に基づくユーザインターフェース（UI）のために使用され得る。例えば、観察者が頭を右に傾けることで、傾いた頭が中央に戻るまでゆっくりと右に画像を回転させるようにシステムに通知することができる。輝度及び焦点は、画像の一部を凝視することで制御され得る。観察者が左目を瞬きするたびに、最大に達するまで輝度が増加されてもよく、それから再び下がり始める。焦点は、右目を瞬きすることによって増加され得る。処理は、一秒間、両目を閉じることによって所望の選択で停止され得る。類似の方法がクロップのために使用され得る。

【0066】

〔付加的詳細〕

一実施形態では、観察者の目が画像の一部に集中している場合、画像のその部分が更なる細部と共に拡張又は拡大し得る。最大の細部又は拡大図を有する画像が代用された後であって、観察者の目がそのエリアに固定されたままのとき、写真家により事前決定された画像のその一部に関係付けられるリンクが、ダウンロードされる他の画像又は動画をもたらし、画像のその部分に代用され得る。写真家が特定のリンクを入れなかった場合、そのセクションに見られる初期画像は、画像が何であるかを決定して更なる情報を供給するか、又は他の写真家によりキャプチャされた類似のコンテンツの他の画像をダウンロードして表示することができるクラウドベースのアナライザに送信され得る。例えば、観察者がワシントンDCのシーンを見ており、ワシントン記念塔を凝視したままであれば、記念塔は、画像のそのセクション用の一連の代替画像によってディスプレイ全体の半分を占めるようにサイズが拡張して細部を示してもよい。観察者が記念塔を見つめ続ける場合、インターネットからダウンロードされた動画は、記念塔の歴史及び建築を記述するスクリーンのそのセクションを再生してもよい（これは写真家により前もって指定されているか又はクラウドベースの画像認識装置により自動的に選択されていてもよい）。

【0067】

「フラッシュライト」

写真家は、観察者が見る画像のどのような部分も即座により高い輝度に修正され得ることを指定し得る。構成者は、x-yエリアに対して「フラッシュライト」を単に指定しさえすればよい。従って、フラッシュライトが、実際に観察者の目が向けられて、シーンに対して作用しているように見えるであろう。〔ヒューマンインターフェーステスト〕広告に関するアプリケーションで、人物が画像のどこを見る傾向が有るかを記録する既存のシステムが有る。一実施形態では、どのくらい頻繁に観察者が画像の特定の照らされた又は

10

20

30

40

50

焦点を合わせられた部分に戻るかに基づいて出力されるログを想像してもよい。また、画像構成者が画像代替の「パス」を指定する場合、即ち、画像Aのサブ部分を見ることが画像Bの代替をもたらし、画像Bの一部を見ることがAに戻り、且つ観察者がA - B - Aのこのパスを反復した場合、ログは、パスの情報及びそれが反復されたことを保存するであろう。このような情報は、（複数の）画像と観察者の相互作用を更に発展させるのに構成者にとって有益であろう。

【0068】

〔他の観察者とのネットワーキング及びゲーム〕

同じ場所に居る他の観察者とのゲームのために1つより多くの視線検出装置が使用され得る。また、システムは、様々なソーシャル目的のためにインターネット又はローカルネットワークに接続され得る。例えば、異なる場所に居る一人より多くの観察者が同じEyetrinary™を見ている場合、大部分の観察者により見られた部分は拡大された部分であってもよく、又はその反対であってもよい。即ち、殆どの人が何かを見なかった場合、Eyetrinary™は全ての観察者がその画像に行くことを強いるように修正されてもよい。第1の人物が適切な場所を見ること、又は適切な場所で武器を作動させること等に基づいて2つ以上の視線検出システムがポイントを与えるゲーム等、他のタイプのソーシャル相互作用が実装され得る。最後に、その部分を見ることによって、画像の部分がインターネットで他者と共有され得る。

10

【0069】

〔構成者へのフィードバック〕

実施形態は、観察者が表示時間にどこを見るかに応じるという構成者の事前決定された意図に関する。しかしながら、観察者は、ネットワークを介してリアルタイムで構成者に有益なフィードバックを送信するためにシステムを使用することができる。例えば、画像の一部、例えば、観察者が構成者に置換してもらいたい複数の人々の写真において特定の人物が有り、観察者はその人物を見ている間に3度瞬きをして、目でその周りに円を描くことができる。次に、観察者は、ネットワーク接続を介して、「私がこの人物を見ると、あなたに別の画像を代用してもらいたい」を言ってもよい。顔のこの場所が音声ファイルと共に構成者に渡されるであろう。従って、構成者は、観察者の入力により良好な又は異なるEyetrinary™を開発することができる。

20

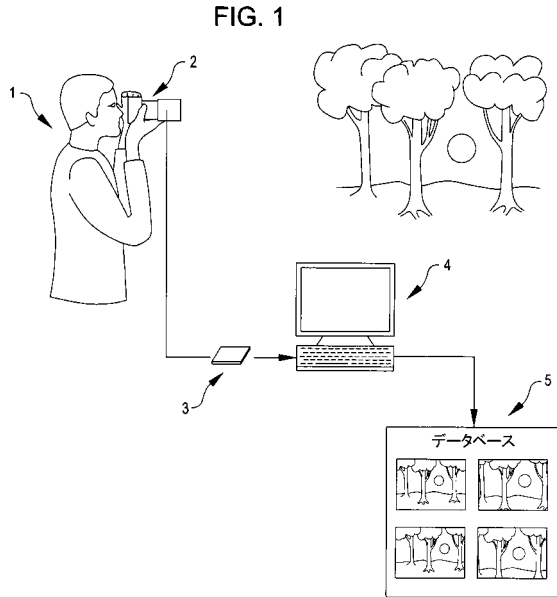
【0070】

以上により、特定の実施形態は例示のために本明細書に記載されているが、本開示の精神及び範囲から逸脱することなく様々な修正が行われ得ることが理解される。更に、特定の実施形態に関して代替手段が開示されている場合、この代替手段は特段の記載が無ければ他の実施形態にも適用され得る。更に、記載された任意の構成要素又は動作が、ハードウェア、ソフトウェア、ファームウェア又はハードウェア、ソフトウェア及びハードウェアの何れか2つ以上の組み合わせで実装され/実行されてもよい。更に、記載された装置又はシステムの1つ以上のコンポーネントは、明確のため又は他の理由で記載から省略されている場合がある。更に、説明に含まれている記載された装置又はシステムの1つ以上のコンポーネントが装置又はシステムから省略されてもよい。

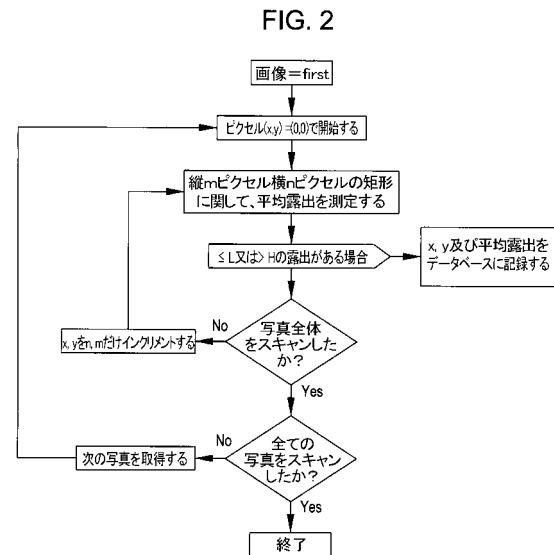
30

40

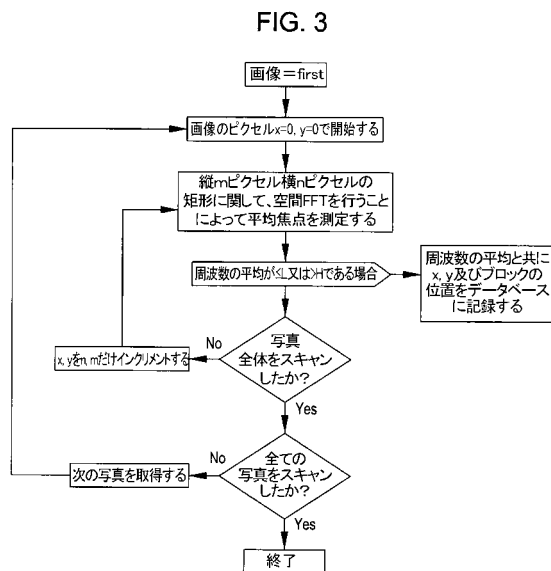
【 図 1 】



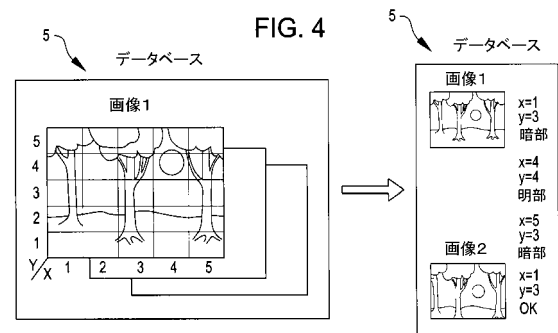
【 図 2 】



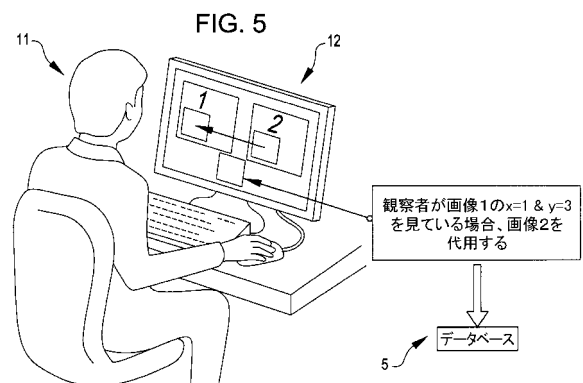
【 図 3 】



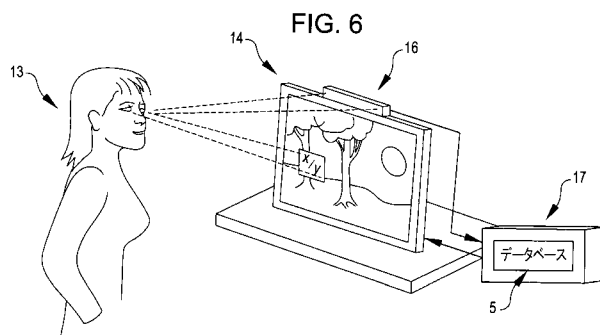
【 図 4 】



【 図 5 】



【図 6】



## フロントページの続き

(51)Int.Cl.	F I			テーマコード (参考)	
<b>G 0 9 G 5/02 (2006.01)</b>	G 0 9 G	5/36	5 2 0 F	5 E 5 5 5	
<b>G 0 9 G 5/38 (2006.01)</b>	G 0 9 G	5/36	5 2 0 G		
<b>H 0 4 N 5/232 (2006.01)</b>	G 0 9 G	5/02	B		
<b>H 0 4 N 5/66 (2006.01)</b>	G 0 9 G	5/00	5 1 0 Q		
<b>G 0 6 F 3/0346 (2013.01)</b>	G 0 9 G	5/38	B		
<b>G 0 6 F 3/0484 (2013.01)</b>	G 0 9 G	5/36	5 1 0 V		
<b>G 0 6 F 3/0488 (2013.01)</b>	H 0 4 N	5/232	2 9 0		
	H 0 4 N	5/66	A		
	H 0 4 N	5/66	D		
	G 0 6 F	3/0346	4 2 3		
	G 0 6 F	3/0484	1 2 0		
	G 0 6 F	3/0484	1 5 0		
	G 0 6 F	3/0488			
	G 0 9 G	5/00	5 1 0 H		

- (72)発明者 ヘイリー、 ジェフリー  
 アメリカ合衆国 9 8 0 0 5 ワシントン州 ベルビュー サウスイースト 2 7 ス プレイス  
 1 3 4 3 4
- (72)発明者 シュウォラー ジュニア、 トーマス レナード  
 アメリカ合衆国 9 8 1 0 9 ワシントン州 シアトル 4 ス アベニュー ノース 1 9 0 8
- (72)発明者 リッチモンド、 マクスウェル ウルフ  
 アメリカ合衆国 9 8 1 9 9 ワシントン州 シアトル パーキンズ レーン ウェスト 3 0 1  
 9
- (72)発明者 リッチモンド、 イーライ ニュートン  
 アメリカ合衆国 9 8 1 9 9 ワシントン州 シアトル パーキンズ レーン ウェスト 3 0 1  
 9
- (72)発明者 リッチモンド、 オウエン セーガン  
 アメリカ合衆国 9 8 1 9 9 ワシントン州 シアトル パーキンズ レーン ウェスト 3 0 1  
 9

F ターム(参考) 5B057 BA02 CA08 CA12 CA16 CB08 CB12 CB16 CC01 CE08  
 5B087 AB04 AD00 BC16 BC32 CC26 CC33 DE01  
 5C058 BA05 BA11 BA17  
 5C122 EA42 EA61 FH01 FH02 FH09 FH11 FH14 HA29 HA86  
 5C182 AB08 AB11 AB12 AB13 AB14 AC03 AC42 AC43 AC46 BA06  
 BA14 BA35 BA44 BA45 BA46 BA54 BA55 BA56 BA66 BA75  
 BC26 CA01 CA02 CA22 CA32 CA33 CB03 CB13 CB14 CB32  
 CB52 CC14 CC15 CC21 CC27  
 5E555 AA08 AA23 AA26 AA44 AA62 AA71 BA01 BA05 BA06 BA18  
 BA53 BA73 BA83 BB01 BB05 BB06 BB18 BC04 BC17 BC18  
 CA02 CA13 CA42 CA44 CA47 CB09 CB19 CB21 CB23 CB45  
 CB64 CB65 CC22 CC24 CC25 DB54 DC36 DC63 DC75 DC84  
 DD07 DD11 EA05 EA07 FA00

【 外国語明細書 】

## TITLE OF THE INVENTION

**COMPOSING AND REALIZING A VIEWER'S INTERACTION WITH DIGITAL MEDIA****Background**

**[0001]** Today's digital cameras and smart phones use computer power to enhance images for viewing either immediately or later on the screens of other devices. One example is HDR, or High Dynamic Range, where the camera quickly takes several pictures at different exposures and builds an image where all portions, even the brightest and darkest, are exposed to bring out all details.

**[0002]** There are also existing systems that use gaze tracking to determine how long a viewer looks at a certain ad on a web page, for the purposes of making it more attractive to the eye, or to measure the value of the advertising in terms of how much time (*e.g.*, average time, aggregate time) how many viewers spend on it. There are also systems that use gaze tracking to see if the viewer is no longer looking at display, such as a smart phone's display, in order to stop playing a video and save battery power or make sure that the viewer does not miss any part of the video.

**Summary**

**[0003]** An embodiment combines these two concepts to enhance the experience of viewing images by tracking where the viewer is looking. The result is to make the experience more like viewing the original scene, or enhance it in new ways beyond the original experience, either automatically, or by interacting with a photographer's previously specified intentions for what should happen when the viewer looks at a particular portion of an image or images taken by that photographer.

**[0004]** An embodiment allows an artist to create an original image for a display screen, or multiple original images for multiple display screens, that changes as a viewer looks at the image, as a function of who the viewer is or what the viewer does. The viewer's action that causes the image to change can be where the viewer directs a gaze or touches the screen, or what the viewer says or how far the viewer is from the display, or how many viewers there are, or the identity of a particular expected viewer, etc.

**[0005]** An embodiment is a method for modifying one or more original images displayed on one or more display screens as a function of data collected via one or more cameras capturing an image of one or more viewers' faces before the one or more display screens. The data collected may be one or more of: where location of gaze of eyes of one or more viewers is directed within the

original image, or facial appearance change of one or more viewers, or amount of ambient light, or distance of a viewer from a display screen, or number of viewers, or identity of a viewer.

**[0006]**      Brightness. The modification may be to change brightness of a part, or all, of the original image. The brightness change can be effected by replacing the original image with a substitute image, where the contents of the substitute image are the same as the original image but with some or all pixels modified such that the brightness of at least a portion of the substitute image is different from the brightness of at least the corresponding portion of the original image. For example, a pair of photographic images can be taken with a camera where the images are essentially the same except for a different exposure. Some or all of one image can replace some, or all, of the other image to change brightness. Alternatively, an algorithm can adjust numerical brightness of some image pixels but not all image pixels as a function of actions by one or more viewers.

**[0007]**      Focus. The modification may be to change focus of all or a part of the original image. The change may be effected by replacement of the original image with a substitute image, where the contents of the substitute image are the same as the original image but with some, or all, pixels modified such that the focus of at least a portion of the substitute image is different from the focus of at least the corresponding portion of the original image. The modification may be to change the focus to objects that appear to be close in a part of, or in all of, the image. The modification may be to change the focus to objects that appear to be far away in a part of, or in all of, the image. The modification may be to change the apparent depth of field of part of, or all of, the image. The modification may be to zoom in or out with the zooming centered on a part of the original image. The modification may include replacement of the original image with a substitute image. The modification may be to zoom in or out in synchrony with changes in measured distance from a display screen to a viewer.

**[0008]**      Color. The modification may be to change the color balance or color saturation or color depth or contrast of part of, or all of, the image. This may be effected by replacement of the original image with a substitute image, where the contents of the substitute image are the same as the original image but with some or all pixels modified such that the color (*e.g.*, color balance or color saturation) of at least a portion of the substitute image is different from the color of at least the corresponding portion of the original image.

**[0009]**     Sharpness. The modification may be to change the sharpness of part of, or all of, the image. This may be effected by replacement of the original image with a substitute image, where the contents of the substitute image are the same as the original image but with some or all pixels modified such that the sharpness of at least a portion of the substitute image is different from the sharpness of the corresponding portion of the original image.

**[0010]**     Sound output. The modification may be to cause or to change playing of sounds that accompany the original image.

**[0011]**     Sprite. The modification may be to cause movement or cessation of movement of a sprite within the original image.

**[0012]**     Animated GIF (Graphics Interchange Format). The modification may be to replace all or part of an image with a small set of images that together form an animated GIF.

**[0013]**     Video Branching. The modification may be to select a branch of a multi-branch video.

**[0014]**     In any of the above-described embodiments, the one or more original images may be still, or video, or still with moving sprites. The original images may be two dimensional or three dimensional.

**[0015]**     In any of the above-described embodiments, the algorithm for modification of the original image or images may be custom determined by an artist or composer or content creator who selected the one or more images. The algorithm for modification may be pre-determined by a company that supplied software for selection and use by an artist or content creator who selected the one or more original images. The algorithm may be pre-determined by a company that supplied software, and may be such that the algorithm works with any image or set of images. The algorithm may allow an artist or content creator to specify a speed at which a gradual-change modification (*e.g.*, gradual change of brightness, gradual change of focus, gradual change of color, expanding a new image from tiny (or partial) to full, or other transition method) will be performed.

**[0016]**     Touch or Mouse Input. Another embodiment is a method for modifying an original one or more images displayed on one or more display screens according to an algorithm that operates as a function of data collected from touch, mouse, or other input device by the viewer, for example, touching, or pointing to, a portion of the screen within the image, where the modification is replacement of the original image with one or more substitute images where the contents of the substitute images are the same as the original image but with some or all pixels modified. The data



collected by mouse might comprise data identified by one or more of: clicking or mousing over or clicking and dragging and moving. The data collected from touch may comprise data identified by one or more of: touch by one or multiple fingers, force of touch, duration of touch, movement of touch, and gestures (*e.g.*, pinching, waving, pointing) in which a user's hands/fingers are close to, but do not contact, the image-display screen. The modification may be any of the modifications described above. The one or more original images may be still, or video, or still with moving sprites, or two dimensional, or three dimensional. The algorithm may comprise any of the elements described above.

**[0017]**      Voice Input. Another embodiment is a method for modifying an original one or more images displayed on one or more display screens according to an algorithm that operates as a function of data collected from voice sounds, where the modification is replacement of the original image with one or more substitute images where the contents of the substitute images are the same as the original image but with some or all pixels modified. The modification may be any of the modifications described above. The one or more original images may be still, or video, or still with moving sprites, or two dimensional, or three dimensional. The algorithm may comprise any of the elements described above.

**[0018]**      Accelerometer Input. Another embodiment is a method for modifying an original image displayed on a hand-held display screen according to an algorithm that operates as a function of data collected from one or more accelerometers embedded in a housing of the display screen, where the modification is to change field of view of the original image by enlarging or contracting each of one or more of the left boundary, the right boundary, the top boundary, or the bottom boundary. The data collected from the one or more accelerometers may be a tilting of a first edge of the display away from a viewer relative to an opposite edge, which tilting causes more of the image near the first edge to come into view. The modification may be to zoom in or to zoom out, with the zooming centered on a part (not necessarily the center) of the original image.

**[0019]**      Authoring Tools. Another embodiment is a method in a system with a server computer and a client computer to receive instructions from an author and, based on those instructions, to generate a data set for showing images that change based on actions of a viewer. The method comprises (a) receiving at the server computer from a client computer specification of a plurality of images in a sequence; (b) receiving at the server computer from the client computer specification of

data to be collected that should trigger a transition from an original image in the sequence of images to a second image in the sequence of images; (c) receiving at the server computer from the client computer a specification of speed of transition from the original image in the sequence of images to the second image in the sequence of images; and (d) assembling on the server computer a data set that can be transmitted to a viewer's computer, where the data set can be viewed on the viewer's computer, and where data input received from actions of the viewer causes modification of the viewed image by replacing the original image in the sequence with the second image at the specified transition speed.

**[0020]** The data input may comprise, but is not limited to, one or more of: location of gaze of eyes of one or more viewers within the original image, or face-appearance change of one or more viewers, or ambient light, or distance of a viewer from a display screen, or number of viewers, or a viewer's identity, or voice input, or touch input.

**[0021]** The modification may comprise, but is not limited to, one or more of to change brightness of a part or all of the original image; to change focus of a part of the original image; to zoom in or out from a part of the original image; to change the color balance or color saturation or color depth or contrast of part or all of the original image; to cause movement or cessation of movement of a sprite within the original image; to select a branch of a multi-branch video.

**[0022]** The original image may be, but is not limited to: still, or video, or still with moving sprites, or two dimensional, or three dimensional.

**[0023]** The client computer and the server computer may each include a respective software program contained within a respective single computer housing, and may be activated by a single user.

### **Brief Description of the Figures**

**[0024]** Figure 1 shows a photographer taking multiple images of a single target, according to an embodiment.

**[0025]** Figure 2 is a flowchart for creating an exposure map for a set of images, according to an embodiment.

**[0026]** Figure 3 is a flowchart for creating a focus map for a set of images, according to an embodiment.

**[0027]** Figure 4 shows how the exposure map for a set of images is stored, according to an embodiment.

**[0028]** Figure 5 shows how a composer determines how an image will change as a viewer looks at different parts of an image, according to an embodiment.

**[0029]** Figure 6 shows the system causes the image to change as a viewer looks at different parts of an image, according to an embodiment.

### **Detailed Description**

**[0030]** One or more embodiments of this invention are to enhance photography and make stills and videos interactive with a viewer in, *e.g.*, dimensions of brightness, focus, depth of field, and other qualities, either automatically, or as determined in advance by the photographer/composer and driven later by, for example, where in an image the viewer is looking.

**[0031]** In its simplest and most limited form, an embodiment recreates how someone would see the subject matter of a scene that was captured in a photograph, where the original scene had a very wide variation in exposure or depth. The following examples will elucidate this concept.

**[0032]** 1. Brightness.

**[0033]** Imagine a band of trees with the sun setting behind them. If a viewer looks towards the sun, her irises contract, thus causing her pupils to decrease in size and let in less light and reduce brightness, and she sees the remaining half disk of the red sun (the portion of the sun not obscured by trees), but sees the trees become black silhouettes. As she looks at a single particular tree, however, her pupils open again and the character and texture of the bark on the tree become visible.

**[0034]** Any normal, conventional photograph of such a scene would be limited to one exposure: either dark enough to see the setting sun's shape, in which case the trees are black; or light enough to see the bark, in which case the background of the sunset becomes a white "blown out" area with no sun.

**[0035]** There is an existing photographic method intended to compensate for this problem, and this method is called "HDR", or High Dynamic Range. In HDR, several exposures are taken in quick succession to avoid movement, and are automatically blended so that every area of the image has the correct exposure. However, an HDR compensated image often looks "fake" or "saccharine" or "unreal" and, therefore is often unsatisfying to the viewer.

**[0036]** Instead, an embodiment recreates the action of the viewer's own pupil-exposure system. As with HDR, it uses a group of photographs that differ only in the exposure used when they were taken. When any one of the photographs is displayed, however, a very different thing (as compared to the HDR technique) happens.

**[0037]** For example, say the photograph with the sun exposed correctly is first displayed brightly on a monitor. If the viewer then looks at a blacker part of the image that was originally underexposed, like a tree in this example, the gaze tracking portion of the system detects the viewer's shift in gaze to the darker part of the image, and the system finds and displays another photograph from the set that has the area of the tree more exposed.

**[0038]** In this way, as the viewer's pupils expand because the viewer is looking at a very dark area, the viewer is rewarded with more image detail blossoming. A look back at the sun does the opposite: the area looks too bright and blown out, but quickly settles down into a red setting sun between the silhouetted trees. In this way, the "feeling" of looking at the original scene is rescued and recreated, and the experience of viewing is enhanced.

**[0039]** An embodiment gives the photographer/composer a way to record a wide range of camera and displaying variations, beyond just brightness, that she or he decides upon when taking a still photograph, and allows the viewer to trigger and explore these variations naturally by, *e.g.*, gaze tracking or touching or mousing on the image display screen. The triggering can be manual, automatic, or according to preset intentions.

**[0040]** Alternatively, instead of image swapping with several images available for the swap, an algorithm can adjust the brightness (*e.g.*, numerically) of some image pixels, but not all image pixels, in an image as a function of actions by one or more viewers.

**[0041]** 2. Focus.

**[0042]** The same methods may be used for focus. For example, if the scene has objects at a variety of distances, no normal photograph can be perfectly focused for all objects, at least given a finite depth of field and the camera that took the photograph having a finite focal length. Accordingly, an embodiment can track which object in the photograph the viewer's eyes are looking at and select, from a set of many different photos each taken with a different focus, a photo that is focused correctly for that object.

**[0043]** 3. "Walking up to a Window"

**[0044]** As a viewer walks towards a display of a photograph or other image that does not take up the entire display, the increasing angle subtended by her eyes can be computed from the changing distance between the viewer and the display, and the changing distance can be measured with a camera mounted on the display, and the outside frame of the photograph can be increased while simultaneously displaying more of the photograph. This technique simulates walking up to an actual window in that more content appears as the viewer gets closer to the actual window, but the viewer may need to look left and right “through” the window to see the additional content.

**[0045]** 4. Post processing effects: Cropping/scrolling

**[0046]** After the camera settings are varied and photographs are taken, the photographer decides on how the image should be displayed. For example, in cropping, the photographer picks a more artistic subset of the image to present to the viewer. However, the photographer may intend for more of the image to be seen if a viewer is interested, in the following way. For example, the image can be scrolled up or down when the viewer’s eye looks down or up. Done correctly, this could provide an experience of looking down into the content of the image; *e.g.*, in a photo of a distant mountain, if the viewer looks down, he would see the trail leading up to the viewpoint. Put another way, the horizon would shift, from the distant mountain, to a few feet forward on the trail so that the portion of the trail in front of the viewer takes up most of the frame. That is, the image can be scrolled up when the viewer looks down, and can be scrolled down when the viewer looks up. Similarly the image can be scrolled right or left when the viewer looks left or right.

**[0047]** The same effect can be achieved with a hand-held display by using accelerometer data to scroll the image when the display is moved up, down, or sideways. That is, the hand-held display can be made to mimic a portable window that the viewer can move to view a desired portion of an image, where the image is larger than the window.

**[0048]** 5. Shutter-time and Movement

**[0049]** Imagine a still photograph showing several maple wing seeds descending. In an embodiment, several photographs of short exposure time could be taken as well as one with a longer time. When the viewer first looks at an individual seed in the photograph with the longer exposure time, she sees its wing as a blurred circle around the center. However, as she stares at it, she is shown, in quick succession, several frames formed by the photographs having the shorter exposure times (or portions of these photographs including the seed), and sees the seed actually turning

quickly and falling slightly and then freezing in place (*e.g.*, when the last of frames formed by the photographs having the shorter exposure times is displayed). While the viewer looks away from the portion of the photograph including the seed, the seed would return to its initial position and blurred appearance (*e.g.*, the photograph, or portion thereof, with the longer exposure time would be displayed again). Then the above-described cycle would repeat when the viewer looks back at the seed.

**[0050]** Alternatively, when a viewer looks at an object that looks like it might be moving, such as the described maple seed, a portion of the still image can be replaced with a video image or a sprite or animated GIF showing the object moving.

**[0051]** This also applies to "Sequential Photography" where a moving gymnast or skier or jumper is captured in several positions against a single background. In an embodiment, only one position would be shown. As the viewer looks at the display, the other shots would be shown, leading the eye through each position across the frame ... and backwards from the last one. This would be valuable for those (*e.g.*, a coach or doctor) who would like to see how the captured subject is actually moving, since a single still image typically reveals less than a sequence of images that depict movement.

**[0052]** 6. Face appearance

**[0053]** Image recognition can be used to determine if a person is smiling or frowning or showing other particular emotions. Such facial changes can cause changes in a part of the image or all of the image. Image recognition can be used to determine identity of a viewer, and the determined identity can cause changes to the image.

**[0054]** 7. Distance of a viewer from the display

**[0055]** A camera mounted on the display can be used to compute distance of a viewer from the display, and the distance can be used to cause changes to part or all of the image. (The method of measuring the distance can be any of several methods used by cameras with autofocus, such as an optical method using infrared pulses for focus detail, or an ultrasound method using sonic pulses). That is, the range-finding mechanism of the camera, which mechanism the camera may use for auto focus, can be exploited to determine the distance of a viewer from a display screen to which the camera is mounted or otherwise associated.

**[0056]** 8. Touch by a viewer

**[0057]** A touch-sensitive input layer on the face of the display can be used to receive touch input from a viewer, which touch input can be used to cause changes to part or all of the image.

**[0058]** 9. Voice sounds from a viewer

**[0059]** A microphone mounted on, or otherwise associated with, the display can be used to receive voice input from a viewer, which voice input can be used to cause changes to part or all of the image. The voice sounds can be used to identify the viewer, and the image can be modified based on the identity of the viewer.

**[0060]** 10. Acceleration of the display

**[0061]** Where the display is hand held, input from accelerometers can be used to receive input from a viewer moving the display, input which can be used to cause changes to part or all of the image.

**[0062]** 11. Ambient light

**[0063]** A camera or other light sensor mounted on the display can be used to determine the amount of ambient light, which can be used to cause changes to part or all of the image, such as to the brightness of the image.

**[0064]** 12. Further capabilities

**[0065]** The preceding examples attempt to recreate, with more fidelity than existing imaging techniques, what the original scene looked like, and thus enhance the experience of viewing the recreating of the original scene. However, there is no reason to limit these enhancement techniques to accurate scene re-creation. These techniques can also serve as an artistic tool for the photographer/composer to lead the viewer through what he or she intended for the piece. For example, the composer may intend that the first time a viewer sees a portion of the picture, there should be a different brightness and focus than the second, or any subsequent time, that the viewer sees the portion of the picture. Or the composer may purposely blur or defocus parts of the image near where the viewer is looking to emphasize a concept or meaning implicit in the portion of the image that the viewer is concentrating on, or to purposely distract the viewer from concentration.

### **Detailed Example**

**[0066]** In this section, it is shown how an embodiment accomplishes its overall objective of enhancing the experience of viewing an image, and giving the photographer a tool to more fully

express her intention regarding the viewing of the image by others. Also shown are the components involved and how they work together. This explanation is shown in detail for a subset of the possible variations in the photographs, but can be easily extended to the other variations as described above using the same basic process. Therefore, this explanation is intended to be a non-limiting example of an embodiment.

**[0067]** Figure 1 shows the image and preprocessing phase in a simple embodiment. A photographer **1** is shown using a camera **2** manually taking, or using an in-camera program to take, a set of images of a scene, with various photographic parameters varied for each image, such as brightness, focus, and shutter time.

**[0068]** The photographs may be transferred to a computer **4**, such as via a Secure Digital (SD) card **3** or any other transfer mechanism (or the capture mechanism and circuitry already is, or already includes, a computer), then labeled and stored in a database **5**. The images are then automatically analyzed and each is associated with expected actions of a viewer. For example, portions of each image that are underexposed or overexposed in one image and exposed normally in another image are automatically noted. "Underexposed" means that the details in the scene photographed are indistinguishable from black, or zero lightness. "Overexposed" means that original details are indistinguishable from white, or 100% lightness. Referring to the earlier example, the bark on the tree in front of a sunset, the images with dark in the areas of bark compared with normally exposed areas in the same place in other images are noted and that information is saved. When later the viewer looks at an underexposed location of an image, the photograph (or portion thereof) with the same location normally exposed will be swapped in. Thus the viewer, looking at an underexposed, dark and black section, where there is no distinguishable detail, would instead see the bark of the tree, with its crevices and bumpy appearance. Conversely, when the viewer is looking at a portion of the picture that is very white and "blown out" in photographer parlance, such as the sky, she would instead see bright white clouds against a slightly less bright blue, with all the shape and detail of the clouds visible.

**[0069]** Figure 2 shows a simple example of how the uncompressed image pixel information can be analyzed, according to an embodiment, to find areas of the image that are either underexposed or overexposed. To find an underexposed area, the average darkness of the pixels in that area is measured and compared to the average within smaller regions within that area to see if



they differ by a predetermined amount. For example, if the average brightness  $B_{avg}$  for the entire area is  $\leq 5\%$  of maximum brightness  $B_{max}$ , and if at least 75 % of the regions of ten square pixels within the area vary by no more than  $\pm 2\%$  of  $B_{max}$  from the average brightness  $B_{avg}$  of the area, then the entire area is considered, and labeled, as being underexposed. A similar determination can be made for overexposed. For example, if the average brightness  $B_{avg}$  for the entire area is  $\geq 95\%$  of maximum brightness  $B_{max}$ , and if at least 75% of the regions of ten square pixels within the area vary by no more than  $\pm 2\%$  of  $B_{max}$  from the average brightness  $B_{avg}$  of the area, then the entire area is considered, and labeled, as being overexposed. The x-y coordinates of these portions of the image are saved in the database 5 associated with that image. These are the portions that will be replaced later, if the viewer's gaze is directed to them, with another image (or portion thereof) in which the same portion is not underexposed or overexposed as defined above. It is noted that although example definitions of "underexposed" and "overexposed" are defined above, it is contemplated that a photographer or another person, or a software application, can define "underexposed" and "overexposed" in any suitable manner. For example, an area of an image can be "underexposed" if the average brightness  $B_{avg}$  of the area is less than  $x\%$  of the maximum brightness  $B_{max}$  that the area can have, and if at least  $t\%$  of  $s\%$  of the regions of  $m \cdot n$  pixels within the area vary by no more than  $\pm v\%$  of  $B_{max}$  from the average brightness  $B_{avg}$  of the area, where  $x$  can have any range such as 5 – 15,  $t$  can have any range such as 50 – 80,  $s$  can have any range such as 50 – 75,  $m \cdot n$  can have any range such as 4 – 500 pixels<sup>2</sup>, and  $v$  can have any range such as 1 - 10. Similarly, an area of an image can be "overexposed" if the average brightness  $B_{avg}$  of the area is greater than  $y\%$  of the maximum brightness  $B_{max}$  that the area can have, and if at least  $t\%$  of  $s\%$  of the regions of  $m \cdot n$  pixels within the area vary by no more than  $\pm v\%$  of  $B_{max}$  from the average brightness  $B_{avg}$  of the area, where  $y$  can have any range such as 85 – 95,  $t$  can have any range such as 50 – 80,  $s$  can have any range such as 50 - 75,  $m \cdot n$  can have any range such as 4 – 500 pixels<sup>2</sup>, and  $v$  can have any range such as 1 - 10. And  $B_{max}$  may be the maximum digital brightness value (e.g., 255 in an 8-bit system) that a pixel can have, and  $B_{min}$  may be the minimum digital brightness value (e.g., 0 in an 8-bit system) that a pixel can have.

**[0070]** Figure 3 shows how a spatial fast Fourier transformation (FFT) is performed for sub-portions of an image to distinguish areas with high spatial variation (crisp focusing) from those with low spatial variation (unfocussed). Areas with no high frequencies of, *i.e.*, of low, spatial

variation (out of focus, underexposed (*e.g.*, black), or overexposed (*e.g.*, white) are marked to be replaced with another image at display time that has high frequencies in the same area. "High spatial variation H" can be predetermined, and, for example, might mean that there is a change in brightness  $\geq \pm a\%$  in the vertical dimension or horizontal dimension of an  $m \cdot n$  pixel block relative to the brightness of the brightest pixel within the  $m \cdot n$  pixel block, where  $a$  can have any suitable range such as 70 – 95, and  $m \cdot n$  can have any suitable range such as 4 – 500 pixels<sup>2</sup>. Similarly, "low spatial variation L" can be predetermined, and, for example, might mean that there is a change in brightness  $\leq \pm c\%$  in the vertical dimension or horizontal dimension of an  $m \cdot n$  pixel block relative to the brightness of the brightest pixel within the  $m \cdot n$  pixel block, where  $c$  can have any suitable range such as 0 – 30 and  $m \cdot n$  can have any suitable range such as 4 – 500 pixels<sup>2</sup>. It is noted that although example definitions of "high spatial variation" and "low spatial variation" are defined above, it is contemplated that a photographer or another person, or a software application, can define "high spatial variation" and "low spatial variation" in any suitable manner.

**[0071]** Figure 4 illustrates how the picture of the sun behind the trees might be broken down into rectangles of different x, y coordinates in the database. The rectangles can be small enough so that each rectangle that includes a respective portion of the tree covers the respective portion of the tree and no, or only a tiny, portion of the space between the trees. In Figure 4, an arbitrary system of numbering the rectangles is used to refer to their location in the picture. In this case the rectangles are numbered starting from the bottom left. Each rectangle consists of many pixels (*e.g.*, 100 – 500 pixels). Rectangle (1,3), where x is 1 and y is 3, is mostly underexposed because it includes a portion of the tree, and rectangle (4,4) is mostly overexposed because it includes a portion of the sun. However, rectangle (3,4) is effectively half underexposed and half overexposed because it includes both a portion of the tree and a portion of the sun; thus the algorithm would not know whether to substitute a more exposed or less exposed picture when the viewer looks at this rectangle. But if smaller rectangles are used, then this problem is less likely to occur, since a photograph interesting to the human eye contains objects that take up more pixels than a rectangle of a certain predetermined size, for example, more than 10 pixels by 10 pixels (unless the photograph is of a checkerboard of black and white squares that are exactly the same size as the rectangles, for which the probability is vanishingly small). For example, in the above situation, the computer can increase the resolution of the algorithm by breaking down the rectangle (3,4) into smaller rectangles

until each rectangle is primarily overexposed or underexposed. Since the overhead of analyzing smaller rectangles with fewer pixels is trivial for many of today's computer integrated circuits (ICs) such as a microprocessor or microcontroller, there is no reason not to use smaller rectangles except that they could be smaller than the resolution of a gaze tracker or mouse or any other way the viewer might indicate where she is looking. This resolution is defined by the maximum resolution of the hardware used and the distance of the viewer from the display. Accordingly, the photographer can either specify in advance the hardware resolution and the distance, or default values will be used.

**[0072]** Other picture parameters can be varied and associated with any x-y area that a viewer will later view. Furthermore, if identities of particular potential viewers can be specified in advance and identified at display time, the system can accept as input which individual viewer is looking at the x-y coordinates and vary its response based on the identify of the viewer.

**[0073]** Steps taken by the image composer.

**[0074]** As shown in Figure 5, using a simple user interface on a computer display **12**, the photographer or other image composer **11** can quickly delineate and specify which portions of an image in the database should trigger which action (if any) when the viewer views them, or in what order, as follows:

**[0075]** The composer **11** is presented with a list of areas of the image likely to have an action associated with them when viewed (also referred to hereafter as "at display time"); for example, the list may include underexposed areas, overexposed areas, areas of high spatial resolution, or areas of low spatial resolution. The composer then specifies the action for each listed area. The action can be specified to be, but is not limited to, some of the following: replace the entire image or just that part of the image with another image that has a different exposure value that is, for example, neither underexposed nor overexposed; or that has a different variation of spatial frequencies that is, for example, the maximum detail for the resolution of the original camera, or is "in focus"; replace with image number x; replace with image number y if this is the nth replacement; replace with image z if area b was just looked at, etc. Different actions can be associated with each image in the database, or a set of actions associated with a primary image (see below). Actions can even be expanded to selecting images or videos available in real time, when the viewer is looking at the image, pulled from the internet or other systems.

**[0076]** A single photograph can be designated as the primary image that will be the image viewed and that will have all directions for when viewing an x-y rectangle within it; or any of the photographs can have these directions for when that photograph is viewed. Note that, in this example, if the composer does not specify an action, the default is "best automatic", as described below.

**[0077]** When the composer is finished composing all the actions that the viewer might take and all the substitutions of all or part of an image that would result, all this information is saved as one "show" or "Eyetrinary™". This show could all be stored online in the cloud, or in any other suitable electronic storage device or location.

**[0078]** When a viewer **13** views an image, as shown in Figure 6, a display **14** displays a photograph from the database **5** stored in and read by a computer **17**. A gaze-tracker hardware device **16** consisting, for example, of one or more cameras mounted on the display **14**, gives the x-y coordinates of where in the photograph the viewer's eyes are looking. The system uses these coordinates to look up the composer's specified action in the database **5**, and the system performs the action by referring to either the primary or current photograph's entry for that x-y rectangle.

**[0079]** If the action is the default "best automatic", the system will, for example, replace the picture the viewer is looking at with the best exposed picture for the x-y rectangle the viewer is currently viewing. "Best exposed" here, as before, is what a set of photographers would agree shows the most detail and is least "blown out" in lightness or black of which standard definitions are available. This replacement can be a quick dissolve, the nature and time of which can be specified in advance to be used when the viewer is looking, or a default value will be used.

**[0080]** It is important to note that the database referred to above might be stored online in the cloud as a page of html that consists of an initial image and a set of links or javascript calls that are associated with portions of that image. The gaze tracker would tell the browser where the viewer's eyes were looking, as if their eyes were a mouse, and the browser would automatically perform an action by either displaying a new image that is the target of a link, or by calling a javascript function to, for example, change some aspect of the image. In this way, the composer might need only to send a link to a viewer that points to the html page and a set of images, all within a directory on a server.

**[0081]** Instead of the composer specifying in advance using a user interface with a display and a pointer, instead he or she could use a display with gaze tracking hardware and specify, while looking at a particular x-y area, which other photograph (or portion thereof) to use by commands to the system through voice, or mouse, or any other method. These other correlations to the other photographs (or portions thereof) would be stored in the database. For example, if the photographer is looking at rectangle (1,3) in Figure 4, she could say, in effect, to a voice-recognition system, "when the viewer is looking at the area I am looking at now, replace the picture shown with the picture with the filename "tree-exposed.jpg".

**[0082] Additional Uses**

**[0083]** An embodiment can be used for more than expressing an artistic intent. For example, a vision therapist might want to increase a patient's ability to coordinate her two eyes by drawing her attention back and forth on a screen. A pilot could be trained with a flight-simulator system with a display to quickly locate threatening air traffic by making such threatening aircraft appear wherever she was not looking, based on a preexisting set of images or sprites. If the viewer allows her eyes to drift apart (looking further away, towards the horizon), an embodiment could substitute a picture taken with the focus further away. Thus a viewer could train herself to focus in and out and "travel" in and out, a form of 3D.

**[0084]** Gaze-tracking image tools. An embodiment can be used for a gaze-tracking-based user interface (UI) for image tools. For example, the viewer tilting her head to the right could inform the system to rotate the image slowly to the right until she tilted her head back to the center. Brightness and focus could be controlled by fixating on a portion of the image. Every time the viewer blinks her left eye, the brightness could increase until it reaches a maximum and then starts down again. Focus could be increased by blinking the right eye. The process could be stopped at the desired selection by closing both eyes for one second. Similar methods could be used for cropping.

**[0085]** Additional detail. In an embodiment, if the viewer's eye concentrated on a portion of the image, that portion of the image could expand or magnify with further detail. After the image with the maximum detail or magnification is substituted, and the viewer's eyes remain fixed on that area, a link associated with that portion of the image, predetermined by the photographer, could result in other images or a video being downloaded and substituted for that portion of the image. If the photographer did not put in a specific link, the initial image looked at in that section could be

transmitted to a cloud-based analyzer that could determine what the image is and supply further information, or at a minimum download and display other images of similar content captured by other photographers. For example, if the viewer was looking at a scene of Washington DC, and let her gaze remain over the Washington Monument, the Monument might expand in size and detail to take up one half of the entire display by means of a succession of substituted images for that section of the image. If the viewer continues to stare at the Monument, a video downloaded from the internet might play on that section of the screen describing the Monument's history and construction (which might either have been specified by the photographer in advance or automatically selected by a cloud-based image recognizer).

**[0086]**     "Flashlighting." The photographer could specify that whatever portion of the image at which the viewer looks could be immediately modified to a higher brightness. The composer would merely have to specify "flashlight" for an x-y area. Thus it would appear that a flashlight, in effect aimed by the viewer's eyes, was playing upon the scene. Human interface testing. There are existing systems that record where a person tends to look in an image, with applications to advertising. In an embodiment, one could imagine a log being output based on how often a viewer returns to a particularly illuminated or focused portion of the image. Also, if the image composer specified a "path" of image substitutions, *i.e.*, looking at a subportion of image A results in a substitution of image B, and looking at a portion of image B goes back to A, and the viewer repeated this path A-B-A, a log would save the information of the path and that it was repeated. Such information would be useful to the composer in further developing the viewer's interaction with the image(s).

**[0087]**     Networking and games with other viewers. More than one gaze-tracking device can be used for games with other viewers in the same location; and the system can be connected to the internet or local network for various social purposes. For example, if more than one viewer in different locations is looking at the same Eyetinerary™, the portions looked at by a majority of viewers could be the ones magnified, or the opposite: if most people have not seen something, the Eyetinerary™ might be modified to force all the viewers to go to that image. Other types of social interaction can be implemented, such as games where two or more gaze-tracking systems award points based on the first person to look in the right place, or setting off weapons in the right place,

etc. Finally, portions of the image can be shared with others on the internet by looking at the portions.

**[0088]**     Feedback to the composer. An embodiment concerns a pre-determined intention of the composer to respond to where a viewer looks at display time. However, a viewer could use the system to send useful feedback to the composer in real time over a network. For example, if there's a portion of the image, say a particular person in a picture of multiple people, that the viewer would like the composer to make a substitution for, he could blink three times while looking at the person, and then drawing a circle around it with his eyes. The viewer might then say, over a network connection, "when I look at this person, I would like you to substitute another image." This location of the face would be passed to the composer along with the voice file. Thus the composer could develop a better or different Eyetinerary™ with viewer input.

**[0089]**     From the foregoing it will be appreciated that, although specific embodiments have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the disclosure. Furthermore, where an alternative is disclosed for a particular embodiment, this alternative may also apply to other embodiments even if not specifically stated. In addition, any described component or operation may be implemented/performed in hardware, software, firmware, or a combination of any two or more of hardware, software, and firmware. Furthermore, one or more components of a described apparatus or system may have been omitted from the description for clarity or another reason. Moreover, one or more components of a described apparatus or system that have been included in the description may be omitted from the apparatus or system.

What is claimed:

1. A method for modifying one or more original images displayed on one or more display screens according to an algorithm that operates as a function of data collected via one or more cameras capturing an image of one or more viewers' faces before the one or more display screens.
2. The method of claim 1 where the data collected is one or more of: where location of gaze of eyes of one or more viewers is directed within the original image, or face appearance change of one or more viewers, or identity of one or more viewers.
3. The method of claim 1 where the data collected is one or more of: ambient light, or distance of a viewer from a display screen, or number of viewers.
4. The method of claim 1 where the modification is to change brightness of a part or all of the original image.
5. The method of claim 4 where the modification includes replacement of the original image with a substitute image where the contents of the substitute image are the same as the original image but with some or all pixels modified.
6. The method of claim 1 where the modification is to change focus of all or a part of the original image.
7. The method of claim 6 where the modification is replacement of the original image with a substitute image where the contents of the substitute image are the same as the original image but with some or all pixels modified.
8. The method of claim 6 where the modification is to change the focus to objects that appear to be close in part or all of the image.
9. The method of claim 6 where the modification is to change the focus to objects that appear to be far away in part or all of the image.
10. The method of claim 6 where the modification is to change the apparent depth of field of part or all of the image.



11. The method of claim 6 where, when a viewer is gazing at the one or more images, processing of data gathered from the one or more cameras determines whether the viewer is causing their eyes to diverge as if looking at a distant object and, if so, the modification is to change focus of parts of the image such that objects that appear close in the image are defocused and objects that appear far away are brought into sharper focus.
12. The method of claim 6 where, when a viewer is gazing at the one or more images, the modification is to defocus parts of the image around a focused area and move the focused area across the image to draw a viewer's gaze across the image.
13. The method of claim 1 where the modification is to zoom in or out with the zooming centered on a part of the original image.
14. The method of claim 13 where the modification includes replacement of the original image with a substitute image.
15. The method of claim 13 where the modification is to zoom in or out in synchrony with changes in measured distance from a display screen to a viewer.
16. The method of claim 13 where the modification is to zoom in with the zooming centered on a location in the original image where a viewer's gaze is fixed.
17. The method of claim 1 where the modification is to change color balance or color saturation or color depth or contrast of part or all of the image.
18. The method of claim 17 where the modification is replacement of the original image with a substitute image where the contents of the substitute image are the same as the original image but with some or all pixels modified.
19. The method of claim 1 where the modification is to cause or change playing of sounds that accompany the original image.
20. The method of claim 1 where the modification is to cause movement or cessation of movement of a sprite within the original image.
21. The method of claim 1 where the modification is to select a branch of a multi-branch video.

22. The method of claim 1 where the one or more original images are still.
23. The method of claim 1 where the one or more original images are video.
24. The method of claim 1 where the one or more original images are still with moving sprites.
25. The method of claim 1 where the one or more original images are two dimensional.
26. The method of claim 1 where the one or more original images are three dimensional.
27. The method of claim 1 where the algorithm is custom determined by an artist who selected the one or more images.
28. The method of claim 1 where the algorithm is pre-determined by a company that supplied software for selection and use by an artist or content creator who selected the one or more original images.
29. The method of claim 1 where the algorithm is pre-determined by a company that supplied software and the algorithm works with any image or set of images.
30. The method of claim 1 where the algorithm allows an artist or content creator to specify a speed at which a gradual change modification will be performed.
31. A method for modifying an original one or more images displayed on one or more display screens according to an algorithm that operates as a function of data collected from touch by the viewer touching a portion of the screen within the image, where the modification is replacement of the original image with one or more substitute images where the contents of the substitute images are the same as the original image but with some or all pixels modified.
32. The method of claim 31 where the data collected from touch comprises one or more of: touch by one or multiple fingers, force of touch, duration of touch, and movement of touch.
33. The method of claim 31 where the modification is to change focus of all or a part of the original image.
34. The method of claim 33 where the modification comprises one or more of:  
to change the focus to objects that appear to be close in part or all of the image, or

to change the focus to objects that appear to be far away in part or all of the image, or  
to change the apparent depth of field of part or all of the image.

35. The method of claim 31 where the modification comprises one or more of:  
to change brightness of a part or all of the original image, or  
to zoom in or out from a part of the original image, or  
to change color balance or color saturation or color depth or contrast of part or all of the original image, or  
to cause movement or cessation of movement of a sprite within the original image, or  
to select a branch of a multi-branch video.

36. The method of claim 31 where the one or more original images are:  
still, or  
video, or  
still with moving sprites, or  
two dimensional, or  
three dimensional.

37. The method of claim 31 where the algorithm comprises one or more of:  
custom determined by an artist who selected the one or more images, or  
pre-determined by a company that supplied software for selection and use by an artist or content creator who selected the one or more original images, or  
pre-determined by a company that supplied software and the algorithm works with any image or set of images, or  
allows an artist or content creator to specify a speed at which a gradual change modification will be performed.

38. A method for modifying an original one or more images displayed on one or more display screens according to an algorithm that operates as a function of data collected from voice sounds, where the modification is replacement of the original image with one or more substitute images where the contents of the substitute images are the same as the original image but with some or all pixels modified.

39. The method of claim 38 where the modification comprises one or more of:

to change brightness of a part or all of the original image;

to change focus of a part or all of the original image;

to zoom in or out from a part or all of the original image;

to change color balance or color saturation or color depth or contrast of part or all of the original image;

to change sharpness of part or all of the original image;

to cause movement or cessation of movement of a sprite within the original image;

to select a branch of a multi-branch video.

40. The method of claim 38 where, from the data collected from voice sounds, a likely identity of a person speaking is determined and the modification is determined at least in part by the likely identity.

41. A method for modifying an original image displayed on a hand-held display screen according to an algorithm that operates as a function of data collected from one or more accelerometers embedded in a housing of the display screen where the modification is to change field of view of the original image by enlarging or contracting each of one or more of the left boundary, the right boundary, the top boundary or the bottom boundary.

42. The method of claim 41 where the data collected from the one or more accelerometers is a tilting of a first edge of the display away from a viewer relative to an opposite edge which causes more of the image near the first edge to come into view.

43. The method of claim 41 where the modification is to zoom in or out with the zooming centered on a part of the original image.

44. A method in a system with a server computer and a client computer to receive instructions from an author and, based on those instructions, generate a data set for showing images that change based on actions of a viewer, comprising:

(a) receiving at the server computer from a client computer specification of a plurality of images in a sequence;

(b) receiving at the server computer from the client computer specification of data to be collected that should trigger a transition from an original image in the sequence of images to a second image in the sequence of images;

(c) receiving at the server computer from the client computer a specification of speed of transition from the original image in the sequence of images to the second image in the sequence of images; and

(d) assembling on the server computer a data set that can be transmitted to a viewer's computer where the data set can be viewed on the viewer's computer and data input received from actions of the viewer causes modification of the viewed image by replacing the original image in the sequence with the second image at the specified transition speed.

45. The method of claim 44 where the data input comprises one or more of:  
location of gaze of eyes of one or more viewers within the original image, or  
face appearance change of one or more viewers, or  
ambient light, or  
distance of a viewer from a display screen, or  
number of viewers, or  
voice input, or  
touch input.

46. The method of claim 44 where the modification comprises one or more of:  
to change brightness of a part or all of the original image;  
to change focus of a part of the original image;  
to zoom in or out from a part of the original image;  
to change the color balance or color saturation or color depth or contrast of part or all of the original image;  
to cause movement or cessation of movement of a sprite within the original image;  
to select a branch of a multi-branch video.

47. The method of claim 44 where the original image is:  
still, or  
video, or

still with moving sprites, or  
two dimensional, or  
three dimensional.

48. The method of claim 44 where the client computer and the server computer are contained within a single computer housing and activated by a single user.

# ABSTRACT

An embodiment combines the concepts of image enhancement and gaze tracking to enhance the experience of viewing images by tracking where the viewer is looking. The result is to make the viewing experience more like viewing the original scene, or to enhance the viewing experience in new ways beyond the original experience, either automatically, or by interacting with a photographer's previously specified intentions for what should happen when the viewer looks at a particular portion of an image or images taken by that photographer.

Representative Drawing

Fig.1

FIG. 1

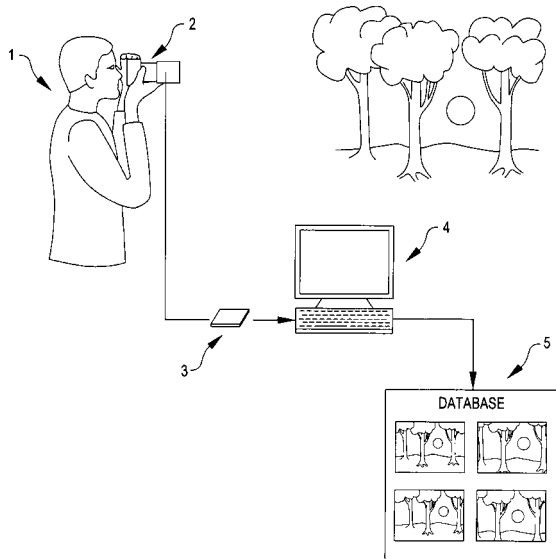


FIG. 2

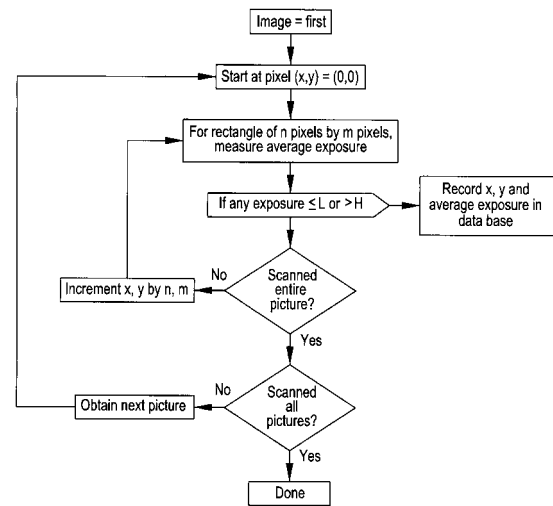


FIG. 3

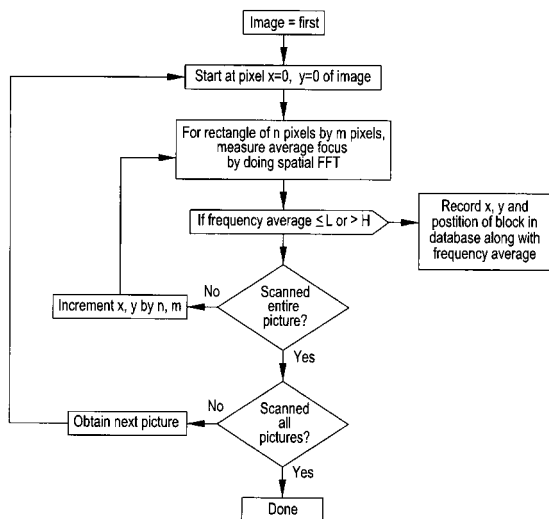


FIG. 4

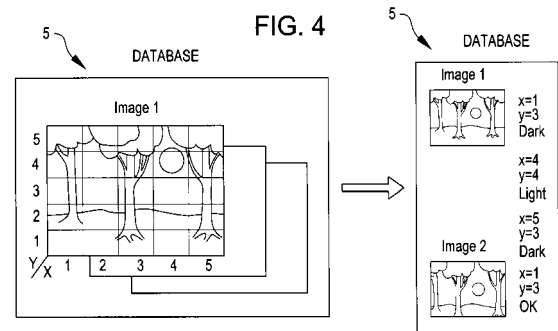


FIG. 5

