

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第4499380号
(P4499380)

(45) 発行日 平成22年7月7日 (2010.7.7)

(24) 登録日 平成22年4月23日 (2010.4.23)

(51) Int.Cl.
H04N 7/15 (2006.01)

F I
H04N 7/15 630Z

請求項の数 58 (全 42 頁)

(21) 出願番号	特願2003-175464 (P2003-175464)	(73) 特許権者	500046438
(22) 出願日	平成15年6月19日 (2003.6.19)		マイクロソフト コーポレーション
(65) 公開番号	特開2004-80750 (P2004-80750A)		アメリカ合衆国 ワシントン州 9805
(43) 公開日	平成16年3月11日 (2004.3.11)		2-6399 レッドモンド ワン マイ
審査請求日	平成18年6月15日 (2006.6.15)		クロソフト ウェイ
(31) 優先権主張番号	10/178,443	(74) 代理人	100077481
(32) 優先日	平成14年6月19日 (2002.6.19)		弁理士 谷 義一
(33) 優先権主張国	米国 (US)	(74) 代理人	100088915
			弁理士 阿部 和夫
		(72) 発明者	ツェンユー ツァン
			アメリカ合衆国 98052 ワシントン
			州 レッドモンド 177 アベニュー
			ノースイースト 10090

最終頁に続く

(54) 【発明の名称】 ホワイトボードおよび音声のキャプチャのためのシステムおよび方法

(57) 【特許請求の範囲】

【請求項 1】

オーディオビジュアルの内容を要約し、インデックスを付けるための方法であって、
電子式でないホワイトボードに書き込まれた内容のイメージシーケンスをカメラを使用してキャプチャするステップと、
前記イメージシーケンスに相関する音声信号を記録するステップと、
前記イメージシーケンスを解析して前記ホワイトボードの内容の重要なポイントを要約するキーフレームを分離するステップ
を含み、
前記解析するステップは、
前記イメージシーケンスの中のすべてのイメージ内の前記ホワイトボードのビューを修正するステップと、
ホワイトボード背景色を抽出するステップと、
セルイメージをそれぞれが表わすセルに前記イメージシーケンスの各イメージを分割するステップと、
時間にわたる各セルに関する前記イメージシーケンス全体のセルイメージをクラスタ化するステップと、
各セルイメージをストローク、前景オブジェクト、またはホワイトボードセルとして分類するステップと、
当該分類結果を使用してキーフレームイメージを抽出するステップ

を含むことを特徴とする方法。

【請求項 2】

前記記録された音声信号を前記キーフレームに相関させるステップをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記音声信号が、前記記録された音声信号と前記イメージシーケンスの両方に付けられたタイムスタンプで前記イメージシーケンスに相関させられることを特徴とする請求項 2 に記載の方法。

【請求項 4】

前記音声信号を前記イメージシーケンスと相関させることは、
前記イメージシーケンスがキャプチャされた時刻に共通クロックで前記イメージシーケンスにタイムスタンプを付けることと、
前記音声信号が記録された時刻に共通クロックで前記音声信号にタイムスタンプを付けることと、
前記共通クロックの前記タイムスタンプを使用して前記イメージシーケンスと前記音声信号を相関させること
を含むことを特徴とする請求項 3 に記載の方法。

10

【請求項 5】

前記イメージシーケンスの中の所望のポイントで前記イメージシーケンス、および前記相関させられた前記音声信号にアクセスすることをさらに含むことを特徴とする請求項 3 に記載の方法。

20

【請求項 6】

前記キーフレームは、前記イメージシーケンスの中の前記所望のポイントを選択するのに使用されることを特徴とする請求項 5 に記載の方法。

【請求項 7】

セルイメージを空間的にフィルタにかけて分類結果を精密にするステップをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 8】

前記セルイメージを時間的にフィルタにかけて分類結果を精密にするステップをさらに含むことを特徴とする請求項 1 に記載の方法。

30

【請求項 9】

前記キーフレームイメージにカラーバランス処理を行ってイメージ品質を向上させることをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 10】

前記ホワイトボードを修正するステップは、
イメージ内の前記ホワイトボードの 4 つの隅を指定することと、
各イメージのホワイトボードでない領域をトリミングすることと、
バイキュービック補間を使用して各ホワイトボード領域に関する長方形の形状に対してバイリニアワープを行い、前記イメージシーケンスのそれぞれのキャプチャされたイメージ内でトリミングされ、修正されたホワイトボードイメージを獲得すること
をさらに含むことを特徴とする請求項 1 に記載の方法。

40

【請求項 11】

前記ホワイトボード背景色を抽出するためのステップは、
最も明るい輝度および最小の変動を有するホワイトボードセルを判別することと、
ホワイトボードの色の中のホールを判別することと、
前記ホールの周囲のホワイトボードセルを探索し、かつ、ホールでない最も近接したセルの色にカラー設定することによって各ホールを埋めること
を含むことを特徴とする請求項 1 に記載の方法。

【請求項 12】

前記ホワイトボードの中のホールを判別するためのステップは、

50

ホワイトボードセルである可能性がある色の範囲を指定するステップと、
色の前記範囲外にあるかどうか各ホワイトボードセルを検査するステップと、
前記範囲外の色を有するセルをホールとして指定するステップ
を含むことを特徴とする請求項1に記載の方法。

【請求項 1 3】

前記ホワイトボード背景色を抽出するステップは、
ホワイトボードイメージ輝度のヒストグラムを作成することと、
ピークのホワイトボード輝度を判定することと、
ピークの輝度に対応する色を初期のホワイトボードの色として指定することと、
誤りのあるデータを表わすホワイトボードの色の外れ値を判別することと、
外れ値にホールとしてマークを付けることと、
ホールの周囲のセルを探索し、ホールでない最も近接したセルの色にカラー設定することによって各ホールを埋めること
を含むことを特徴とする請求項1に記載の方法。

10

【請求項 1 4】

各ホールを埋めた後、ホワイトボードのカラーイメージをフィルタリングすることをさらに含むことを特徴とする請求項1 3に記載の方法。

【請求項 1 5】

最小 2 乗平均法を使用して外れ値が検出されることを特徴とする請求項1 4に記載の方法。

20

【請求項 1 6】

前記最小 2 乗平均法は、
2 乗された誤りの中央値を最小化することによって輝度 Y または R G B の空間に平面を合わせるステップと、
上記モデルに従わないセルを外れ値として指定するステップ
を含むことを特徴とする請求項1 5に記載の方法。

【請求項 1 7】

前記ホワイトボードのカラーイメージをフィルタリングすることは、平面をレッド、グリーン、ブルー (R G B) の色空間に合わせることを含むことを特徴とする請求項1 4に記載の方法。

30

【請求項 1 8】

前記セルに各イメージを分割するステップは、各イメージを長方形のセルに分割することを含むことを特徴とする請求項1に記載の方法。

【請求項 1 9】

前記セルに各イメージを分割するステップは、セルサイズが前記ホワイトボード上の単一の書き込まれた文字のサイズとほぼ同じサイズであるセルに各イメージを分割することを含むことを特徴とする請求項1に記載の方法。

【請求項 2 0】

前記セルのサイズは、前記ホワイトボードにおける縦 1 . 5 インチ (3 . 8 センチメートル)、横 1 . 5 インチ (3 . 8 センチメートル) の領域に対応することを特徴とする請求項1 9に記載の方法。

40

【請求項 2 1】

前記セルのサイズは、カメラ解像度およびホワイトボードサイズで決まることを特徴とする請求項1 8に記載の方法。

【請求項 2 2】

前記セルのサイズは、縦 1 画素、横 1 画素であることを特徴とする請求項1 4に記載の方法。

【請求項 2 3】

前記セルイメージをクラスタ化するステップは、時間を経て同じであると見なされるセルイメージをクラスタ化することを含むことを特徴とする請求項1に記載の方法。

50

【請求項 2 4】

前記時間を経て同じであると見なされるセルイメージをクラスタ化することは、正規化された相互相関法を使用して一度に 2 つのセルを比較することにより、セルイメージが同じであるかどうかを判定する動作を含むことを特徴とする請求項 2 3 に記載の方法。

【請求項 2 5】

前記相互相関のスコアは、2 つのイメージが全く似通っていないことを表わす - 1 から、2 つの全く同一のイメージを表わす 1 までの範囲内にあることを特徴とする請求項 2 4 に記載の方法。

【請求項 2 6】

マハラノビス距離試験を適用して 2 つのセルが同じであるかどうかを判定するステップをさらに含むことを特徴とする請求項 2 5 に記載の方法。

10

【請求項 2 7】

前記マハラノビス距離試験は、

【数 1】

$$d = |\bar{I} - \bar{I}'| / (\sigma + \sigma')$$

によって与えられ、ただし、 \bar{I} は、第 1 のセルイメージであり、かつ、 \bar{I}' は、第 2 のセルイメージであり、

【数 2】

$$\bar{I}$$

20

は、前記第 1 のセルイメージの平均の色であり、

【数 3】

$$\bar{I}'$$

は、前記第 2 のセルイメージの平均の色であり、

【数 4】

$$\bar{I}$$

30

からの標準偏差であり、かつ、

【数 5】

$$\bar{I}'$$

からの標準偏差であって、

\bar{I} および \bar{I}' は、 $d < T_d$ かつ $c > T_c$ である場合にだけ、同一であると見なされ、かつ、1 つの特定の実施形態では、 $T_d = 2$ かつ $T_c = 0.707$ であることを特徴とする請求項 2 6 に記載の方法。

【請求項 2 8】

前記各セルイメージを分類するステップは、

40

R G B 値がほぼ同一である場合、セルイメージをホワイトボードのセルとして指定することと、

前記セルが大体において白またはグレーであり、1 つまたは 2 つの原色が混じっている場合、セルイメージをストロークとして指定することと、

セルイメージがホワイトボードセルの特性、またはストロークセルの特性を有さない場合、セルイメージを前景セルとして指定すること

をさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 2 9】

前記各セルイメージを分類するステップは、

現行のセルイメージの色分布を判定することと、

50

対応するホワイトボードセルの色分布を判定することと、
 前記現行のセルイメージの前記色分布と前記対応するホワイトボードセルの色分布が同一であるか、同一ではないが強い類似性を有するか、または全く異なっているかを判定すること
 をさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 30】

前記セルイメージは、

【数 6】

$$|\bar{I} - \bar{I}_w| / (\sigma + \sigma_w) < T_w$$

10

であり、かつ、 $\sigma / \sigma_w < T$ である場合にだけ、ホワイトボードセルとして分類され、ただし、

【数 7】

$$\bar{I}_w$$

は、ホワイトボードの色であり、 σ_w は、

【数 8】

$$\bar{I}_w$$

20

からの標準偏差であり、

【数 9】

$$\bar{I}$$

は、現行のセルの平均の色であり、かつ、 σ は、

【数 10】

$$\bar{I}$$

からの標準偏差であり、 $T_w = 2$ かつ $T = 2$ であり、それ以外の場合、前景オブジェクトセルとして分類されることを特徴とする請求項 28 に記載の方法。

30

【請求項 31】

前記セルイメージを空間的にフィルタにかけるステップは、

孤立した前景セルを特定するステップと、

孤立した前景セルをストロークセルとして分類しなおすステップと、

前景セルに結合したストロークセルを特定するステップと、

前景セルに結合した前記ストロークセルを前景セルとして分類しなおすステップ

を含むことを特徴とする請求項 7 に記載の方法。

【請求項 32】

前記セルイメージを空間的にフィルタにかけるステップは、

時間にわたって各セルを評価して、前記各セルが、前景セルとして分類された前後に同一のストロークセルを含んでいたかどうかを判定するステップと、

40

セルが、前景セルとして分類された前後に同一のストロークセルとして分類されていた場合に、前記セルをストロークセルとして分類するステップ

を含むことを特徴とする請求項 8 に記載の方法。

【請求項 33】

前記キーフレームイメージを抽出するステップは、

前記分類結果を入力することと、

前記イメージシーケンスの中の各フレームに関するストロークセルを計数することと、

前記フレームに関する前記ストロークカウントを使用してストロークカウントのピークおよびくぼみを判定することと、

50

それぞれの隣接するピークとくぼみの間の差が、所定のしきい値を超えた場合、くぼみとくぼみの間のデータを区分として指定し、各区分の中の前記ピークを前記区分を代表するキーフレームとして指定すること

を含むことを特徴とする請求項 1 に記載の方法。

【請求項 3 4】

前記所定のしきい値は、合計のストロークカウントの 20 パーセントであることを特徴とする請求項 3 3 に記載の方法。

【請求項 3 5】

前記キーフレームイメージを再構成するステップをさらに含むことを特徴とする請求項 3 3 に記載の方法。

【請求項 3 6】

前記キーフレームイメージを再構成するステップは、

前記イメージシーケンスの分類されたセルイメージ、およびセルイメージに分割されたキーフレームを入力するステップと、

キーフレームセルイメージが、ホワイトボードイメージまたはストロークイメージとして分類された場合、ホワイトボードまたはストロークのイメージがレンダリングされるステップと、

キーフレーム前景セルイメージが、ストロークの範囲内にある場合、そのセルイメージがシーケンスの中の隣接するイメージからのストロークセルイメージを使用してレンダリングされるステップと、

前記キーフレームセルイメージが、ホワイトボードイメージとしても、ストロークイメージとしても、またはストロークの範囲内の前景セルとしても分類されない場合、ホワイトボードイメージとしてレンダリングされるステップ

をさらに含むことを特徴とする請求項 3 5 に記載の方法。

【請求項 3 7】

前記キーフレームイメージにカラーバランス処理を行ってイメージ品質を向上させることは、

平均のホワイトボードの色を使用してセル内の各画素の色を調整することによって背景を一様に白くし、ペンストロークの色の彩度を高めるステップと、

イメージノイズを低減するステップ

をさらに含むことを特徴とする請求項 9 に記載の方法。

【請求項 3 8】

会議の音声の内容および映像の内容をキャプチャするためのシステムであって、

ホワイトボードに書き込まれたデータのイメージシーケンス、および会議中に生じたサウンドに対応する音声信号をキャプチャするキャプチャシステムと、

前記ホワイトボードに書き込まれたキーデータフレームを抽出し、前記音声信号を前記キーデータフレームに関連させる、前記イメージシーケンスを解析するためのアナリシスサーバと、

前記解析された会議キーデータフレーム、および関連させられた音声を閲覧するためのブラウズモジュール

を含み、

前記アナリシスサーバは、前記キーデータフレームを、

前記イメージシーケンスの中のすべてのイメージの中で前記ホワイトボードのビューを修正すること、

ホワイトボード背景色を抽出すること、

前記イメージシーケンスの各イメージをセルイメージのセルに分割すること、

時間にわたって各セルに関して前記イメージシーケンス全体で似通っているセルイメージをクラスタ化すること、

各セルイメージをストローク、前景オブジェクト、またはホワイトボードセルとして分類すること、および、

10

20

30

40

50

前記分類結果を使用してキーフレームイメージを抽出することによって特定することを特徴とするシステム。

【請求項 3 9】

前記キャプチャシステムは、
イメージシーケンスをキャプチャするように配置されたカメラと、
前記音声信号を記録するマイクロホンと、
前記イメージシーケンスおよび前記音声信号を記録するためのコンピュータ
をさらに含むことを特徴とする請求項 3 8 に記載のシステム。

【請求項 4 0】

前記カメラは、スチールカメラ、および、ビデオカメラの少なくともどちらかであること
を特徴とする請求項 3 9 に記載のシステム。 10

【請求項 4 1】

前記カメラは、解像度を最大化するように可能な限り前記ホワイトボードに接近してズームインされることを特徴とする請求項 3 9 に記載のシステム。

【請求項 4 2】

前記カメラは、シーン深度を最小化するように可能な限り前記ホワイトボードに平行に揃えられることを特徴とする請求項 3 9 に記載のシステム。

【請求項 4 3】

前記アナリシスサーバは、前記キーデータフレームを、
前記イメージシーケンスの中のすべてのイメージの中で前記ホワイトボードのビューを
修正すること、
ホワイトボード背景色を抽出すること、
時間にわたって各セルに関して前記イメージシーケンス全体で似通っている画素をクラスタ化すること、
各画素をストローク、前景オブジェクト、またはホワイトボードセルとして分類すること、および、
前記分類結果を使用してキーフレームイメージを抽出すること
によって特定することを特徴とする請求項 3 8 に記載のシステム。 20

【請求項 4 4】

記録が開始する前に、1名以上のユーザがキャプチャユニットにおいてユーザ識別子を登録し、
少なくとも1名のユーザが、前記キャプチャユニットで登録した場合、前記アナリシスサーバはイベントの記録および解析の後にアクセストークンを生成し、
前記アクセストークン、および解析された会議データのコンピュータメモリロケーションが前記登録されたユーザ識別子に提供され、さらに、
前記1名以上のユーザが、前記解析されたイベントデータの前記コンピュータメモリロケーションにアクセスして前記解析された会議データを閲覧すること
を特徴とする請求項 3 8 に記載のシステム。 30

【請求項 4 5】

前記ユーザ識別子は電子メールアドレスであることを特徴とする請求項 4 4 に記載のシステム。 40

【請求項 4 6】

前記解析されたイベントデータの前記コンピュータメモリロケーションはインターネットWebサイトのアドレスであることを特徴とする請求項 4 4 に記載のシステム。

【請求項 4 7】

前記イメージシーケンスの部分、および、前記音声の部分の少なくともどちらかをユーザが消去することがかできるようにするプライバシー機構を前記キャプチャユニット内部にさらに含むことを特徴とする請求項 3 8 に記載のシステム。

【請求項 4 8】

前記プライバシー機構は、グラフィカルユーザインターフェースのボタン、または物理 50

的ボタンを押すことによって起動されることを特徴とする請求項4 7に記載のシステム。

【請求項 4 9】

前記キャプチャシステムはポータブルであることを特徴とする請求項3 8に記載のシステム。

【請求項 5 0】

イベントの内容を取り出すための方法であって、

電子式でないホワイトボードに書き込まれた内容のイメージシーケンスをカメラを使用してキャプチャするステップと、

前記イメージシーケンスに相関する音声信号を記録するステップと、

前記イメージシーケンスを解析して前記ボードの内容の要点を、

前記イメージシーケンスの中の各イメージ内の前記ホワイトボードのビューを修正すること、

ホワイトボード背景色を抽出すること、

前記イメージシーケンスの各イメージをセルに分割すること、

時間にわたって各セルに関して前記イメージシーケンス全体のセルイメージをクラスタ化すること、

各セルイメージをストローク、前景オブジェクト、またはホワイトボードセルとして分類すること、および、

前記分類結果を使用してキーフレームイメージを抽出すること

によって要約するキーフレームを分離するステップ

を含むことを特徴とする方法。

【請求項 5 1】

前記解析して分離するステップは、光学式文字認識技術を使用してキーフレームイメージから編集可能なテキストを提供するステップをさらに含むことを特徴とする請求項5 0に記載の方法。

【請求項 5 2】

前記解析して分離するステップは、音声認識技術を使用して音声をテキストに変換することをさらに含むことを特徴とする請求項5 0に記載の方法。

【請求項 5 3】

ジェスチャ認識技術を使用することをさらに含むことを特徴とする請求項5 0に記載の方法。

【請求項 5 4】

前記キャプチャするステップは、

既知の色特性を有する色パッチを前記ホワイトボードの上方に取り付けることと、

前記イメージシーケンスの中のイメージ内で前記色パッチをキャプチャすることと、

前記イメージ内の前記色パッチが示す色を判別することと、

前記イメージシーケンスの中の次のイメージ内で前記色パッチをキャプチャすることと

、

前記次のイメージ内の前記色パッチが示す色を判別することと、

前記パッチが示す色が前記イメージと前記次のイメージの間で所定のしきい値を超えて変化しているかどうかを判定することと、

示された前記色が、前記しきい値を超えている場合に、設定されていたとすれば、前記次のイメージ内の前記色パッチの色を前記イメージ内で示される前記色パッチの色と一致させたであろう前記カメラの露出設定に、前記カメラの露出パラメータを調整することをさらに含むことを特徴とする請求項5 0に記載の方法。

【請求項 5 5】

前記調整することは、キャプチャされたすべてのイメージに対して行われることを特徴とする請求項5 4に記載の方法。

【請求項 5 6】

会議の内容を取り出すためのシステムであって、

ホワイトボードに書き込まれたデータシーケンスをキャプチャし、ペン位置を追跡してホワイトボードに書き込まれた内容を推定し、かつ、ホワイトボードに書き込まれた前記内容に相関する音声信号を記録するキャプチャシステムと、

前記ホワイトボードに書き込まれたキーデータフレームを抽出し、前記音声信号を前記キーデータフレームに相関させるイメージシーケンスを解析するためのアナリシスサーバを含み、

前記アナリシスサーバは、

前記ホワイトボードの各領域をセルに分割する動作と、

時間にわたって各セルに関して書き込まれた前記データシーケンス全体で同じであるセルをクラスタ化する動作と、

各セルをストロークセルまたはホワイトボードセルとして分類すると、

前記分類結果を使用してキーフレームイメージを抽出する動作
を行うことを特徴とするシステム。

【請求項 57】

前記ホワイトボードセルは、1つの書き込まれた文字とほぼ同じサイズのセルに分割されることを特徴とする請求項 56 に記載のシステム。

【請求項 58】

会議中に記録されたデータに対する非線形アクセスを提供するためのコンピュータ実行可能命令を有するコンピュータ可読記憶媒体であって、

前記コンピュータ実行可能命令は、

非電子式のホワイトボードに書き込まれた内容のイメージシーケンスをカメラを使用してキャプチャするためのプログラムモジュールと、

前記イメージシーケンスに相関する音声信号を記録するためのプログラムモジュールと、

前記イメージシーケンスを解析して前記ホワイトボードの内容の要点を、

前記イメージシーケンスの中の各イメージ内の前記ホワイトボードのビューを修正すること、

ホワイトボード背景色を抽出すること、

前記イメージシーケンスの各イメージをセルに分割すること、

時間にわたって各セルに関して前記イメージシーケンス全体のセルイメージをクラスタ化すること

各セルイメージをストローク、前景オブジェクト、またはホワイトボードセルとして分類すること、および、

前記分類結果を使用してキーフレームイメージを抽出すること、
によって要約するキーフレームを分離するためのプログラムモジュール
を含むことを特徴とする媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、会議を記録するためのシステムおよび方法を対象とする。より詳細には、本発明は、会議のホワイトボードの内容と音声とともにキャプチャするためのシステムおよび方法を対象とする。

【0002】

【従来の技術】

会議は、多くの勤労者の労働時間の大きな部分を成している。会議で費やされる時間をより効率的に使用することにより、生産性が大幅に向上することになる。

【0003】

多くの会議のシナリオは、ブレインストーミング (brainstorming) セッション、講演、プロジェクト計画会議、特許情報公開等のために頻繁にホワイトボードを使用する。ボード上に書かれたものをノートに取り、コピーすることが、しばしば、この会

10

20

30

40

50

議中の多くの参加者の積極的な貢献および関与の妨げとなる。この結果、何らかの自動化された仕方ホワイトボードの内容をキャプチャしようとするいくつかの試みがなされてきた。

【 0 0 0 4 】

ホワイトボードの内容を自動的にキャプチャするいくつかの技術が開発されている。最早期のものの1つである、ホワイトボード複写機は、組み込まれた複写機を備えた特殊なホワイトボードである。ボタンのクリックで、ホワイトボードの内容が走査され、印刷される。ホワイトボードの内容は、紙に印刷されると、複写する、ファックス送信する、ファイルキャビネットに収納する、または走査してデジタル形式にすることができる。

【 0 0 0 5 】

より新しい技術は、最初からデジタル形式でホワイトボードの内容をキャプチャしようとする。この技術は、2つのカテゴリに入る。すなわち、ホワイトボードのイメージをキャプチャする技術、およびペンの位置を追跡し、その位置からホワイトボードの内容を推測する技術である。

【 0 0 0 6 】

第1のカテゴリのデバイスは、ホワイトボードのイメージを直接にキャプチャする。ナショナルテレビジョンシステム委員会(National Television System Committee)(NTSC)解像度ビデオカメラが、しばしば、廉価であるために使用される。このカメラは、普通、通常の会議室サイズのホワイトボード上に書かれたものを鮮明にキャプチャするのに十分な解像度を有していないため、いくつかの映像フレームを一緒につなぎ合わせて単一のホワイトボードイメージを生成しなければならない。第1のカテゴリの別のデバイスは、デジタルスチールカメラ(still camera)である。高解像度のデジタルカメラがより安価になるにつれ、ボードのスナップショットをデジタルカメラで撮ることが一般的な選択になっている。

【 0 0 0 7 】

第2のカテゴリのデバイスは、高い頻度でホワイトボード上に書き込むのに使用されるペンの位置を追跡し、ペン座標の履歴からホワイトボードの内容を推測する。ときとして、このデバイスは、従来のホワイトボードの側部に取り付けられた付加デバイスを含み、ドライインク(dry-ink)ペンとイレーザ(eraser)のための特殊なケースを使用する。それぞれのペンが、ボードに押しつけられると超音波パルスを発する。付加デバイスにおける2つのレシーバが、音声パルスの到着時刻の差を使用してペン座標を三角法で測定する。ペン座標の履歴がキャプチャされるので、任意の所与の時点におけるホワイトボード上の内容を後に再構成することができる。このタイプのホワイトボード記録のユーザは、映画のようにホワイトボードの内容を再生することができる。内容は、ベクトル形式でキャプチャされるため、低帯域幅および小さいストレージの要件で伝送し、アーカイブする(archive)ことができる。

【 0 0 0 8 】

また、電子ホワイトボードは、ペン追跡技術も使用する。電子ホワイトボードは、ホワイトボードを対話型デバイスにすることで、前述した付加デバイスを使用するシステムより一歩前進している。ユーザは、コンピュータによって追跡される特殊なスタイラス(stylus)を使用してモニタ上に書き込む。コンピュータが、スタイラスがスクリーンに触れたところであればどこでも、スクリーン上にストローク(stroke)を、あたかもインクがスタイラスによって付着させられたかのようにレンダリングする。ストロークは、コンピュータによって生成されたものであるため、編集すること、引き直す(re-flow)こと、および動画化することが可能である。また、ユーザは、コンピュータにジェスチャコマンドを発行し、同じスクリーン上に他のコンピュータアプリケーションを示すこともできる。

【 0 0 0 9 】

ただし、電子ホワイトボードは、現在、高価であり、サイズが小さい(電子ホワイトボードのサイズが、対角線長で6フィート(1メートル83センチメートル)を超えることはめったにない)ため、設置ベースがまだ限られている。さらに、ペン追跡デバイスを備え

10

20

30

40

50

たシステムは、次の欠点を有する。すなわち、１）システムがオンになっていない、またはユーザが、特殊ペンを使用せずに書き込んだ場合、内容をデバイスが回復することができない。２）多くの人々は、ホワイトボード上の小さい誤りを訂正するのに、特殊イレーザではなく、自分の指を使うことを好む。この、よくある行動によりキャプチャされた内容には余分なストロークが現れることになる。３）人々が、特殊なドライインクペンアダプタを使用しなければならず、これにより、一部のデバイスに関して、ペンがより太く、押さえづらくなる。４）ペン追跡の不正確さにより、ときとして、隣接したペンストロークの位置ずれ（mis-registration）が生じる。

【００１０】

ホワイトボードキャプチャ方法に関して前述した努力の他に、特に講演および会議に関して、マルチメディア体験のキャプチャ、統合、およびアクセスに関して大量の研究がなされている。手書きのノート、ホワイトボードの内容、スライド、または手書きの注釈を使用して、容易なアクセスのために記録された映像および音声にインデックスを付ける技術およびシステムが開発されている。

10

【００１１】

例えば、Classroom 2000 プロジェクトと呼ばれるプロジェクトで、Abowd 他が、電子ホワイトボードを使用してインクストロークにタイムスタンプを付けて、見る人（学生）が、そのインクストロークを記録されている映像および音声に対するインデックスとして使用することができるようにしている。キーフレーム（ページと呼ばれる）が、電子ホワイトボードによって提供された消去イベントに基づいて計算されている。ただし、Classroom 2000 プロジェクトは、電子ホワイトボードを必要とした。電子ホワイトボードを使用すると、エンドユーザの視点から多くの欠点が存在する。第１に、ほとんどのオフィスおよび会議室には、電子ホワイトボードが設置されていない。第２に、人々は、電子ホワイトボードより、通常のホワイトボードを使用する方がはるかに自然であると感じることが示されている。第３に、カメラを使用してキャプチャされたイメージの方が、誰が書込みを行い、どのトピックが議論されていたか（通常、手による指し示しにより）などの、はるかに多くのコンテキスト情報を提供する。以上の欠点に加え、電子ホワイトボードは、高価である可能性があり、したがって、容易に入手することができない。

20

【００１２】

【発明が解決しようとする課題】

本発明は、ホワイトボードの内容および関連する音声をキャプチャするためのシステムにおける前述した限界を克服するシステムおよびプロセスを対象とする。

30

【００１３】

【課題を解決するための手段】

本ホワイトボードキャプチャシステムは、キャプチャされたイメージに対して解析を行うことによってペンストロークのタイムスタンプ、およびキーフレームを計算することで、ホワイトボードのイメージを直接にキャプチャする前述したシステムとは区別される。キーフレームイメージは、ホワイトボード上の重要な内容のすべてを含み、記録の要約の役割をする。キーフレームイメージは、他のドキュメントにカットアンドペースト（cut and paste）すること、またはメモとして印刷することができる。タイムスタンプおよびキーフレームは、記録された音声に対する効果的なインデックスである。さらに、本ホワイトボードキャプチャシステムの発明は、一部の他のホワイトボードキャプチャシステムのように電子ホワイトボードではなく、通常のホワイトボードを使用する。したがって、このシステムは、変更することなしに任意の既存のホワイトボードで 사용할ことができる。

40

【００１４】

本ホワイトボードキャプチャシステムは、非電子式のホワイトボード上に書き込まれた内容のイメージのシーケンスをカメラを使用してキャプチャする。本ホワイトボードキャプチャシステムは、会議の音声信号を同時に記録する。記録が完了すると、イメージシーケン

50

スが解析されて、ホワイトボード上に書き込まれた内容の重要なポイントを要約するキーフレームが分離される。音声記録は、記録された音声とイメージシーケンスの両方に関連付けられたタイムスタンプでキーフレーム上でペンストロークに相関させられる。このタイムスタンプは、イメージ解析を介して計算される。

【0015】

キーフレームを獲得するための一般的な解析プロセスには、イメージのシーケンスの中のすべてのイメージの中でホワイトボードビューを修正することが関与する。また、ホワイトボード背景色も抽出され、イメージのシーケンスの各イメージが、セルに分割される。時間を経て同一であるセルイメージが、後により詳細に説明するとおり、一緒にクラスタ化される。次に、各セルイメージが、ストローク、前景オブジェクト、またはホワイトボードセルとして分類される。次に、この分類結果を使用してキーフレームイメージが抽出される。セルイメージは、キーフレーム抽出に先立って空間的および時間的にフィルタにかけて分類結果を精密にすることができる。さらに、キーフレームイメージは、抽出されると、カラーバランス処理 (color balance) を行ってイメージ品質を向上させることができる。

10

【0016】

より具体的には、ホワイトボードビューを修正することには、各イメージのホワイトボードでない領域をトリミングすることが関与する。次に、各イメージの中でホワイトボードの4つの隅が指定される。各イメージに関して、バイキュービック (bi-cubic) 補間を使用してバイリニアワープ (bi-linear warp) を行って、それぞれのキャプチャされたイメージの中でトリミングされ、修正されたホワイトボードイメージを獲得する。

20

【0017】

ホワイトボード背景色を抽出するために2つの方法を使用することができる。第1の方法は、最も明るい輝度および最も小さい変動を有するホワイトボードセルを判定することに関する。最も明るい輝度および最も小さい変動を有する色をホワイトボード背景色として指定する。このようにして、ホワイトボード背景色が決定されると、ホワイトボードの色におけるどんなホール (holes) も見出され、各ホールの周囲のホワイトボードセルを探索することによって埋められる。次に、各ホールの色は、ホールでない最も近接したセルの色に設定される。

【0018】

ホワイトボード背景色を抽出するための第2の方法は、ホワイトボードイメージ輝度のヒストグラムを作成して、ピークのホワイトボード輝度を判定することに関する。ピークの輝度に対応する色が、初期のホワイトボードの色として指定される。次に、あらゆるホワイトボードの色の外れ値 (誤ったデータ) が、最小2乗平均法を使用して判定される。この外れ値 (outliers) は、ホールとしてマークが付けられ、前述したホワイトボードの色を決定する第1の方法と同じやり方で埋められる。各ホールを埋めた後、ホワイトボードのカラーイメージをフィルタリングすることができる。

30

【0019】

入力シーケンスの中の各イメージをセルに分割するプロセスにより、解析処理速度が向上する。通常、各イメージは、セルサイズが、ボード上の単一の文字とほぼ同じサイズになるようにセルに分割される。これは、縦1.5インチ (3.8センチメートル)、横1.5インチ (3.8センチメートル)、または通常の会議サイズのホワイトボードの場合、縦25画素、横25画素に等しい。ただし、別法では、すべての解析を画素ごとに行うことができる。

40

【0020】

入力イメージのシーケンスが修正され、ホワイトボードの色が決定された後、セルイメージがクラスタ化される。時間を経て同一であると見なされるセルイメージが、グループと一緒にクラスタ化される。正規化された相互相関法およびマハラノビス (Mahalanobis) 距離試験を使用して、2つのセルが同一であるかどうかを判定する。

【0021】

50

セル分類プロセスにより、セルイメージが、ホワイトボードセルであるか、ストロークであるか、または前景オブジェクトであるかが判定される。セルイメージは、レッド、グリーン、ブルー（RGB）の値がほぼ同一である場合、ホワイトボードセルとして指定される。あるいは、セルイメージは、そのセルが大体において白またはグレーであり、1つまたは2つの原色が混じっている場合、ストロークセルとして指定される。そうでない場合、セルイメージは、前景セルとして指定される。セル分類プロセスにより、現行のセルイメージの色分布、および対応するホワイトボードセルの色分布が判定される。次に、セルは、現行のセルイメージおよび対応するホワイトボードセルの色分布が同一であるか、同一ではないが、強い類似性を有しているか、または全く異なっているかに基づいて分類される。

10

【0022】

前述した分類手続きは、単一のセル内の色情報を使用するだけである。セルグループ間の空間的關係および時間的關係を利用することによってより正確な結果を得ることができる。空間的にフィルタをかける際、すべてのホワイトボードイメージに対して2つの動作を行うことができる。第1に、孤立した前景セルを識別し、ストロークとして分類しなおす。第2に、いくつかの前景セルに直結しているストロークセルを前景セルとして分類しなおす。時間的にフィルタをかけることに関して、基本的な知見は、ストロークが消去された後、まったく同一の位置に同じストロークを書き込むのは、実質的に不可能であるということである。言い換えれば、任意の所与のセルに関して、2つの異なるフレームのセルイメージが、同一のストロークを有する場合、その2つのフレームの間のすべてのセルイメージは、そのセルを遮る前景オブジェクトが存在しない限り、同一のストロークを有するはずである。時間的にフィルタをかけるステップで、このセルは、前景オブジェクトがこのセルを遮った前後にカメラに晒されている限り、ストロークとして分類される。

20

【0023】

次に、キーフレームを抽出することができる。この目的で、イメージのシーケンスの中の各イメージまたは各フレームに関して、分類結果が使用され、ストロークセルが計数される。ストロークカウントのピーク（peak）とくぼみ（valley）が判定される。ストロークカウントのそれぞれ隣接するピークとくぼみの間の差が、規定されたしきい値を超えた場合、くぼみとくぼみの間のデータが、区分（chapter）（各区分は、異なるトピックを意味する）として指定され、また各区分の中のピークが、その区分を代表するキーフレームとして指定される。

30

【0024】

次に、キーフレームイメージが再構成される。これには、分類されたセルイメージ、およびセルイメージに分割されたキーフレームを入力することが関与する。キーフレームのセルイメージが、ホワイトボードイメージまたはストロークイメージとして分類される場合、キーフレームのイメージは、それぞれ、ホワイトボードイメージまたはストロークイメージとしてレンダリングされる。あるいは、キーフレームの前景セルイメージが、ストロークの範囲内にある場合、そのセルイメージは、シーケンスの中の隣接するイメージからのストロークのセルイメージとともにレンダリングされる。キーフレームのセルイメージは、ホワイトボードイメージ、ストロークイメージ、またはストロークの範囲内の前景セルとして分類されない場合、ホワイトボードイメージとしてレンダリングされる。

40

【0025】

次に、カラーバランス処理を使用して、平均のホワイトボードの色を使用してセル内の各画素の色を調整する（scale）ことによって背景を一様に白くし、ペンストロークの色の彩度を高めることによってキーフレームのイメージ品質を向上させる。また、イメージノイズも低減される。

【0026】

アナリシスサーバが、イメージシーケンスを処理し、インデックスおよびキーフレームイメージを生成した後、処理済みの記録のユニフォームリソースロケータ（Uniform Resource Locator）（URL）（Webサイトまたはその他のインターネットサービスの「アド

50

レス」またはロケーション)を伴う電子メールを登録されたセッション参加者に送信する。ユーザは、そのURLをクリックして、ブラウザソフトウェアを立ち上げることができる。ブラウザにより、ユーザは、キーフレームイメージを閲覧し、特定のトピックに関連する音声に迅速にアクセスすることが可能になる。

【0027】

ブラウザソフトウェアのユーザインターフェース(UI)は、様々な構成要素を有する。ブラウザUIの主要な要素には、キーフレームサムネールが表示されるキーフレームペイン(pane)、並びにカメラからの生のイメージと現行のキーフレームイメージの合成を示すブラウザのメイン表示ペインが含まれる。

【0028】

また、キーフレームペインには、ユーザが、メイン表示ペイン内に表示されたイメージを調整して生の入力イメージからキーフレームイメージにすることを可能にする背景透明スライダ(transparency slider)も組み込まれている。会議再生タイムラインにおいて既書き込まれているストロークである現行のペンストロークが、将来のストロークよりも濃く、より明確にレンダリングされる。参加者が将来に書き込むことになるペンストロークは、ゴースト(ghost)のようなスタイルで示される。この映像化の技法は、次のプロセスを使用して実現される。現行のホワイトボードの内容が、現行の区分のキーフレームイメージ、およびタイムスタンプ情報を使用してレンダリングされる。次に、将来のストロークが、ガウスフィルタを使用してグレースケールに変換され、ぼかされてレンダリングされる。次に、以上の2つのイメージが合わせられ、もたらされたイメージが、入力シーケンスからの修正されたイメージにアルファブレンド(alpha-blend)される。ユーザは、レンダリングされたキーフレームホワイトボードイメージだけを示す0から、元のイメージそのものを示す1まで、GUIスライダを使用してアルファ値を制御することができる。

【0029】

VCR-標準タイムラインコントロールが、メイン表示ペインの下でブラウザUIの左下隅に提供される。VCR-標準タイムラインコントロールにより、ユーザは、通常のビデオカセットレコーダ(VCR)上に見られるコントロールと同じように、イメージ/音声シーケンスをゆっくりと、または迅速に戻すこと、または先に進めること、または停止させることが可能になる。タイムラインバーが、音声/イメージシーケンスの長さをバーとしてグラフィックスで表示し、会議再生の開始時刻、終了時刻、および現在時刻の数値を提供する。このバーの上のポイントを選択し、タイムラインバーに沿って前後にドラッグして、イメージ/音声シーケンスを直線的に先に進めること、および戻すことができる。

【0030】

記録された音声に対する2つのレベルの非線形アクセスが、ビジュアル式インデックス付け(visual indexing)のコンテキストで提供される。第1のレベルの非線形アクセスは、キーフレームサムネールの使用を介する。ユーザは、キーフレームサムネールをクリックして、対応するキーフレームに関する音声の開始ポイント(例えば、区分の先頭)にジャンプすることができる。各キーフレームには、その特定のキーフレームに関連する時間範囲を判定する際にユーザを助ける時間範囲が関連付けられている。記録された音声に対する第2のレベルの非線形アクセスは、各キーフレーム内のペンストロークの使用を介する。カーソルが、メインウィンドウの中でペンストロークセル(現在のストロークセルまたは将来のストロークセル)の上を行き来(hover)しているとき、そのカーソルが、そのセルが選択可能(例えば、マウスで「クリック可能」)であることを示す「手」の記号に変化する。マウス、またはその他の入力デバイスでセルをダブルクリックすることにより、アプリケーションが、音声再生モードにされる。再生は、クリックされたストロークセルが書き込まれたセッションの時点から開始する。ユーザは、それでも、他のストロークセルをクリックしてセッションの他の部分にジャンプすることができる。VCR-標準タイムラインコントロール1514と併せて、以上の2つのレベルのビジュアル式インデックス付けにより、ユーザは、非常に効率的な仕方では会議をブラウズすることが可能にな

10

20

30

40

50

る。

【0031】

前述したとおり、キーフレームイメージのサムネールは、キーフレームペインの中にリストされる。サムネールの1つを選択することにより、対応するキーフレームイメージが、左側でメインウィンドウにもたらされ、アプリケーションが、イメージビューイングモードにされ、このモードで、ユーザが、ズーム制御ボタンを使用してズームインおよびズームアウトを行うこと、イメージの中のテキストおよび図を読むこと、またはイメージの一部分を他のドキュメントにカットアンドペーストすることができる。さらに、キーフレーム全体を他のドキュメントにカットアンドペーストすること、またはメモとして印刷することができる。

10

【0032】

本ホワイトボードキャプチャシステムでは、会議参加者が、会議記録の始めにキャプチャソフトウェアに登録するように求められる。すべての記録されたセッションが、Webサーバ上に常駐する。誰も登録しなかった場合、会議は、公衆アクセス可能なWebページに掲示される。少なくとも1名の参加者が登録した場合、会議の記録および解析の後、アクセストークンが生成される。トークンは、固有の会議識別子を含む長いランダムに生成されたストリングである。そのトークンを含むURLが、登録された参加者に電子メールで送信される。受け手は、そのURLに行ってWebブラウザソフトウェアを立ち上げて会議を綿密に見る(review)。受け手は、その会議に出席しなかった人々にそのURLを転送することもできる。

20

【0033】

前述した基本的なホワイトボードキャプチャシステムを多くの他の技術およびデバイスと組み合わせて代替の実施形態を提供することも可能である。1つのそのような実施形態では、キーフレームに対して従来の光学式文字認識(Optical Character Recognition)(OCR)が行われて、ドキュメントまたはプレゼンテーションビューグラフ(viewgraph)を作成するのに簡単に使用される編集可能なテキストが提供される。

【0034】

別の実施形態では、従来の音声認識ソフトウェアを使用してキャプチャされたデータの音声部分がテキストに変換される。これにより、会議の議事録およびその他のドキュメントを容易に作成することが可能になる。また、会議情報を聴覚障害者に提供する比較的安価な手立ても提供される。

30

【0035】

また、本ホワイトボードキャプチャシステムは、例えば、三脚の上に装着されたマイクロホンおよびカメラを備えたノートコンピュータを使用してポータブルにすることも可能である。この構成は、ホワイトボードに対するカメラの位置を決める追加の初期校正だけを必要とする。この校正は、イメージの中のパネルの4つの隅を手作業で決めることによって手作業で、またはエッジ検出などの方法を使用することによって自動的に行うことが可能である。

【0036】

また、ホワイトボードの内容を推測するのにペン追跡を使用するホワイトボードキャプチャシステムとともに本ホワイトボードキャプチャシステムの解析ソフトウェアを使用して、キーフレームを決定することも可能である。そのようなシステムとともに本ホワイトボードキャプチャシステムの解析ソフトウェアを使用することにより、解析プロセスが単純化される。ホワイトボード背景色の決定、またはホワイトボード領域修正の必要性はなく、空間的にフィルタをかけること、および時間的にフィルタをかけることも必要なく、またセルイメージがストロークまたはホワイトボードであるため、ホワイトボードセルの分類がより簡単である。というのは、ホワイトボード上に書き込まれたテキストを妨げる前景イメージが存在しないからである。

40

【0037】

さらに、より高いフレームレートを実現するため、スチールカメラの代わりに、HDTV

50

カメラなどの高解像度のビデオカメラを使用することができる。

【 0 0 3 8 】

さらに別の実施形態では、本ホワイトボードキャプチャシステムは、ジェスチャ認識を組み込んでジェスチャコマンドを使用する。例えば、ホワイトボード上のどこかにコマンドボックスが書き込まれていることが可能である。ユーザが動いたとき、またはボックスを指し示したとき、システムは、ジェスチャ認識を使用してそのジェスチャが行われた特定の時刻におけるイメージにタイムスタンプを付ける。

【 0 0 3 9 】

本ホワイトボードキャプチャシステムにより、会議参加者は、日常的なノートをとる作業を免れ、会議中に意見を出すこと、および意見を吸収することに集中できるようになる。ホワイトボードの内容を要約するキーフレームイメージ、および音声に対する構造化されたビジュアル式インデックス付けを提供することにより、システムは、参加者が、後の時点で会議を綿密に見ることを助ける。さらに、会議に出席しなかった人々が、しばしば、瞬時に会議の要点を理解することができる。

【 0 0 4 0 】

本発明の具体的な特徴、態様、および利点は、以下の説明、頭記の特許請求の範囲、および添付の図面に関連してよりよく理解される。

【 0 0 4 1 】

【発明の実施の形態】

本発明の好ましい実施形態の以下の説明では、説明の一部を成す添付の図面を参照する。図面では、例として、本発明を実施することができる特定の実施形態を示している。本発明の範囲を逸脱することなく、他の実施形態を利用することができ、また構造上の変更を加えることができることを理解されたい。

【 0 0 4 2 】

1 . 0 例示的な動作環境

図 1 は、本発明を実施することができる適切なコンピューティングシステム環境 1 0 0 の例を示している。コンピューティングシステム環境 1 0 0 は、適切なコンピューティング環境の一例に過ぎず、本発明の使用または機能の範囲に対する何らかの限定を示唆するものではない。また、コンピューティング環境 1 0 0 は、例示的な動作環境 1 0 0 に示した構成要素のいずれか 1 つ、またはいずれかの組合せに関連する依存性または要件を有するものと解釈してはならない。

【 0 0 4 3 】

本発明は、多数の他の汎用または特殊目的のコンピューティングシステム環境またはコンピューティングシステム構成で機能する。本発明で使用するのに適していることが可能な周知のコンピューティングシステム、コンピューティング環境、および/またはコンピューティング構成の例には、パーソナルコンピュータ、サーバコンピュータ、ハンドヘルドデバイスまたはラップトップデバイス、マルチプロセッサシステム、マイクロプロセッサベースのシステム、セットトップボックス、プログラマブル家庭用電化製品、ネットワーク PC、ミニコンピュータ、メインフレームコンピュータ、前述したシステムまたはデバイスの任意のものを含む分散コンピューティング環境等が含まれるが、以上には限定されない。

【 0 0 4 4 】

本発明は、コンピュータによって実行される、プログラムモジュールなどのコンピュータ実行可能命令の一般的なコンテキストで説明することができる。一般に、プログラムモジュールには、特定のタスクを行う、または特定の抽象データタイプを実装するルーチン、プログラム、オブジェクト、構成要素、データ構造等が含まれる。また、本発明は、通信網を介してリンクされた遠隔の処理デバイスによってタスクが行われる分散コンピューティング環境において実施することもできる。分散コンピューティング環境では、プログラムモジュールは、メモリ記憶デバイスを含むローカルのコンピュータ記憶媒体と遠隔のコンピュータ記憶媒体の両方の中に配置されることが可能である。

【 0 0 4 5 】

図 1 を参照すると、本発明を実施するための例示的なシステムが、コンピュータ 1 1 0 の形態で汎用コンピューティングデバイスを含んでいる。コンピュータ 1 1 0 の構成要素には、処理ユニット 1 2 0、システムメモリ 1 3 0、並びにシステムメモリから処理ユニット 1 2 0 までを含む様々なシステム構成要素を結合するシステムバス 1 2 1 が含まれることが可能であるが、以上には限定されない。システムバス 1 2 1 は、様々なバスアーキテクチャの任意のものを使用するメモリバスまたはメモリコントローラ、周辺バス、およびローカルバスを含むいくつかのタイプのバス構造の任意のものであることが可能である。例として、限定としてではなく、そのようなアーキテクチャには、インダストリスタンドアークテクチャ (Industry Standard Architecture) (I S A) バス、マイクロチャネルアークテクチャ (Micro Channel Architecture) (M C A) バス、エンハンスド I S A (Enhanced ISA) (E I S A) バス、ビデオエレクトロニクススタンダーズアソシエーション (Video Electronics Standards Association) (V E S A) ローカルバス、およびメザニン (Mezzanine) バスとしても知られるペリフェラルコンポーネントインターコネクト (Peripheral Component Interconnect) (P C I) バスが含まれることが可能である。

10

【 0 0 4 6 】

コンピュータ 1 1 0 は、通常、様々なコンピュータ可読媒体を含む。コンピュータ可読媒体は、コンピュータ 1 1 0 によってアクセス可能な任意の利用可能な媒体であることが可能であり、揮発性の媒体と不揮発性の媒体、取外し可能な媒体と取外し不可能な媒体がともに含まれる。例として、限定としてではなく、コンピュータ可読媒体は、コンピュータ記憶媒体および通信媒体を含むことが可能である。コンピュータ記憶媒体には、コンピュータ可読命令、データ構造、プログラムモジュール、またはその他のデータなどの情報を記憶するための任意の方法または任意の技術で実装される揮発性の媒体および不揮発性の媒体、取外し可能な媒体および取外し不可能な媒体がともに含まれる。コンピュータ記憶媒体には、R A M、R O M、E E P R O M、フラッシュメモリまたは他のメモリ技術、C D - R O M、デジタルバーサタイルディスク (D V D) または他の光ディスクストレージ、磁気カセット、磁気テープ、磁気ディスクストレージまたは他の磁気記憶デバイス、あるいは所望の情報を記憶するのに使用することができ、コンピュータ 1 1 0 がアクセスすることができる任意の他の媒体が含まれるが、以上には限定されない。通信媒体は、通常、搬送波などの変調されたデータ信号、または他のトランスポート機構でコンピュータ可読命令、データ構造、プログラムモジュール、またはその他のデータを実体化し、あらゆる情報配信媒体が含まれる。「変調されたデータ信号」という用語は、信号に情報を符号化するような仕方で特徴の 1 つまたは複数設定または変更されている信号を意味する。例として、限定としてではなく、通信媒体には、有線ネットワークまたは直接配線接続などの有線媒体、並びに音響媒体、R F 媒体、赤外線媒体、およびその他の無線媒体などの無線媒体が含まれる。また、前述したもののいずれかの組合せも、コンピュータ可読媒体の範囲に含まれるべきものである。

20

30

【 0 0 4 7 】

システムメモリ 1 3 0 は、読取り専用メモリ (R O M) 1 3 1 およびランダムアクセスメモリ (R A M) 1 3 2 などの揮発性メモリおよび / または不揮発性メモリの形態でコンピュータ記憶媒体を含む。始動中など、コンピュータ 1 1 0 内部の要素間で情報を転送するのを助ける基本ルーチンを含む基本入力 / 出力システム 1 3 3 (B I O S) が、通常、R O M 1 3 1 の中に記憶されている。R A M 1 3 2 は、通常、処理ユニット 1 2 0 によって即時にアクセス可能であり、かつ / または現在、処理されているデータおよび / またはプログラムモジュールを含む。例として、限定としてではなく、図 1 は、オペレーティングシステム 1 3 4、アプリケーションプログラム 1 3 5、その他のプログラムモジュール 1 3 6、およびプログラムデータ 1 3 7 を示している。

40

【 0 0 4 8 】

また、コンピュータ 1 1 0 は、他の取外し可能な / 取外し不可能な、揮発性 / 不揮発性の

50

コンピュータ記憶媒体を含むことも可能である。単に例として、図 1 は、取外し不可能な不揮発性の磁気媒体に対して読取りおよび書込みを行うハードディスクドライブ 141、取外し可能な不揮発性の磁気ディスク 152 に対して読取りおよび書込みを行う磁気ディスクドライブ 151、および CD-ROM または他の光媒体などの取外し可能な不揮発性の光ディスク 156 に対して読取りおよび書込みを行う光ディスクドライブ 155 を示している。例示的な動作環境において使用することができる他の取外し可能な / 取外し不可能な揮発性 / 不揮発性のコンピュータ記憶媒体には、磁気テープカセット、フラッシュメモリカード、デジタルバーサタイルディスク、デジタルビデオテープ、ソリッドステート RAM、ソリッドステート ROM 等が含まれるが、以上には限定されない。ハードディスクドライブ 141 は、通常、インターフェース 140 のような取外し不可能なメモリアンターフェースを介してシステムバス 121 に接続され、また磁気ディスクドライブ 151 および光ディスクドライブ 155 は、通常、インターフェース 150 のような取外し可能なメモリアンターフェースでシステムバス 121 に接続される。

【0049】

前述し、図 1 に示すドライバおよび関連するコンピュータ記憶媒体により、コンピュータ可読命令、データ構造、プログラムモジュール、およびその他のデータのストレージがコンピュータ 110 に提供される。図 1 では、例えば、ハードディスクドライブ 141 が、オペレーティングシステム 144、アプリケーションプログラム 145、他のプログラムモジュール 146、およびプログラムデータ 147 を記憶しているものとして示されている。以上の構成要素は、オペレーティングシステム 134、アプリケーションプログラム 135、他のプログラムモジュール 136、およびプログラムデータ 137 と同じであることも、異なることも可能であることに留意されたい。オペレーティングシステム 144、アプリケーションプログラム 145、他のプログラムモジュール 146、およびプログラムデータ 147 に、ここでは、少なくともそれらが異なるコピーであることを示すために異なる符号を付けている。ユーザは、キーボード 162、および一般にマウス、トラックボール、またはタッチパッドと呼ばれるポインティングデバイス 161 などの入力デバイスを介してコンピュータ 110 にコマンドおよび情報を入力することができる。他の入力デバイス（図示せず）には、マイクロホン、ジョイスティック、ゲームパッド、サテライトディッシュ、スキャナ等が含まれる可能性がある。以上の入力デバイスおよび他の入力デバイスは、しばしば、システムバス 121 に結合されたユーザ入力インターフェース 160 を介して処理ユニット 120 に接続されるが、パラレルポート、ゲームポート、またはユニバーサルシリアルバス（Universal Serial Bus）（USB）などの他のインターフェースおよびバス構造で接続してもよい。また、モニタ 191 または他のタイプの表示デバイスも、ビデオインターフェース 190 などのインターフェースを介してシステムバス 121 に接続される。モニタに加えて、コンピュータは、出力周辺インターフェース 195 を介して接続することができるスピーカ 197 やプリンタ 196 などの他の周辺出力デバイスも含むことが可能である。本発明に特に重要なこととして、一続きのイメージ 164 をキャプチャすることができるカメラ 163（デジタル / 電子式のスチールカメラまたはビデオカメラ、またはフィルム / 写真スキャナ）も、パーソナルコンピュータ 110 に対する入力デバイスとして含まれることが可能である。さらに、1つのカメラだけを描いているが、複数のカメラが、パーソナルコンピュータ 110 に対する入力デバイスとして含まれることも可能である。1つまたは複数のカメラからのイメージ 164 が、適切なカメラインターフェース 165 を介してコンピュータ 110 に入力される。このインターフェース 165 は、システムバス 121 に接続され、これにより、イメージを RAM 132、またはコンピュータ 110 に関連する他のデータ記憶デバイスの 1 つにルーチングして、その中に記憶することができるようになる。ただし、イメージデータは、カメラ 163 の使用を必要とすることなしに、前述したコンピュータ可読媒体の任意のものからもコンピュータ 110 に入力することができることに留意されたい。

【0050】

コンピュータ 110 は、リモートコンピュータ 180 のような 1 つまたは複数のリモート

10

20

30

40

50

コンピュータに対する論理接続を使用するネットワーク化された環境で動作することも可能である。リモートコンピュータ 180 は、パーソナルコンピュータ、サーバ、ルータ、ネットワーク PC、ピアデバイス、または他の一般的なネットワークノードであることが可能であり、通常、コンピュータ 110 に関連して前述した要素の多く、またはすべてを含むが、メモリ記憶デバイス 181 だけを図 1 に示している。図 1 に描いた論理接続は、ローカルエリアネットワーク (LAN) 171 およびワイドエリアネットワーク (WAN) 173 を含むが、他のネットワークを含むことも可能である。そのようなネットワーキング環境は、オフィス、企業全体のコンピュータネットワーク、イントラネット、およびインターネットで一般的である。

【0051】

LAN ネットワーキング環境で使用されるとき、コンピュータ 110 は、ネットワークインターフェースまたはネットワークアダプタ 170 を介して LAN 171 に接続される。WAN ネットワーキング環境で使用されるとき、コンピュータ 110 は、通常、インターネットなどの WAN 173 を介して通信を確立するためのモデム 172、または他の手段を含む。内部にあることも、外部にあることも可能なモデム 172 は、ユーザ入力インターフェース 160 または他の適切な機構を介してシステムバス 121 に接続することが可能である。ネットワーク化された環境では、コンピュータ 110 に関連して描いたプログラムモジュール、またはプログラムモジュールの部分を遠隔のメモリ記憶デバイスの中に記憶することができる。例として、限定としてではなく、図 1 は、リモートアプリケーションプログラム 185 をメモリデバイス 181 上に常駐するものとして示している。図示したネットワーク接続は、例示的なものであり、コンピュータ間で通信リンクを確立する他の手段を使用してもよいことが認められよう。

【0052】

以上、例示的な動作環境について述べたので、この説明のセクションの以下の部分は、本発明を実現するプログラムモジュールの説明に向けられる。

【0053】

2.0 ホワイトボードキャプチャのシステムおよび方法

2.1 システムアーキテクチャ

概念上、本ホワイトボードキャプチャシステムは、図 2 に示すとおり、3 つの主要な構成要素、すなわち、キャプチャユニット 202、解析 / 処理サーバ 204、およびブラウザソフトウェア 206 から成る。

【0054】

1. キャプチャユニット：キャプチャユニットは、ホワイトボードの内容のイメージをキャプチャし、ホワイトボードの内容の生成に関連する音声を記録するのに使用される。キャプチャユニットは、会議が行われる部屋の中に設置される。図 3 に示すとおり、キャプチャユニットは、デジタルカメラ 302、マイクロホン 304、およびパーソナルコンピュータ (PC) 306 を含む。キャプチャユニットは、ホワイトボード 308 のイメージを取り込み、PC 306 に記憶される音声をマイクロホン 304 を介して記録する。取り込まれたイメージと音声の両方にタイムスタンプが付けられる。イメージおよび音声サンプルは、通常、システムクロックである共通クロックで獲得される。その共通クロックのタイミングが、イメージおよび音声サンプルに関連付けられ、イメージおよび音声サンプルのタイムスタンプとして記憶される。

【0055】

2. アナリシスサーバ：アナリシスサーバ 204 は、中心的な場所に配置され、記録されたイメージデータを解析して記憶する。一実施形態では、ユーザが、キャプチャユニットにおける記録を停止した後、解析プログラムが自動的に立ち上げられる。記録されたデータを処理した後、会議記録が用意されている URL を含む電子メールが、登録済みの参加者に送信される。登録済みのユーザが存在しない場合、会議記録は、公衆アクセス可能な Web サイトに掲示されることが可能である。

【0056】

3. ブラウズソフトウェア：ブラウズソフトウェア 2.0.6 により、ユーザは、記録され、解析された会議データを閲覧すること、および再生することができるようになる。ブラウズソフトウェア 2.0.6 は、好ましくは、会議記録を閲覧することを望むユーザによってインストールされる Web プラグインとして提供される。インストールされると、ユーザは、前述した URL をクリックしてそのソフトウェアを立ち上げてアナリシスサーバ上のデータにアクセスすることができる。

【0057】

2.2 イメージ獲得

本ホワイトボードキャプチャシステムへの入力は、1組の静止デジタルイメージである。図4は、例示的な1組のそのようなイメージを示している。このイメージシーケンスを解析して、ユーザがボード上でいつどこに書込みを行ったかを判別し、セッション全体のホワイトボードの内容を要約する1組のキーフレームイメージを取り出す。

【0058】

コンピュータによるカメラ制御を可能にする任意の比較的高解像度のカメラをイメージ獲得のために使用することができる。カメラは、好ましくは、会議室の側方または後方に設置される。カメラは、可能な限りホワイトボードに近接してズームインして有効解像度を最大化する。カメラは、設置後、固定であり、ホワイトボードは動かないことが想定され、したがって、ホワイトボードイメージは、キャプチャされたシーケンス全体で静止したものである。

【0059】

使用されるカメラが自動焦点モードだけしか有さない場合、ホワイトボードの手前の被写体が、カメラの自動焦点機構の注意を引き起こしたとき、ホワイトボードに焦点が合わなくなる可能性がある。この問題は、カメラのイメージ平面を可能な限りホワイトボードに平行に揃えてシーン深度 (scene depth) を最小化し、かつ/または絞りを最小化して被写界深度を大きくすることによって緩和することができる。実際には、本ホワイトボードキャプチャシステムの実用の実施形態では、1~2%のフレームだけで、焦点が合っていないことが見られた。

【0060】

カメラは、できる限り迅速に写真を撮影し、そのイメージを、好ましくはUSB接続を介して、PCに転送する。本ホワイトボードキャプチャシステムの実用の実施形態では、およそ5秒毎に1つのJPEGイメージが獲得された。露出およびホワイトバランス (white-balance) パラメータは、通常、一定に保たれる。照明セッティングが1つのセッション中に変化しないと想定すると、ホワイトボード背景色は、シーケンスの中で一定に留まるはずである。

【0061】

わずかに露出不足のイメージの方が、より良好な色の彩度をもたらす、これにより、以下に述べるストローク抽出プロセスがより正確になる。カラーバランス処理のステップを記録の後に行って、グレーがかったホワイトボードイメージをより見栄えのよいものにすることができる。

【0062】

2.3 イメージシーケンス解析

ボードに書込みをしている人物は、デジタルカメラとホワイトボードの間の照準線上にいますので、しばしば、ホワイトボードのある部分を覆い隠し、別の部分に影を落とす。ストロークから、前景オブジェクト (例えば、ボード上に書込みをしている人物) およびホワイトボードを区別することが必要である。分類結果が分かると、キーフレームイメージおよびインデックスをブラウズソフトウェアが使用することができる。

【0063】

画素ごとのレベルでイメージを解析するのではなく (それを行うことは可能であるが)、ホワイトボード領域を長方形のセルに分割して計算コストを下げる。セルサイズは、ボード上の単一の文字の予期されるサイズ (実用の実施形態では、縦 1.5 インチ (3.8 セ

10

20

30

40

50

ンチメートル)、横 1.5 インチ (3.8 センチメートル)、または縦 2.5 画素、横 2.5 画素) とほぼ同じになるように選択する。セル格子により、入力シーケンスの中の各フレームがセルイメージに分割されるので、入力は、セルイメージの 3 次元マトリックス (例えば、x、y、時間) として考えることができる。各イメージをセルに分割することは、通常、入力イメージが修正された後に行われる。

【0064】

図 5 および以下に示すとおり、入力イメージシーケンスを解析する手続きの一般的なプロセスステップは、以下のとおりである。

1. シーケンスの中のすべてのイメージのホワイトボード領域を修正する (プロセスステップ 502)
2. ホワイトボード背景色を抽出する (プロセスステップ 504)
3. シーケンスの中のすべてのイメージを対応するセルイメージ (プロセスステップ 506) に分割した後、同じセルに関してシーケンス全体のセルイメージをクラスタ化する。2 つのセルイメージが時間を経て同じであると考えられる場合は、2 つのセルイメージは同じグループにクラスタ化される
4. 各セルイメージをストローク、前景オブジェクト、またはホワイトボードとして分類する (プロセスステップ 508)
5. セルイメージを空間的にも、時間的にもフィルタにかけて分類結果を精密にする (プロセスステップ 510)
6. 分類結果を使用してキーフレームイメージを抽出する (プロセスステップ 512)
7. キーフレームイメージのカラーバランス処理を行う (プロセスステップ 514)

【0065】

以下の段落では、図 4 に示した実用例を使用して入力イメージシーケンス解析手続きを例示する。

【0066】

2.3.1 ホワイトボードイメージを修正すること

イメージシーケンスをストローク抽出プロセスに送り込む前に、ホワイトボードでない領域をトリミングして、イメージを修正する。実用の実施形態で使用されるカメラのレンズの径方向の歪みは相当に小さいので、ホワイトボードの 4 つの隅を特定するだけでよい (歪みが小さくない場合、イメージを修正することに先立って、従来の方法で径方向の歪みを補正する必要がある可能性がある)。これは、1 回限りの較正ステップ中にキャプチャされたイメージの中でホワイトボードの 4 つの隅の位置をクリックすることにより、手作業で行われるが、自動的に (例えば、エッジ検出を使用することにより) 行うことも可能である。その 4 つの隅で、バイキュービック補間を使用して各イメージに関して簡単な従来のバイリニアワープを行い、それぞれのキャプチャされたイメージの中でトリミングされ、修正されたビューを獲得する。

【0067】

2.3.2 ホワイトボードの色を計算すること

セルの分類のため、各セルに関してホワイトボードの色がどのようなものであるか (つまり、何も書き込まれていないホワイトボード自体の色) を知る必要がある。また、ホワイトボードの色は、キーフレームを生成する際のホワイトバランス処理のためにも使用され、したがって、正確に推定してキーフレームイメージの高い品質を確実にしなければならない。

【0068】

ホワイトボードの色を計算するために 2 つの戦略が使用されてきた。図 6 A に概略を述べた第 1 の戦略は、ホワイトボードセルが、時間の経過で見ると最も明るい輝度を有し、小さい変動 (すなわち、各セル内でほとんど一様である) という前提に基づいている。この前提は、ストロークの色 (レッド、グリーン、ブルー、または黒) が輝度を低下させるので、妥当である。プロセスステップ 602 に示したとおり、最も明るい輝度および最も小さい変動 (variance) を有するホワイトボードセルが計算される。ただし、これにより、

最終のホワイトボードのカラーイメージにホールを生じさせる可能性がある。例えば、セルが、ストロークを含むか、またはシーケンス全体で前景オブジェクトによって遮られている場合、このセルに関して計算されたホワイトボードの色は正しくなくなる（このセルが、ホワイトボードの残りの部分とは異なって見え、したがって、ホールのように見える）。このため、プロセスステップ604に示すとおり、ホワイトボードのカラーイメージにおけるホールが存在すれば、最小2乗平均法と呼ばれる技法（次の段落で説明する外れ値検出方法と同様の）を使用してそのホールが検出される。次に、そのホールが埋められる（プロセスステップ606）。ホールを埋めるため、ホールの近辺が探索され、ホワイトボードの色は、ホールでない最も近接したセルの色に設定される。この戦略は、通常、極めてうまく行くが、人物が白いTシャツを着ており、かつ／または白い紙を手を持っている場合、失敗する。図7の左側のイメージは、フレームのいくつかの中で人物が白い紙を手を持っている場合の、図4の入力シーケンスから計算されたホワイトボードのカラーイメージの結果を示している。計算されたホワイトボードの色が、白い紙によって損なわれているのを見て取ることができる。

10

【0069】

ホワイトボードの色を判定する第2の戦略を図6Bに示しており、この戦略は、より高度である。前提は、時間の経過で見ると各セル内の画素の相当な部分が、ホワイトボードに属するということである。各セルに関する輝度のヒストグラムを作成することにより、高い輝度値を有するピークに対応する色が、そのセルに関するホワイトボードの色である可能性が非常に高い。したがって、第1のステップは、各セルに関してヒストグラムを作成し、ピークの輝度を計算して（プロセスステップ610ないし614）、このやり方で初期のホワイトボードの色を計算することである。この技法は、セルが、シーケンス全体でストロークを含む場合でもうまく行くが、人物が、白いTシャツを着ており、かつ／または白い紙を手を持っている場合、またはセルが、人々またはその他の物体で常に隠れている場合、失敗する。そのような場合、計算されたホワイトボードのカラーイメージは、外れ値を含む。次の動作は、外れ値があれば、それを検出することである（プロセスステップ616）。外れ値検出は、最小2乗平均法と呼ばれる堅牢な技法に基づく。色が、ホワイトボード全体で平滑に変化するものと想定すると、2乗誤差の平均を最小化することによって輝度YまたはRGBの空間に平面が合わせられる。色がこのモデルに従わないセルは、外れ値と見なされ、したがって、拒否される。すなわち、ホールとしてマークが付けられる（プロセスステップ618）。関心のある読者は、この技法の詳細については、付録を参照されたい。次に、プロセスステップ620に示すとおり、第1のホワイトボードの色を計算する戦略と同じ手続きを使用することによってホールが埋められる（プロセスステップ620）。最後に、結果をさらに向上させるため、平面をRGB空間に局所的に合わせることによってホワイトボードのカラーイメージをフィルタリングすることができる（プロセスステップ622）。関心のある読者は、やはり詳細については、付録を参照されたい。同一の例に対するこの新しい技法を使用して得られた結果を図7の中央のイメージに示している。左側に示した第1の戦略を使用して得られた結果に優る明らかな向上が見られる。実際の空白のホワイトボードも、比較のために右側のイメージで示している。

20

30

40

【0070】

2.3.3 時間にわたってセルイメージをクラスタ化すること
会議中、各セルの内容は、通常、時間とともに変化する。各セルに関して、時間シーケンスの中のすべてのセルイメージを、時間を経て同じであると見なされるセルイメージをそれぞれが含むグループにクラスタ化することが望まれる。変更された正規化された相互相関アルゴリズムを使用して、2つのセルイメージが同じであるか否かを判定する。以下に、正規化された相互相関法をイメージの1つの色成分を使用して説明するが、この技法は、RGB成分のすべてに適用される。

【0071】

2つのセルイメージIおよびI'を考慮されたい。

50

【 0 0 7 2 】

【数 1 1】

 \bar{I}

【 0 0 7 3 】

および

【 0 0 7 4 】

【数 1 2】

 \bar{I}'

10

【 0 0 7 5 】

を平均の色とし、 および ' を標準偏差とする。正規化された相互相関スコアは、

【 0 0 7 6 】

【数 1 3】

$$c = \frac{1}{N\sigma\sigma'} \sum_i (I_i - \bar{I})(I'_i - \bar{I}')$$

【 0 0 7 7 】

によって与えられ、ただし、総和は、すべての画素 i を対象とし、 N は、画素の総数である。スコアは、2つのイメージが全く類似していない場合の -1 から、2つのイメージがまったく同一である場合の 1 までの範囲にある。このスコアは、平均の色を引いた後に計算されるので、2のイメージが非常に異なる平均の色を有している場合でも、高い値を与える可能性がある。したがって、マハラノビス距離に基づいて平均の色の差に対してさらなる試験を使用し、この試験は、

20

【 0 0 7 8 】

【数 1 4】

$$d = |\bar{I} - \bar{I}'| / (\sigma + \sigma')$$

【 0 0 7 9 】

によって与えられる。要約すると、2つのセルイメージ I および I' は、同一であると見なされ、したがって、 $d < T_d$ かつ $c > T_c$ である場合にだけ、同じグループに入れられるべきである。本ホワイトボードキャプチャシステムの実用の実施形態では、 $T_d = 2$ および $T_c = 0.707$ が成功裏に使用された。

30

【 0 0 8 0 】

2.3.4 セルを分類すること

セルを分類するプロセスステップにより、セルイメージがホワイトボードであるか、ストロークであるか、または前景オブジェクトであるかが判定される。次のヒューリスティックが使用される。すなわち、1) ホワイトボードセルの色は、一様であり、グレーまたは白である(すなわち、RGB値は、ほぼ同じである)、2) ストロークセルは、大体において白またはグレーであり、1つまたは2つの主要な色が混じっている、3) 前景オブジェクトは、以上の特性を有さない。したがって、分類により、現行のセルイメージの色分布とホワイトボードの色分布が同じであるか、または同じではないが、強い重なり合いを有しているか、または全く異なるかが判別される。この場合も、以下に説明するとおり、マハラノビス距離が使用される。

40

【 0 0 8 1 】

ホワイトボードの色は、前述したとおり、既に計算済みであることに留意されたい。この場合も、RGBのどれかの色成分を例として使用する。

【 0 0 8 2 】

【数 1 5】

50

\bar{I}_w

【 0 0 8 3 】

をホワイトボードの色とし、また σ_w を標準偏差とする（これは、ホワイトボードセルがほぼ一様であるため、小さい値である）。

【 0 0 8 4 】

【 数 1 6 】

 \bar{I}

【 0 0 8 5 】

および σ を現行のセルイメージの平均および標準偏差とする。セルイメージは、

【 0 0 8 6 】

【 数 1 7 】

$$|\bar{I} - \bar{I}_w| / (\sigma + \sigma_w) < T_w$$

【 0 0 8 7 】

かつ $\sigma / \sigma_w < T$ である場合にだけ、ホワイトボードセルとして分類され、

【 0 0 8 8 】

【 数 1 8 】

$$|\bar{I} - \bar{I}_w| / (\sigma + \sigma_w) < T_w$$

【 0 0 8 9 】

かつ

【 0 0 9 0 】

【 数 1 9 】

$$\sigma / \sigma_w \geq T_\sigma$$

【 0 0 9 1 】

である場合にだけ、ストロークセルとして分類され、それ以外の場合、前景オブジェクトセルとして分類される。本ホワイトボードキャプチャシステムの実用の実施形態では、 $T_w = 2$ および $T = 2$ が成功裏に使用された。

【 0 0 9 2 】

2 . 3 . 5 セル分類をフィルタリングすること

前述した分類手続きは、単一のセル内の色情報を使用するだけである。セルグループ間の空間的關係および時間的關係を利用することによってより正確な結果を得ることができる。

【 0 0 9 3 】

2 . 3 . 5 . 1 空間的にフィルタにかけること

空間的にフィルタにかけることに関して、基本的な知見は、人物が、通常、ホワイトボードの連続した領域を遮るので、前景セルは、空間的に孤立して現れないはずであるということである。空間的にフィルタにかける際、図 8 に示すとおり、すべてのホワイトボードイメージに対して 2 つの動作が行われる。第 1 に、孤立した前景セルが識別され、ストロークとして分類しなおされる（プロセスステップ 8 0 2）。第 2 に、いくつかの前景セルに直結したストロークセルが、前景セルとして分類しなおされる（プロセスステップ 8 0 4）。第 2 の動作の 1 つの主な目的は、前景オブジェクトの境界にあるセルに対処することである。そのようなセルがストロークを含む場合、第 2 の動作は、そのセルを前景オブジェクトとして誤って分類する。ただし、幸い、以下の時間的なフィルタリングにより、そのような誤りが生じた場合、それが訂正される。

【 0 0 9 4 】

2 . 3 . 5 . 2 時間的にフィルタにかけること

10

20

30

40

50

時間的にフィルタにかけることに関して、基本的な知見は、ストロークを消去した後、同じストロークをまったく同じ位置に書き込むことが実質的に不可能であるということである。言い換えれば、任意の所与のセルに関して、2つの異なるフレームのセルイメージが同じストロークを含む場合、その2つのフレームの間のすべてのセルイメージは、そのセルを遮る前景オブジェクトが存在しない限り、同一のストロークを有するはずである。この知見は、前景オブジェクトを分けて除くのに非常に役立つ。前景オブジェクトの境界にあるストロークセルが、前景セルとして誤って分類される前のセクションの例を考慮されたい。時間的にフィルタにかけるステップで、そのセルは、前景オブジェクトがそのセルを遮った前後にカメラに晒されている限り、ストロークとして分類される。

【0095】

図9は、ストロークがグリーンであり、前景が黒であり、かつホワイトボードが白である場合の図4のサンプルイメージに関する分類結果を示している。

【0096】

2.3.6 キーフレームイメージ抽出

キーフレームイメージは、ホワイトボード上のすべての重要な内容を含み、記録の要約の役割をする。ユーザは、キーフレームイメージが、次のプロパティを有することを予期しているはずである。すなわち、2) キーフレームイメージは、ボード上のすべての重要な内容をキャプチャしていなければならない、2) キーフレームの数は、最小限に抑えられなければならない、3) キーフレームイメージは、ペンストロークおよびホワイトボードだけを含み、手前に人物を含んでいてはならない、4) キーフレームイメージは、容易なカットアンドペーストおよび印刷のため、一様の白い背景、および飽和度に達したペンの色を有していなければならない。

【0097】

キーフレーム抽出手続きは、前述したプロセスステップからのセルイメージ分類結果を使用する。この手続きは、まず、シーケンスの中のどのフレームがキーフレームとして選択されるべきかを決め、次に、キーフレームイメージを再構成する。これを以下により詳細に説明する。

【0098】

2.3.6.1 キーフレーム選択

キーフレームを選択する際、会議を要約するのに唯一のやり方が存在しないのと同様に、唯一の解決策は存在しない。最も一般的な意味で、図11を参照すると、ストローク、前景、またはホワイトボードとして分類されている入力イメージセルが使用されている(プロセスステップ1102)。会議は、まず、いくつかの区分(トピック)に分割される(プロセスステップ1104)。ボードの内容の大部分の消去により、普通、トピックの変化が示され、したがって、この消去が、区分の仕切りとして使用される。次に、ホワイトボードの内容を代表するキーフレームイメージが、その区分に関して生成される(プロセスステップ1106)。大部分の消去が始まる直前のフレームが、キーフレームとして選択され、これにより、そのフレーム内のコンテンツが保存されることが確実になる。図12に示した詳細な手続きは、以下のとおり行われる。

【0099】

1. シーケンスの中の各フレームに関するストロークセルの数が計数される(プロセスステップ1202)。1つのストロークセルイメージが、複数のフレームにまたがる、すなわち、そのフレームのそれぞれに関してカウントに含まれることが可能である。図10は、例のセッション(図4)におけるフレーム数に対してプロットされたストロークセルカウントを示している。プロットの上昇により、より多くのストロークがボード上に書き込まれたことが示され、プロットの下降(dip)により、いくつかのストロークが消去されたことが示される。このグラフには、極めてノイズが多い。次の2つの理由が存在する。すなわち、1) ユーザは、ボード上で常に小さい調整を行っている、2) 分類結果は、小さい誤りを含む。

【0100】

10

20

30

40

50

2. 様々なフレームに関するストロークカウントを使用して、ピークおよびくぼみが判別される（プロセスステップ1204）。各下降でキーフレームが生成された場合、数十のキーフレームがもたらされる。キーフレームの数を最小限に抑えるため、データをフィルタにかけて有意な消去イベントだけを留めておく。この手続きは、隣接するピークとくぼみの間の差が、あるしきい値を超えない限り、データの変動を無視する（プロセスステップ1206）。本システムの実用の実施形態では、最大ストロークカウントの20パーセントが成功裏に使用された。

【0101】

3. 次に、データにおけるくぼみを使用して、セッションが区分に分割される（プロセスステップ1208）。区分の中のピークを含むフレームが、その区分を代表するキーフレームとなるように選択される。

10

【0102】

2.3.6.2 イメージ再構成

フレームが選択されると、その時点でホワイトボードが呈していた様相に対応するイメージを再構成する必要がある。ただし、入力シーケンスからの生のイメージを単に使用することはできない。というのは、生のイメージは、前景オブジェクトを含む可能性があるためである。イメージは、フレーム内のセルイメージを集めることによって再構成される。図13を参照すると、セルイメージに分割されたフレーム、およびセルイメージに分割されたキーフレームが入力されている（プロセスステップ1302）。セル分類に応じて3つのケースが存在する。

20

【0103】

1. キーフレームセルイメージがホワイトボードまたはストロークである場合、そのホワイトボードまたはストロークのイメージ自体が使用される（プロセスステップ1304、1306）。

【0104】

2. キーフレーム前景セルイメージが、ストロークの範囲内にある場合（すなわち、人物が、ボード上のストロークを覆い隠している。これは、解析段階中に時間的にフィルタにかけることで判定される）、そのセルイメージは、隣接するフレームからのストロークセルイメージで置き換えられる（プロセス1308、1310）。

【0105】

30

3. それ以外の場合、プロセスステップ1312および1314に示すとおり、前景オブジェクトが、そのセル内のホワイトボード背景を覆っているはずであり、前述したとおり計算されたホワイトボードの色で埋められる。

【0106】

2.3.7 キーフレームカラーバランス

再構成プロセスにより、ホワイトボードイメージから人物が除去されるが、それでもイメージは、入力シーケンスからの生のイメージのように見える。すなわち、グレーがかっていて、ノイズが多い。イメージにカラーバランス処理を行ってより良好なイメージを生成することができる。このプロセスは、図14のとおり以下の2つのステップから成る。

【0107】

40

1. 背景を一樣に白くして、ペンストロークの色の彩度を高める。各セルに関して、前述したとおり計算されたホワイトボードの色、

【0108】

【数20】

\bar{I}_w

【0109】

を使用してセル内の各画素の色を調整する。

【0110】

【数21】

50

$$I_{out} = \min\left(255, \frac{I_{in}}{I_w} \cdot 255\right)$$

【 0 1 1 1 】

である（プロセスステップ 1 4 0 2）。

【 0 1 1 2 】

2．イメージノイズを低減する。キーフレームの中の各画素の各色チャネルの値を S 字型の曲線に従ってマップしなおす（プロセスステップ 1 4 0 4）。2 5 5 / 2 より低い光度を 0 の方にスケールダウン（scale down）し、2 5 5 / 2 より高い光度を 2 5 5 の方にスケールアップ（scale up）する。

10

【 0 1 1 3 】

区分の開始時刻および終了時刻、および区分のキーフレームイメージのファイル名が、ストロークのタイムスタンプとともにインデックスに保存される。ストロークのタイムスタンプは、そのストロークが出現する最初のフレームである。この情報は、セクション 2．3．3 で計算済みである。

【 0 1 1 4 】

2．4 ブラウザの動作およびユーザインターフェース

2．4．1 概要

アナリシスサーバは、イメージシーケンスを処理して、インデックスおよびキーフレームイメージを生成した後、処理済みの記録に対する URL を伴う電子メールを登録済みのセッション参加者に送信する。ユーザは、その URL をクリックしてブラウズソフトウェアを立ち上げる。ブラウズソフトウェアの目的は、ユーザが、キーフレームイメージを閲覧することができ、特定のトピックに関連する音声に迅速にアクセスできるようにすることである。

20

【 0 1 1 5 】

ブラウズソフトウェアのユーザインターフェース（UI）を図 1 5 に示している。UI の主要な領域は、キーフレームサムネール 1 5 0 2（キーフレームイメージのグラフィックス表示）が表示されるキーフレームペイン 1 5 0 4、並びにカメラからの生のイメージ 1 5 1 2、および現行のキーフレームイメージ 1 5 0 2 の合成を示すブラウザのメイン表示ペインを含む。また、キーフレームペイン 1 5 0 4 には、ユーザが、メイン表示ペイン内に表示されたイメージを調整して生の入力イメージからキーフレームイメージにすることを可能にする背景透明スライダも組み込まれている。会議再生タイムラインにおいて既に書き込まれているストロークである現行のペンストローク 1 5 1 0 が、会議再生タイムラインにおいてまだ書き込まれていない将来のストローク 1 5 0 8 よりも濃く、より明確にメイン表示ペイン内でレンダリングされる。参加者が将来に書き込むことになるペンストローク 1 5 0 8 は、ゴーストのようなスタイルで示される。この映像化の技法は、後により詳細に説明する。

30

【 0 1 1 6 】

VCR - 標準タイムラインコントロール 1 5 1 4 が、メイン表示ペイン 1 5 0 6 の下でブラウザ UI の左下隅に提供される。VCR - 標準タイムラインコントロール 1 5 1 4 により、ユーザは、通常のビデオカセットレコーダ（VCR）上に見られるコントロールと同じように、イメージ / 音声シーケンスをゆっくりと、または迅速に戻すこと、または先に進めること、または停止させることが可能になる。タイムラインバー 1 5 1 8 が、音声 / イメージシーケンスの長さをバーとしてグラフィックスで表示し、会議再生の開始時刻、終了時刻、および現在時刻の数値を提供する。このバー 1 5 1 8 の上のポイント 1 5 2 0 を選択し、前後にドラッグして、イメージ / 音声シーケンスを直線的に先に進めること、および戻すことができる。

40

【 0 1 1 7 】

前述した UI 要素のいくつかの場所を提示しているが、これは限定することを意図するものではないことに留意されたい。以上の UI 要素は、ディスプレイ上の任意の場所に、単

50

独で、または他の要素と組み合わせて提供することが可能である。

【 0 1 1 8 】

2 . 4 . 2 会議データに対する非線形アクセス

記録済みの音声に対する2つのレベルの非線形アクセスが、ビジュアルのインデックス付けのコンテキストで提供されている。

【 0 1 1 9 】

第1のレベルの非線形アクセスは、キーフレームサムネール1502の使用を介するものである。各キーフレームサムネールには、ディスプレイ上である時間範囲が関連付けられている。ユーザは、キーフレームサムネールをクリックして、対応するキーフレームに関する音声の開始ポイント（例えば、区分の先頭）にジャンプすることができる。

10

【 0 1 2 0 】

記録済みの音声に対する第2のレベルのアクセスは、各キーフレーム内のペンストロークの使用を介するものである。カーソルが、メインウィンドウ1506内でペンストロークセル（現行のストローク1510または将来のストローク1508）の上を行き来しているとき、そのカーソルは、そのセルが選択可能（例えば、マウスで「クリック可能」）であることを示す「手」の記号に変化する。マウスまたは他の入力デバイスでそのセルをダブルクリックすることにより、アプリケーションが音声再生モードにされる。再生は、クリックされたストロークセルが書き込まれたセッションの時点から開始する。クリックされたストロークが書き込まれた時点は、シーケンスの中で同一のパターンのセルイメージが出現した最も早期の時点である。メインウィンドウは、その時点でイメージを示すことを開始する。ユーザは、それでも、他のストロークセルをクリックしてセッションの別の部分にジャンプすることができる。

20

【 0 1 2 1 】

VCR - 標準タイムラインコントロール1514と併せて、以上の2つのレベルのビジュアル式インデックス付けにより、ユーザは、非常に効率的な仕方では会議をブラウズすることが可能になる。

【 0 1 2 2 】

2 . 4 . 3 イメージビューイング

図15に示すとおり、キーフレームイメージのサムネール（例えば、1502）が、キーフレームペイン1504内にリストされる。サムネール1502の1つをマウスのカーソルまたはその他の入力デバイスで選択することにより、対応するキーフレームイメージが、左側でメインウィンドウ1506にもたらされ、アプリケーションが、イメージビューイングモードにされ、このモードで、ユーザは、ズーム制御ボタン1522を使用してズームインおよびズームアウトを行うこと、イメージの中のテキストおよび図を読み取ること、またはイメージの一部を他のドキュメントにカットアンドペーストしてすることができる。さらに、キーフレーム全体を他のドキュメントにカットアンドペーストすること、またはメモとして印刷することが可能である。

30

【 0 1 2 3 】

2 . 4 . 4 ホワイトボードの内容の映像化

キーフレームイメージおよびタイムスタンプ情報を所与として、任意の所与の時点におけるホワイトボードの内容に対応するイメージを再構成することができる。毎フレームのイメージが、タイムラインコントロール1514を使用する音声再生時刻に従ってレンダリングされた場合、メインウィンドウは、ホワイトボードの内容を映画のように再生する。この手法を使用することで、ユーザは、セッションに対する聴覚のコンテキストと視覚のコンテキストをともに有する。ただし、ユーザは、ユーザを先の時刻に進めるペンストローク（将来のストローク1508）をクリックすることはできない。というのは、そのストロークが、メインウィンドウ内でまだレンダリングされていないからである。

40

【 0 1 2 4 】

本ホワイトボードキャプチャシステムの初期の実施形態では、将来のストロークは、色あせた（washed out）モードで示されていた。しかし、短い試用期間の後、ブラウザのユー

50

ザは、しばしば、将来のストロークをきれいに消去されていないストロークと混同した。このインターフェースに関する別の苦情は、ユーザは、手前に人物のいないホワイトボードイメージを好んだが、ときとして、誰がストロークを書き込んだかを知ることが望んだということである。

【 0 1 2 5 】

数回の設計を繰り返した後、前述した懸念のすべてに対処する図 1 6 に示した以下の映像化プロセスが決められた。このプロセスのプロセスステップは、以下のとおりである。

【 0 1 2 6 】

1 . 現行の区分のキーフレームイメージおよびタイムスタンプ情報を使用して現行のホワイトボードの内容をレンダリングする (プロセスステップ 1 6 0 2) 。

10

【 0 1 2 7 】

2 . 将来のストロークをレンダリングし、結果をグレースケールに変換し、ガウスフィルタを使用してぼかす (プロセスステップ 1 6 0 4) 。

【 0 1 2 8 】

3 . ステップ 1 およびステップ 2 からのイメージを合わせる (プロセスステップ 1 6 0 6) 。

【 0 1 2 9 】

4 . ステップ 3 からのイメージを入力シーケンスからの修正済みのイメージにアルファブレンドする (プロセスステップ 1 6 0 8) 。修正済みのイメージは、入力シーケンスからの対応するイメージ (図 4 に示した) であるが、ホワイトボードでない領域がトリミングされた後、長方形の形状にマップしなおされたものである。ユーザは、レンダリングされたキーフレームホワイトボードイメージだけを示す 0 から、元の修正済みイメージそのものを示す 1 まで、GUI スライダ (図 1 5 の 1 5 1 6) を使用してアルファ値を制御することができる。レンダリングされたキーフレームホワイトボードイメージは、前景オブジェクトが除去され、前景オブジェクトが遮っていたストロークで置き換えられたキーフレームイメージである。

20

【 0 1 3 0 】

以上は、1) 現在のストロークと将来のストロークがともに、レンダリングされたホワイトボードイメージ上に示され、ユーザが、過去にジャンプして戻ることも、将来にジャンプして進むこともできるようにするため、また 2) 修正済み入力イメージをキーフレームと融合させることにより、前景オブジェクトが追加されて、より多くのコンテキストが与えられるため、非常に有用な映像化の仕方であると考えられる。アルファ = 0 . 8 を使用するそのような映像化の例に関して、図 1 5 を参照されたい。

30

【 0 1 3 1 】

2 . 5 セキュリティ

会議参加者は、機密に関わる情報が意図しない人々によって見られる可能性があるため、会議を記録することに関して、通常、危惧の念を抱いている。会議参加者には、記録されたデータをセキュアに保つことは、懸念すべき問題である。この問題に対処するため、簡単なトークンベースのアクセスセキュリティモデルが開発された。このプロセスのプロセスステップを図 1 7 に示している。

40

【 0 1 3 2 】

本ホワイトボードキャプチャシステムでは、会議参加者は、会議記録の開始時にキャプチャソフトウェアに登録するように求められる (プロセスステップ 1 7 0 2) 。会議参加者は、コンピュータスクリーン上のダイアログボックスの中で自らの電子メールの別名を記入するか、またはプロセスをスピードアップするため、自らの企業 ID カードをスマートカード読取装置に挿入して登録することができる。

【 0 1 3 3 】

すべての記録済みセッションは、Web サーバ上に存在する。誰も登録しなかった場合、会議は、公衆アクセス可能な Web ページ上に掲示される (プロセスステップ 1 7 0 4 、1 7 0 6) 。少なくとも 1 名の参加者が登録した場合、会議の記録および解析の後、アク

50

セストークンが生成される（プロセスステップ1708）。トークンは、固有の会議識別子を含む長いランダムに生成されたストリングである。トークンを含むURLが、登録済み参加者に電子メールで送信される（プロセスステップ1710）。受け手は、そのURLに行ってWebブラウザソフトウェアを立ち上げて会議を綿密に見る（1712）。また、受け手は、そのURLを会議に出席しなかった人々に転送することもできる。

【0134】

この簡単な隠蔽によるセキュリティ（Security by Obscurity）モデルは、うまく機能するようである。ただし、その他のセキュリティ対策を使用することも可能である。

【0135】

本ホワイトボードキャプチャシステムの前述したセキュリティ機構に加えて、会議を記録している間のプライバシーモデルも用意されている。会議参加者が、記録されることを望まないことを何か言った、または書いた場合、前の15秒（別の所定の期間を使用することも可能であるが）のイメージと音声データの両方を消去する機構が存在する。この消去は、物理的なボタン、またはGUIのボタンを押すことによって開始される。

【0136】

2.6 代替の実施形態

前述した基本的なホワイトボードキャプチャシステムを多数の他の技術およびデバイスと組み合わせて代替の実施形態を提供することが可能である。以下に述べる様々な実施形態は、単独で、または組合せで使用することが可能である。

【0137】

1つのそのような実施形態では、キーフレームに対して従来の光学式文字認識（OCR）を行って、ドキュメントまたはプレゼンテーションビュウグラフを作成するのに容易に使用される編集可能なテキストが提供される。

【0138】

別の実施形態では、従来の音声認識ソフトウェアを使用してキャプチャされたデータの音声部分がテキストに変換される。これにより、会議議事録およびその他のドキュメントの容易な作成が可能になる。また、会議情報を聴覚障害者に提供する比較的安価なやり方も提供される。

【0139】

また、本ホワイトボードキャプチャシステムは、例えば、三脚上のマイクロホンおよびカメラを備えたノートブックコンピュータを使用することによってポータブルにすることができる。この構成は、ホワイトボードに対するカメラの位置を決める追加の初期校正だけを必要とする。この校正は、イメージの中のパネルの4つの隅を手作業で決めることによって手作業で、またはエッジ検出などの従来の方法を使用することによって自動的に行うことが可能である。

【0140】

また、本ホワイトボードキャプチャシステムの解析ソフトウェアを使用して、ホワイトボードの内容を推定するのにペン追跡を使用するホワイトボードキャプチャシステムでキーフレームを決定することも可能である。ペン座標の履歴は、通常、そのシステムにおいてベクトル形式でキャプチャされるので、任意の所与の時点におけるホワイトボード上の内容を後に再構成することができる。そのようなシステムで本ホワイトボードキャプチャシステムの解析ソフトウェアを使用することにより、解析プロセスが単純化される。セルイメージは、ストロークまたはホワイトボードであるため、ホワイトボード背景色の判定、またはホワイトボード領域修正は必要なく、空間的および時間的にフィルタにかけることも必要なく、またホワイトボードセルの分類もより簡単である。というのは、ホワイトボード上に書き込まれたテキストを妨げる前景オブジェクトが存在しないからである。次に、セル「イメージ」が、ホワイトボード領域上のペン位置で推定された内容から導出される。本発明のこの実施形態は、基本的に、図5のプロセスステップ506で述べたとおり、セル「イメージ」をクラスタ化し、前景セルが存在しないことを除けば、プロセスステップ508と同様に各セルをストロークまたはホワイトボードセルとして分類し、その分

10

20

30

40

50

類結果を使用してキーフレームイメージを抽出する（プロセスステップ512）。その結果を低帯域幅および小さいストレージの要件で伝送し、アーカイブすることができる。さらに、この実施形態でも、OCRを使用してキャプチャ済みのキーフレームを複写することが可能である。

【0141】

さらに、本ホワイトボードキャプチャシステムの実用の実施形態では、システムのフレームレートは、市販のスチールカメラのフレームレートによって制限される。より高いフレームレートを実現するため、HDTVカメラなどの高解像度ビデオカメラを使用することが可能である。

【0142】

さらに別の実施形態では、本ホワイトボードキャプチャシステムは、ジェスチャ認識を組み込んでジェスチャコマンドを使用する。例えば、ホワイトボード上のどこかにコマンドボックスが書き込まれていることが可能である。ユーザが動いたとき、またはボックスを指し示したとき、システムは、ジェスチャ認識を使用してそのジェスチャが行われた特定の時刻におけるイメージにタイムスタンプを付ける。

【0143】

基本的な適用例では、解析プロセスは、ホワイトボード背景色が、入力シーケンスの中で一定であるものと想定している。ただし、誰によってもカメラから隠される可能性がないホワイトボードの最上部に既知の色パッチを取り付けることができる。ソフトウェアは、この容易に検出可能なパッチの既知の色特性に基づき、フレームごとに異なる照明条件に合わせてカメラ露出パラメータを調整することができる。これは、次のとおり行われる。露出パラメータが一定のままである場合、室内の異なる照明条件でキャプチャされたイメージにおいてパッチの色が異なる可能性がある。カメラは、前のフレームの中でパッチがどのような色であるかということから、露出パラメータを調整することができる。パッチの色が指定された範囲内にとどまり、ホワイトボード領域も同様であることが可能である。

【0144】

3.0 システムのパフォーマンスおよび使用

3.1 背景

本ホワイトボードキャプチャシステムの設計目標は、本ホワイトボードキャプチャシステムが、1) 任意の既存のホワイトボードで機能し、2) ホワイトボードの内容を自動的に信頼の置けるようにキャプチャし、3) ホワイトボードの内容をビジュアル式インデックスとして使用して、本システムを使用して記録された会議を効率的にブラウズすることである。

【0145】

感知機構または電子ホワイトボードを使用するホワイトボードキャプチャシステムと比べて、本ホワイトボードキャプチャシステムは、1組の独特な技術的課題も有している。第1に、ホワイトボード背景色は、通常、あらかじめ較正されない（例えば、空白のホワイトボードの写真を撮影することにより）。というのは、各部屋が、セッションごとに異なる可能性があるいくつかの照明セッティングを有するからである。第2に、しばしば、人々は、デジタルカメラとホワイトボードの間で動き、この前景オブジェクトが、ホワイトボードのいくらかの部分の覆い隠し、ホワイトボード上に影を落とす。シーケンスの中で、全く覆い隠されていないフレームが存在しない可能性がある。タイムスタンプを計算し、キーフレームを抽出するため、以上の問題に対処しなければならなかった。

【0146】

3.2 システム構成要素

本ホワイトボードキャプチャシステムの設計中、試作品のシステムを構築し、繰り返し改良を行った。3つの会議室に本ホワイトボードキャプチャシステムを備えた。この3つの部屋に関する情報を以下のテーブル1にリストしている。サンプルイメージ（ボード上で80×80画素、およそ96ポイントのフォント）を図18Aに示している（イメージは

10

20

30

40

50

それぞれ、左から右に、部屋 1、部屋 2、および部屋 3 に対応する)。

【 0 1 4 7 】

【表 1】

表 1 3つの設置場所に関する情報

	部屋1	部屋2	部屋3
ボード寸法 (フィート (センチメートル))	4×3 (121×91)	8×5 (244×152)	12×5 (366×152)
キーフレームイメージ寸法 (画素)	1200×900	2400×1500	2400×1000
解像度 (dpi)	25	25	16.7

10

【 0 1 4 8 】

以上の部屋内のホワイトボードのサイズは異なり、生成されたキーフレームイメージの品質も異なっていた。サンプルイメージ (図 1 8 A) から見て取ることができるとおり、1 2' × 5' (3 6 6 センチメートル × 1 5 2 センチメートル) のボード上の書込みは、4 メガピクセルの入力イメージで解像度が限度に達したため、その他の 2 つのボード上の書込みよりも不明瞭である (一番右)。それでも、その書込みは、やはり極めて読みやすい。1 2' × 5' (3 6 6 センチメートル × 1 5 2 センチメートル) のホワイトボードを使用するセッションからのいくつかの選択されたフレーム (図 1 8 B)、および対応するキーフレームイメージ (図 1 8 C および 1 8 D) も示している。

20

【 0 1 4 9 】

本システムは、特殊なペンおよびイレーザを必要とすることなしに、任意の既存のホワイトボードで機能しなければならないので、ホワイトボードの内容をキャプチャするのに、直接キャプチャデバイスであるスチールカメラが選択された。本ホワイトボードキャプチャシステムの例示的な実用の実施形態では、4 メガピクセルを有するキャノン (登録商標) Power Shot G 2 デジタルスチールカメラが使用された。このカメラは、2 2 7 2 画素 × 1 7 0 4 画素、つまり 6' × 4' (1 8 3 センチメートル × 1 2 1 センチメートル) のボードの場合、3 1 . 6 dpi に等しいイメージを提供する。このカメラが選択された 1 つの重要な理由は、P C からカメラを制御するようにカスタマイズされたソフトウェアソリューションを書くことを可能にするソフトウェア開発キットが入手できることである。このソフトウェアは、実質的にすべてのカメラパラメータをショットごとに指定することを可能にする。本システムは、ホワイトボードの写真を直接に撮影するので、ペンのストロークの位置ずれは存在しない。ユーザが、消去を行う前にシステムをオンにする限り、内容は保存される。

30

40

【 0 1 5 0 】

アナリシスサーバは、ペンティアム (登録商標) I I I 8 0 0 M H z のデュアル C P U P C 上で動作する。解析プロセスは、セッション時間の毎時間におよそ 2 0 分間かかる。1 6 ビット 1 1 K H z のモノラルの音声に関する記憶要件は、M P 3 符号化を使用しておよそ毎時 1 5 M b を要する。入力イメージシーケンスは、モーション J P E G 圧縮を使用しておよそ毎時間 3 4 M b を要する。

【 0 1 5 1 】

3 つの会議室内に設置されたシステムは、様々なチームによって頻繁に使用された。6 カ月の間に、合計 4 8 時間の 1 0 8 回のセッションが記録された。すなわち、平均して 1 回のセッション当り 2 7 分間であり、1 週間当り 4 . 5 回のセッションになる。1 回のセッ

50

ション当りのキーフレームの平均数は、2.7であった。キーフレームイメージは、J P E G形式で保存した。平均イメージサイズは、51.8 K bであった。サイズは、17 K bから150 K bまでの範囲内にあった。J P E G圧縮は、一様に白い背景に対して極めてよく機能したため、イメージサイズは、イメージの寸法よりも、どれだけ多くユーザがボード上に書き込みを行ったかの方により強く関連していた。

【0152】

本システムのすべてのユーザが、本システムは、ホワイトボードを頻繁に使用する会議に非常に役立つと考えた。キーフレームイメージ、およびビジュアル式インデックス付けの機能により、参加者が、後の時点で会議を綿密に見る（review）ことができるようになるだけでなく、会議に出席しなかったユーザが、実際の会議時間の数分の1の時間で会議の要点を理解することができるようになる。

10

【0153】

一部のユーザは、当初、意図されていなかった本システムの新しい用途を見出した。ホワイトボード上に書き込むことを通常、必要としなかった旧態依然の会議（status meeting）を例にとろう。人々は、やはり、ホワイトボードキャプチャシステムをオンにした。誰かが発言する順番になったとき、主任が、その人の名前をボード上に書き込んで、その発言のセグメントを後に記録された音声の中で、そのキーフレームイメージの中の名前をクリックすることによって容易に見出すことができるようにした。別の例は、ブレンストーミングセッション中であり、誰かが、よい考えを思い付いたとき、その人が、ボードの脇に星印を書き込み、その考えを声に出して言った。音声は、後にその星印をクリックすることによって取り出すことができる。

20

【0154】

本発明の以上の説明は、例示および説明の目的で提示した。以上の説明は、すべてを網羅する、または本発明を開示した形態そのものに限定することを意図するものではない。以上の教示に鑑みて、多数の変更形態および変形形態が可能である。本発明の範囲は、以上の詳細な説明によってではなく、本明細書に頭記の特許請求の範囲によって限定されるものとする。

【0155】

付録：平面ベースのホワイトボードの色の推定

カラーイメージの1つの成分だけを考慮するが、以下に説明する技法は、すべての成分（R、G、B、またはY）に適用される。各セル i が、イメージ座標 (x_i, y_i) で定義される。セルの色は、 z_i （ $z = R, G, B, \text{または} Y$ ）で指定される。色は、セクション2.3.2に説明したとおり計算され、したがって、ノイズが多く、誤りさえある。会議室での体験では、ホワイトボードの色は、しばしば、一様ではない。ホワイトボードの色は、通常、上部ではるかに明るく、下部に向かって暗くなるか、あるいは上隅の1つではるかに明るく、反対側の下隅に向かって暗くなる。これは、照明が天井に対して設置されているためである。したがって、ある局所領域（例えば、 7×7 セル）に関して、平面を色に正確に適合させることができ、イメージ全体に関して、平面の適合（plane fitting）はそれでも、非常に適切であり、セルの色が外れ値であるかどうかの堅牢な指示を提供する。

30

40

【0156】

平面は、 $ax + by + c - z = 0$ で表わすことができる。 z_i にだけノイズを有する1組の3Dポイント $\{(x_i, y_i, z_i) \mid i = 1, \dots, n\}$ が与えられる。平面パラメータ $p = [a, b, c]^T$ は、次の目的関数を最小化することによって推定することができる。

【0157】

【数22】

$$F = \sum_i f_i^2$$

【0158】

50

ただし、 $f_i = a x_i + b y_i + c - z_i$ である。最小2乗の解が、以下のとおり与えられる。

【0159】

$$p = (A^T A)^{-1} A^T z、$$

ただし、

【0160】

【数23】

$$A = \begin{bmatrix} x_1 & y_1 & 1 \\ \cdots & \cdots & \cdots \\ x_n & y_n & 1 \end{bmatrix}$$

10

【0161】

であり、かつ $z = [z_1, \dots, z_n]^T$ である。平面パラメータが決まると、セル i の色を

【0162】

【数24】

$$\hat{z}_i = ax_i + by_i + c$$

【0163】

で置き換える。

【0164】

最小2乗法は、誤りのあるデータ（外れ値）に対して堅牢ではない。前述したとおり、最初に計算されたホワイトボードの色は、外れ値を含まない。外れ値を検出し、拒否するため、平面をホワイトボードイメージ全体に適合させる堅牢な技法を使用する。データの半分近くが外れ値であることを許容することができる非常に堅牢な技法である最小2乗平均法 [11] を使用する。考え方は、2乗された誤りの総和ではなく、中央値を最小化することによってパラメータ、すなわち、

【0165】

【数25】

$$\min_p \text{median}_i f_i^2$$

【0166】

を推定することである。3つのポイントの最初の m 個のランダムなサブサンプルが取られる（3は、平面を定義する最小数である）。各サブサンプルにより、平面の推定値が与えられる。数 m は、 m 個のサブサンプルの少なくとも1つが良好である確率が1に近く、例えば、99%になるだけ十分に大きくなければならない。データの半分が外れ値である可能性があるものと想定した場合、 $m = 35$ であり、したがって、ランダムなサンプリングを非常に効率的に行うことができる。各サブサンプルに関して、平面パラメータ、および2乗された誤り

40

【0167】

【数26】

$$f_i^2$$

【0168】

の中央値を計算する。Mで表わされる2乗された誤りの最小中央値を与える平面パラメータが保持される。次に、いわゆる堅牢な標準偏差

【0169】

【数27】

$$\sigma = 1.4826\sqrt{M}$$

【 0 1 7 0 】

(この係数を使用して、外れ値が存在しないときと同じ効率が実現される)を計算する。ポイント i は、ポイント i の誤りが $|f_i| > 2.5$ である場合、外れ値と見なし、廃棄する。最後に、前述した最小 2 乗法を使用してその良好なポイントに平面を適合させる。外れ値のセル i の色を

【 0 1 7 1 】

【数 2 8】

$$\hat{z}_i = ax_i + by_i + c$$

10

【 0 1 7 2 】

で置き換える。

【図面の簡単な説明】

【図 1】本発明を実施するための例示的なシステムを構成する汎用コンピューティングデバイスを描いた図である。

【図 2】本ホワイトボードキャプチャシステムの 3 つの主な構成要素、すなわち、キャプチャユニット、アナリシスサーバ、およびブラウズソフトウェアを描いている、試作品のホワイトボードキャプチャシステムを使用してキャプチャされた図である。

【図 3】本発明によるホワイトボードキャプチャシステムを示す概略図である。

20

【図 4】入力イメージシーケンスからの選択されたフレームを示すイメージシーケンスを示す図である。

【図 5】本発明によるシステムおよび方法のイメージ解析プロセスを描いた流れ図である。

【図 6 A】ホワイトボードの色を計算する第 1 の技術を示す図である。

【図 6 B】ホワイトボードの色を計算する第 2 の技術を示す図である。

【図 7】左側のイメージが、ホワイトボードの色を計算する第 1 の戦略の結果であり、中央のイメージが、ホワイトボードの色を計算する第 2 の戦略の結果であり、また右側のイメージが、実際の空白のホワイトボードイメージを示しているホワイトボード色抽出結果を示すイメージシーケンスを示す図である。

30

【図 8】本発明によるシステムおよび方法のセル分類プロセスを描いた流れ図である。

【図 9】上のイメージが、トリミングおよび修正の後の図 5 のイメージに対応する分類結果の一連のサンプルを示す図である。

【図 10】図 4 のシーケンスに関するストローク数対時間のプロットを示す図である。

【図 11】入力イメージのシーケンスからキーフレームを選択するのに使用される一般的なプロセスを描いた流れ図である。

【図 12】本発明によるシステムおよび方法において区分およびキーフレームを識別するプロセスを描いた流れ図である。

【図 13】本発明によるシステムおよび方法においてキーフレームイメージを再構成するプロセスを描いた流れ図である。

40

【図 14】本発明によるシステムおよび方法においてキーフレームイメージのカラーバランス処理を行うプロセスを描いた流れ図である。

【図 15】各キーフレームイメージが、記録の中の重要な時点のホワイトボードの内容を代表する、本ホワイトボードキャプチャシステムのブラウザインターフェースを描いたイメージを示す図である。

【図 16】本発明によるシステムおよび方法における現在のペンストロークおよび将来のペンストロークを表示するプロセスを描いた流れ図である。

【図 17】本発明によるシステムおよび方法において使用されるセキュリティ処理を描いた流れ図である。

【図 18 A】本発明の実用の実施形態の 3 つの設置場所において取り込まれたホワイトボ

50

ードの内容のサンプルイメージを提供する図である。

【図 1 8 B】本発明の実用の実施形態の入力キーフレームイメージを描いた図である。

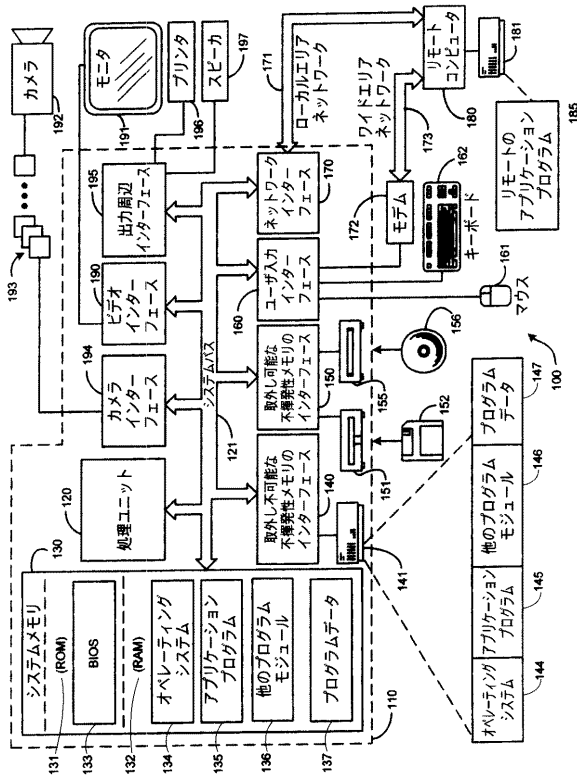
【図 1 8 C】本発明の実用の実施形態の出力キーフレームイメージを描いた図である。

【図 1 8 D】本発明の実用の実施形態の出力キーフレームイメージを描いた図である。

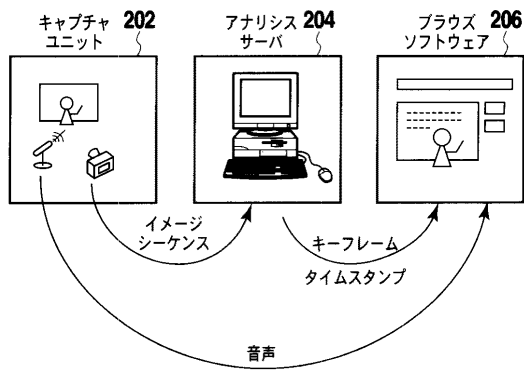
【符号の説明】

1 2 0	処理ユニット	
1 2 1	システムバス	
1 3 0	システムメモリ	
1 3 4	オペレーティングシステム	
1 3 5	アプリケーションプログラム	10
1 3 6	他のプログラムモジュール	
1 3 7	プログラムデータ	
1 4 0	取外し不可能な不揮発性メモリのインターフェース	
1 4 4	オペレーティングシステム	
1 4 5	アプリケーションプログラム	
1 4 6	他のプログラムモジュール	
1 4 7	プログラムデータ	
1 5 0	取外し可能な不揮発性メモリのインターフェース	
1 6 0	ユーザ入力インターフェース	
1 6 1	マウス	20
1 6 2	キーボード	
1 7 0	ネットワークインターフェース	
1 7 1	ローカルエリアネットワーク	
1 7 2	モデム	
1 7 3	ワイドエリアネットワーク	
1 8 0	リモートコンピュータ	
1 8 5	リモートアプリケーションプログラム	
1 9 0	ビデオインターフェース	
1 9 1	モニタ	
1 9 2	カメラ	30
1 9 4	カメラインターフェース	
1 9 5	出力周辺インターフェース	
1 9 6	プリンタ	
1 9 7	スピーカ	
2 0 2	キャプチャユニット	
2 0 4	アナリシス / 処理サーバ	
2 0 6	ブラウズソフトウェア	
1 5 0 2	キーフレームサムネール	
1 5 0 8	将来のストローク	
1 5 1 0	現在のストローク	40
1 5 1 2	生のイメージ	
1 5 1 4	V C R - タイムラインコントロール	

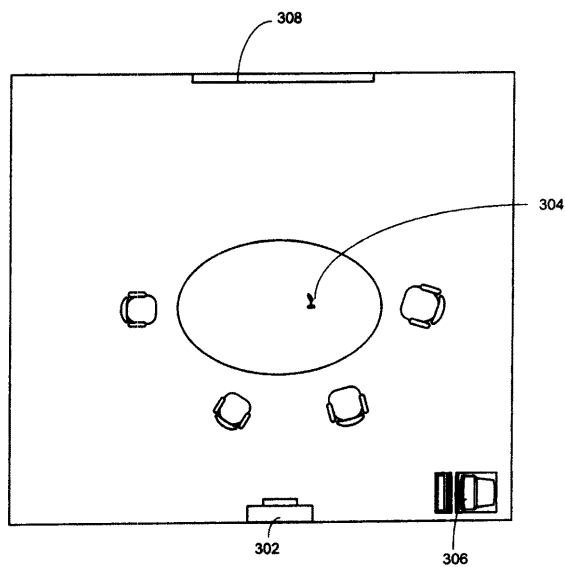
【図 1】



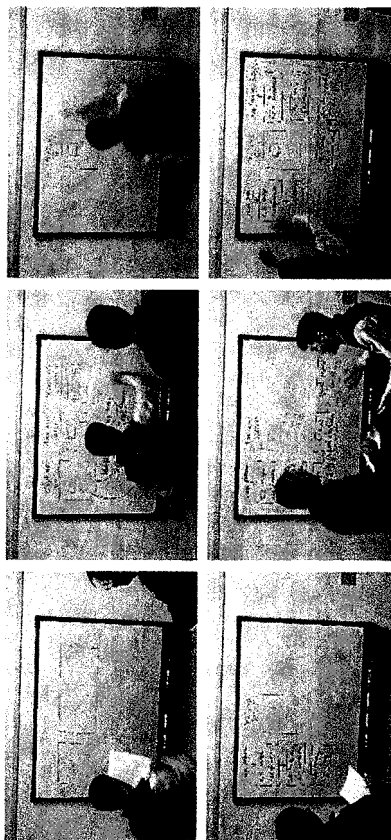
【図 2】



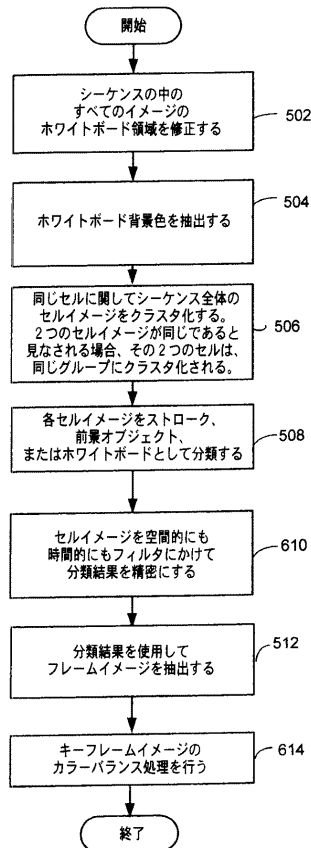
【図 3】



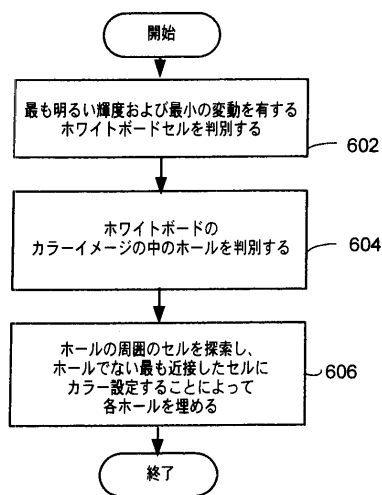
【図 4】



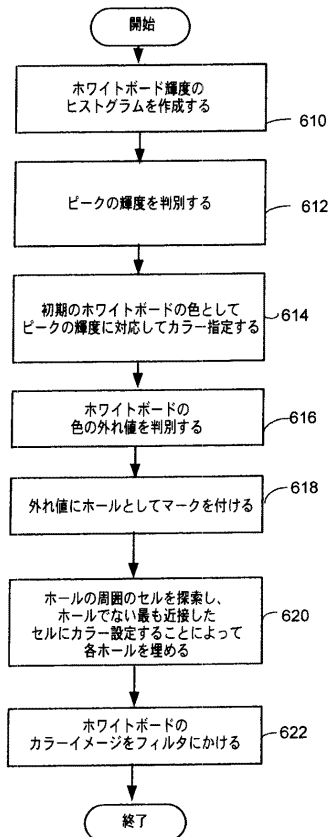
【図 5】



【図 6 A】



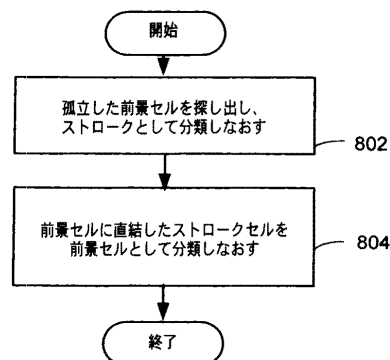
【図 6 B】



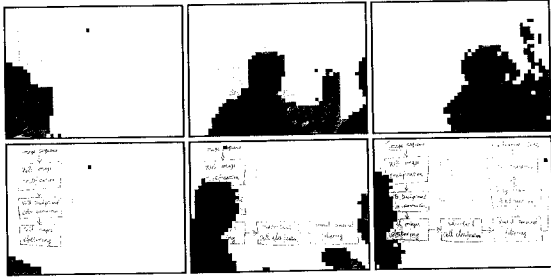
【図 7】



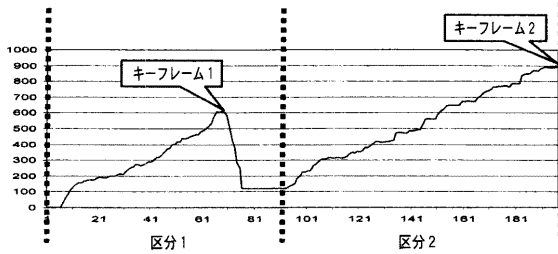
【図 8】



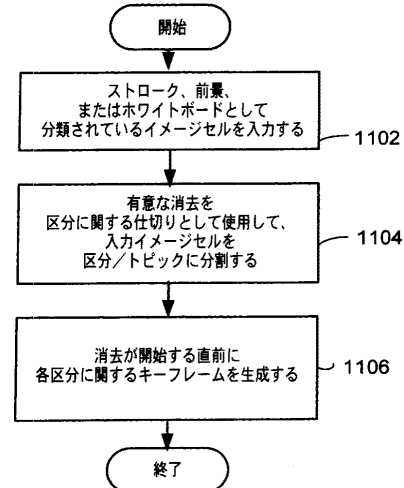
【図 9】



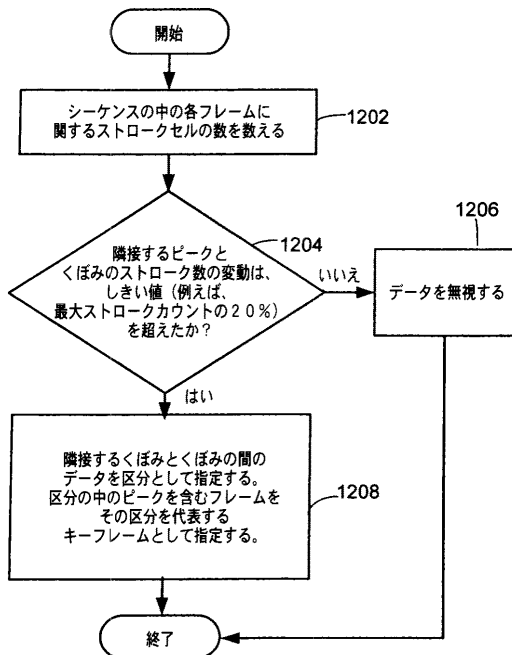
【図 10】



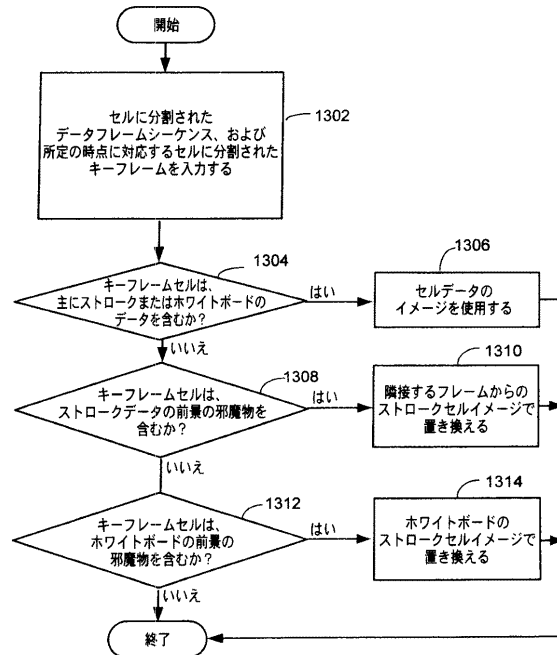
【図 11】



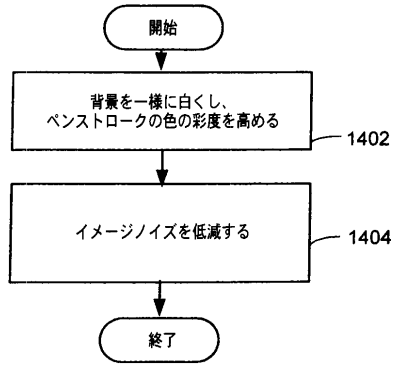
【図 12】



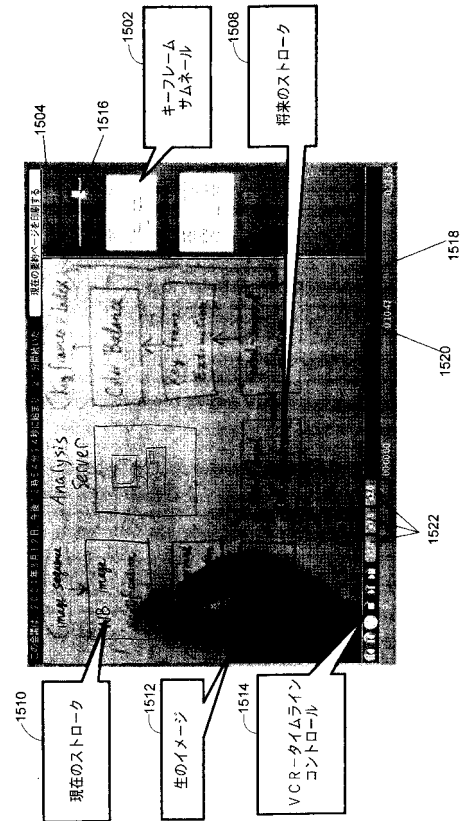
【図 13】



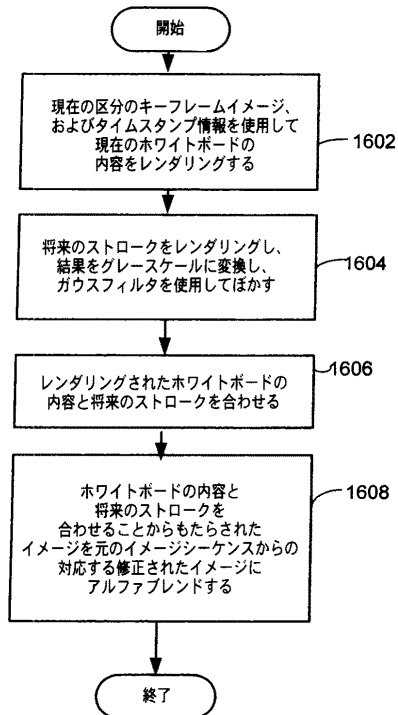
【図 14】



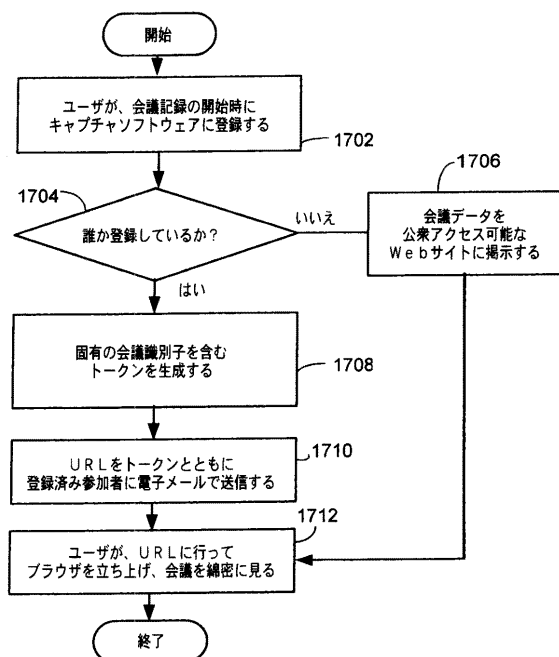
【図 15】



【図 16】



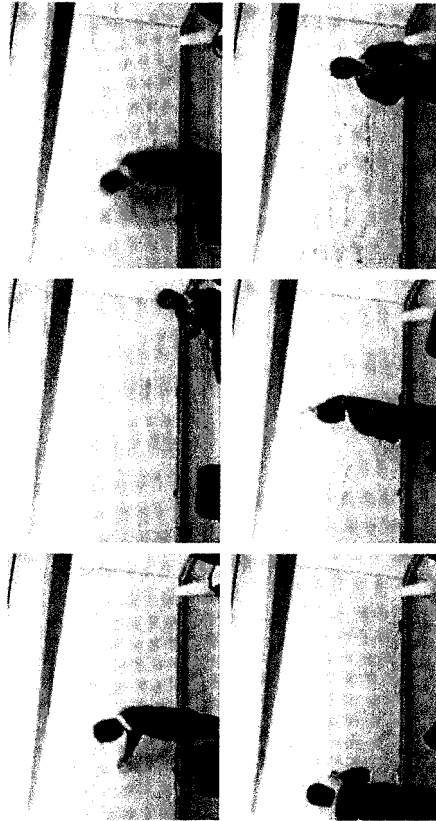
【図 17】



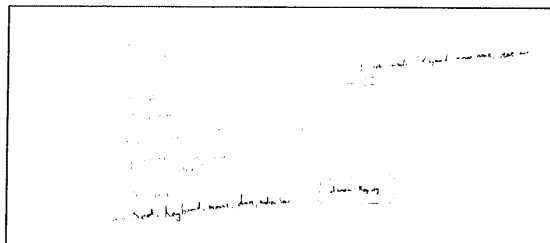
【図 18 A】

net	middle l	Total
1 min	1 min	30%

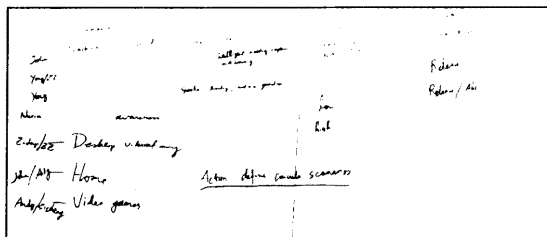
【図 18 B】



【図 18 C】



【図 18 D】



フロントページの続き

(72)発明者 アヌーブ グプタ

アメリカ合衆国 98072 ワシントン州 ウッディンビル ノースイースト 129 ストリート 19908

(72)発明者 ロス カトラー

アメリカ合衆国 98019 ワシントン州 ドボル 277 プレイス ノースイースト 16031

(72)発明者 ジチェン リウ

アメリカ合衆国 98006 ワシントン州 ベルビュー サウスイースト 63 ストリート 14743

(72)発明者 リ - ウェイ ヘ

アメリカ合衆国 98052 ワシントン州 レッドモンド ウェスト レイク サマミッシュパークウェイ ノースイースト 4221

審査官 伊東 和重

(56)参考文献 国際公開第00/027125 (WO, A1)

特開平10-214270 (JP, A)

特表平09-510852 (JP, A)

特表2002-529975 (JP, A)

(58)調査した分野(Int.Cl., DB名)

H04N 7/15

H04N 7/18

H04N 5/262