

(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2010-141897

(P2010-141897A)

(43) 公開日 平成22年6月24日 (2010.6.24)

(51) Int.Cl. F I テーマコード (参考)
 HO4N 17/02 (2006.01) HO4N 17/02 Z 5C061

審査請求 未請求 請求項の数 2 O L 外国語出願 (全 31 頁)

(21) 出願番号	特願2009-283345 (P2009-283345)	(71) 出願人	391002340
(22) 出願日	平成21年12月14日 (2009.12.14)		テクトロニクス・インコーポレイテッド
(31) 優先権主張番号	61/122, 221		TEKTRONIX, INC.
(32) 優先日	平成20年12月12日 (2008.12.12)		アメリカ合衆国 オレゴン州 97077
(33) 優先権主張国	米国 (US)		-0001 ビーバートン サウスウエスト
(31) 優先権主張番号	12/635, 456		ト カール・ブラウン・ドライブ 141
(32) 優先日	平成21年12月10日 (2009.12.10)		50
(33) 優先権主張国	米国 (US)	(74) 代理人	100090376
			弁理士 山口 邦夫
		(74) 代理人	100124109
			弁理士 山口 隆史
		(72) 発明者	ケビン・エム・ファーガソン
			アメリカ合衆国 オレゴン州 97008
			ビーバートン サウスウエスト コット
			ンテール・レーン 13315
		Fターム(参考)	5C061 BB01 BB03 CC05

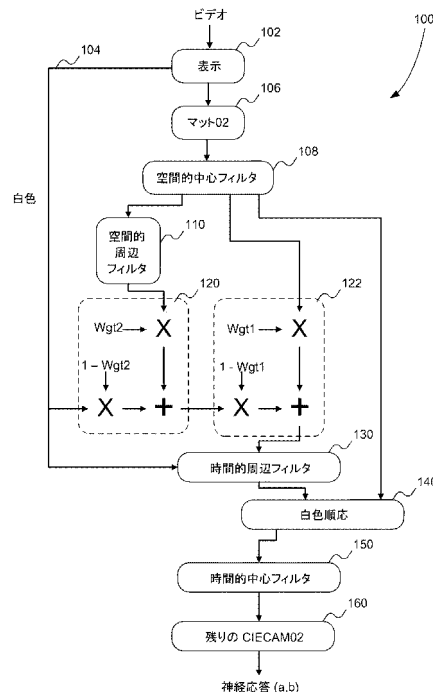
(54) 【発明の名称】 ビデオ内の色視覚差予測方法及び移動イメージ色見えモデル用システム

(57) 【要約】

【課題】 移動イメージ色見えをモデル化する。

【解決手段】 入力ビデオ信号を受け、白色発生器がビデオ信号から未ろ波白色信号を生成する。コーン応答発生器は、入力に結合されて、ビデオ信号からの1組のコーン応答を発生する。空間中心フィルタは、1組のコーン応答の1つの空間中心を発生する。空間的周辺フィルタは、1組のコーン応答の1つの空間周辺を発生する。時間的周辺フィルタは、空間中心、空間周辺及び未ろ波白色信号から時間的周辺を発生する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

被分析ビデオ信号を受け、
ビデオ表示及び観察環境に応じた未ろ波白色信号を受け、
上記ビデオ信号から空間的周辺を生成し、
上記ビデオ信号から空間的中心を生成し、
1つ以上の重み付けファクタを用いて、上記空間的周辺、上記空間的中心及び上記未ろ波白色信号の組合せから上記ビデオ信号の時間的周辺を生成する
ビデオ内の色視覚差予測方法。

【請求項 2】

ビデオ信号を受けるとして構成された入力と、
上記ビデオ信号から未ろ波白色信号を生成するように構成された白色発生器と、
上記入力に結合され、上記ビデオ信号から1組のコーン応答を発生するように構成されたコーン応答発生器と、
上記1組のコーン応答の1つの空間的中心を発生するように構成された空間的中心フィルタと、
上記1組のコーン応答の1つの空間的周辺を発生するように構成された空間的周辺フィルタと、
上記空間的中心、空間的周辺及び未ろ波白色信号から時間的周辺を発生するように構成された時間的周辺フィルタと
を具えた移動イメージ色見えモデル用システム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、ビデオの画像品質用のビデオ試験測定機器に関し、特に、ビデオ内の色視覚差予測方法及び移動イメージ色見えモデル用システムに関する。

【背景技術】**【0002】**

ビデオ又はイメージのストリームを操作するとき、例えば、ビデオを圧縮してDVDに蓄積するか又はテレビジョン・チャンネルにより伝送するとき、または、ビデオをあるフォーマットから他のフォーマットに変換するとき、この操作は、圧縮又は変換からのアーティファクトを結果のビデオに導入する。人間の目の生理及び現象により、また、目が光及び動きを知覚する方法により、ときどき、人間の観察者にはこれらアーティファクトが目に見えないか又は知覚できない。他のときには、アーティファクトがはっきりと見える。結果のビデオ・データを単に分析することにより、操作されたビデオをどのように知覚するかを判断することがほぼ可能である。

【0003】

この代わりには、ビデオ品質を評価するのに集まった人々の統計的大多数を調査することにより、ビデオ品質を最良に判断する。視聴者は、ビデオを観察し、所定の品質尺度でこのビデオを評価する。しかし、かかる聴衆の評価は、ほとんどのアプリケーションに対して実際的ではなくコスト的に高くなるので、自動的な画像品質評価の方法が開発されてきた。

【0004】

フル基準(FR)画像品質分析器は、2つのイメージ又は2つのイメージ・ストリームを比較し、人間がビデオ変換の品質をどのように知覚するか、又はオリジナルの試験ストリーム及び変更した試験ストリームの間の一致をどのように知覚するかの「スコア」又は他の尺度を発生する。高いスコアは、高品質の変換を示すが、低いスコアは、結果のビデオがそのオリジナルの劣化したものであることを示す。

【0005】

断片的であるが、最先端においては、画像品質分析は、1)ビデオ品質の計算に色分析

10

20

30

40

50

を含んでいないか、又は、2)色PSNR(ピーク信号対ノイズ比)の如き簡単なオブジェクト測定、若しくはブロックング又はリングングの如き色欠陥測定を含んでいるが、時空間的なコンテキストに応じた色変化の検知にて大きな変化を起こす人間の視覚システムの適合メカニズムを考慮しない。

【0006】

ビデオ品質での計算において色分析を含んでいない画像品質測定装置の例は、テクトロニクスPQA500型である。分析をルミナンス(光の強さ)値のみで行う。色見えモデルにフィルタを追加する提案が公表されているが、今までのところ、これらは(静的ガウシアン空間フィルタの如き)固定モデルである。PQA500型のルミナンス知覚モデルにおいて、順応応答は、「周辺」時空間的フィルタの出力として示される統合順応点に関連し、時間及び空間における高い分解能が「中心」時空間フィルタにより設定される。輝度の知覚できるエラーを検出する最先端であるこのルミナンスだけの方法は、ルミナンス・エラー及びクロマ・エラーが密接に関連しているという主な傾向により、ほとんどのビデオ伝送システムの評価アプリケーションには充分であった。しかし、エンコーダ及び他のビデオ処理装置における柔軟性と、より複雑なシステムでは、色エラーとして主に見える欠陥の機会が増える。例えば、ビデオ放送及びウェブキャストにおいて、入力ビデオと1つ以上の処理又は表示コンポーネントの色度測定が不一致となり、全域の色の外をクリップし、いくつかの場合には、全域の色の外を反対の色にオーバーフローさせる。ルミナンス分析は、これらのエラーの大部分の検出を失敗するので、品質評価が不正確になる。

10

【先行技術文献】

20

【特許文献】

【0007】

【特許文献1】特開2001-197527号公報

【特許文献2】特開2003-199127号公報

【発明の概要】

【発明が解決しようとする課題】

【0008】

本発明の実施例では、従来技術のこれらの及びその他の制限を解決する。

【課題を解決するための手段】

【0009】

30

上述の課題を解決するために、本発明は次のように構成される。

(1)被分析ビデオ信号を受け;ビデオ表示及び観察環境に応じた未ろ波白色信号を受け;上記ビデオ信号から空間的周辺を生成し;上記ビデオ信号から空間的中心を生成し;1つ以上の重み付けファクタを用いて、上記空間的周辺、上記空間的中心及び上記未ろ波白色信号の組合せから上記ビデオ信号の時間的周辺を生成するビデオ内の色視覚差予測方法。

(2)上記未ろ波白色信号は、1つ以上の周囲の白色信号、表示白色信号及び絶対白色信号の組合せである概念1の方法。

(3)更に、第1組合せ重み付けファクタに応じて上記周囲白色信号及び上記表示白色信号を組合せて、中間白色信号を生成し;第2組合せ重み付けファクタに応じて上記中間白色信号及び上記全体白色信号を組合せて上記未ろ波白色信号を生成する概念2の方法。

40

(4)時間的周辺を生成することは、可変レート・フィルタによりデータを時間的にろ波することである観点1の方法。

(5)時間的周辺を生成することは、適応可能フィルタによりデータを時間的にろ波することである観点1の方法。

(6)時間的周辺を生成することは、非対称フィルタによりデータを時間的にろ波することである観点1の方法。

(7)時間的周辺を生成することは、循環フィルタによりデータを時間的にろ波することである観点1の方法。

(8)上記ストリームを時間的に処理することは、第1及び第2ロウパス・フィルタによ

50

りデータをろ波することである観点6の方法。

(9) 上記第2ロウパス・フィルタの1組の動作パラメータの少なくとも1つは、色動き方向に応じて変化する観点7の方法。

(10) 更に、上記第1及び第2ロウパス・フィルタの出力を加算して上記時間的周辺を発生する概念8の方法。

(11) ビデオ信号を受けるように構成された入力と；上記ビデオ信号から未ろ波白色信号を生成するように構成された白色発生器と；上記入力に結合され、上記ビデオ信号から1組のコーン応答を発生するように構成されたコーン応答発生器と；上記1組のコーン応答の1つの空間的中心を発生するように構成された空間的中心フィルタと；上記1組のコーン応答の1つの空間的周辺を発生するように構成された空間的周辺フィルタと；上記空間的中心、空間的周辺及び未ろ波白色信号から時間的周辺を発生するように構成された時間的周辺フィルタとを具えた移動イメージ色見えモデル用システム。

10

(12) 第1重み付けファクタに応じて、上記未ろ波白色信号を上記空間的周辺と組合せるように構成された第1混合制御器を更に具えた概念11のシステム。

(13) 第2重み付けファクタに応じて、上記第1混合制御器の出力を上記空間的中心と組合せるように構成された第2混合制御器を更に具えた概念11のシステム。

(14) 上記未ろ波白色信号が、周囲白色信号、表示白色信号及び絶対白色信号の組合せである概念11のシステム。

(15) 上記時間的周辺フィルタは、第1ロウパス・フィルタと、第2ロウパス・フィルタと、上記第2ロウパス・フィルタの入力及び出力に結合された比較器とを更に具える概念11のシステム。

20

(16) 上記1組のコーン応答の各々に対して、1組の空間的中心フィルタ、1組の空間的周辺フィルタ、1組の時間的周辺フィルタを更に具えた概念11のシステム。

(17) 上記時間的周辺フィルタが可変レート・フィルタである概念11のシステム。

(18) 上記時間的周辺フィルタが適応可能フィルタである概念11のシステム。

(19) 上記時間的周辺フィルタが非対称フィルタである概念11のシステム。

(20) 上記時間的周辺フィルタが循環フィルタである概念10のシステム。

【図面の簡単な説明】

【0010】

【図1】本発明の実施例による移動イメージ用色分析を実現するシステムの機能ブロック図である。

30

【図2】本発明の実施例による図1のシステムの時間周辺処理コンポーネントの機能ブロック図である。

【図3】図2に示す時間的周辺処理コンポーネントのフィルタの1つのディラック・デルタ・インパルス応答を示すグラフである。

【図4A】図2に示す時間的周辺フィルタの1つの正ディラック・デルタ・インパルス応答を示すグラフである。

【図4B】図4Aに示した正応答用の同じフィルタの負ディラック・デルタ・インパルス応答を示すグラフである。

【図5】図2の時間的処理コンポーネントを用いた応答を示すグラフである。

40

【図6A】CIE 1931 {x、y}平面における試験刺激を示す図である。

【図6B】図2に示す時間的フィルタを用いて処理した場合と処理しない場合の応答を示す図である。

【発明を実施するための形態】

【0011】

周辺又は環境が変化すると、又は、他の色を見た後に観察すると、人間の目の色知覚適合により、所定の一定の物理的光刺激を異なる色に見える。一般的には、ダイナミックな光刺激により、空間及び/又は時間の変化により、所定の静止光刺激の知覚が一般には変化したように見える。順応による色見えのこの明らかな変化における時間的経過は、多くの刺激状態に応じて決まる。

50

【 0 0 1 2 】

イルミネーション及び白色点の認識したディスカウントにより、人間は、赤色光のみで照明された白色背景に対してリングを観察できると共に、リングが赤色か緑色か、又は背景が赤色か白色かを依然判断できる。赤色で照明される限りそれが赤色として見えるが、白色背景が白色であると知ることが、認識のディスカウントの例である。いくつかのアプリケーションにおいて、全体的な色のシフト（この例では、照明色の方向に）のこの形式のディスカウントが望ましいが、他の場合には、ディスカウントが望ましくない。白色バランスには、映画又はテレビ・ショーの場面の間を一致させることが一般的には望ましく、空での太陽の位置のシフト又はスタジオ照明の不注意な変化による色の違いは、望ましくない。一方、インターネット・ブラウザにより表示される V E S A 標準色度測定（青みがかった白色点）によるコンピュータ・モニタに対して、標準色度測定により民生液晶テレビ上のコーデック関連欠陥の知覚を評価することが望ましいならば、画像品質評価において、色標準でこれらの違いを含むことは望ましくない。したがって、画像品質測定を行うときに、これら 2 つの場合及びこれらの間の他の場合もカバーするように、知覚ディスカウントのゼロからフルまでの連続的な制御が重要である。

10

【 0 0 1 3 】

凝視の変能力に依りて、知覚色順応が非常に局部的になる。例えば、比較的長時間にわたって固定された点を凝視すると、色の知覚が低下する。代わりに、色順応は、主に、白色でないイルミネーション又はイメージからの他の色によるバイアスの如きバイアスに、全体的に基づいている。固定点での長時間の凝視による局部的な色順応効果により、色残像となり、これは、効果において、ある色に対する目の感度抑圧の結果である。いくつかのアプリケーションにおいて、ほとんど劇的であるこれら局部的順応の影響を予測することは、ビデオ品質を予測するために重要である。他のアプリケーションにおいて、たぶん、全体的な効果のみが関心事である。例えば、約 2 ダースの人数の個人のビデオ品質の意見を数値的にスコア化した値の平均をとることにより（即ち、ITU-R BT. 500 のように）、DMOS（差平均意見スコア）の如きビデオ品質のスコアを予測するとき、「Systems and Methods for Predicting Video Location of Attention Focus Probability Trajectories Due to Distractions」という名称の米国特許出願公開第 2 0 0 8 / 0 2 6 7 4 4 2 号明細書に記載のように、注意の焦点の正確な位置を予測することは、予測精度を改善する。かかる場合、統計的にもっともありそうな固定の軌跡をシミュレーションするために、焦点の小さな領域内での局部的色順応が適切であり、関連した局部的な順応が生じる。反対の例では、他の場合で、例えば、プロダクションが Blu-ray DVD を処理するとき、注意の焦点が全ての場所に等しい全てのスポットにおける任意の予測可能な色変化を評価することが望ましく、固定がランダムになる。この場合において、順応をシミュレーションするために、全体的な総計、近似平均、色を用いる。さらに、これら極端な例の間にある場合もあるので、最も望ましいのは、全ての場合に順応する連続的な制御である。

20

30

【 0 0 1 4 】

正式なフランス名が Commission internationale de l'éclairage であるため C I E と略称されるイルミネーションの国際委員会は、最初の C I E Luv が 1 9 7 6 年に導入以来、いくつかの改善された色見えモデル（C A M）を考案した。現在の最も正確な C A M は、参照文献として一体に考慮する C I E C A M 0 2 として知られている国際標準 C A M である。C I E C A M 0 2 開発チームの一部は、C I E C A M 0 2 に基づいた静止イメージ色見えモデル i C A M 及び i C A M 0 6 を開発した。これら両方は、イメージの空間的処理を含んでおり、後者は、高いダイナミックのイルミネーション・レンジを扱う余分な処理がある。依然として国際標準がないが、空間応答に影響する距離又は類似の要因を観察する際に、変化を考慮していない。一方、C I E C A M 0 2 は、そのオリジナルの用途を広く受け入れられることがわかった。これは、異なる照明環境において、主に、一致する塗料の如き反射光アプリケーション用である。反射光アプリケーションは、彩度及び輝度レベルにおける最大値を除いて直接光よりも小さな全域を示す。しかし、現在の放送ビデオ標

40

50

準と、最も商業的な表示技術の最大輝度出力とにとって、ビデオ全域は、CIECAM02ターゲットのほとんどと重なっている。CIECAM02は、色適合を考慮するが、適合のレベルの入力を必要とし、入力又は出力へのいかなる時間的次元も含んでいない。すなわち、静止周辺及び制御し照明を有する静止色パッチとして機能するようにCIECAM02が設計されている。

【0015】

高度に順応可能な移動画像色見えモデルは、主観的なビデオ品質の予測方法を改善して、似ていない表示、イメージ・サイズ、観察環境、フレーム・レート及びビデオ品質クラスの比較が可能となる。順応時空輝度応答とCIECAM02「静止パッチ」予測色見えモデルとの組合せは詳細であり、追加の技術は、新たな時間的周辺フィルタを含む色中心周辺時空フィルタを用いる。例示のモデル時間的周辺応答試験の結果は、人間の視覚の知覚応答と良好な比較を行う。

10

【0016】

図1は、本発明の実施例により、移動イメージ用に、人間の視覚系のマクロ行動シミュレーションによる色分析を実現するシステム100の機能ブロック図である。一般に、ビデオ信号をシステム100に供給する。このシステムは、神経系応答と呼ばれる値、又は一連の値を発生する。これら神経応答の試験及び基準の空間的及び時間的統計の間の差で変換し正規化したものは、顕著な差及び/又は知覚的に等しい増分の如き主観的分類知覚応答に反映する。これら主観的分類知覚応答は、試験ビデオ信号を評価する人が、オリジナルで基準ビデオ信号に関連したビデオ信号の品質をどのように評価するかに影響する。

20

【0017】

システム100は、CIECAM02処理に一般的に従い、少なくとも2チャンネル{a, b} (そして、実質的な利点が見えないが「直交」{ac, bc})の色応答を求める。標準処理と異なり、システム100は、表示モデルを含んでおり、移動イメージの色知覚を考慮して順応空間及び時間の処理に適切なCIECAM02が必要とする入力用の光単位に標準デジタル・ビデオを変換し、更に、認識のディスカウント及び順応局部化のレベルの制御を含んでいる。

【0018】

まず、ビデオ信号をシミュレーション表示102に供給する。この表示は、いくつかの実施例において、CRT、LCD、DLPなどをシミュレーションする表示の形式に基づいて、3刺激のCIE1031XYZ出力を発生する。シミュレーション表示102からのシミュレーションされた出力の一部は、「白色」成分値104を含んでいる。これは、標準白色値と周囲白色と組合せられる表示102からの白色値のファクタを含んでいる。白色成分値を、詳細に後述する時間的周辺フィルタ130に供給すると共に、詳細に後述する混合制御器120にも供給する。

30

【0019】

CIECAM02標準に基づくMAT02マトリクスの如き変換マトリクス106により、シミュレーション表示102の出力を、予め順応させたRGBコーン応答に変換する。これを次に空間中心フィルタ108に供給する。個別のコーン応答の各々が別々にろ波されるので、図1に示すシステム100には3個の空間的中心フィルタ108があり、各々が赤、緑及び青野コーン応答用である。

40

【0020】

空間的中心フィルタ108を設定して、2次線形(非適応)ミラーIIRフィルタ応答を発生する。実施例では、名称がAdaptive Spatio-Temporal Filter for Human Vision System Modelsである米国特許第6907143号の適応空間中心(sc)フィルタを用いて、適応制御を0に設定し、フィルタ係数を以下の表1に示すように設定する。良好な設定は、各フィルタ・パラメータ用の第1値に対応する。その範囲は、通常の視覚である人間の例として異なる合理的な範囲を反映する。

【0021】

表1

50

空間中心フィルタ：2次線形（非適応）ミラー I I R フィルタ：

@ 32サンプル/角度：

-校正された係数の範囲(3つ可能な設定)：

$b0rsc = b0gsc = 0.357071$, $b0bsc = 0.214243$

又は

$b0rsc = 0.357071$ 又は 0.321364 又は 0.321364

$b0gsc = 0.357071$ 又は 0.257091 又は 0.257091

$b0bsc = 0.214243$ 又は 0.205673 又は 0.143971

【 0 0 2 2 】

空間中心フィルタ 1 0 8 を通過したコーン応答を次に空間周辺フィルタ 1 1 0 への入力として用いる。また、3つの個別の空間周辺フィルタ 1 1 0 があり、各々が赤色、緑色及び青色の信号用である。上述の混合制御器 1 2 0 と同様に、空間中心フィルタ 1 0 8 からの出力を混合制御器 1 2 2 に供給する。これは、上述の混合制御器 1 2 0 と連携して、どの程度の重み係数を時間周辺フィルタ 1 3 0 の入力に供給するかを制御する能力をユーザに与える。これは、図 2 を参照して詳述する。

【 0 0 2 3 】

図 1 のシステム 1 0 0 において、人間の視覚系の各解剖学的成分の空間的及び時間的応答の組合せの近似としての統計的又は「集中」空間的及び / 又は時間的応答として、フィルタ 1 0 8、1 1 0、1 3 0 及び 1 5 0 を適用する。線形システムと異なり、フィルタが適用される 1 0 0 内の処理チェーン内の点が重要であることに留意されたい。線形システムにより、他の処理の前又は後に波を行って、システム応答内に影響させない。しかし、色見えの正確な予測を行うのに必要な C I E C A M 0 2 での非線形性のために、時空的フィルタの統計での空間的及び時間的な中心及び周辺成分は、好ましくは、1 0 0 内の重要な点に配置される。統計フィルタの各々を 1 0 0 内の処理フロー内に注意深く配置して、中間処理の修正の如き望ましくないアーティファクトを生成することから、その後の C I E C A M 0 2 処理の非線形性を最小にする。オプションとして処理フロー内に配置する他に、完全に透明な時空的応答のために、フィルタ係数（上述で参照した 1 4 3 特許の $b0$ ）を 1 . 0 に設定するならば、C I E C A M 0 2 の静的応答を変化なしに維持できる。

【 0 0 2 4 】

図 1 に示すシステムの主要フィルタ 1 0 8、1 1 0、1 3 0 からの出力は、白色順応回路 1 4 0 によりさらに変調され、各色チャンネル用の白色に順応したコーン応答を発生する。これらコーン応答は、最終ろ波ステップ用に時間的中心フィルタ 1 5 0 に供給される。いくつかの実施例において、空間的中心フィルタ 1 0 8 及び時間的中心フィルタ 1 5 0 を共通フィルタに組合せてもよい。1 6 0 に示すように、次に、システム 1 0 0 は、C I E C A M 0 2 処理の残りを適用して、所望の神経性応答を最終的に発生できる。

【 0 0 2 5 】

混合制御器 1 2 0、1 2 2 を戻り参照する。これら制御器により、組合せ出力信号としての 1 対の入力信号からどの程度の効果を提供できるかをユーザが正確に制御できる。例えば、周辺対表示白色点及び周囲（室内光）ファクタ（順応、認識のディスカウントなど）の影響を制御するために、混合回路 1 2 0 は、表示 1 0 2 からの白色成分値の入力と共に、空間周辺フィルタ 1 1 0 からの出力とを含んでいる。混合制御器 1 2 0 への「白色」入力は、組合せた表示白色及び周囲イルミネーション X Y Z 光組合せへのプリ順応 R G B コーン応答を表す。重み付け値「 $wgt2$ 」用の値を設定することにより、どの程度の空間的周辺フィルタ 1 1 0 からのろ波済み応答を混合制御器 1 2 2 に送るかをユーザが判断する。ユーザは、混合制御器 1 2 0 を設定して、白色値 1 0 4 を 1 0 0 % ドミナントとし、空間周辺フィルタ 1 1 0 の出力を 0 % ドミナントにできる。純粋な表示及び周囲白色点順応にとっては $wgt2=0$ であり、純粋な空間周辺（各チャンネルが Y_w に正規化）順応にとっては $wgt2=1$ である。他の設定により、ユーザは、順応重み付け値 $wgt2$ を選択して、任意の組合せ値を選択できる。

【 0 0 2 6 】

10

20

30

40

50

同様に、混合制御器 122 を基準として、局部的対全体的な色順応を制御するために、ユーザは、重み付け値 $wgt1$ を設定して、空間的中心フィルタ 108 の出力と混合制御器 120 からの以前に設定した混合出力からの出力との混合値を選択し、時間的周辺フィルタ 130 への制御された混合入力を生じさせる。空間的中心フィルタ出力 108 は、空間的周辺フィルタ 130 及び 2 つの他の入力の両方に繋がる。これら 2 つの他の入力は、(色順応に応答可能な) CIECAM02 の白色順応回路 140 部分へのプリ順応コーン応答入力と、混合回路 120 で重み付け値 $wgt1$ による重み付けメカニズムとであり、順応の局部的対全体的な空間的概念をどの程度用いるかを制御する。よって、 $wgt1$ は、空間の色順応の局所化を制御する。これが、停止した凝視対移動中の凝視の効果を反映する。局部的色順応(詳細な残像)を最小(無)にするためには $wgt1=0$ であり、最大のためには $wgt1=1$ である。 10

【0027】

光源、表示白色、周囲イルミネーションなどを補正する機能である認識のディスカウントは、色を識別するとき、図 1 に示されていないが、「 $wgt3$ 」と呼ばれるだい 3 の重み付けにより制御できる。この補正のために、 $wgt3=1$ と設定することにより、制御器 120 への「白色」{ Rwd 、 Gwd 、 Bwd } 入力は、 $Rw' = Gw' = Bw' = Yw$ (CIECAM02 での順応パイプライン $D = 0$ に等化) へのクロフフェードでもよい。ここで、 $Rw' = wgt3 * Yw + (1 - wgt3) * Rwd$ であり、 Gwd 及び Bwd も同様である。さらに、周囲白色と、図 1 に示していない追加の補正制御ファクタ $wgt4$ は、図 1 の表示モデル 102 からの「表示」白色信号との間の混合の重み付けを制御できる。よって、混合制御器 120 への「白色」信号入力は、周囲白色、表示白色、絶対即ち「標準」白色の混合でもよく、この混合は、重み付けファクタ $wgt3$ 及び $wgt4$ により制御される。 20

【0028】

CIECAM02 の従来用途は、 $wgt1=0$ 、 $wgt2=0$ 、周囲 = 0、表示白色 = 輝度、 $wgt3=0$ に対応する。

【0029】

時間的周辺フィルタ 130 は、いくつかの理由に対して有用である。まず、CIECAM02 は、順応のパーセントが与えられた「周辺」及び / 又は光源に関連した固定の順応を考慮するメカニズムを常に有する。部分的に順応を制御する CIECAM02 への 1 つの入力は、140 に入力する「順応白色」{ Rw 、 Gw 、 Bw } であり、これを用い、順応の瞬時レベルに応じてコーン応答を変調する。同様に、CIECAM02 と、他の参考文献でありここで参考文献として一体に考慮する 2007 年 1 月に Proceedings of the Third International Workshop Processing and Quality Metrics for Consumer Products で発行されたケビン・ファーガソンによる「An Adaptable Human Vision Model for Subjective Video Quality Rating Prediction Among CIF, SD, HD, AND E-CINEMA」とにおいて、周辺チャンネルを用いて、中心チャンネルを変調して、主要順応時空的応答を生じる。CIECAM02 静的応答を変化しないようにするために、図 1 のシステム 100 において、時間的処理を順応白色処理 140 に入力に適用して、全体的な周辺チャンネルとして扱う。さらに、上述のマトリクス 106 にてイメージ光刺激が 3 つのプリ順応 (RGB) コーン応答に変換された後に、時間的順応フィルタ・モデル(と共に全ての空間的及び時間的な波)が CIECAM02 に適用される点を想起されたい。 30 40

【0030】

上述の各チャンネルに対して、複合時間的周辺フィルタは、上述のフィルタを元にした適応積分器の平行な組合せである。

【0031】

図 2 は、更に詳細な説明をするための図 1 の例の時間的周辺フィルタ 130 (3 チャンネルの 1 つ) の機能ブロック図である。

【0032】

周辺フィルタ 200 は、色順応、順応光源及び認識ディスカウントを考慮するための (図 1 を参照して説明したような) 重み付け入力の和により生成された「コーン空間周辺混 50

合」入力を含んでおり、これは、(R、G及びBの各チャンネルに対する) 1組の重み付けされたコーン応答である。

【 0 0 3 3 】

「白色」Yw入力は、(C I E C A M 0 2による) 白色入力のY部分である。Ywは、表示モデル104 (図1) からの直接入力、又は、白色RGBからのYwに戻り変換されたものでもよく、そうでなければ、例えば、C I E C A M 0 2にてRw、Gw、Bwとして知られている。この「白色」Yw入力は、「直流利得」ファクタ、例えば0.48により重み付けされ、白色プロセッサ210にて3つの「コーン空間的周辺混合」から抽出される。この結果は、プリ時間的周辺る波された差、又は「コーン空間的周辺混合」応答の「交流」成分である。各チャンネルは、輝度尺度の「順応白色」からの各コーン応答での差を表す。

10

【 0 0 3 4 】

この応答を2つのロウパス・フィルタ (L P F) 220、230に供給する。これらは、IIRロウパス・フィルタに基づく適応積分器でもよい。これらフィルタ220、230を並列に用いて、追加の色信号をろ波する。L P F 220は、低速の第1 L P Fであり、上述で参照したファーガソンの文献に一般的に用いられているのと同じ又は類似の適応時間的周辺フィルタで実現してもよく、1秒当たり60サンプルの公称サンプル・レートに対して0.999516に設定された公称期間係数a1及びゼロの適合用の更新されたパラメータのみを有する。L P F 230は、L P F 220に類似した高速L P Fであるが、L P F 230は、その入力が出力よりも小さいか大きいかに応じる可変係数を含んでいる。比較器240は、マルチプレクサ250により選択する係数を決定する。量 (L P F 230の入力マイナスL P F 230の出力) が正ならばalp=0.9854であり、負ならばalp=0.99370である。本発明の要旨を逸脱することなく、他の係数を用いてもよい。

20

【 0 0 3 5 】

図2に示すように、tfw1の値を選択することにより、重み付け平均値を用いて、コンバイナ260にてL P F 220及びL P F 230の出力を組合せる。いくつかの実施例において、tfw1を約0.5に等しくできる。組合せた出力を生成し、その結果が複合る波済み差信号である。上述の白色プロセッサ210が生成した輝度縮尺した白色輝度信号212を加減することにより、この結果の複合る波済み差信号を絶対信号に再生する。

30

【 0 0 3 6 】

時間的周辺フィルタを生成するのに用いるいくつかの成分の設計詳細は、上述のファーガソンの参考文献や、ここでの参考文献として一体に考慮するファーガソンの米国特許第6976045号及び第6907143号に記載されている。

【 0 0 3 7 】

L P F 220及びL P F 230のフィルタの詳細な例を表2に示す。

【 0 0 3 8 】

表2

L P F 1: 時間的周辺とする米国特許第6907143号の適応輝度フィルタ

順応パラメータ = 0 及び

$$a1 = 0.999516, b0 = 1 - a1$$

60サンプル/秒のサンプル・レートに対して

40

L P F 2: 時間的周辺とする米国特許第6907143号の適応輝度フィルタ

順応パラメータ = 0 及び

L P F 2出力 > L P F 2入力ならa1 = a1n、そうでなければa1 = a1p:

$$a1n = 0.99370$$

$$a1p = 0.9854$$

60サンプル/秒のサンプル・レートに対して

【 0 0 3 9 】

特に入力用のL P F 220のディラック・デルタ・インパルス応答を図3に示す。2つの曲線310及び320は、1秒当たり10及び60のサンプルを夫々表す。これは、2

50

つの異なるフレーム・レートでのフィルタ応答を維持する例を示し、上述のファーガソンの参考文献の適応積分器を用いることにより、時間的サンプル・レートの順応要求が合う。

【0040】

主に、LPF230により考慮された正対負の変化の応答の間の差により、空間的周辺フィルタ200用の全体的な応答が非線形である。これは、図4A及び4Bにてわかる。図4Aは、実線及び点線で夫々示す10及び60サンプル/秒の2つのモデル・シミュレーション・サンプル・レートにおける正方向の応答変化を示し、図4Bは、負方向の応答変化を示す。これら応答は、互いを反映しておらず、色が白色、灰色又は黒色（無色）より大きい又は小さい変化をするかに応じた異なる人間の応答があることを意味する点に留意されたい。後述の如く、応答におけるこの変化の例は、周辺での突然の時間的変化と、無色の応答に対する刺激の変化の追跡を含む。

10

【0041】

[試験及び確認]

中心及び周辺フィルタの校正と、2度パッチ用のCIECAM02の適応応答の回帰試験のために、視覚科学研究での実験をシミュレーションにて反復して、適切な応答をチェックする。各データ設定は、空間的及び時間的周波数並びに期待応答に対応する刺激形式により大雑把に分類される。なお、期待応答は、近似知覚しきい値、又はほぼ一定の（大きさが等しい）しきい値上応答である。

【0042】

フィルタを追加した後のCIECAM02確認及び回帰試験の両方に対して、種々の異なる既知のデータ設定を用いた。時間的中心フィルタ校正と共に空間的中心フィルタ校正用に、特定の既知のデータ設定を用いた。試験結果は、時空的組合せと時間的周辺も確認した。空間的周辺データは、最低周波数における既知のデータ設定の組合せから収集した一部である。統計において、これら刺激も角度及び時間的範囲をサンプルした。

20

【0043】

収集された時間的周辺フィルタ・データで、1つの実験は、人間の色知覚の長期間の時間的応答の詳細を示すのに特に有用であった。フェアチャイルド及びレニフは、以前、実験を行って、空間的周辺の変化に対する時間的ステップ応答を追跡した。この実験において、主題は、周辺での色の時間的ステップ変化に追従する中心での静止ターゲット・パッチの無色（灰色）見えのタスクであった。よって、入力ターゲット色（CIE1931xY、XYZなどでの定義されるように）が時間に伴って著しく変化しても、この実験の刺激による結果は、本発明を用いて無色の応答である。

30

【0044】

フェアチャイルド及びレニフ試験での無色の応答を維持するのに要求される色平面（CIExy又はCIECAM02abなど）での入力試験ターゲット色での変化の方向は、周辺での変化の方向に依存する。3つの方向（各コーン応答のピークを近似する赤色、緑色、青色）及びこれらのCIEXYZの反対を、全体では6つの周辺色ステップに対して用いられた。よって、この実験で、6つの時間過程（時間的軌道）が得られた。

【0045】

本発明の実施例により、CIECAM02及び時間的周辺フィルタを用いて、6つの刺激をシミュレーションした。例としての時間的クロマ応答を図5に示すが、これは、CIECAM02{a,b}大きさの等価値対時間（時間周辺フィルタを用いた及び用いない場合）を示す。図5は、フィルタに対する青色から黄色の周辺ステップ応答を示しており、フィルタがあるとき（低振幅の青色曲線）とないとき（高振幅の赤色曲線）との時間的周辺サブJND最大差がある。公称しきい値及び「ガードバンド・マーカ」（しきい値プラス1標準偏差）を波線及び1点鎖線で夫々示す。図5において、波線510は、無色（色なし）からの公称色差しきい値（ ~ 1 JND）を示し、1点鎖線520は、しきい値プラス1標準偏差を示す。図5に示すピーク・クロマ応答は、6つの周辺ステップの全ての色相クロマ・プロットを示す図6A及び6Bの最大軌跡に対応する。図6Aは、CIE1

40

50

9 3 1 { x、y } 平面でのステップ刺激を示し、図 6 B は、C I E C A M 0 2 { 彩度、色相 } 平面を示す。原点は、期待される無色（色なし）応答を表す。図 6 B のダイヤモンドは、いかなる時間的処理もない C I E C A M 0 2 応答を表し、X は、時間周辺処理が追加されたときの応答を示す。この場合（図 5 及び 6 を含む）、時間的中心フィルタが除去されると、初期の過渡状態が X のいくつかをその原点からいくらかそらす。これは、図 5 の応答の短いピークでも判る。時間的中心フィルタを追加することで、これらの過度状態を緩和する。これら最大軌跡が周辺ステップ遷移の瞬間に生じること留意されたい。よって、本発明の実施例により時間的中心フィルタを追加することにより、無色応答からのこれらの偏差が大幅に緩和される。さらに、上述の空間フィルタは、6 つのクロマ応答の全てをわずかに低減させる。

10

【 0 0 4 6 】

特定の実施例について上述したが、本発明の原理はこれらの実施例に限定されないことは明らかである。以下の請求項に記載の本発明の原理を逸脱することなく、変形変更が可能である。

【 符号の説明 】

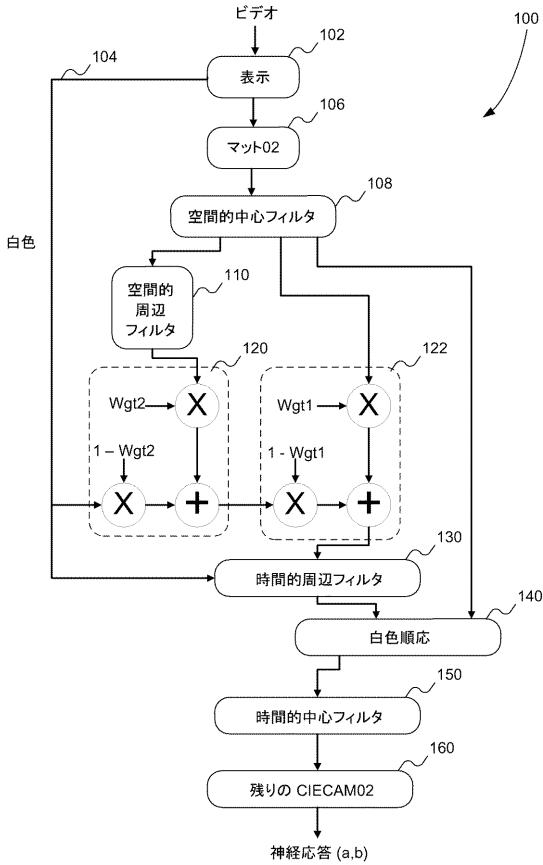
【 0 0 4 7 】

- 1 0 2 表示
- 1 0 6 マトリクス
- 1 0 8 空間中心フィルタ
- 1 1 0 空間周辺フィルタ
- 1 2 0 混合制御器
- 1 2 2 混合制御器
- 1 3 0 時間的周辺フィルタ
- 1 4 0 白色順応
- 1 5 0 時間的中心フィルタ
- 1 6 0 残りの C I E C A M 0 2
- 2 1 0 白色プロセッサ
- 2 2 0 ロウパス・フィルタ 1
- 2 3 0 ロウパス・フィルタ 2
- 2 4 0 比較器
- 2 5 0 マルチプレクサ
- 2 6 0 コンバイナ

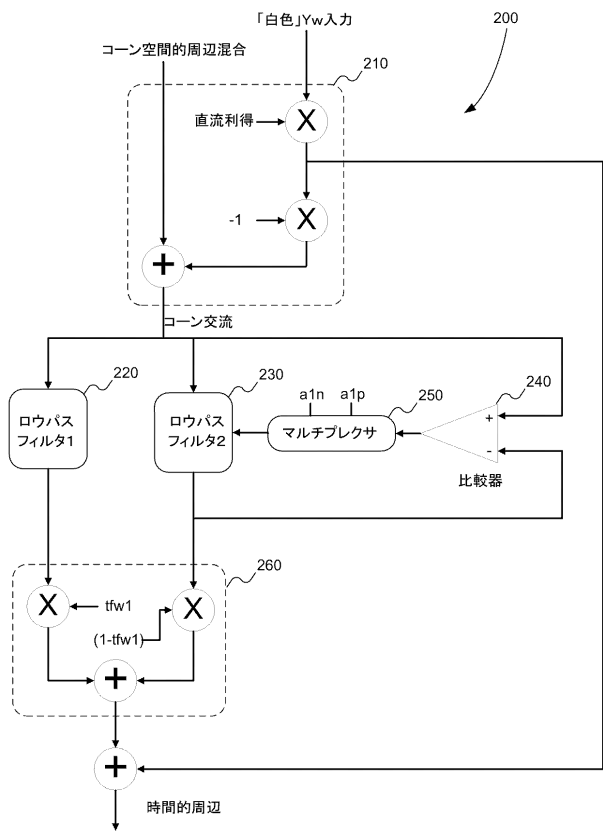
20

30

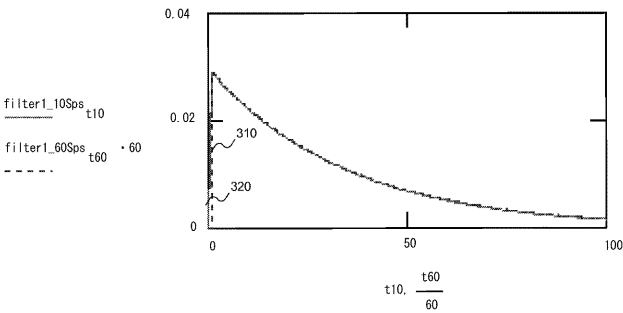
【 図 1 】



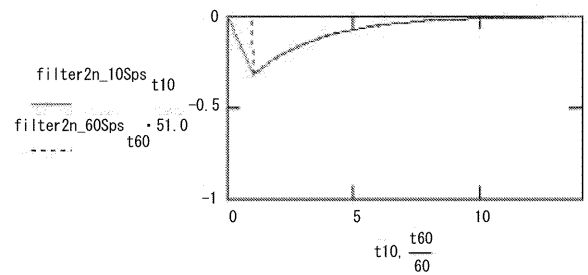
【 図 2 】



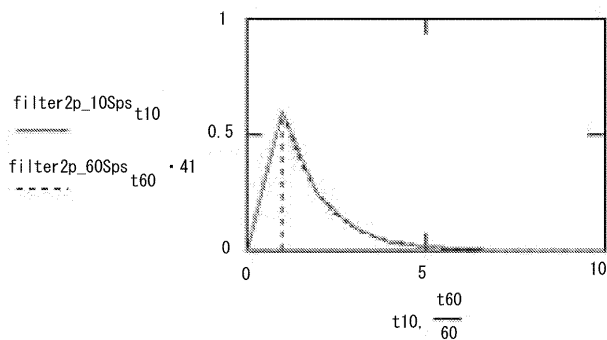
【 図 3 】



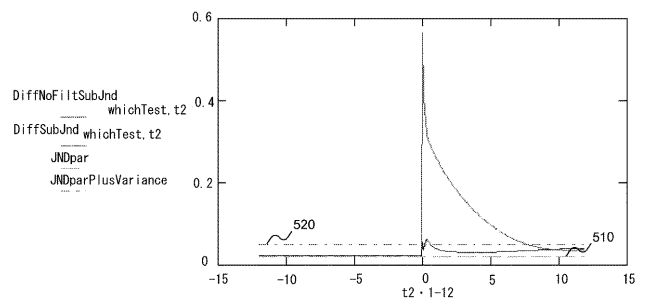
【 図 4 B 】



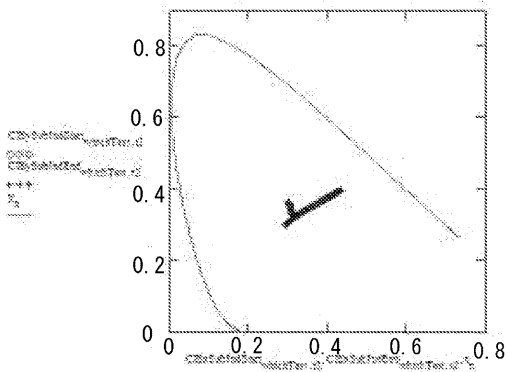
【 図 4 A 】



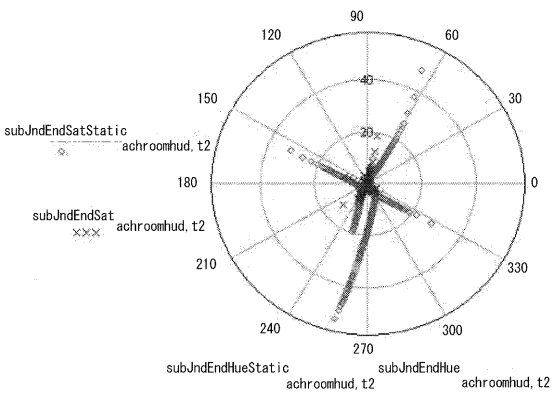
【 図 5 】



【 図 6 A 】



【 図 6 B 】



【 外国語明細書 】**【 Title of the Invention 】**

A Method of Predicting Color Visual Differences in Video and A System for Moving Image Color Appearance Modeling

【 Field of the Invention 】**【 0 0 0 1 】**

The present invention is related to video test and measurement equipment for picture quality measurements for video, and more particularly to a method of predicting color visual differences in video and a system for moving image color appearance modeling.

【 Background of the Art 】**【 0 0 0 2 】**

When video or image stream is manipulated, for example when a video is compressed so that it may be stored on a DVD or transmitted over a television channel, or when video is transcoded from one format to another, the manipulations introduce artifacts from the compression or conversion into the resultant video. Sometimes these artifacts are invisible or imperceptible to human viewers, due to physiology and phenomena of the human eye and the way it perceives light and motion. Other times the artifacts are plainly visible. It is nearly impossible to determine how manipulated video will be perceived by simply analyzing the resultant video data.

【 0 0 0 3 】

Instead, video quality is best determined by a survey of a statistically significant number of people gathered to grade the quality of the video. Audiences watch the video and grade it on a predetermined quality scale. Such audience surveys, however, are impractical and cost prohibitive for most applications, and thus methods of automating picture quality rating were developed.

【 0 0 0 4 】

Full reference (FR) picture quality analyzers are machines that compare two images, or two image streams, and generate a “score” or other measure of how humans would perceive the quality of the video conversion, or the match between the original and the modified test streams. A high score indicates a high-quality conversion, whereas a low score indicates that the resultant video is a poor representation of its original.

【 0 0 0 5 】

A disconnect exists, though, in that the state of the art picture quality analyzers either 1) do not include color analysis in their calculations on video quality or 2) include simple objective measures such as color PSNR (peak signal to noise ratio) or color impairment measurements such as blocking or ringing, but do not take into account the human vision system's adaptation

mechanisms causing drastic changes in sensitivity to color changes depending on spatiotemporal contexts.

【 0 0 0 6 】

An example picture quality measurement device that does not include color analysis in their calculations on video quality is the Tektronix PQA500. Analysis is made on luminance (light intensity) values only. Although there have been published proposals of adding filters to color appearance models, they so far have been fixed models (such as static Gaussian spatial filters). In the luminance perception model of the PQA500, adaptive response is relative to an integrated adaptation point represented as the output of a "surround" spatiotemporal filter with upper resolution in time and space set by a "center" spatiotemporal filter. This luminance only method, being state of the art for detecting perceptible errors in intensity, has been quite sufficient for most video transmission system assessment applications, mostly due to the tendency for luminance error and chroma error to correlate well. However, increasingly complex systems and flexibility in encoders and other video processing devices has lead to increased chances of impairments primarily seen as color errors. For example, in video broadcast and webcast, oftentimes colorimetry of input video and one or more processing or display components is mismatched; out of gamut colors may be clipped, and in some cases out of gamut colors may overflow into an opposite color. Luminance analysis fails to detect the majority of these errors, and consequently the quality rating becomes inaccurate.

【Document of Prior Art】

【Patent Document】

【 0 0 0 7 】

【Patent Document 1】 Patent Publication No. P2001-197527A

【Patent Document 2】 Patent Publication No. P2003-199127A

【Summary of the Invention】

【Problem to be Solved by the Invention】

【 0 0 0 8 】

Embodiments of the invention address these and other limitations in the prior art.

【Means for Solving the Problem】

【 0 0 0 9 】

In order to solve the above problems, the present invention is configured as follows.

- (1) A method of predicting color visual differences in video, comprising: accepting a video signal to be analyzed; receiving an unfiltered white signal corresponding to the video display and viewing environment; creating a spatial surround from the video signal; creating a spatial center from the video signal; and creating a temporal surround of the video signal from a combination of the spatial surround, the spatial center, and the unfiltered white signal using one or more weighting factors.
- (2) The method of aspect 1, in which the unfiltered white signal is a combination of one or more of an

ambient white signal, a display white signal, and an absolute white signal.

- (3) The method of aspect 2, further comprising: combining the ambient white signal and the display white signal according to a first combinational weighting factor to create an intermediate white signal; and creating the unfiltered white signal by combining the intermediate white signal and the absolute white signal according to a second combinational weighting factor.
- (4) The method of aspect 1 in which creating a temporal surround comprises temporally filtering data through a variable rate filter.
- (5) The method of aspect 1 in which creating a temporal surround comprises temporally filtering data through an adaptable filter.
- (6) The method of aspect 1 in which creating a temporal surround comprises temporally filtering data through an asymmetric filter.
- (7) The method of aspect 1 in which creating a temporal surround comprises temporally filtering data through a recursive filter.
- (8) The method of aspect 6, in which temporally processing the stream further comprises: filtering data through a first and a second low pass filter.
- (9) The method of aspect 7, in which at least one of a set of operating parameters of the second low pass filter varies depending on a color movement direction.
- (10) The method of aspect 8, further comprising adding outputs of the first and second low pass filters to produce the temporal surround.
- (11) A system for moving image color appearance modeling, comprising: an input structured to accept a video signal; a white generator structured to create an unfiltered white signal from the video signal; a cone response generator coupled to the input and structured to generate a set of cone responses from the video signal; a spatial center filter structured to generate a spatial center of one of the set of cone responses; a spatial surround filter structured to generate a spatial surround of the one of the set of cone responses; and a temporal surround filter structured to generate a temporal surround from the spatial center, spatial surround, and the unfiltered white signal.
- (12) The system of aspect 11, further comprising: a first mix controller structured to combine the unfiltered white signal with the spatial surround according to a first weighting factor.
- (13) The system of aspect 11, further comprising: a second mix controller structured to combine an output from the first mix controller with the spatial center according to a second weighting factor.
- (14) The system of aspect 11, in which the unfiltered white signal comprises a combination of an ambient white signal, a display white signal, and an absolute white signal.
- (15) The system of aspect 11, in which the temporal surround filter further comprises a first low pass filter, a second low pass filter, and a comparator coupled to an input and an output of the second low pass filter.
- (16) The system of aspect 11, further comprising a set of spatial center filters, a set of spatial surround filters, and a set of temporal surround filters for each of the set of cone responses.
- (17) The system of aspect 11 in which the temporal surround filter is a variable rate filter.
- (18) The system of aspect 11 in which the temporal surround filter is an adaptable filter.
- (19) The system of aspect 11 in which the temporal surround filter is an asymmetric filter.
- (20) The system of aspect 10 in which the temporal surround filter is a recursive filter.

【Simple Explanation of Drawings】

【 0 0 1 0 】

【Figure 1】 is a functional block diagram of a system that implements a color analysis for moving images according to embodiments of the invention.

【Figure 2】 is a functional block diagram of a temporal surround processing component of the system of Figure 1 according to embodiments of the invention.

【Figure 3】 is a graph that illustrates a Dirac delta impulse response of one of the filters of the temporal surround processing component illustrated in Figure 2.

【Figure 4A】 is a graph that illustrates the positive Dirac delta impulse response for one of the filters of the temporal surround filter illustrated in Figure 2.

【Figure 4B】 is a graph that illustrates the negative Dirac delta impulse response for the same filter for which the positive response was illustrated in Figure 4A.

【Figure 5】 is a graph illustrating responses using the temporal processing component of Figure 2.

【Figure 6A】 is a plot illustrating test stimuli in an CIE 1931 {x,y} plane.

【Figure 6B】 is a plot illustrating responses with and without processing using the temporal filter illustrated in Figure 2.

【Configuration for Implementing the Invention】

【 0 0 1 1 】

Color perception adaptation in human eyes causes a given constant physical light stimulus to appear to have a different color when the surround or environment changes, or when viewed after seeing other colors. In general, dynamic light stimuli generally causes perceptions of given static light stimuli to appear to change due to changes in space and/or time. This apparent change of color appearance due to adaptation has a time course that depends on a number of stimulus conditions.

【 0 0 1 2 】

Cognitive discounting of illumination and white point allows humans to view an apple against a white background illuminated by red light only and still determine if the apple is red or green and if the background is red or white. Knowing that the white background is white, even though it appears as red as the red illuminant, is an example of cognitive discounting. In some applications, it is desirable to discount this type of overall color shift (towards the illuminant color in this example), while in other cases discounting is undesirable. It is commonly desirable for white balance matching between scenes in a movie or TV show, so differences in colors due to shifts of the sun's position in the sky or inadvertent changes in studio lighting are undesirable. On the other hand, if it is desired to assess the perceptibility of codec related impairments on a consumer LCD TV with standard colorimetry versus a computer monitor with a VESA standard colorimetry (bluer white point) displayed via an internet browser, it is undesirable to include

these differences in color standard in the picture quality measurement. Therefore, in order to cover these two cases and others in between when making picture quality measurements, it is important to have continuous control cognitive discounting from none to full.

【 0 0 1 3 】

Depending on the agility of one's gaze, perceptual color adaptation can become very localized, for example if one stares at a fixed point for a relatively long time the sensation of color is reduced. Alternatively, color adaptation can become primarily global based on the biases such as the non-white illumination or other color bias from the image. Localized color adaptation effects from long gazes at fixed points can cause color after-images that in effect are a result of desensitization of the eye to certain colors. In some applications the prediction of these localized adaptation affects, which can be most dramatic, are important for predicting video quality. In other applications, perhaps only the global effect is of interest. For example, when predicting video quality scores such as DMOS (difference mean opinion score) made by taking the mean numerically scored video quality opinions of approximately two dozen individuals (i.e. as per ITU-R BT.500), it has been shown that predicting the precise locations of focus of attention improves prediction accuracy, such as described in US Publication 2008/0267442 entitled Systems and Methods for Predicting Video Location of Attention Focus Probability Trajectories Due to Distractions. In such a case, localized color adaptation in smaller areas of focus would be appropriate in order to simulate the most probable fixation trajectories in aggregate, and the associated local adaptation that occurs. For a counter example, in other cases, for example when a production house is mastering a Blu-ray DVD, it may be desired to assess any possible perceptible color change in every spot as if focus of attention is equal everywhere, and thus fixation is randomized. In this case, only the overall aggregate, approximately average, color would be used for simulating adaptation. And, as there are cases in between these extreme examples, it is most preferable to have a continuous control to accommodate all cases.

【 0 0 1 4 】

The International Commission on Illumination, abbreviated as CIE because of its official French name of Commission internationale de l'éclairage, has created several improved color appearance models (CAMs) since the original CIE Luv was introduced in 1976. The current, and most accurate CAM, is the international standard CAM known as CIECAM02, which is incorporated by reference herein. A subset of the CIECAM02 development team has developed still image color appearance models based on CIECAM02: iCAM and iCAM06, both of which include spatial processing for images, the latter with extra processing to handle high dynamic

luminance range. These are not yet international standards, however, and do not account for changes in viewing distance or similar factors affecting spatial response. Meanwhile, CIECAM02 has found wide acceptance for its original use, which was primarily for reflected light applications such as matching paints in different lighting environments. The reflected light application represents a smaller gamut than direct light, excluding the extremes in saturation and luminance levels. For current broadcast video standards and the maximum luminance output of most commercial display technologies, however, the video gamut mostly overlaps the CIECAM02 target. While CIECAM02 takes into account color adaptation, it requires an input of the level of adaptation and does not include any temporal dimension to the input or output. In other words, CIECAM02 is designed to work with static color patches having static surround and static lighting.

【 0 0 1 5 】

A highly adaptable moving picture color appearance model improves methods of predicting subjective video quality, allowing the comparison of dissimilar displays, image sizes, viewing environments, frame rates and video quality classes. The combination of adaptable spatiotemporal luminance response and the CIECAM02 "static patch" predicting color appearance model is detailed, with additional techniques using color center-surround spatiotemporal filtering, including a new temporal surround filter. Example model temporal surround response test results compare well with human vision perceptual response.

【 0 0 1 6 】

Figure 1. is a functional block diagram of a system 100 that implements a color analysis through macro-behavioral simulation of the human vision system for moving images according to embodiments of the invention. In general, a video signal is applied to the system 100 which generates a value, or series of values called a Neural Response. The transformed, normalized differences between the test and reference spatial and temporal aggregates of these neural responses reflect Subjectively Classified Perceptual Responses such as just noticeable differences and/or perceptually equal increments. These Subjectively Classified Perceptual Responses reflect how a person rating the test video signal would rate the quality of the video signal relative to an original, reference video signal.

【 0 0 1 7 】

The system 100 generally follows a CIECAM02 process to obtain at least two channels {a,b} (& "orthogonal" {ac,bc}), though no net advantage was seen) color response. Different from the

standard process, the system 100 includes a display model to convert standard digital video into light units for input as required by CIECAM02, appropriate adaptive spatial and temporal processing to account for color perception of moving images, and further includes controls for levels of cognitive discounting and color adaptation localization.

【 0 0 1 8 】

Initially, a video signal is applied to a simulated display 102, which, in some embodiments, produces a tristimulus CIE1931 XYZ output based on the type of display it is simulating, such as CRT, LCD, DLP, etc. A portion of simulated output from the simulated display 102 includes a “white” component value 104, which includes factors for the white values from the display 102 combined with a standard white value as well as ambient white. The white component value is provided to a temporal surround filter 130, described in detail below, as well as to a mix controller 120, also described below.

【 0 0 1 9 】

The output from simulated display 102 is converted through a conversion matrix 106, such as an MAT02 matrix as per the CIECAM02 standard, into pre-adapted RGB cone responses, which are then fed to a spatial center filter 108. Because each of the separate cone responses is separately filtered, there are three spatial center filters 108 in the system 100 depicted in Figure 1, one each for the Red, Green, and Blue cone responses.

【 0 0 2 0 】

The spatial center filters 108 may be set to produce a 2nd order linear (non-adaptive) mirrored IIR filter response. An example implementation uses the adaptive spatial center (sc) filter of US Patent 6,907,143, entitled Adaptive Spatio-Temporal Filter for Human Vision System Models with adaptation control set to 0 and filter coefficients set as illustrated below in Table 1. The preferred set corresponds to the first value for each filter parameter. The range is given to reflect different reasonable values for examples of humans with normal vision.

【 0 0 2 1 】

TABLE 1

Spatial Center Filter: 2nd order linear (non-adaptive) mirrored IIR filter:

@ 32 Samples/deg:

Range of ~calibrated coefficients (3 possible sets):

b0rsc = b0gsc = 0.357071, b0bsc = 0.214243

OR

b0rsc = 0.357071 OR 0.321364 OR 0.321364

b0gsc = 0.357071 OR 0.257091 OR 0.257091

$$b0b_{sc} = 0.214243 \text{ OR } 0.205673 \text{ OR } 0.143971$$

【 0 0 2 2 】

The cone responses passed through the spatial center filter 108 are then used as inputs to a spatial surround filter 110. Again, there are three separate spatial surround filters 110, one each for the Red, Green and Blue signals. Similar to the mix controller 120, mentioned above, outputs from the spatial center filters 108 are provided to a mix controller 122, which, in conjunction with the previously mentioned mix controller 120, provide a user with the ability to control how much weighting effect is provided to the inputs of the temporal surround filter 130, which is described in detail with reference to Figure 2.

【 0 0 2 3 】

In the system 100 of Figure 1, the filters 108, 110, 130 and 150 are applied as aggregate or "lumped" spatial and/or temporal response as an approximation of the combination of spatial and temporal response of each anatomical component of the human vision system. Note that unlike linear systems, the point in the processing chain in 100 where filters are applied is important. With linear systems, filtering may be placed before or after the other processes with no effect in system response. However, because of nonlinearities in CIECAM02 that are required to produce accurate predictions of color appearance, the aggregate spatial and temporal center and surround components of the spatiotemporal filtering are preferably placed at key points within 100. Each of the aggregate filters have been carefully placed in the process flow in 100 to minimize non-linearities of subsequent CIECAM02 processing from creating unwanted artifacts such as rectification of intermediate responses. In addition to being optimally placed in the process flow, the static response of CIECAM02 can remain unchanged if the filter coefficients (b_0 's of the '143 patent reference above) are set to 1.0 for completely transparent spatiotemporal response.

【 0 0 2 4 】

Output from the main filters 108, 110, 130 in the system illustrated in Figure 1 are further modified by a white adaptation circuit 140, which produces a white adapted cone response for each color channel. These cone responses are then fed to a temporal center filter 150 for a final filtering step. In some embodiments the spatial center filter 108 and the temporal center filter 150 may be combined into a common filter. As illustrated by 160, the system 100 may then apply the remainder of the CIECAM02 processes to finally produce the desired Neural Response.

【 0 0 2 5 】

With reference back to the mix controllers 120, 122, these controllers allow a user to precisely control how much effect to apply from a pair of input signals as a combined output signal. For instance, to control effects of the surround vs. display white point and ambient (room light) factors (adaptation, cognitive discounting, etc.), the mix controller 120 includes inputs of the white component values 104 from the display 102, as well as an output from the spatial surround filter 110. The "White" input to the mix controller 120 represents the pre-adapted RGB cone responses to the combined display white and ambient illumination XYZ light combination. By setting the values for a weighting value "wgt2," the user determines how much of a filtered response from the spatial surround filter 110 is sent to the mix controller 122. The user may set the mix controller 120 such that the white value 104 is 100% dominant and the output from the spatial surround filter 110 is 0% dominant. For pure display and ambient white point adaptation $wgt2 = 0$, while for pure spatial surround (each channel normalized to Y_w) adaptation, $wgt2=1$. Other settings allow the user to select any combination value, simply by selecting the appropriate weight value $wgt2$.

【 0 0 2 6 】

Similarly, with reference to the mix controller 122, to control local vs. global color adaptation, the user may set the weighting value $wgt1$ to select the mix value of the output from the spatial center filter 108 and the output from the previously set mix output from the mix controller 120, to produce a controlled mixed input to the temporal surround filter 130. The spatial center filter output 108 connects to both the spatial surround filter 130 and two other inputs: the pre-adapted cone response input to the white adaptation circuit 140 portion of CIECAM02 (responsible for color adaptation), and the weighting mechanism, via the weighting value $wgt1$, in the mix controller 120, to control how much of local vs. global spatial aspect of the adaptation is used. Thus $wgt1$ controls localization of spatial color adaptation. This reflects effects of fixated vs. roaming gaze. For minimum (no) local color adaptation (detailed after-images), $wgt1 = 0$, while for maximum $wgt1 = 1$.

【 0 0 2 7 】

Cognitive discounting, which is the ability to compensate for the illuminant, display white, ambient illumination, etc. when identifying colors, may be controlled by a third weighting, termed "wgt3" but not illustrated in Figure 1. For this compensation, the "white" $\{R_{wd}, G_{wd}, B_{wd}\}$ input to the controller 120 may be cross-faded to $R_w' = G_w' = B_w' = Y_w$ (equivalent to adaptation parameter $D=0$ in CIECAM02) by setting $wgt3 = 1$, where $R_w' = wgt3 * Y_w + (1 - wgt3) * R_{wd}$, and likewise for G_w' and B_w' . Further, an additional

compensation control factor, wgt4, also not illustrated in Figure 1, may control the weighting of a blend between the ambient white with the “display” white signal from the display model 102 of Figure 1. Thus, the “white” signal input to the mix controller 120 may be a blend of ambient white, display white, and absolute or “standard” white, the mix of which is controlled by weighting factors wgt3 and wgt4.

【 0 0 2 8 】

Conventional use of CIECAM02 would correspond to wgt1=0, wgt2=0, ambient=0, display white = illuminant and wgt3=0.

【 0 0 2 9 】

The temporal surround filter 130 is useful for several reasons. First, CIECAM02 already has a mechanism to take into account fixed adaptation relative to "surround" and/or the illuminant, given the percentage of adaptation. One input to CIECAM02 that in part controls adaptation is "adaptation white" {Rw,Gw,Bw} input to 140, used to modify the cone responses according to instantaneous level of adaptation. Likewise in both CIECAM02 and in another reference entitled “An Adaptable Human Vision Model for Subjective Video Quality Rating Prediction Among CIF, SD, HD, AND E-CINEMA,” by Kevin Ferguson, published in Proceedings of the Third International Workshop on Video Processing and Quality Metrics for Consumer Products, January, 2007, and incorporated by reference herein, the surround channel is used to modify the center channel to produce the primary adapted spatiotemporal response. So as to leave CIECAM02 static response unchanged, in the system 100 of Figure 1, the temporal processing is applied to the input for adaptation white processing 140, and thus is treated as the overall surround channel. Recall, too, that the temporal adaptation filter model (as well as all spatial and temporal filtering) is applied to CIECAM02 after image light stimulus is converted into the three pre-adapted (RGB) cone responses in the matrix 106 above.

【 0 0 3 0 】

For each channel, described above, the composite temporal surround filter is a parallel combination of the adaptive integrator based filters mentioned above.

【 0 0 3 1 】

Figure 2 is a functional block diagram of an example (for one of the three channels of the) temporal surround filter 130 of Figure1 giving the description in further detail.

【 0 0 3 2 】

A surround filter 200 includes a "Cone Spatial Surround Mix" input, which was created by the sums of weighted inputs (as described with reference to Figure 1) to account for localization of color adaptation, the adaptation illuminant and cognitive discounting, is a set of weighted cone responses (a channel each of R,G and B).

【 0 0 3 3 】

A "White" Yw input is the Y portion of the white input (as per CIECAM02). Yw may be either directly input from the display model 104 (Figure 1) or may be converted back to Yw from the white RGB, otherwise known for example in CIECAM02 as Rw,Gw,Bw. This "White" Yw input is weighted by a "DC gain" factor, for example 0.48, and subtracted from each of the three "Cone Spatial Surround Mix" response inputs in a white processor 210. The result is pre-temporal surround filtered differential or "AC" components of the "Cone Spatial Surround Mix" response. Each channel represents the difference in respective cone response from intensity scaled "adaptation white."

【 0 0 3 4 】

This response is fed to two low-pass filters (LPF), 220, 230, which may be adaptive integrator based IIR low-pass filters. These filters 220, 230 are used in parallel to filter the differential color signals. The LPF 220 is a slow first LPF, which may be an instantiation of the same or similar adaptive temporal surround filter generally used in the Ferguson reference incorporated above, only having updated parameters for zero adaptation and nominal feedback coefficient a_{1p} set to 0.999516 for a nominal sample rate of 60 samples per second. The LPF 230 is a faster LPF, similar to the LPF 220, except that the LPF 230 includes variable coefficients depending on whether its input is less than or greater than its output. A comparator 240 determines which coefficient to select through a multiplexer 250: $a_{1p} = 0.9854$, if the quantity (LPF 230 input minus LPF 230 output) is positive, or $a_{1n} = 0.99370$ if negative. Other coefficients may be used without deviating from the scope of the invention.

【 0 0 3 5 】

The outputs from the LPF 220 and LPF 230 are combined in a combiner 260 using a weighted average by selecting a value for $tfw1$, as illustrated in Figure 2. In some embodiments $tfw1$ can be approximately equal to 0.5. The combined output is created, resulting in the composite filtered differential signal. This resulting composite filtered differential signal is restored to an absolute signal by adding back the intensity scaled white luminance signal 212, which was created by the white processor 210 described above.

【 0 0 3 6 】

Design details of some of the components that may be used to create the temporal surround filter may be found in the Ferguson reference incorporated above, as well as in US Patents 6,976,045 and US 6,907,143 to Ferguson, both of which are incorporated herein by reference.

【 0 0 3 7 】

Example filter details for LPF 220 and LPF 230 are given in Table 2.

【 0 0 3 8 】

TABLE 2

LPF1: Adaptive Lum Filter of US 6,907,143, referred to as temporal surround:
with adaptation parameters = 0 and

$$a1 = 0.999516, b0 = 1-a1$$

for sample rate of 60Samples/second.

LPF2: Adaptive Lum Filter of US 6,907,143, referred to as temporal surround:
with adaptation parameters = 0 and

$$a1 = a1n \text{ if LPF2 output} > \text{LPF2 input, else } a1 = a1p:$$

$$a1n = 0.99370$$

$$a1p = 0.9854$$

for sample rate of 60Samples/second.

【 0 0 3 9 】

The Dirac delta impulse response of LPF 220 for particular inputs is shown in Figure 3. The two curves 310 and 320 represent 10 and 60 samples per second, respectively. This illustrates an example of maintaining filter response at two different frame rates, the temporal sample rate adaptation requirement met by using the adaptive integrator of the above-incorporated Ferguson reference.

【 0 0 4 0 】

The overall response for the temporal surround filter 200 is nonlinear primarily due to the difference between responses to positive vs. negative changes accounted for by LPF 230. This can be seen in Figures 4A and 4B. Figure 4A indicates a positive direction response change in two model simulation sample rates, 10 and 60 samples/second shown in solid and dashed lines, respectively, while Figure 4B illustrates a negative direction response change. Note that the responses are not reflections of one another, which means that there is a different human response depending on whether a color is changing to be more or less white, gray or black (achromatic). An example of this difference in response involves abrupt temporal changes in surround and tracking stimulus changes for achromatic response, as described below.

【 0 0 4 1 】

[Tests and verification]

For purposes of calibrating center and surround filters as well as regression testing of the adapted response of CIECAM02 for 2 degree patches, experiments conducted in vision science research were replicated in simulation to check for appropriate responses. Each data set was roughly categorized by stimuli type corresponding to spatial and temporal frequencies and expected responses: near perceptual threshold or relatively constant (equal magnitude) suprathreshold responses.

【 0 0 4 2 】

For both CIECAM02 validation and regression testing after adding filters, a variety of different, well-known data sets were used. Specific well-known data sets were used for temporal center filter calibration, as well as for spatial center filter calibration. Testing results also verified the spatiotemporal combination and temporal surround. Spatial surround data was in part gathered from the combination of the well-known data sets at lowest frequencies. In aggregate, these stimuli also sampled angular and temporal extent.

【 0 0 4 3 】

Of the temporal surround filter data gathered, one experiment was particularly useful for indicating details of longer term temporal response of human color perception. Fairchild and Renif had previously conducted an experiment to track temporal step responses to spatial surround changes. In this experiment, subjects were given the task of maintaining achromatic (grey) appearance of a stationary target patch in the center following a temporal step change in color in the surround. Thus, simulation of this experiment should result in achromatic response using embodiments of the invention, even though input target color (as defined in CIE 1931 xyY, XYZ, etc.) is changing significantly over time.

【 0 0 4 4 】

The direction of changes in input test target color in the color plane (CIE xy or CIECAM02 ab, etc.) required for maintaining achromatic response in the Fairchild and Renif test depend on the direction of the change in surround. Three directions (red, green, blue approximating respective cone response peaks) and their CIE XYZ opposites, for a total of six surround color steps were used. Thus, six time courses (temporal trajectories) were obtained in the experiment.

【 0 0 4 5 】

The six stimuli were simulated using CIECAM02 and the temporal surround filter according to embodiments of the invention. An example temporal chromatic response is shown in Figure 5,

depicting CIECAM02 {a,b} magnitude equivalent vs. time (with and without the temporal surround filter). Figure 5 illustrates a Blue to yellow surround step response for filters: temporal surround sub-JND maximum difference with (low amplitude blue curve) and without filters (high amplitude red curve). Nominal threshold and "guard-band markers" (threshold plus one standard deviation) are illustrated by dashed and dash-dot lines respectively. In Figure 5, the dashed line 510 represents nominal color difference threshold (~1 JND) from achromatic (no color) whereas the dot-dash line 520 represents the threshold plus 1 standard deviation. The peak chromatic responses shown in Figure 5 correspond to the maximum excursions of Figures. 6A and 6B, which illustrate hue-chroma plots for all six surround steps. Figure 6A shows step stimuli in CIE 1931 {x,y} plane, while Figure 6B shows responses in CIECAM02 {Saturation, Hue} plane. The origin represents the expected achromatic (colorless) response. Diamonds in Figure 6B represent CIECAM02 response without any temporal processing, while X's show response once temporal surround processing is added. In this case (including Figures 5 and 6), the temporal center filter has been omitted so initial transients cause some of the X's to deviate from the origin somewhat, as can also be seen in the short peak of the response in Figure 5. Adding the temporal center filter mitigates these transients. Note that these maximum excursions occur at the instant of surround step transition. Therefore, the addition of the temporal center filter according to embodiments of the invention greatly mitigates these deviations from achromatic response. In addition, the spatial filter described above to slightly reduces all six chromatic responses.

【 0 0 4 6 】

Although particular embodiments have been described, it will be appreciated that the principles of the invention are not limited to those embodiments. Variations and modifications may be made without departing from the principles of the invention as set forth in the following claims.

【Explanation of Reference Numbers】

【 0 0 4 7 】

102 display
 106 matrix
 108 spatial center filter
 110 spatial surround filter
 120 mix controller
 122 mix controller
 130 temporal surround filter
 140 white adaptation
 150 temporal center filter
 160 remaining CIECAM02
 210 white processor

220 low pass filter 1
230 low pass filter 2
240 comparator
250 multiplexer
260 combiner

【Claim 1】 A method of predicting color visual differences in video, comprising:
accepting a video signal to be analyzed;
receiving an unfiltered white signal corresponding to the video display and viewing environment;
creating a spatial surround from the video signal;
creating a spatial center from the video signal; and
creating a temporal surround of the video signal from a combination of the spatial surround, the spatial center, and the unfiltered white signal using one or more weighting factors.

【Claim 2】 A system for moving image color appearance modeling, comprising:
an input structured to accept a video signal;
a white generator structured to create an unfiltered white signal from the video signal;
a cone response generator coupled to the input and structured to generate a set of cone responses from the video signal;
a spatial center filter structured to generate a spatial center of one of the set of cone responses;
a spatial surround filter structured to generate a spatial surround of the one of the set of cone responses; and
a temporal surround filter structured to generate a temporal surround from the spatial center, spatial surround, and the unfiltered white signal.

【Abstract】

【Problem】 A moving image color appearance is modeled.

【Means for Solving Problem】 An input accepts a video signal and a white generator creates an unfiltered white signal from the video signal. A cone response generator is coupled to the input and generates a set of cone responses from the video signal. A spatial center filter generates a spatial center of one of the set of cone responses. A spatial surround filter generates a spatial surround of the one of the set of cone responses. A temporal surround filter generates a temporal surround from the spatial center, spatial surround, and the unfiltered white signal.

【Representative Drawing】 Figure 1

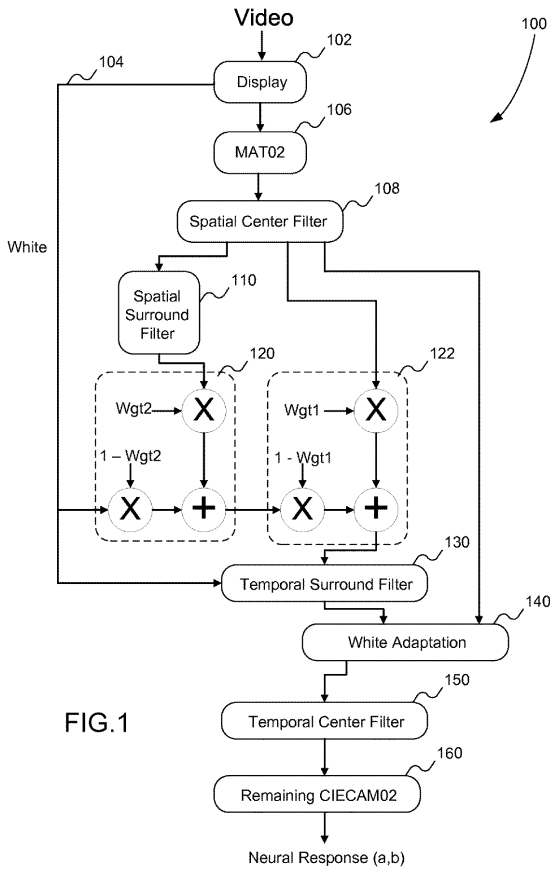


FIG. 1

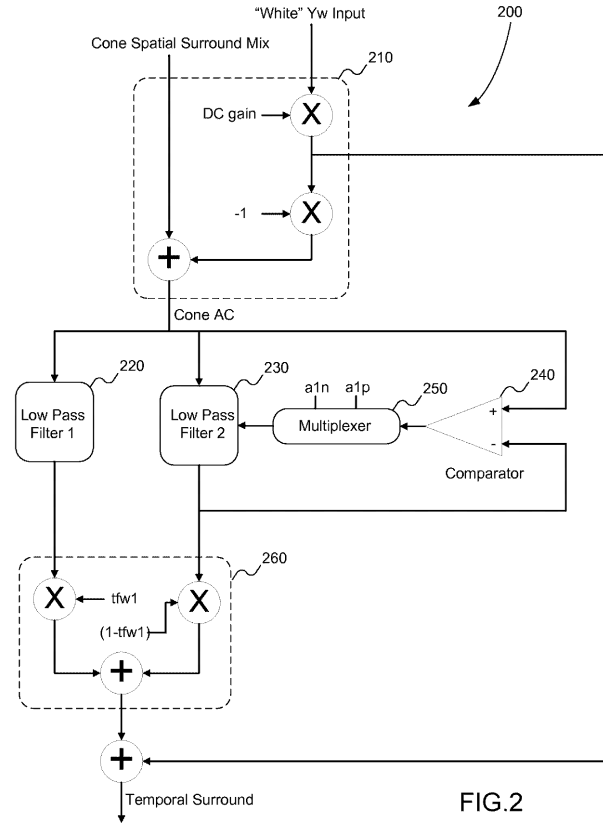


FIG. 2

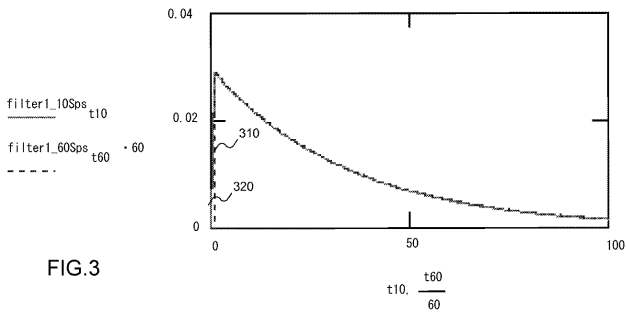


FIG. 3

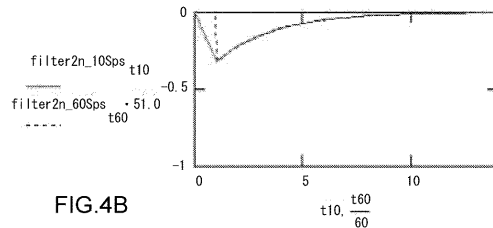


FIG. 4B

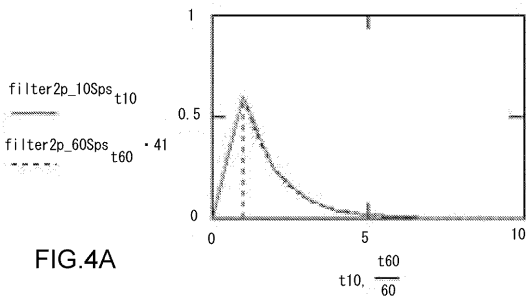


FIG. 4A

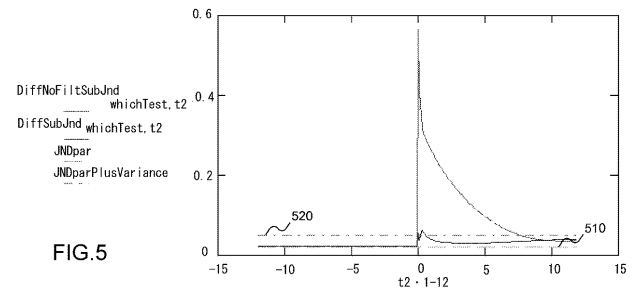


FIG. 5

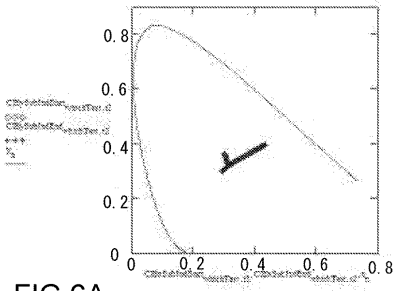


FIG.6A

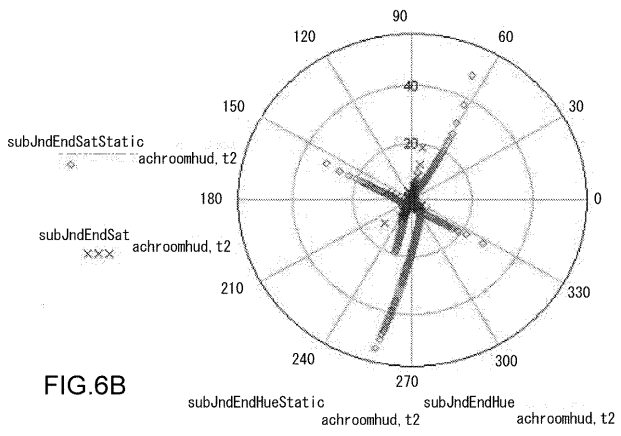


FIG.6B