

(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2009-3615  
(P2009-3615A)

(43) 公開日 平成21年1月8日(2009.1.8)

(51) Int.Cl.  
G06T 7/20 (2006.01)

F I  
G06T 7/20

テーマコード (参考)  
5 L096

審査請求 未請求 請求項の数 9 O L (全 29 頁)

(21) 出願番号 特願2007-162477 (P2007-162477)  
(22) 出願日 平成19年6月20日 (2007.6.20)

(71) 出願人 000004226  
日本電信電話株式会社  
東京都千代田区大手町二丁目3番1号  
(74) 代理人 100064908  
弁理士 志賀 正武  
(74) 代理人 100108453  
弁理士 村山 靖彦  
(72) 発明者 木村 昭悟  
東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内  
(72) 発明者 柏野 邦夫  
東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 注目領域抽出方法、注目領域抽出装置、コンピュータプログラム、及び、記録媒体

(57) 【要約】

【課題】注目度の時間的変動を考慮して注目領域を抽出する。

【解決手段】入力映像のあるフレームの中で顕著な特性を持つ空間領域を表示した基礎注目度画像を抽出する過程と、現在より1つ前のフレームから抽出された基礎注目度画像の基礎注目度が最も大きい領域である最大基礎注目度領域を、現在フレームから抽出された基礎注目度画像において抑制し、瞬時抑制後注目度画像を抽出する過程、または、現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、基礎注目度画像、または、瞬時遷移注目度画像において、抽出した領域を強調するとともに他の領域を抑制して漸進抑制後注目度画像を抽出する過程の一方、または、両方の過程と、瞬時抑制後注目度画像または漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する過程とを有する。

【選択図】 図 1



## 【特許請求の範囲】

## 【請求項 1】

対象となる入力映像から、その入力映像の中で顕著な特性を持つ時空間領域を表示した映像である注目度映像を抽出する注目領域抽出方法であって、

入力映像を構成するあるフレームから、当該フレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出する基礎注目度画像抽出過程と、

前記基礎注目度画像抽出過程により前記入力映像の現在より1つ前のフレームから抽出された基礎注目度画像において、各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を、前記基礎注目度画像抽出過程により前記入力映像の現在フレームから抽出された基礎注目度画像において抑制することにより、瞬時抑制後注目度画像を抽出する注目度瞬時抑制過程、

10

または、

前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、前記基礎注目度画像抽出過程により前記入力映像の当該フレームから抽出された基礎注目度画像、または、前記注目度瞬時抑制過程により前記入力映像の当該フレームから抽出された瞬時遷移注目度画像において、抽出した前記領域を強調するとともに他の領域を抑制することにより、漸進抑制後注目度画像を抽出する注目度漸進抑制過程

の一方、または、両方の過程と、

前記基礎注目度画像抽出過程と、注目度瞬時抑制過程または注目度漸進抑制過程の一方または両方の過程とを、前記入力映像の各フレームに対して順に繰り返して実行することにより、前記瞬時抑制後注目度画像または前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する注目度映像出力過程と

20

を有することを特徴とする注目領域抽出方法。

## 【請求項 2】

前記基礎注目度画像抽出過程は、

前記入力映像のあるフレームから、基礎特徴画像を複数種類抽出する基礎特徴画像抽出過程と、

前記基礎特徴画像抽出過程により抽出された基礎特徴画像の各種類について、その多重解像度表現である多重解像度画像を抽出する多重解像度画像抽出過程と、

30

前記多重解像度画像抽出過程により抽出された多重解像度画像の各種類について、解像度の異なる画像の間の差分である解像度差分画像を複数抽出する解像度差分画像抽出過程と、

前記解像度差分画像抽出過程により抽出された解像度差分画像の各種類について、解像度の異なる解像度差分画像を統合することにより、顕著度画像を抽出する顕著度画像抽出過程と、

前記顕著度画像抽出過程により抽出された顕著度画像について、複数種類の顕著度画像を統合することにより、基礎注目度画像を抽出する顕著度画像統合過程とからなり、

前記注目度漸進抑制過程は、基礎注目度画像に代えて漸進抑制画像を用いて顕著度画像を抽出する

40

ことを特徴とする請求項 1 に記載の注目領域抽出方法。

## 【請求項 3】

さらに、

前記基礎注目度画像抽出過程により抽出された基礎注目度画像から前記最大基礎注目度領域を抽出し、複数種類の顕著度画像について前記最大基礎注目度領域に対応する領域の中の値を算出し、その値の大きさから、対応する顕著度画像の重みである顕著度画像統合比率を決定する顕著度画像統合比率算出過程と

を有し、

前記基礎注目度画像抽出過程は、前記入力映像の1つ前のフレームについて前記顕著度画像統合比率算出過程により算出された顕著度画像統合比率によって顕著度画像を重み付

50

けして統合することにより、基礎注目度画像を抽出する

ことを特徴とする請求項 1 または請求項 2 に記載の注目領域抽出方法。

【請求項 4】

前記顕著度画像統合比率算出過程は、

前記入力映像の 1 つの前のフレームについて算出された前記顕著度画像統合比率を初期値とし、各顕著度画像について算出された前記最大基礎注目度領域の中の値を当該初期値に対しての差分値として新たな顕著度画像統合比率を更新する

ことを特徴とする請求項 3 に記載の注目領域抽出方法。

【請求項 5】

前記注目度瞬時抑制過程は、

前記基礎注目度画像抽出過程により前記入力映像の現在より 1 つ前のフレームから抽出された基礎注目度画像から前記最大基礎注目度領域を抽出する最大基礎注目度領域検出過程と、

前記最大基礎注目度領域検出過程により抽出された前記最大基礎注目度領域を遮蔽する画像である最大基礎注目度領域遮蔽画像を抽出する最大基礎注目度領域遮蔽画像抽出過程と、

前記最大基礎注目度領域遮蔽画像抽出過程により抽出された最大基礎注目度領域遮蔽画像によって遮蔽されている領域について、その領域における遮蔽を低減する画像である注目度漸進回復画像を抽出する注目度漸進回復画像抽出過程と、

前記最大基礎注目度領域遮蔽画像抽出過程により抽出された最大基礎注目度領域遮蔽画像、及び、前記注目度漸進回復画像抽出過程により抽出された注目度漸進回復画像を統合することにより注目度瞬時抑制画像を生成する注目度瞬時抑制画像生成過程と、

前記注目度瞬時抑制画像生成過程により生成された注目度瞬時抑制画像、及び、基礎注目度画像抽出過程により抽出された前記入力映像の現在のフレームの基礎注目度画像を統合することにより瞬時抑制後注目度画像を生成する瞬時抑制後注目度画像生成過程とからなる

ことを特徴とする請求項 1 から請求項 4 のいずれかの項に記載の注目領域抽出方法。

【請求項 6】

前記注目度漸進抑制過程は、

前記基礎注目度画像抽出過程により抽出された基礎注目度画像を漸進的に遮蔽する画像である注目度漸進遮蔽画像を生成する注目度漸進遮蔽画像生成過程と、

前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像もしくは顕著度画像について、時間軸方向で顕著な値を持つ領域を抽出し、その領域に対応する前記基礎注目度画像もしくは前記瞬時抑制後注目度画像における領域の基礎注目度の抑制を解除させる画像である注目度瞬時回復画像を生成する注目度瞬時回復画像生成過程と、

前記注目度漸進遮蔽画像生成過程により生成された注目度漸進遮蔽画像、及び、前記注目度瞬時回復画像生成過程により生成された注目度瞬時回復画像を統合することにより注目度漸進抑制画像を生成する注目度漸進抑制画像生成過程と、

前記基礎注目度画像抽出過程により抽出された基礎注目度画像、もしくは、前記注目度瞬時抑制過程により抽出された瞬時抑制後注目度画像と、前記注目度漸進抑制画像生成過程に生成された注目度漸進抑制画像とを統合することにより漸進抑制後注目度画像を生成する漸進抑制後注目度画像生成過程とからなる

ことを特徴とする請求項 1 から請求項 5 のいずれかに記載の注目領域抽出方法。

【請求項 7】

対象となる入力映像から、その入力映像の中で顕著な特性を持つ時空間領域を表示した映像である注目度映像を抽出する注目領域抽出装置であって、

入力映像を構成するあるフレームから、当該フレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出する基礎注目度画像抽出部と、

前記基礎注目度画像抽出部により前記入力映像の現在より 1 つ前のフレームから抽出さ

10

20

30

40

50

れた基礎注目度画像において、各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を、前記基礎注目度画像抽出部により前記入力映像の現在フレームから抽出された基礎注目度画像において抑制することにより、瞬時抑制後注目度画像を抽出する注目度瞬時抑制部、

または、

前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、前記基礎注目度画像抽出部により前記入力映像の当該フレームから抽出された基礎注目度画像、または、前記注目度瞬時抑制部により前記入力映像の当該フレームから抽出された瞬時遷移注目度画像において、抽出した前記領域を強調するとともに他の領域を抑制することにより、漸進抑制後注目度画像を抽出する注目度漸進抑制部

10

の一方、または、両方と、

前記入力映像の各フレームに対して、注目度瞬時抑制部が抽出した前記瞬時抑制後注目度画像、または、前記注目度漸進抑制部が抽出した前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する注目度映像出力部と

を備えることを特徴とする注目領域抽出装置。

#### 【請求項 8】

対象となる入力映像から、その入力映像の中で顕著な特性を持つ時空間領域を表示した映像である注目度映像を抽出する注目領域抽出装置として用いられるコンピュータに、

20

入力映像を構成するあるフレームから、当該フレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出する基礎注目度画像抽出過程と、

前記基礎注目度画像抽出過程により前記入力映像の現在より1つ前のフレームから抽出された基礎注目度画像において、各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を、前記基礎注目度画像抽出過程により前記入力映像の現在フレームから抽出された基礎注目度画像において抑制することにより、瞬時抑制後注目度画像を抽出する注目度瞬時抑制過程、

または、

前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、前記基礎注目度画像抽出過程により前記入力映像の当該フレームから抽出された基礎注目度画像、または、前記注目度瞬時抑制過程により前記入力映像の当該フレームから抽出された瞬時遷移注目度画像において、抽出した前記領域を強調するとともに他の領域を抑制することにより、漸進抑制後注目度画像を抽出する注目度漸進抑制過程

30

の一方、または、両方の過程と、

前記基礎注目度画像抽出過程と、注目度瞬時抑制過程または注目度漸進抑制過程の一方または両方の過程とを、前記入力映像の各フレームに対して順に繰り返して実行することにより、前記瞬時抑制後注目度画像または前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する注目度映像出力過程と

を実行させることを特徴とするコンピュータプログラム。

40

#### 【請求項 9】

対象となる入力映像から、その入力映像の中で顕著な特性を持つ時空間領域を表示した映像である注目度映像を抽出する注目領域抽出装置として用いられるコンピュータに、

入力映像を構成するあるフレームから、当該フレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出する基礎注目度画像抽出過程と、

前記基礎注目度画像抽出過程により前記入力映像の現在より1つ前のフレームから抽出された基礎注目度画像において、各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を、前記基礎注目度画像抽出過程により前記入力映像の現在フレームから抽出された基礎注目度画像において抑制することにより、瞬時抑制後注目度画像を抽出する注目度瞬時抑制過程、

50

または、

前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、前記基礎注目度画像抽出過程により前記入力映像の当該フレームから抽出された基礎注目度画像、または、前記注目度瞬時抑制過程により前記入力映像の当該フレームから抽出された瞬時遷移注目度画像において、抽出した前記領域を強調するとともに他の領域を抑制することにより、漸進抑制後注目度画像を抽出する注目度漸進抑制過程

の一方、または、両方の過程と、

前記基礎注目度画像抽出過程と、注目度瞬時抑制過程または注目度漸進抑制過程の一方または両方の過程とを、前記入力映像の各フレームに対して順に繰り返して実行することにより、前記瞬時抑制後注目度画像または前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する注目度映像出力過程と

を実行させるコンピュータプログラムを記録したコンピュータ読み取り可能な記録媒体

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、映像信号における注目領域を抽出する注目領域抽出方法、注目領域抽出装置、コンピュータプログラム、及び、記録媒体に関する。

【背景技術】

【0002】

デジタルカメラの普及、記録媒体の大容量化、低価格化、大容量ネットワークの普及により、多様な画像・映像を蓄積して検索・利用する技術が必要となってきた。中でも、言葉では表現しにくい色・形状・構図・模様などの概念から、全体として類似した画像・映像を高速に引き出してくる技術は、ショッピングサイトや動画共有サイトなど、幅広い分野で開発が求められている。

【0003】

上記のような検索技術の実現において、ユーザの検索意図を反映させるために、画像・映像の中で人間が重要と感じる度合いを注目度として抽出し、この抽出した注目度に基づいて画像の類似性を判断する方法が有望である。前記注目度を算出する技術として非特許文献1及び非特許文献2に記載の方法が、また前記注目度に基づいて画像の類似性を判断する技術として特許文献1に記載の方法が知られている。

【特許文献1】特開2006-338313号公報

【非特許文献1】L. Itti et al., "A model of saliency-based visual attention for rapid scene analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 20, Number11, pp. 1254-1259, November 1998.

【非特許文献2】L. Itti et al., "Realistic avatar eye and head animation using a neurobiological model of visual attention", Proceedings of SPIE International Symposium on Optical Science and Technology, Volume 5200, pp. 64-78, August 2003.

【発明の開示】

【発明が解決しようとする課題】

【0004】

しかしながら、上記文献に記載の方法では、注目度が時間的に変動する側面についてはほとんど考慮されていない。本発明は、このような事情に鑑みてなされたものであり、その目的は、時間的変動を考慮した新たな注目度の算出方法を導入した注目領域抽出方法、注目領域抽出装置、コンピュータプログラム、及び、記録媒体を提供することにある。

【0005】

すなわち、本発明においては、以下の2点により、注目度の時間的変動を考慮した注目度算出方法を実現する。

10

20

30

40

50

## 【 0 0 0 6 】

( 1 ) 注目度が最大となる領域の瞬時的な移動の導入。

人間の初期視覚において、できるだけ早く視野全体の情報を把握するために、注視点が高速に移動する「視覚探索」という現象が知られている ( R.M. Klein: "Inhibition of return," Trends in Cognitive Sciences, Vol.4, No.4, pp.138-147, April 2000. ; 非特許文献 3 )。本発明ではこの「視覚探索」を注目度の算出に導入する。

## 【 0 0 0 7 】

( 2 ) 変動の小さい視覚的刺激に対する注目度の抑制の導入。

人間の初期視覚において、できるだけ少ない注視点移動で視野全体の情報を把握するために、視覚的刺激が小さい領域に対する注目の度合いが時間経過と共に低下する「視覚適応」という現象が知られている ( S. Martinez-Conde, S. L. Macknik and D. H. Hubel: "The role of fixational eye movements in visual perception," Nature Reviews, Vol.5, pp.229-240, March 2004. ; 非特許文献 4 )。本発明では、これの「視覚適応」を注目度の算出に導入する。

【 課題を解決するための手段 】

## 【 0 0 0 8 】

前記課題を解決するために、本発明は、対象となる入力映像から、その入力映像の中で顕著な特性を持つ時空間領域を表示した映像である注目度映像を抽出する注目領域抽出方法であって、入力映像を構成するあるフレームから、当該フレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出する基礎注目度画像抽出過程と、前記基礎注目度画像抽出過程により前記入力映像の現在より 1 つ前のフレームから抽出された基礎注目度画像において、各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を、前記基礎注目度画像抽出過程により前記入力映像の現在フレームから抽出された基礎注目度画像において抑制することにより、瞬時抑制後注目度画像を抽出する注目度瞬時抑制過程、または、前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、前記基礎注目度画像抽出過程により前記入力映像の当該フレームから抽出された基礎注目度画像、または、前記注目度瞬時抑制過程により前記入力映像の当該フレームから抽出された瞬時遷移注目度画像において、抽出した前記領域を強調するとともに他の領域を抑制することにより、漸進抑制後注目度画像を抽出する注目度漸進抑制過程の一方、または、両方の過程と、前記基礎注目度画像抽出過程と、注目度瞬時抑制過程または注目度漸進抑制過程の一方または両方の過程とを、前記入力映像の各フレームに対して順に繰り返して実行することにより、前記瞬時抑制後注目度画像または前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する注目度映像出力過程とを有することを特徴とする注目領域抽出方法である。

## 【 0 0 0 9 】

上記発明では、注目度瞬時抑制過程によって、注目度が最大となる領域の瞬時的な移動を導入し、「視覚探索」現象を信号処理により実現する。また、注目度漸進抑制過程によって、変動の小さい視覚的刺激に対する注目度の抑制を導入し、「視覚適応」現象を信号処理により実現する。

## 【 0 0 1 0 】

また、本発明は、上述する注目領域抽出方法であって、前記基礎注目度画像抽出過程は、前記入力映像のあるフレームから、基礎特徴画像を複数種類抽出する基礎特徴画像抽出過程と、前記基礎特徴画像抽出過程により抽出された基礎特徴画像の各種類について、その多重解像度表現である多重解像度画像を抽出する多重解像度画像抽出過程と、前記多重解像度画像抽出過程により抽出された多重解像度画像の各種類について、解像度の異なる画像の間の差分である解像度差分画像を複数抽出する解像度差分画像抽出過程と、前記解像度差分画像抽出過程により抽出された解像度差分画像の各種類について、解像度の異なる解像度差分画像を統合することにより、顕著度画像を抽出する顕著度画像抽出過程と、前記顕著度画像抽出過程により抽出された顕著度画像について、複数種類の顕著度画像を

統合することにより、基礎注目度画像を抽出する顕著度画像統合過程とからなり、前記注目度漸進抑制過程は、基礎注目度画像に代えて漸進抑制画像を用いて顕著度画像を抽出することを特徴とする。

【0011】

また、本発明は、上述する注目領域抽出方法であって、さらに、前記基礎注目度画像抽出過程により抽出された基礎注目度画像から前記最大基礎注目度領域を抽出し、複数種類の顕著度画像について前記最大基礎注目度領域に対応する領域の中の値を算出し、その値の大きさから、対応する顕著度画像の重みである顕著度画像統合比率を決定する顕著度画像統合比率算出過程とを有し、前記基礎注目度画像抽出過程は、前記入力映像の1つ前のフレームについて前記顕著度画像統合比率算出過程により算出された顕著度画像統合比率によって顕著度画像を重み付けして統合することにより、基礎注目度画像を抽出することを特徴とする。

10

【0012】

また、本発明は、上述する注目領域抽出方法であって、前記顕著度画像統合比率算出過程は、前記入力映像の1つ前のフレームについて算出された前記顕著度画像統合比率を初期値とし、各顕著度画像について算出された前記最大基礎注目度領域の中の値を当該初期値に対しての差分値として新たな顕著度画像統合比率を更新することを特徴とする。

【0013】

また、本発明は、上述する注目領域抽出方法であって、前記注目度瞬時抑制過程は、前記基礎注目度画像抽出過程により前記入力映像の現在より1つ前のフレームから抽出された基礎注目度画像から前記最大基礎注目度領域を抽出する最大基礎注目度領域検出過程と、前記最大基礎注目度領域検出過程により抽出された前記最大基礎注目度領域を遮蔽する画像である最大基礎注目度領域遮蔽画像を抽出する最大基礎注目度領域遮蔽画像抽出過程と、前記最大基礎注目度領域遮蔽画像抽出過程により抽出された最大基礎注目度領域遮蔽画像によって遮蔽されている領域について、その領域における遮蔽を低減する画像である注目度漸進回復画像を抽出する注目度漸進回復画像抽出過程と、前記最大基礎注目度領域遮蔽画像抽出過程により抽出された最大基礎注目度領域遮蔽画像、及び、前記注目度漸進回復画像抽出過程により抽出された注目度漸進回復画像を統合することにより注目度瞬時抑制画像を生成する注目度瞬時抑制画像生成過程と、前記注目度瞬時抑制画像生成過程により生成された注目度瞬時抑制画像、及び、基礎注目度画像抽出過程により抽出された前記入力映像の現在のフレームの基礎注目度画像を統合することにより瞬時抑制後注目度画像を生成する瞬時抑制後注目度画像生成過程とからなることを特徴とする。

20

30

【0014】

また、本発明は、上述する注目領域抽出方法であって、前記注目度漸進抑制過程は、前記基礎注目度画像抽出過程により抽出された基礎注目度画像を漸進的に遮蔽する画像である注目度漸進遮蔽画像を生成する注目度漸進遮蔽画像生成過程と、前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像もしくは顕著度画像について、時間軸方向で顕著な値を持つ領域を抽出し、その領域に対応する前記基礎注目度画像もしくは前記瞬時抑制後注目度画像における領域の基礎注目度の抑制を解除させる画像である注目度瞬時回復画像を生成する注目度瞬時回復画像生成過程と、前記注目度漸進遮蔽画像生成過程により生成された注目度漸進遮蔽画像、及び、前記注目度瞬時回復画像生成過程により生成された注目度瞬時回復画像を統合することにより注目度漸進抑制画像を生成する注目度漸進抑制画像生成過程と、前記基礎注目度画像抽出過程により抽出された基礎注目度画像、もしくは、前記注目度瞬時抑制過程により抽出された瞬時抑制後注目度画像と、前記注目度漸進抑制画像生成過程に生成された注目度漸進抑制画像とを統合することにより漸進抑制後注目度画像を生成する漸進抑制後注目度画像生成過程とからなることを特徴とする。

40

【0015】

また、本発明は、対象となる入力映像から、その入力映像の中で顕著な特性を持つ時空間領域を表示した映像である注目度映像を抽出する注目領域抽出装置であって、入力映像

50

を構成するあるフレームから、当該フレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出する基礎注目度画像抽出部と、前記基礎注目度画像抽出部により前記入力映像の現在より1つ前のフレームから抽出された基礎注目度画像において、各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を、前記基礎注目度画像抽出部により前記入力映像の現在フレームから抽出された基礎注目度画像において抑制することにより、瞬時抑制後注目度画像を抽出する注目度瞬時抑制部、または、前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、前記基礎注目度画像抽出部により前記入力映像の当該フレームから抽出された基礎注目度画像、または、前記注目度瞬時抑制部により前記入力映像の当該フレームから抽出された瞬時遷移注目度画像において、抽出した前記領域を強調するとともに他の領域を抑制することにより、漸進抑制後注目度画像を抽出する注目度漸進抑制部の一方、または、両方と、前記入力映像の各フレームに対して、注目度瞬時抑制部が抽出した前記瞬時抑制後注目度画像、または、前記注目度漸進抑制部が抽出した前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する注目度映像出力部とを備えることを特徴とする注目領域抽出装置である。

10

## 【0016】

また、本発明は、対象となる入力映像から、その入力映像の中で顕著な特性を持つ時空間領域を表示した映像である注目度映像を抽出する注目領域抽出装置として用いられるコンピュータに、入力映像を構成するあるフレームから、当該フレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出する基礎注目度画像抽出過程と、前記基礎注目度画像抽出過程により前記入力映像の現在より1つ前のフレームから抽出された基礎注目度画像において、各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を、前記基礎注目度画像抽出過程により前記入力映像の現在フレームから抽出された基礎注目度画像において抑制することにより、瞬時抑制後注目度画像を抽出する注目度瞬時抑制過程、または、前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、前記基礎注目度画像抽出過程により前記入力映像の当該フレームから抽出された基礎注目度画像、または、前記注目度瞬時抑制過程により前記入力映像の当該フレームから抽出された瞬時遷移注目度画像において、抽出した前記領域を強調するとともに他の領域を抑制することにより、漸進抑制後注目度画像を抽出する注目度漸進抑制過程の一方、または、両方の過程と、前記基礎注目度画像抽出過程と、注目度瞬時抑制過程または注目度漸進抑制過程の一方または両方の過程とを、前記入力映像の各フレームに対して順に繰り返して実行することにより、前記瞬時抑制後注目度画像または前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する注目度映像出力過程とを実行させることを特徴とするコンピュータプログラムである。

20

30

## 【0017】

また、本発明は、対象となる入力映像から、その入力映像の中で顕著な特性を持つ時空間領域を表示した映像である注目度映像を抽出する注目領域抽出装置として用いられるコンピュータに、入力映像を構成するあるフレームから、当該フレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出する基礎注目度画像抽出過程と、前記基礎注目度画像抽出過程により前記入力映像の現在より1つ前のフレームから抽出された基礎注目度画像において、各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を、前記基礎注目度画像抽出過程により前記入力映像の現在フレームから抽出された基礎注目度画像において抑制することにより、瞬時抑制後注目度画像を抽出する注目度瞬時抑制過程、または、前記入力映像の現在のフレーム以前のいくつかのフレームから算出された基礎注目度画像について、時間軸方向で顕著な値を持つ領域を抽出し、前記基礎注目度画像抽出過程により前記入力映像の当該フレームから抽出された基礎注目度画像、または、前記注目度瞬時抑制過程により前記入力映像の当該フレームから抽出された瞬時遷移注目度画像において、抽出した前記領域を強調するとともに他の領域を抑

40

50

制することにより、漸進抑制後注目度画像を抽出する注目度漸進抑制過程の一方、または、両方の過程と、前記基礎注目度画像抽出過程と、注目度瞬時抑制過程または注目度漸進抑制過程の一方または両方の過程とを、前記入力映像の各フレームに対して順に繰り返して実行することにより、前記瞬時抑制後注目度画像または前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する注目度映像出力過程とを実行させるコンピュータプログラムを記録したコンピュータ読み取り可能な記録媒体である。

【発明の効果】

【0018】

本発明によれば、映像から注目領域を抽出する際に、注目度が最大となる領域の瞬時的な移動の導入、変動の小さい視覚的刺激に対する注目度の抑制を行うことにより、映像の注目度の時間的変更を考慮し、人間が重要と感じる部分に近い注目領域を抽出することが可能となる。

10

【発明を実施するための最良の形態】

【0019】

以下、図面を用いて本発明の実施形態を説明する。

【0020】

[第1の実施形態]

図1は、本発明の第1の実施形態による注目領域抽出装置の機能ブロック図である。

同図に示す注目領域抽出装置は、基礎注目度画像抽出部1と、注目度瞬時抑制部3と、注目度映像出力部5とにより構成され、注目度抽出の対象となる映像のデータである入力映像を入力し、入力画像の中で注目度の高い顕著な特性を持つ時空間領域を表示した映像である注目度映像を出力する。映像は、複数のフレームの画像からなる。

20

【0021】

基礎注目度画像抽出部1は、入力映像のあるフレームである入力画像を入力し、そのフレームの中で顕著な特性を持つ空間領域を表示した画像である基礎注目度画像を抽出し、その基礎注目度画像を出力する。

基礎注目度画像の抽出方法は特に限定されるものではないが、本実施形態においては、基礎注目度画像抽出部1が、基礎特徴画像抽出部11と、多重解像度画像抽出部12と、解像度差分画像抽出部13と、顕著度画像抽出部14と、顕著度画像統合部15とによって構成される場合の特徴抽出方法について述べる。

30

【0022】

基礎特徴画像抽出部11は、入力画像を入力し、複数の特徴抽出方法により入力画像からそれぞれ基礎特徴画像を抽出し、それら基礎特徴画像の集合を出力する。

基礎特徴画像の抽出方法は特に限定されるものではないが、本実施形態においては、基礎特徴画像抽出部11が、輝度特徴画像抽出部111と、色特徴画像抽出部112と、方向特徴画像抽出部113と、点滅特徴画像抽出部114と、運動特徴画像抽出部115とによって構成される場合の特徴抽出方法について述べる。

【0023】

輝度特徴画像抽出部111は、入力画像を入力し、入力画像の輝度成分を表現する輝度特徴画像を出力する。輝度特徴画像は、入力画像のRGB (Red, Green, Blue) 成分の平均として、以下の(式1)のように求める。

40

【0024】

【数1】

$$I(i) = \frac{r(i) + g(i) + b(i)}{3} \quad (\text{式1})$$

【0025】

ここで、 $r(i)$ 、 $g(i)$ 、 $b(i)$ は、それぞれ*i*番目の入力画像 (= 入力映像の

50

i 番目のフレーム)における R (赤色) , G (緑色) , B (青) の各成分であり、画素値はそれぞれ 0 から 1 の実数値で表されている。また、I ( i ) は、i 番目の入力画像から計算される輝度特徴画像である。

【 0 0 2 6 】

色特徴画像抽出部 1 1 2 は、入力画像を入力し、入力画像の色成分を表現する色特徴画像を出力する。色特徴画像抽出部 1 1 2 では、赤色・緑色・青色・黄色にそれぞれ対応する 4 種類の色特徴画像を抽出する。これらの色特徴画像はそれぞれ、以下の ( 式 2 ) ~ ( 式 4 ) のように求める。ただし、max は、括弧 { } 内の値のうちの最大値を示す。

【 0 0 2 7 】

【 数 2 】

$$R(i)_{(x,y)} = \max \left\{ 0, r(i)_{(x,y)} - \frac{g(i)_{(x,y)} + b(i)_{(x,y)}}{2} \right\} \quad (式 2)$$

$$G(i)_{(x,y)} = \max \left\{ 0, g(i)_{(x,y)} - \frac{r(i)_{(x,y)} + b(i)_{(x,y)}}{2} \right\} \quad (式 3)$$

$$B(i)_{(x,y)} = \max \left\{ 0, b(i)_{(x,y)} - \frac{r(i)_{(x,y)} + g(i)_{(x,y)}}{2} \right\} \quad (式 4)$$

$$Y(i)_{(x,y)} = \max \left\{ 0, \frac{r(i)_{(x,y)} + g(i)_{(x,y)}}{2} - \frac{|r(i)_{(x,y)} - g(i)_{(x,y)}|}{2} - b(i)_{(x,y)} \right\} \quad (式 5)$$

【 0 0 2 8 】

ここで、R ( i ) , G ( i ) , B ( i ) , Y ( i ) は、それぞれ赤色・緑色・青色・黄色に対応する、i 番目の入力画像から計算される色特徴画像である。また、R ( i ) ( x , y ) は、座標 ( x , y ) における色特徴画像 R ( i ) の画素値である。以降の説明において、必要な場合を除いて、下添字の ( x , y ) は省略するものとする。

【 0 0 2 9 】

方向特徴画像抽出部 1 1 3 は、入力画像を入力し、入力画像の方向成分を表現する方向特徴画像を出力する。方向特徴画像は、輝度特徴画像抽出部 1 1 1 によって求めた前記輝度特徴画像 I ( i ) にガボールフィルタを作用させることによって、以下の ( 式 6 ) のように求められる。なお、ガボールフィルタとは、画像の局所的な濃淡情報を取り出すものである。

【 0 0 3 0 】

【 数 3 】

$$O_{\phi}(i) = g_{\phi} * I(i) \quad (式 6)$$

【 0 0 3 1 】

ここで、g は回転角 を持つガボールフィルタ、\* は畳み込み (関数を平行移動しながらもう一方の関数を重ね足し合わせる) を表現する演算子である。また、O ( i ) は i 番目の入力画像から計算される、回転角 に対応する方向特徴画像である。方向特徴画像は、n 通りの回転角について抽出される。このとき回転角 は、 = 1 8 0 ° を均等に n 分割するように、以下の ( 式 7 ) のように選択される。

【 0 0 3 2 】

【 数 4 】

$$\phi = 0, \frac{\pi}{n_{\phi}}, \dots, \frac{(n_{\phi} - 1)\pi}{n_{\phi}} \quad (式 7)$$

【 0 0 3 3 】

点滅特徴画像抽出部 1 1 4 は、入力画像を入力し、入力画像の点滅成分を表現する点滅

10

20

30

40

50

特徴画像を出力する。点滅特徴画像は、輝度特徴画像抽出部 1 1 1 により、現在及びそれ以前のいくつかの入力画像から計算される前記輝度特徴画像  $I(i)$  を用い、以下の(式 8) のように計算される。

【 0 0 3 4 】

【 数 5 】

$$F(i) = \left| I(i) - \sum_{k=1}^{n_F} 2(n_F - k + 1) \frac{I(i - k)}{n_F(n_F + 1)} \right| \quad (\text{式 8})$$

【 0 0 3 5 】

ここで、 $n_F$  は点滅特徴画像を抽出する際に参照する輝度特徴画像の数(ただし現在の入力画像から抽出された輝度特徴画像は除く)、 $F(i)$  は  $i$  番目及びそれ以前の入力画像から計算される点滅特徴画像である。 $n_F = 1$  としたとき、 $F(i) = |I(i) - I(i - 1)|$  となり、非特許文献 2 に記載の方法と一致する。

【 0 0 3 6 】

運動特徴画像抽出部 1 1 5 は、入力画像を入力し、入力画像の運動成分を表現する運動特徴画像を出力する。運動特徴画像の抽出方法は特に限定されるものではないが、本実施形態においては、輝度特徴画像の各点におけるオプティカルフローを求めることによって抽出する。オプティカルフローの導出方法は特に限定されるものではないが、本実施形態においては、例えば、一般に Lucas-Kanade 法と呼ばれる画像勾配に基づく方法を用いて、以下の(式 9) ~ (式 10) のように計算する。

【 0 0 3 7 】

【 数 6 】

$$M_x(i)_{(x,y)} = \frac{\sum_{(x',y') \in A(x,y)} \frac{\partial x}{\partial I(i)} \Big|_{(x',y')} \cdot \{I(i)_{(x',y')} - I(i-1)_{(x',y')}\}}{\sum_{(x',y') \in A(x,y)} \left( \frac{\partial x}{\partial I(i)} \Big|_{(x',y')} \right)^2} \quad (\text{式 9})$$

$$M_y(i)_{(x,y)} = \frac{\sum_{(x',y') \in A(x,y)} \frac{\partial y}{\partial I(i)} \Big|_{(x',y')} \cdot \{I(i)_{(x',y')} - I(i-1)_{(x',y')}\}}{\sum_{(x',y') \in A(x,y)} \left( \frac{\partial y}{\partial I(i)} \Big|_{(x',y')} \right)^2} \quad (\text{式 10})$$

【 0 0 3 8 】

ここで、 $A(x, y)$  は座標  $(x, y)$  の近傍領域、 $M_x(i)$ 、 $M_y(i)$  はそれぞれ運動の水平成分・垂直成分に対応する運動特徴画像である。

【 0 0 3 9 】

上記の通り、基礎特徴画像抽出部 1 1 は、前記輝度特徴画像、前記色特徴画像、前記方向特徴画像、前記点滅特徴画像、及び前記運動特徴画像を、それぞれ基礎特徴画像とし、それら基礎特徴画像の集合を出力する。

【 0 0 4 0 】

多重解像度画像抽出部 1 2 は、前記基礎特徴画像の集合を入力し、各基礎特徴画像について、その多重解像度表現である多重解像度画像を抽出し、多重解像度画像の集合を出力する。

本実施形態においては、いずれの基礎特徴画像についても同様の処理を行うため、輝度特徴画像を例に処理の説明を行うものとする。輝度特徴画像についての多重解像度表現である輝度多重解像度画像は、輝度特徴画像にガウシアンフィルタを繰り返し作用させることによって以下の(式 11) のように抽出される。ガウシアンフィルタとは、画像のノイ

10

20

30

40

50

ズ除去するための平滑化フィルタのひとつであり、注目画素からの距離に基づく重み付けにガウス関数を用いる。

【 0 0 4 1 】

【 数 7 】

$$I(i; l) = G_{\sigma} * I(i; l-1) \quad (i = 1, 2, \dots, n_l) \quad (\text{式 11})$$

【 0 0 4 2 】

ここで、 $G$  は分散  $\sigma$  を持つガウシアンフィルタ、 $I(i; l)$  は輝度特徴画像  $I(i)$  から抽出した第  $l$  レベルの輝度多重解像度画像、 $n_l$  は多重解像度画像のレベル数である。第 0 レベルの輝度多重解像度画像は輝度特徴画像そのもの、すなわち  $I(i; 0) = I(i)$  とする。

10

【 0 0 4 3 】

他の基礎特徴画像についても同様にして多重解像度画像を抽出することができる。このとき、輝度多重解像度画像が  $n_l$  枚抽出されるのに対して、色特徴画像  $R(i)$ ,  $G(i)$ ,  $B(i)$ ,  $Y(i)$  を用いた場合の色多重解像度画像  $R(i; l)$ ,  $G(i; l)$ ,  $B(i; l)$ ,  $Y(i; l)$  は合計  $4n_l$  枚、方向特徴画像  $O(i)$  を用いた場合の方向多重解像度画像  $O(i; l)$  は合計  $n_l$  枚、点滅特徴画像  $F(i)$  を用いた場合の点滅多重解像度画像  $F(i; l)$  は  $n_l$  枚、運動特徴画像  $M_x(i)$ ,  $M_y(i)$  を用いた場合の運動多重解像度画像  $M_x(i; l)$ ,  $M_y(i; l)$  は合計  $2n_l$  枚、それぞれ抽出される。

20

【 0 0 4 4 】

上記の通り、多重解像度画像抽出部 1 2 は、前記輝度多重解像度画像、前記色多重解像度画像、前記方向多重解像度画像、前記点滅多重解像度画像、及び前記運動多重解像度画像をそれぞれ多重解像度画像とし、それら多重解像度画像の集合を出力する。

【 0 0 4 5 】

解像度差分画像抽出部 1 3 は、多重解像度画像抽出部 1 2 が出力した前記多重解像度画像の集合を入力し、多重解像度画像の各種類（輝度・色など）について、解像度レベルの異なる画像の間の差分画像である解像度差分画像を抽出し、解像度差分画像の集合を出力する。

30

本実施形態においては、以下の（式 1 2）～（式 1 7）のようにして解像度差分画像を抽出する。

【 0 0 4 6 】

【 数 8 】

$$RS_I(i; c, s) = |I(i; c) - I(i; s)| \quad (\text{式 12})$$

$$RS_{RG}(i; c, s) = |(R(i; c) - G(i; c)) - (R(i; s) - G(i; s))| \quad (\text{式 13})$$

$$RS_{BY}(i; c, s) = |(B(i; c) - Y(i; c)) - (B(i; s) - Y(i; s))| \quad (\text{式 14})$$

40

$$RS_O(i; \phi; c, s) = |O_{\phi}(i; c) - O_{\phi}(i; s)| \quad (\text{式 15})$$

$$RS_F(i; c, s) = |F(i; c) - F(i; s)| \quad (\text{式 16})$$

$$RS_{M_k}(i; c, s) = |M_k(i; c) - M_k(i; s)| \quad (c \in L_c, s \in L_s) \quad (\text{式 17})$$

【 0 0 4 7 】

ここで、 $RS_I(i; c, s)$  は第  $c$  レベルと第  $s$  レベルの輝度多重解像度画像から得られる輝度解像度差分画像であり、以降、 $(c, s)$  レベル輝度解像度差分画像と呼ぶこ

50

とにする。同様に、 $RS_{RG}(i; c, s)$  及び  $RS_{BY}(i; c, s)$  はそれぞれ  $(c, s)$  レベル  $RG$  色解像度差分画像及び  $(c, s)$  レベル  $BY$  色解像度差分画像、 $RS_{\theta}(i; c, s)$  は回転角  $\theta$  の  $(c, s)$  レベル方向解像度差分画像、 $RS_F(i; c, s)$  は  $(c, s)$  レベル点滅解像度差分画像、

【0048】

【数9】

$$RS_{M_k}(i; c, s)$$

【0049】

10

(以下、「 $RS_{M_k}(i; c, s)$ 」と記載)は  $k$  方向の  $(c, s)$  レベル運動解像度差分画像である。また、 $L_c, L_s$  は輝度解像度差分画像を抽出する際に考慮する解像度レベルの集合であり、それぞれ中心解像度レベル集合、周辺解像度レベル集合と呼ぶ。

このとき、輝度解像度差分画像が  $|L_c \cdot L_s|$  枚抽出されるのに対して、色解像度差分画像は  $4 |L_c \cdot L_s|$  枚、方向解像度差分画像は  $n |L_c \cdot L_s|$  枚、点滅解像度差分画像は  $|L_c \cdot L_s|$  枚、運動解像度差分画像は  $2 |L_c \cdot L_s|$  枚、それぞれ抽出される。

【0050】

上記の通り、解像度差分画像抽出部13は、前記輝度解像度差分画像、前記色解像度差分画像、前記方向解像度差分画像、前記点滅解像度差分画像、及び前記運動解像度差分画像をそれぞれ解像度差分画像とし、それら解像度差分画像の集合を出力する。

20

【0051】

顕著度画像抽出部14は、解像度差分画像抽出部13により出力された前記解像度差分画像の集合を入力し、解像度差分画像の各種類(輝度・色など)について解像度差分画像を統合した画像である顕著度画像を抽出し、それら顕著度画像の集合を出力する。

顕著度画像の抽出方法は特に限定されるものではないが、本実施形態においては、顕著度画像抽出部14が解像度差分画像正規化部141と、正規化解像度差分画像積算部142とによって構成される場合の顕著度画像抽出方法について述べる。

【0052】

解像度差分画像正規化部141は、解像度差分画像抽出部13により出力された前記解像度差分画像の集合を入力し、各解像度差分画像について正規化処理を施した画像である正規化解像度差分画像を抽出し、正規化解像度差分画像の集合を出力する。

30

本実施形態においては、いずれの解像度差分画像についても同様の処理を行うため、ある  $c \in L_c, s \in L_s$  を選んだときの  $(c, s)$  レベル輝度解像度差分画像  $RS_I(i; c, s)$  を例に処理の説明を行うものとする。解像度差分画像に対する正規化処理は、以下の(式18)~(式20)ようにして行われ、正規化解像度差分画像  $N(RS_I(i; c, s))$  を得る。

【0053】

【数10】

40

$$N(RS_I(i; c, s)) = \frac{\{m^*(RS_I(i; c, s)) - \overline{m}(RS_I(i; c, s))\}^2}{m^*(RS_I(i; c, s))} RS_I(i; c, s) \quad (\text{式18})$$

$$m^*(RS_I(i; c, s)) = \max_{(x,y)} \{RS_I(i; c, s)_{(x,y)}\} \quad (\text{式19})$$

$$\overline{m}(RS_I(i; c, s)) = \frac{1}{n_A - 1} \left( \sum_{(x,y) \in A} \max \{RS_I(i; c, s)_{(x,y)}\} - m^*(RS_I(i; c, s)) \right) \quad (\text{式20})$$

【0054】

ここで、 $m^*(RS_I(i; c, s))$  は、 $(c, s)$  レベル輝度解像度差分画像の中

50

の最大画素値である。

【 0 0 5 5 】

【 数 1 1 】

$$\overline{m}(RS_I(i; c, s))$$

(以下、 $m(RS_I(i; c, s))$ と記載)は、 $(c, s)$ レベル輝度解像度差分画像の局所領域における最大値の平均、 $A$ は、 $m(RS_I(i; c, s))$ を計算する際の局所領域、 $n_A$ はそのような局所領域の総数である。各局所領域は、例えば、画像を格子状に分割したときの1つの格子(領域)とする。なお、以下では、文字の上に  $\sim$  が付いた文字は、 $\sim$ を文字の前に記載して表す(例えば、上述する「 $m$ 」)。

【 0 0 5 6 】

他のレベルの輝度解像度差分画像、及び他の種類の輝度解像度差分画像についても同様にして正規化処理を行うことができ、それぞれの正規化解像度差分画像、すなわち、正規化輝度解像度差分画像、正規化RG色解像度差分画像、正規化BY色解像度差分画像、正規化方向解像度差分画像、正規化方向解像度差分画像、正規化点滅解像度差分画像、正規化運動解像度差分画像を得る。

【 0 0 5 7 】

正規化解像度差分画像積算部142は、解像度差分画像正規化部141により得られた前記正規化解像度差分画像の集合を入力し、それら正規化解像度差分画像を各種類(輝度・色など)について積算することにより顕著度画像を抽出し、顕著度画像の集合を出力する。

本実施形態において、正規化解像度差分画像は、以下の(式21)～(式25)のようにして積算される。

【 0 0 5 8 】

【 数 1 2 】

$$CM_I(i) = \sum_{c \in L_c} \sum_{s \in L_s} N(RS_I(i; c, s)) \quad (\text{式 21})$$

$$CM_C(i) = \sum_{c \in L_c} \sum_{s \in L_s} \{N(RS_{RG}(i; c, s)) + N(RS_{BY}(i; c, s))\} \quad (\text{式 22})$$

$$CM_O(i) = \sum_{k=0}^{n_\phi-1} N \left( \sum_{c \in L_c} \sum_{s \in L_s} N(RS_O(i; k\pi/n_\phi; c, s)) \right) \quad (\text{式 23})$$

$$CM_F(i) = \sum_{c \in L_c} \sum_{s \in L_s} N(RS_F(i; c, s)) \quad (\text{式 24})$$

$$CM_M(i) = \sum_{k=x,y} N \left( \sum_{c \in L_c} \sum_{s \in L_s} N(RS_{M_k}(i; c, s)) \right) \quad (\text{式 25})$$

【 0 0 5 9 】

ここで、 $CM_I(i)$ 、 $CM_C(i)$ 、 $CM_O(i)$ 、 $CM_F(i)$ 、 $CM_M(i)$ はそれぞれ輝度顕著度画像、色顕著度画像、方向顕著度画像、点滅顕著度画像、運動顕著度画像であり、正規化解像度差分画像 $N(RS_I(i; c, s))$ 、正規化RG色解像度差分画像 $N(RS_{RG}(i; c, s))$ 及び正規化BY色解像度差分画像 $N(RS_{BY}(i; c, s))$ 、正規化方向解像度差分画像 $N(RS_O(i; ; c, s))$ 、正規化方向解像度差分画像 $N(RS_F(i; c, s))$ 、正規化点滅解像度差分画像 $N(RS_F(i; c, s))$ 、正規化運動解像度差分画像 $N(RS_{M_k}(i; c, s))$ を用いて同様に求められる。

【 0 0 6 0 】

上記の通り、顕著度画像抽出部14は、前記輝度顕著度画像、前記色顕著度画像、前記

方向顕著度画像、前記点滅顕著度画像、及び前記運動顕著度画像を、それぞれ顕著度画像とし、それら顕著度画像の集合を出力する。

【0061】

顕著度画像統合部15は、顕著度画像抽出部14により出力された前記顕著度画像の集合を入力し、顕著度画像を統合した画像である基礎注目度画像を抽出し、その基礎注目度画像を出力する。

基礎注目度画像の抽出方法は特に限定されるものではないが、本実施形態においては、顕著度画像統合部15が、顕著度画像正規化部151と、正規化顕著度画像積算部152とによって構成される場合の基礎注目度画像抽出方法について述べる。

【0062】

顕著度画像正規化部151は、顕著度画像抽出部14により出力された顕著度画像の集合を入力し、各顕著度画像（輝度顕著度画像、色顕著度画像、方向顕著度画像、点滅顕著度画像、運動顕著度画像）それぞれについて正規化処理を施した画像である正規化顕著度画像を抽出し、正規化顕著度画像の集合を出力する。

顕著度画像に対する正規化処理は、前記解像度差分画像正規化部141による正規化処理と同様である。

【0063】

正規化顕著度画像積算部152は、顕著度画像正規化部151により出力された正規化顕著度画像の集合を入力し、それら正規化顕著度画像を積算することにより基礎注目度画像を抽出し、その基礎注目度画像を出力する。

本実施形態において、正規化顕著度画像 $N(CM_j(i))$  ( $j = I, C, O, F, M$ ) は、以下の(式26)のようにして積算される。ただし、 $\sim j$ は、式中で $j$ と区別するために用いているが、 $j$ と同様、 $\sim j = I, C, O, F, M$ である。

【0064】

【数13】

$$S(i) = \sum_{j=I,C,O,F,M} \frac{w_j(i-1)}{\sum_{\tilde{j}} w_{\tilde{j}}(i-1)} N(CM_{\tilde{j}}(i)) \quad (\text{式 26})$$

【0065】

ここで、 $S(i)$ は $i$ 番目の入力画像から抽出される基礎注目度画像、 $w_j(i)$ は後述の第3の実施形態の顕著度画像統合比率算出部2により抽出される顕著度画像 $CM_j(i)$ に対応する顕著度画像統合比率であるが、本実施形態及び後述する第2の実施形態では、顕著度画像統合比率算出部2を用いないため、 $w_j(i) = 1/5$  ( $i, j$ )とする。(は任意を示す。)

【0066】

上記の通り、基礎注目度画像抽出部1は、前記基礎注目度画像を抽出し、これを出力する。

【0067】

注目度瞬時抑制部3は、基礎注目度画像抽出部1により、現在及び1時点前の入力画像から算出された前記基礎注目度画像を入力し、1時点前の入力画像から算出された基礎注目度画像について、その基礎注目度画像の各画素の値である基礎注目度が最も大きい領域である最大基礎注目度領域を抑制することにより、瞬時抑制後注目度画像を抽出し、その瞬時抑制後注目度画像を出力する。

【0068】

瞬時抑制後注目度画像の抽出方法は特に限定されるものではないが、本実施形態においては、注目度瞬時抑制部3が、最大基礎注目度領域検出部31と、最大基礎注目度領域遮蔽画像抽出部32と、注目度漸進回復画像抽出部33と、注目度瞬時抑制画像生成部34と、瞬時抑制後注目度画像生成部35とによって構成される場合の瞬時抑制後注目度画像

10

20

30

40

50

抽出方法について述べる。

【0069】

最大基礎注目度領域検出部31は、1時点前の入力画像から算出された前記基礎注目度画像を入力し、入力された基礎注目度画像における前記最大基礎注目度領域を抽出し、この最大基礎注目度領域を出力する。

最大基礎注目度領域の抽出方法は特に限定されるものではないが、本実施形態においては、最大基礎注目度領域MSR(i-1)を、基礎注目度画像S(i-1)の画素値(基礎注目度)が最も大きい箇所((x), (y))を中心とした半径の円によって抽出する。すなわち、以下の(式27)のように表現される。ただし、arg maxは、直後の項を最大化する値を返す演算子である。例えば、(式27)においては、S(i-1)(x', y')を最大化する(x', y')を返す。

10

【0070】

【数14】

$$MSR(i-1) = \left\{ (x, y) : (x - \bar{x})^2 + (y - \bar{y})^2 \leq \epsilon^2, (\bar{x}, \bar{y}) = \arg \max_{(x', y')} S(i-1)_{(x', y')} \right\} \quad (\text{式 } 27)$$

【0071】

最大基礎注目度領域遮蔽画像抽出部32は、現在の入力画像から算出された前記基礎注目度画像、現在及びそれ以前の入力画像から算出された前記最大基礎注目度領域、及び1時点前における注目度瞬時抑制画像生成部34の出力である注目度瞬時抑制画像を入力し、最大基礎注目度領域を新たに遮蔽する画像である最大基礎注目度領域遮蔽画像を抽出し、この最大基礎注目度領域遮蔽画像を出力する。

20

最大基礎注目度領域遮蔽画像の抽出方法は特に限定されるものではないが、本実施形態においては、以下の方法によって抽出する。

【0072】

1時点前の注目度瞬時抑制画像生成部34における処理により、注目度瞬時抑制画像ID(i-1)がすでに得られているものとする。

ここで、現在の入力画像がi番目の入力画像であるとする。このとき、1時点前の最大基礎注目度領域MSR(i-1)を新たに遮蔽するようにID(i-1)を更新することで、現在の最大基礎注目度領域遮蔽画像ID<sub>1</sub>(i)を以下の(式28)~(式29)のように生成する。ただし、modは、除算の余りを示す。

30

【0073】

【数15】

$$ID_1(i) = \begin{cases} \max\{0, ID(i-1) - \mu\} & (x, y) \in MSR(i) \\ ID(i-1) & (x, y) \notin MSR(i) \end{cases} \quad (\text{if } i \equiv 0 \pmod{\Delta t_I}) \quad (\text{式 } 28)$$

$$ID_1(i) = ID(i-1) \quad (\text{if } i \not\equiv 0 \pmod{\Delta t_I}) \quad (\text{式 } 29)$$

40

【0074】

ここで、μ(0 < μ < 1)は、最大基礎注目度領域を遮蔽する度合を表現する係数であり、μ = 1のときには、その領域における基礎注目度の大きさによらずその領域を完全に遮蔽する。また、t<sub>I</sub>は最大基礎注目度領域による遮蔽の間隔を制御する定数である。特に、t<sub>I</sub> = 1のときには、すべてのフレームにおいて最大基礎注目度領域による遮蔽を行う。

【0075】

別の実施形態として、1時点前の最大基礎注目度領域遮蔽画像ID<sub>1</sub>(i-1)によって遮蔽されている領域を、その領域の動きに追従して移動させることも可能である。k番目の入力画像から算出された最大基礎注目度領域MSR(k)の時点i(i > k)での位

50

置  $MSR(i; k)$  は、以下の (式 30) ~ (式 32) のようにして算出される。

【 0 0 7 6 】

【 数 1 6 】

$$MSR(i; k) = \{(x + \Delta M_x(i; k), y + \Delta M_y(i; k)) : (x, y) \in MSR(i-1; k)\} \quad (\text{式 30})$$

$$\Delta M_x(i; k) = \sum_{(x', y') \in MSR(i-1; k)} M_x(i-1)_{(x', y')} \quad (\text{式 31})$$

$$\Delta M_y(i; k) = \sum_{(x', y') \in MSR(i-1; k)} M_y(i-1)_{(x', y')} \quad (\text{式 32})$$

10

【 0 0 7 7 】

ただし、 $MSR(i; i) = MSR(i)$  とする。このとき、最大基礎注目度領域遮蔽画像  $ID_1(i)$  は、以下の (式 33) ~ (式 35) のようにして算出される。

【 0 0 7 8 】

【 数 1 7 】

$$ID_1(i)_{(x,y)} = \begin{cases} \max\{0, ID(i-1)_{(x,y)} - \mu\} & \text{if } (x, y) \in MSR(i) \\ ID(i-1)_{(x', y')} & \text{else if } (x, y) \in MSR(i; k) \quad (k = i-1, \dots) \\ 1 & \text{else if } (x, y) \in MSR(i-1; k) \\ ID(i-1)_{(x,y)} & \text{otherwise} \end{cases} \quad (\text{式 33})$$

20

$$x' = x - \Delta M_x(i; k), \quad (\text{式 34})$$

$$y' = y - \Delta M_y(i; k). \quad (\text{式 35})$$

【 0 0 7 9 】

上記 (式 33) の第 1 行は 1 時点前の最大基礎注目度領域を遮蔽する操作、第 2 行は既に遮蔽されている最大基礎注目度領域を前記運動特徴画像の画素値を利用して移動させる操作、第 3 行は第 2 行で移動する前の領域の遮蔽を解除する操作である。

【 0 0 8 0 】

注目度漸進回復画像抽出部 33 は、最大基礎注目度領域遮蔽画像抽出部 32 の算出した前記最大基礎注目度領域遮蔽画像によって遮蔽されている領域における遮蔽を低減させる画像である注目度漸進回復画像を抽出し、この注目度漸進回復画像を出力する。

30

本実施形態において、注目度漸進回復画像  $ID_2(i)$  は、全ての画素値が (0 1) である画像とする。

【 0 0 8 1 】

【 数 1 8 】

$$ID_2(i) = \alpha \quad (\text{式 36})$$

【 0 0 8 2 】

注目度瞬時抑制画像生成部 34 は、最大基礎注目度領域遮蔽画像抽出部 32 の算出した前記最大基礎注目度領域遮蔽画像及び注目度漸進回復画像抽出部 33 の出力した注目度漸進回復画像を入力し、これら画像を統合することにより注目度瞬時抑制画像を生成し、この注目度瞬時抑制画像を出力する。

40

本実施形態において、注目度瞬時抑制画像  $ID(i)$  は、以下の (式 37) のようにして得られる。ただし、 $\min$  は、括弧  $\{ \}$  内の値のうちの最小値を示す。

【 0 0 8 3 】

【数 19】

$$ID(i) = \min\{1, ID_1(i) + ID_2(i)\} \quad (式 37)$$

【0084】

瞬時抑制後注目度画像生成部 35 は、前記注目度瞬時抑制画像及び前記基礎注目度画像を入力し、これら画像を統合することにより瞬時抑制後注目度画像を生成し、この瞬時抑制後注目度画像を出力する。

本実施形態において、瞬時抑制後注目度画像  $S_I(i)$  は、以下の(式 38)のようにして得られる。

【0085】

【数 20】

$$S_I(i) = (ID(i))^{\omega_I(i)} \cdot S(i) \quad (式 38)$$

【0086】

ここで、 $\omega_I(i)$  は注目度瞬時抑制画像  $ID(i)$  に対する重みを表現する係数である。

【0087】

上記の通り、注目度瞬時抑制部 3 は、前記瞬時抑制後注目度画像を抽出し、これを出力する。

【0088】

注目度映像出力部 5 は、前記基礎注目度画像抽出部 1 ~ 注目度瞬時抑制部 3 を、各入力画像について順に繰り返して実行することにより抽出された前記瞬時抑制後注目度画像の時系列である瞬時抑制後注目度映像を抽出し、これを注目度映像として出力する。

【0089】

図 2 に、本実施形態の動作例を示す。

同図において、上段(a)は入力画像、中段の(b)は注目度瞬時抑制画像、下段の(c)は瞬時抑制後注目度画像であり、それぞれ左から時系列順に整列している。

【0090】

[第 2 の実施形態]

図 3 は、本発明の第 2 の実施形態による注目領域抽出装置の機能ブロック図である。

本実施形態に示す注目領域抽出装置は、基礎注目度画像抽出部 1 と、注目度瞬時抑制部 3 と、注目度漸進抑制部 4 と、注目度映像出力部 5 とで構成され、注目度抽出の対象となる入力映像を入力し、入力画像の中で注目度の高い領域を表示した映像である注目度映像を出力する。なお、注目領域抽出装置を、注目度瞬時抑制部 3 を用いずに、基礎注目度画像抽出部 1 と、注目度漸進抑制部 4 と、注目度映像出力部 5 とで構成することも可能である。同図において、第 1 の実施形態と同様の構成は同じ符号を付し、説明を省略する。基礎注目度画像抽出部 1、及び注目度瞬時抑制部 3 は、第 1 の実施形態と同様である。

【0091】

注目度漸進抑制部 4 は、現在及びそれ以前の入力画像のうちいくつかから算出された前記顕著度画像、及び前記基礎注目度画像もしくは前記瞬時抑制後注目度画像を入力し、顕著度画像について時間軸方向で顕著な値を持つ領域を抽出し、その領域に対応する基礎注目度画像もしくは瞬時抑制後注目度画像における領域を強調し、そうではない領域を抑制することにより、漸進抑制後注目度画像を抽出し、その漸進抑制後注目度画像を出力する。

【0092】

漸進抑制後注目度画像の抽出方法は特に限定されるものではないが、本実施形態においては、注目度漸進抑制部 4 が、注目度漸進遮蔽画像生成部 41 と、注目度瞬時回復画像生成部 42 と、注目度漸進抑制画像生成部 43 と、漸進抑制後注目度画像生成部 44 とによ

10

20

30

40

50

って構成される場合の漸進抑制後注目度画像抽出方法について述べる。

【0093】

注目度漸進遮蔽画像生成部41は、前記基礎注目度画像を漸進的に遮蔽する画像である注目度漸進遮蔽画像を生成し、この注目度漸進遮蔽画像を出力する。

注目度漸進遮蔽画像の生成方法は、特に限定されるものではないが、本実施形態においては、基礎注目度画像 $S(i)$ の全ての画素値を1時点ごとに $(0 < \beta < 1)$ ずつ減少させることで以下の(式39)のように注目度漸進遮蔽画像 $GD_1(i)$ を生成する。

【0094】

【数21】

$$GD_1(i) = 1 - i\beta \quad (\text{式 39})$$

10

【0095】

注目度瞬時回復画像生成部42は、現在及びそれ以前の入力画像のうちいくつかの入力画像から算出された前記顕著度画像を入力し、それら顕著度画像について、時間軸方向で顕著な値を持つ領域を抽出し、その領域に対応する前記基礎注目度画像もしくは前記瞬時抑制後注目度画像における領域の基礎注目度の抑制を解除させる画像である注目度瞬時回復画像を生成し、この注目度瞬時回復画像を出力する。

注目度瞬時回復画像の生成方法は、特に限定されるものではないが、本実施形態においては、注目度瞬時回復画像生成部42が、時間顕著度画像生成部421と、時間顕著度画像二値化部422とによって構成される場合の注目度瞬時回復画像生成方法について述べる。

20

【0096】

時間顕著度画像生成部421は、現在及びそれ以前の入力画像のうちいくつかの入力画像から算出された前記顕著度画像を入力し、それら顕著度画像について、時間軸方向で顕著な値を持つ領域を示す画像である時間顕著度画像を生成し、この時間顕著度画像を出力する。

時間顕著度画像の生成方法は、特に限定されるものではないが、本実施形態においては、以下、2通りの方法について記述する。

【0097】

第1の時間顕著度画像生成方法では、L. Itti and P. Baldi, "A principled approach to detecting surprising events in video," in Proc. Conference on Computer Vision and Pattern Recognition (CVPR), pp. 631-637, June 2005. (非特許文献5)に記載の方法に基づき、既定の確率モデルとの一致度を利用して時間顕著度画像を生成する。

30

【0098】

以下の(式40)に示すガンマ分布の確率密度関数 $P(\lambda; \rho_1, \rho_2)$ について、1時点前の入力画像を用いた本処理部における出力により、この確率密度関数のパラメータであるガンマ分布係数 $(\rho_1, \rho_2) = (\rho_1(i-1), \rho_2(i-1))$ が決定されているものとする。

【0099】

【数22】

$$P_\gamma(\lambda; \rho_1, \rho_2) = \frac{\rho_1(\rho_1\lambda)^{\rho_2-1} \exp(-\rho_1\lambda)}{\Gamma(\rho_2)} \quad (\text{式 40})$$

40

【0100】

ここで、 $\Gamma(\cdot)$ はガンマ関数である。

【0101】

第1の時間顕著度画像生成方法の基本的な考え方は、各解像度差分画像の各画素に対応する形でそれぞれ上記のガンマ分布係数 $\rho_1, \rho_2$ が保持されており、各時点の解像度差

50

分画像の画素値に応じてガンマ分布係数を更新しながら、時間顕著度画像を生成する枠組である。以下、簡単のため、ある  $(c, s)$  レベル輝度解像度差分画像  $RS_I(i; c, s)$  を例として処理を説明する。

【0102】

まず、輝度解像度差分画像  $RS_I(i; c, s)$  の各画素  $(x, y)$  について、その画素におけるガンマ分布の入力の推定値であるガンマ分布推定入力係数  $\rho_{1,I}(i)(x, y)$  を、輝度解像度差分画像における(空間的・時間的双方の意味を含めた)注目画素周辺の画素値及び注目画素に対応するガンマ分布係数  $\rho_{1,I}(i-1)(x, y)$ 、 $\rho_{2,I}(i-1)(x, y)$  から導出する。次に、導出されたガンマ分布推定入力係数  $\rho_{1,I}(i)(x, y)$  を用いてガンマ分布係数  $\rho_{1,I}(i-1)(x, y)$ 、 $\rho_{2,I}(i-1)(x, y)$  を以下の(式41)、(式42)のように更新する。

10

【0103】

【数23】

$$\rho_{1,I}(i) = \zeta \cdot \rho_{1,I}(i) + \lambda_I(i) \quad (\text{式41})$$

$$\rho_{2,I}(i) = \zeta \cdot \rho_{2,I}(i) + 1 \quad (\text{式42})$$

【0104】

ここで、 $\zeta$  ( $0 < \zeta < 1$ ) は係数更新における履歴の反映率を表現する係数である。上記の処理により導出されたガンマ分布係数から、輝度成分についての時間方向の顕著性を表す画像である  $(c, s)$  レベル輝度時間顕著度画像  $SP_I(i; c, s)$  を以下のように導出する。

20

【0105】

【数24】

$$SP_I(i; c, s) = D(P_\gamma(\lambda(i); \rho_1(i-1), \rho_2(i-1)) \| P_\gamma(\lambda(i); \rho_1(i), \rho_2(i))) \quad (\text{式43})$$

$$= \rho_2(i) \log \frac{\rho_1(i-1)}{\rho_1(i)} + \log \frac{\Gamma(\rho_2(i))}{\Gamma(\rho_2(i-1))} + \rho_1(i) \frac{\rho_2(i-1)}{\rho_1(i-1)} + (\rho_2(i-1) - \rho_2(i)) \Psi(\rho_2(i-1)) \quad (\text{式44})$$

30

【0106】

ここで、 $D(P \| P')$  は確率密度関数  $P$  と  $P'$  との間のKullback-Liebler divergenceである。また、 $\Psi(x)$  はdigamma functionと呼ばれ、以下の(式45)で表される。

【0107】

【数25】

$$\Psi(x) = \frac{d}{dx} \ln \Gamma(x) \quad (\text{式45})$$

40

【0108】

なお、表記の簡略化のため、輝度成分を表す  $I$  及び注目画素を示す  $(x, y)$  は省略している。

【0109】

このようにして導出された輝度時間顕著度画像  $SP_I(i)$  と、同様の処理によって導出された色時間顕著度画像  $SP_C(i)$ ・方向時間顕著度画像  $SP_O(i)$ ・点滅時間顕著度画像  $SP_F(i)$ ・運動時間顕著度画像  $SP_M(i)$  を統合して、時間顕著度画像  $SP(i)$  を生成する。具体的には、以下の(式46)～(式51)のようにして生成する。

【0110】

50

【数 2 6】

$$SP(i) = \sum_{j=I,C,O,F,M} SP_j(i) \quad (式 46)$$

$$SP_I(i) = \sum_{c \in L_c} \sum_{s \in L_s} N(SP_I(i; c, s)) \quad (式 47)$$

$$SP_C(i) = \sum_{c \in L_c} \sum_{s \in L_s} \{N(SP_{RC}(i; c, s)) + N(SP_{BY}(i; c, s))\} \quad (式 48)$$

$$SP_O(i) = \sum_{k=0}^{n_\phi-1} N \left( \sum_{c \in L_c} \sum_{s \in L_s} N(SP_O(i; k\pi/n_\phi; c, s)) \right) \quad (式 49)$$

$$SP_F(i) = \sum_{c \in L_c} \sum_{s \in L_s} N(SP_F(i; c, s)) \quad (式 50)$$

$$SP_M(i) = \sum_{k=x,y} N \left( \sum_{c \in L_c} \sum_{s \in L_s} N(SP_{M_k}(i; c, s)) \right) \quad (式 51)$$

10

【0 1 1 1】

これに対し、第 2 の時間顕著度画像生成方法では、解像度差分画像抽出部 1 3 と類似の方法に基づき、時間軸における重み付き差分処理を利用して時間顕著度画像を生成する。

まず、現在及びそれ以前  $n_T$  時点前までの各種類（輝度・色など）の顕著度画像  $CM_j(i-t)$  ( $0 \leq t \leq n_T$ ) から、ガウス分布によって重み付けされた時間軸方向の差分処理によって、時間応答画像  $T_j(i)$  を (式 5 2) のように算出する。

【0 1 1 2】

【数 2 7】

$$T_j(i) = \sum_{c \in L_c} \sum_{s \in L_s} \sum_{t=0}^{n_T} \{\eta_{c/2}(t) - \eta_{s/2}(t)\} \cdot CM_j(i-t) \quad (式 52)$$

30

【0 1 1 3】

ここで、 $\eta(t)$  は、平均 0、分散  $\sigma^2$  を持つガウス分布の確率密度関数である。上記時間応答画像を統合することで、時間顕著度画像  $SP(i)$  を (式 5 3) のように算出する。

【0 1 1 4】

【数 2 8】

$$SP(i) = \sum_{j=I,C,O,F,M} T_j(i) \quad (式 53)$$

40

【0 1 1 5】

時間顕著度画像二値化部 4 2 2 は、前記時間顕著度画像を入力し、時間顕著度画像を二値化することによって注目度瞬時回復画像を生成し、この注目度瞬時回復画像を出力する。

本実施形態において、時間顕著度画像  $SP(i)$  は以下のように二値化され、注目度瞬時回復画像  $GD_2(i)$  が (式 5 4) のように生成される。

【0 1 1 6】

【数 2 9】

$$GD_2(i) = \begin{cases} 1 & SP(i) \geq \theta \\ 0 & \text{otherwise} \end{cases} \quad (\text{式 54})$$

【0 1 1 7】

ここで、 $\theta$  は時間顕著度画像の二値化における閾値であり、予め決められる。

【0 1 1 8】

上記の通り、注目度瞬時回復画像生成部 4 2 は、注目度瞬時回復画像  $GD_2(i)$  を生成し、これを出力する。 10

【0 1 1 9】

注目度漸進抑制画像生成部 4 3 は、前記注目度漸進遮蔽画像及び前記注目度瞬時回復画像を入力し、これら画像を統合することにより注目度漸進抑制画像を生成し、この注目度漸進抑制画像を出力する。

本実施形態において、注目度漸進抑制画像  $GD(i)$  は、以下の(式 5 5)のようして得られる。

【0 1 2 0】

【数 3 0】

$$GD(i) = \min\{1, GD_1(i) + GD_2(i)\} \quad (\text{式 55}) \quad 20$$

【0 1 2 1】

漸進抑制後注目度画像生成部 4 4 は、前記注目度漸進抑制画像、及び前記基礎注目度画像または前記瞬時抑制後注目度画像を入力し、これら画像を統合することにより漸進抑制後注目度画像を生成し、この漸進抑制後注目度画像を出力する。

漸進抑制後注目度画像の生成方法は、特に限定されるものではないが、本実施形態においては、瞬時抑制後注目度画像  $S_I(i)$  と注目度漸進抑制画像  $GD(i)$  とを統合して、漸進抑制後注目度画像  $S_L(i)$  を(式 5 6) ~ (式 5 7)のよう生成する。

【0 1 2 2】

【数 3 1】

$$S_L(i) = (GD(i))^{\omega_G(i)} \cdot S_I(i) \quad (\text{式 56})$$

$$= (ID(i))^{\omega_I(i)} \cdot (GD(i))^{\omega_G(i)} \cdot S(i) \quad (\text{式 57})$$

【0 1 2 3】

ここで、 $\omega_G(i)$  は注目度漸進抑制画像  $GD(i)$  に対する重みを表現する係数である。上記に示すように、注目度瞬時抑制部 3 の出力する瞬時抑制後注目度画像  $S_I(i)$  を用いた式を、基礎注目度画像抽出部 1 の出力する基礎注目度画像  $S(i)$  を用いた式により算出することが可能である。ここで、 $S_I(i) = 0$  のときには、注目度瞬時抑制部 3 を用いない場合と等価になる。 40

【0 1 2 4】

上記の通り、注目度漸進抑制部 4 は、前記漸進抑制後注目度画像を抽出し、これを出力する。

【0 1 2 5】

注目度映像出力部 5 は、前記基礎注目度画像抽出部 1 ~ 注目度漸進抑制部 4 を、各入力画像に対して順に繰り返して実行することにより抽出された前記漸進抑制後注目度画像の時系列である漸進抑制後注目度映像を抽出し、これを注目度映像として出力する。

【0 1 2 6】

50

図4は、本実施形態の動作例を示す。同図において、(a)は入力画像、(b)は注目度漸進抑制画像、(c)は漸進抑制後注目度画像であり、それぞれ左から時系列順に整理している。

【0127】

[第3の実施形態]

図5は、本発明の第3の実施形態による注目領域抽出装置の機能ブロック図である。

本実施形態に示す注目領域抽出装置は、基礎注目度画像抽出部1と、顕著度画像統合比率算出部2と、注目度瞬時抑制部3と、注目度漸進抑制部4と、注目度映像出力部5とで構成され、注目度抽出の対象となる入力映像を入力し、入力画像の中で注目度の高い領域を表示した映像である注目度映像を出力する。同図において、第1、第2の実施形態と同様の構成は同じ符号を付し、説明を省略する。基礎注目度画像抽出部1、注目度瞬時抑制部3、注目度漸進抑制部4、及び注目度映像出力部5は、第1もしくは第2の実施形態と同様である。

10

【0128】

顕著度画像統合比率算出部2は、それぞれ現在の入力画像から抽出された前記顕著度画像の集合及び前記基礎注目度画像を入力し、前記基礎注目度画像から最大基礎注目度領域を抽出する。そして、各顕著度画像について抽出した最大基礎注目度領域に対応する領域の中の値を算出し、その値の大きさから、対応する顕著度画像の重みである顕著度画像統合比率を決定し、その顕著度画像統合比率の集合を出力する。

【0129】

顕著度画像統合比率の計算方法は、特に限定されるものではないが、本実施形態においては、V. Navalpakkam and L. Itti: "Optimal cue selection strategy," in Advances in Neural Information Processing Systems (NIPS), pp.987-994, December 2005. (非特許文献6)に記載の方法に基づき、最大基礎注目度領域における各顕著度画像の画素値を用いて顕著度画像統合比率を逐次更新していく。

20

【0130】

1時点前の入力画像を用いた本処理部における処理により、各顕著度画像 $CM_j(i-1)$  ( $j = I, C, O, F, M$ )に対応する顕著度画像統合比率 $w_j(i-1)$ が既に得られているものとする。まず、基礎注目度画像 $S(i)$ から最大基礎注目度領域 $MSR(i)$ を抽出する。抽出方法は、前述の最大基礎注目度領域検出部31における抽出方法と同様である。次に、各顕著度画像 $CM_j(i)$ に対応する顕著度画像統合比率 $w_j(i)$ を、以下のように決定する。

30

【0131】

【数32】

$$w_j(i) = w_j(i-1) + \delta \cdot \Delta w_j(i) \quad (\text{式58})$$

$$\Delta w_j(i) = \frac{\sum_{(x,y) \in MSR(i)} CM_j(i)(x,y)}{\sum_{(x,y) \notin MSR(i)} CM_j(i)(x,y)} \quad (\text{式59})$$

40

【0132】

ここで、 $\delta$ は重み更新における履歴の反映比率を与える定数である。

【0133】

続いて、本発明の実施形態を用いた実験データを以下に示す。

入力映像として、大きさ $640 \times 480$ ピクセル、長さ $8 \sim 15$ 秒の映像6種類を用意した。また、各実施形態に示した記号の実際の数値として、以下の値を用いた。

【0134】

$n = 4$ ,  $n_F = 3$ ,  $\alpha = 1.25$ ,  $n_L = 8$ ,  $L_C = \{2, 3, 4\}$ ,  $L_S = \{c+3, c+4\}$  ( $c \in L_C$ ),  $n_A = 32 \times 24 = 768$ ,  $\beta = 2.5$ ,  $\mu$

50

$$= 1.0, \quad t_I = 10, \quad = 1/t_I = 0.1, \quad = 0.0025, \quad t_T = 8, \\ = 0.25 \max(x, y) SP(i)(x, y), \quad = 0.1$$

【0135】

本発明の効果を確認するため、本発明の実施形態及び既知の方法によって得られる注目度映像がどの程度人間の視覚特性を模擬できているかを比較した。人間の視覚特性を表現する数量として、人間が実際に入力映像を見ている際の視線の位置を採用した。5名の被験者に入力映像を提示し、既存の視線測定装置を用いて各被験者の入力映像中の注視位置を逐次測定した。各被験者について、1種類の入力映像を2回提示した。これにより、各被験者・各入力映像について、注視位置の時系列を2本獲得した。この注視位置の時系列を、時刻の整合性を保ちながら入力映像の各フレーム（すなわち入力画像）に対応付けることで、各被験者・各入力画像について2通りの注視位置を獲得した。

10

【0136】

人間の視覚特性を模擬できているかどうかの評価尺度として、被験者の注視位置における注目度を採用した。本発明の方法もしくは既知の方法により入力映像から抽出した注目度映像について、入力画像  $IN(i; k)$  ( $k = 1, 2, \dots, 6$ : 映像の種類に対応) に対応する注目度映像のフレームである注目度画像  $S_F(i; k)$  を考える。本発明の第1の実施形態においては瞬時抑制後注目度画像が注目度画像と等しい、すなわち  $S_F(i; k) = S_I(i; k)$  であり、本発明の第2の実施形態においては漸進抑制後注目度画像が注目度画像と等しい、すなわち  $S_F(i; k) = S_I(i; k)$  である。入力画像  $IN(n; k)$  における被験者  $n$  ( $n = 1, 2, \dots, 5$ ) の注視位置を  $(x(i; k, n), y(i; k, n))$  と表記するとき、被験者  $n$  を「教師」とする注目度画像  $S_F(i; k)$  の評価値  $V(i; k, n)$  を以下の(式60)で定めた。

20

【0137】

【数33】

$$V(i; k, n) = \frac{S_F(i)(x(i; k, n), y(i; k, n))}{\sum_{(x', y')} S_F(i)(x', y')} \quad (\text{式 60})$$

【0138】

上記(式60)の右辺において、分母は、注目度画像を正規化するために付与されている。注目度映像  $\{S_F(i; k)\}_i$  の評価値  $V(k)$  は、(式61)に示すように、各注目度画像についての評価値を総計し、さらに被験者について平均を取ることを得る。 $n_E$  は被験者の数（すなわち、本評価例においては「5」）である。

30

【0139】

【数34】

$$V(k) = \frac{1}{n_E} \sum_{n=1}^{n_E} \sum_i V(i; k, n) \quad (\text{式 61})$$

40

【0140】

上記の評価値により、本発明の第1～第3の実施形態による方法と既知の方法とを比較した。既知の方法として、非特許文献1に記載の方法、及び非特許文献2に記載の方法を用いた。図6は、非特許文献1に記載の方法、非特許文献2に記載の方法、第1の実施形態による方法（本発明の実施形態1）、第2の実施形態による方法（本発明の実施形態2）、第3の実施形態による方法（本発明の実施形態3）を用いた場合の入力映像（Video 1～6）ごとの評価値（図中ではNETR Value）を比較したグラフ、図7に各方法毎に全入力映像について平均を取った評価値を比較したグラフを示す。図7では、左から非特許文献1に記載の方法（Still image algorithm）、非特許文献2に記載の方

50

法 (Moving algorithm)、第 1 の実施形態による方法 (Case 1)、第 2 の実施形態による方法 (Case 2)、第 3 の実施形態による方法 (Case 3) の平均評価値を示している。ここで、本発明の第 1 ~ 第 3 の実施形態による方法では、漸進抑制後注目度画像生成部 4 4 に注目した、3 通りの設定について評価している。第 1 の設定 (図中、本発明の第 1 の実施形態による方法) では、漸進抑制後注目度画像生成部 4 4 において  $(I(i), G(i)) = (1, 0)$   $i$  とした。第 2 の設定 (図中、本発明の第 2 の実施形態による方法) では、漸進抑制後注目度画像生成部 4 4 において  $(I(i), G(i)) = (0, 1)$   $i$  とした。第 3 の設定 (図中、本発明の第 3 の実施形態による方法) では、漸進抑制後注目度画像生成部 4 4 において  $(I(i), G(i)) = (1, 1)$   $i$  とした。図 7 からわかるように、入力映像について平均したときに、本発明の第 2 の実施形態の設定が最も良い評価値を示した。また、図 6 からわかるように、本発明の第 2 の実施形態の設定は、いずれの入力映像についてもその他の方法を上回る評価値を示した。さらに、図 6 からわかるように、本発明の第 1 の及び第 3 の実施形態の設定は、いくつかの入力映像について既知の方法を上回る評価値を示した。

#### 【0141】

なお、上述の注目領域抽出装置は、内部にコンピュータシステムを有している。そして、注目領域抽出装置の基礎注目度画像抽出部 1、顕著度画像統合比率算出部 2、注目度瞬時抑制部 3、注目度漸進抑制部 4、及び、注目度映像出力部 5 の動作の過程は、プログラムの形式でコンピュータ読み取り可能な記録媒体に記憶されており、このプログラムをコンピュータシステムが読み出して実行することによって、上記処理が行われる。ここでいうコンピュータシステムとは、CPU 及び各種メモリや OS、周辺機器等のハードウェアを含むものである。

#### 【0142】

また、「コンピュータシステム」は、WWW システムを利用している場合であれば、ホームページ提供環境 (あるいは表示環境) も含むものとする。

また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM 等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間の間、動的にプログラムを保持するもの、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含むものとする。また上記プログラムは、前述した機能の一部を実現するためのものであっても良く、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであっても良い。

#### 【図面の簡単な説明】

#### 【0143】

【図 1】本発明の第 1 の実施形態による注目領域抽出装置の構成を示すブロック図である。

【図 2】図 1 に示す注目領域抽出装置の動作例を示す図である。

【図 3】本発明の第 2 の実施形態による注目領域抽出装置の構成を示すブロック図である。

【図 4】図 3 に示す注目領域抽出装置の動作例を示す図である。

【図 5】本発明の第 3 の実施形態による注目領域抽出装置の構成を示すブロック図である。

【図 6】本発明と従来技術を用いた方法による入力映像ごとの評価値を示すグラフである。

【図 7】図 6 に示す各方法の評価値の平均を示すグラフである。

#### 【符号の説明】

#### 【0144】

10

20

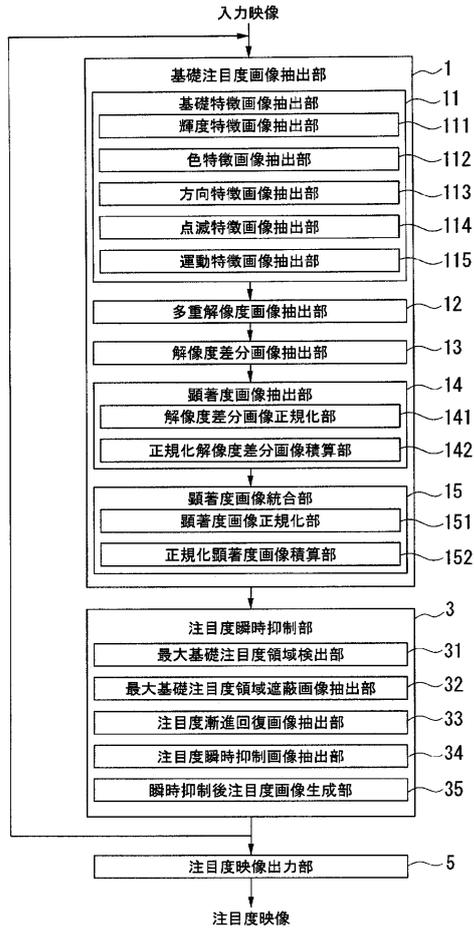
30

40

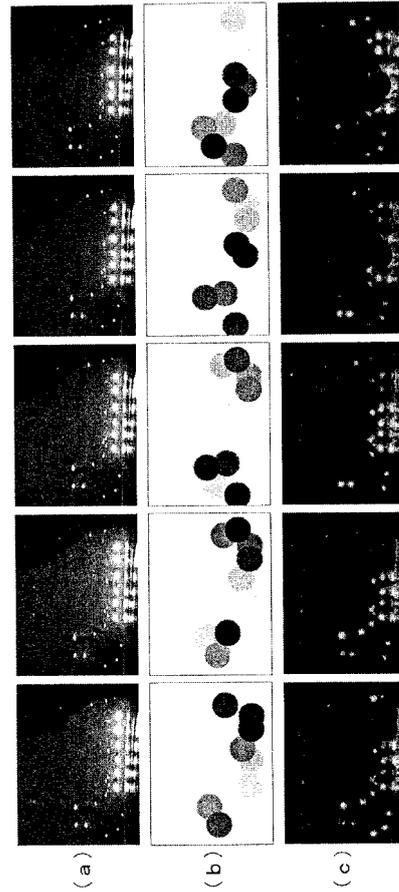
50

1 ... 基礎注目度画像抽出部	
1 1 ... 基礎特徴画像抽出部	
1 1 1 ... 輝度特徴画像抽出部	
1 1 2 ... 色特徴画像抽出部	
1 1 3 ... 方向特徴画像抽出部	
1 1 4 ... 点滅特徴画像抽出部	
1 1 5 ... 運動特徴画像抽出部	
1 2 ... 多重解像度画像抽出部	
1 3 ... 解像度差分画像抽出部	
1 4 ... 顕著度画像抽出部	10
1 4 1 ... 解像度差分画像正規化部	
1 4 2 ... 正規化解像度差分画像積算部	
1 5 ... 顕著度画像統合部	
2 ... 顕著度画像統合比率算出部	
3 ... 注目度瞬時抑制部	
3 1 ... 最大基礎注目度領域検出部	
3 2 ... 最大基礎注目度領域遮蔽画像抽出部	
3 3 ... 注目度漸進回復画像抽出部	
3 4 ... 注目度瞬時抑制画像生成部	
3 5 ... 瞬時抑制後注目度画像生成部	20
4 ... 注目度漸進抑制部	
4 1 ... 注目度漸進遮蔽画像生成部	
4 2 ... 注目度瞬時回復画像生成部	
4 3 ... 注目度漸進抑制画像生成部	
4 4 ... 漸進抑制後注目度画像生成部	
5 ... 注目度映像出力部	

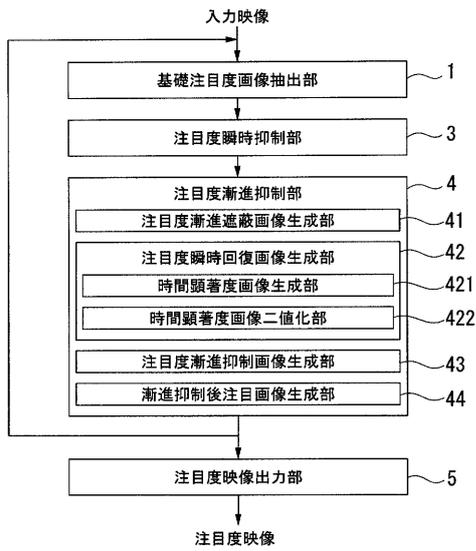
【 図 1 】



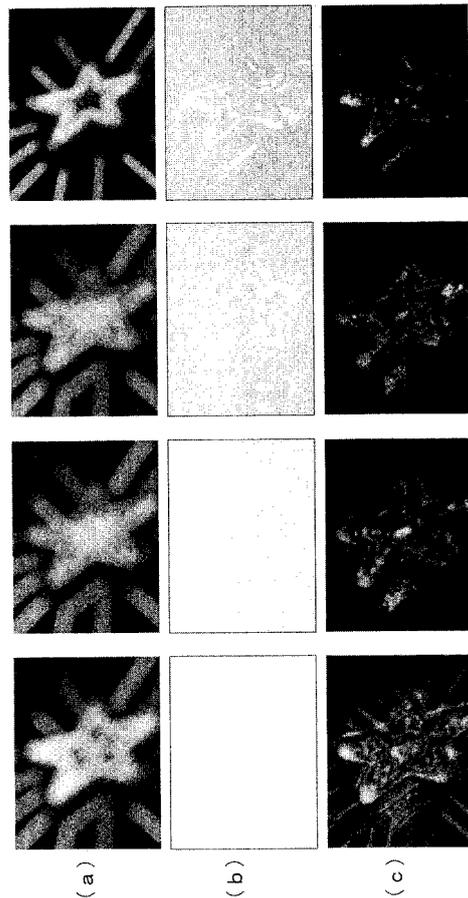
【 図 2 】



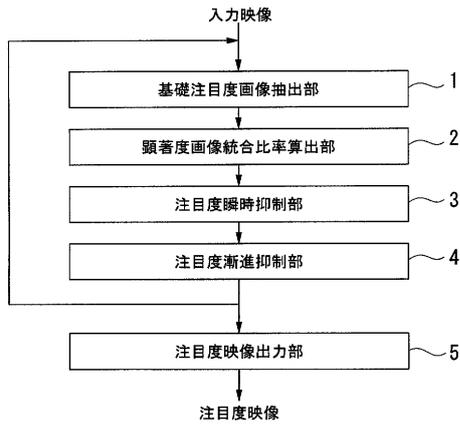
【 図 3 】



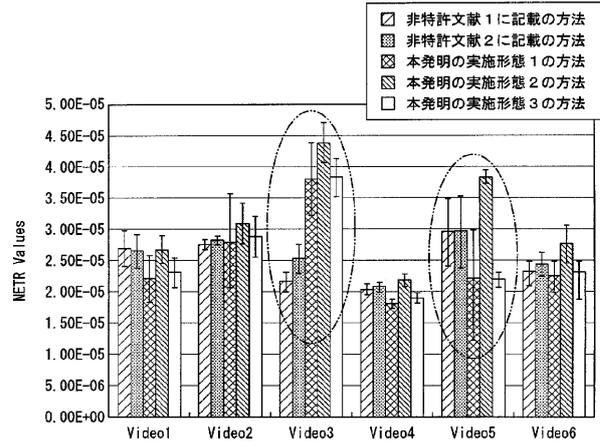
【 図 4 】



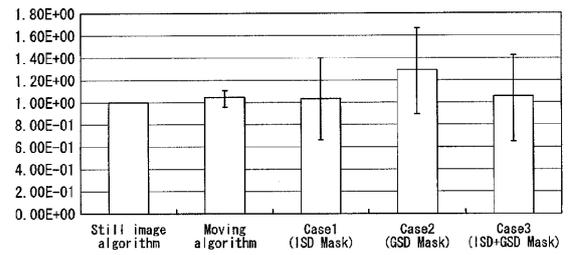
【 図 5 】



【 図 6 】



【 図 7 】



---

フロントページの続き

(72)発明者 竹内 龍人

東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

(72)発明者 クレメント レオン

カナダ ブイ6エヌ 2ダブリュー7 プリティッシュコロンビア州 バンクーバー ダブリュー  
37アベニュー 4098

Fターム(参考) 5L096 AA02 AA06 CA04 HA02