

(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2009-44422  
(P2009-44422A)

(43) 公開日 平成21年2月26日(2009.2.26)

(51) Int. Cl. F I テーマコード (参考)  
**HO4N 7/32 (2006.01)** HO4N 7/137 Z 5C059  
 5C159

審査請求 未請求 請求項の数 15 O L (全 32 頁)

<p>(21) 出願番号 特願2007-206777 (P2007-206777)</p> <p>(22) 出願日 平成19年8月8日(2007.8.8)</p> <p>特許法第30条第1項適用申請有り 平成19年2月9日 国立大学法人 電気通信大学主催の「電気通信大学修士論文発表会」に文書をもって発表</p>	<p>(71) 出願人 504133110 国立大学法人 電気通信大学 東京都調布市調布ヶ丘1丁目5番地1</p> <p>(71) 出願人 000201113 船井電機株式会社 大阪府大東市中垣内7丁目7番1号</p> <p>(74) 代理人 110000626 特許業務法人 英知国際特許事務所</p> <p>(72) 発明者 森田 啓義 東京都調布市調布ヶ丘1丁目5番地1 国立大学法人 電気通信大学内</p> <p>(72) 発明者 坂井 敦典 東京都調布市調布ヶ丘1丁目5番地1 国立大学法人 電気通信大学内</p> <p style="text-align: right;">最終頁に続く</p>
--	---

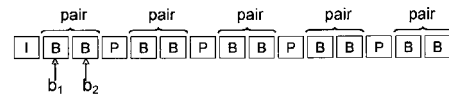
(54) 【発明の名称】 カット部検出システム及びショット検出システム並びにシーン検出システム、カット部検出方法

(57) 【要約】

【課題】 リプレイショットの先頭部分又は最後部分であるリプレイカット部の検出し、リプレイショットおよびその他の特定のショットを精度良く判別でき、更にはシーンを検出する可能なシステムを提供する。

【解決手段】 マクロブロックタイプ情報を含む動画圧縮データから、少なくともBフレーム毎にマクロブロックタイプ情報を抽出する手段と、この抽出されたマクロブロックタイプ情報からフレーム毎のイントラ符号化マクロブロックの数を求める手段と、Bフレームに含まれるイントラ符号化マクロブロックの頻度的特徴を判別条件にして、このBフレーム又はこのBフレームを含むグループオブピクチャを、リプレイショットの先頭部分又は最後部分であるリプレイカット部と判別するリプレイカット部判別手段とを含む。

【選択図】 図13



<瞬時カット点のFrameTypeパターン>

- ・b1: 0 or b2: 3
- ・pair: (1,1), (1,4), (4,4), (1,2), (1,3), (3,4), (3,4)

<リプレイカット部のFrameTypeパターン>

- ・b1 or b2: \*
- ・b1 or b2: \*となるpairが2GOP間で3pair以上ある

**【特許請求の範囲】****【請求項 1】**

マクロブロックタイプ情報を含む動画圧縮データから、少なくとも B フレーム毎にマクロブロックタイプ情報を抽出する手段と、

この抽出されたマクロブロックタイプ情報からフレーム毎のイントラ符号化マクロブロックの数を求める手段と、

B フレームに含まれるイントラ符号化マクロブロックの頻度的特徴を判別条件にして、この B フレーム又はこの B フレームを含むグループオブピクチャを、リプレイショットの先頭部分又は最後部分であるリプレイカット部と判別する手段と、を含むことを特徴とするカット部検出システム。

10

**【請求項 2】**

前記判別条件には、連続する二つの B フレームの内の少なくとも一方におけるイントラ符号化マクロブロックの数が所定数以上であるという条件が含まれていることを特徴とする請求項 1 記載のカット部検出システム。

**【請求項 3】**

前記判別条件には、連続する二つの B フレームの内の少なくとも一方におけるイントラ符号化マクロブロックの数が所定範囲内であって、このパターンとなる前記二つの B フレームが、複数のグループオブピクチャ中で所定数以上あるという条件が含まれていることを特徴とする請求項 1 又は 2 記載のカット部検出システム。

**【請求項 4】**

上記マクロブロックタイプ情報からフレーム毎の順方向予測符号化マクロブロックの数およびフレーム毎の逆方向予測符号化マクロブロックの数を求める手段を備え、

前記判別条件に加える条件として、

順方向予測符号化マクロブロックの数が最頻度で、且つ逆方向予測符号化マクロブロックの数とイントラ符号化マクロブロックの数との合計が所定数以上であるという条件と、

逆方向予測符号化マクロブロックの数が最頻度で、且つ順方向予測符号化マクロブロックの数とイントラ符号化マクロブロックの数との合計が所定数以上であるという条件と、

イントラ符号化マクロブロックの数が最頻度であるという条件と、

の三つ条件の内の何れかを含むことを特徴とする請求項 1 乃至 3 何れか 1 項記載のカット部検出システム。

20

30

**【請求項 5】**

前記カット部検出システムより検出される複数のリプレイカット部の内、連続する二つのリプレイカット部の間を、先の映像の繰り返しを映しているリプレイショットと判別することを特徴とする請求項 1 乃至 4 何れか 1 項記載のカット部検出システムを用いたショット検出システム。

**【請求項 6】**

前記カット部検出システムより検出される複数のリプレイカット部の内、連続する二つのリプレイカット部が、連続する二つの瞬時カット点の間に位置する場合に、これら連続する二つのリプレイカット部の間を、先の映像の繰り返しを映しているリプレイショットと判別することを特徴とする請求項 1 乃至 4 何れか 1 項記載のカット部検出システムを用いたショット検出システム。

40

**【請求項 7】**

前記連続する二つのリプレイカット部の内の一方と、この一方のリプレイカット部に連続する瞬時カット点との間に、所定数以上のグループオブピクチャを有するという条件を、前記リプレイショットの判別条件に加えたことを特徴とする請求項 5 又は 6 記載のショット検出システム。

**【請求項 8】**

前記連続する二つのリプレイカット部の間に所定数以上のグループオブピクチャを有するという条件を、前記リプレイショットの判別条件に加えたことを特徴とする請求項 5 乃至 7 何れか 1 項記載のショット検出システム。

50

## 【請求項 9】

動画圧縮データから抽出されたショットに対し、フレーム毎に動きベクトル情報を抽出する手段と、

この動きベクトル情報からフレーム毎の動きベクトル量を求める手段と、

前記抽出されたショットについて、その1ショット分の前記動きベクトル量をグループオブピクチャ数で割った平均の動きベクトル量を求める手段とを備え、

前記平均の動きベクトル量が所定範囲内であることを判別条件にして、前記抽出されたショットを特定のショットとして判別することを特徴とする請求項 1 乃至 8 何れか1項記載のショット検出システム。

## 【請求項 10】

動画圧縮データから抽出されたショットに対し、フレーム毎に動きベクトル情報を抽出する手段と、

各フレームのマクロブロック毎に動きベクトルの方向を所定数の方向に量子化する手段と、

前記量子化により分類された動きベクトルの方向について、その分散度をフレーム毎に求める手段と、

前記抽出されたショットについて、その1ショット分の前記分散度をグループオブピクチャ数で割った平均の分散度を求める手段とを備え、

前記分散度が所定範囲内であることを判別条件にして、前記抽出されたショットを特定のショットとして判別することを特徴とする請求項 1 乃至 9 何れか 1 項記載のショット検出システム。

## 【請求項 11】

動画圧縮データから、Iフレーム毎に輝度成分 Y、色差成分 C<sub>b</sub>、色差成分 C<sub>r</sub> を抽出する手段と、

各Iフレームを平行な複数のラインに分割し、そのライン毎に、輝度成分 Y、色差成分 C<sub>b</sub>、色差成分 C<sub>r</sub>、色差成分の比  $r = C_b / C_r$  が、それぞれ所定範囲内であるという色条件を満たすマクロブロックの数を求める手段と、

Iフレーム毎に、前記色条件を満たすマクロブロック数が所定範囲以内であるラインの数の平均値を求める手段とを備え、

前記平均値が所定範囲内であるIフレームを含むショットを、特定のショットとして判別することを特徴とする請求項 1 乃至 10 何れか1項記載のショット検出システム。

## 【請求項 12】

Iフレーム毎に、前記ラインの変化に対する前記マクロブロック数の変化の度合を求め、その度合の最大値が所定範囲内であるIフレームを含むショットを、特定のショットとして判別することを特徴とする請求項 1 乃至 11 何れか1項記載のショット検出システム。

## 【請求項 13】

前記ショット検出システムにより検出された各ショットに対し、その種類に応じたショット記号を付与することでショット記号列を作成する手段と、

検出対象のシーンに対応するように複数の前記ショット記号を配列してなるモデルパターン記号列を記憶する手段と、

前記ショット記号列中から、前記モデルパターン記憶手段のモデルパターン記号列に対応する記号列を検出する検出手段とを備え、

前記検出手段により検出された記号列に対応する前記動画像データの連続ショットを、特定のシーンとして認識するようにしたことを特徴とする請求項 5 乃至 12 何れか 1 項記載のショット検出システムを用いたシーン検出システム。

## 【請求項 14】

前記検出手段は、前記ショット記号列に対し前記モデルパターン記号列を照合させて、前記ショット記号列中から前記モデルパターン記号列に略合致する記号列を検出するパターンマッチング処理を行うことを特徴とする請求項 13 記載のシーン検出システム。

10

20

30

40

50

**【請求項 15】**

コンピュータが、マクロブロックタイプ情報を含む動画圧縮データから、少なくとも B フレーム毎にマクロブロックタイプ情報を抽出するステップと、

コンピュータが、この抽出されたマクロブロックタイプ情報からフレーム毎のイントラ符号化マクロブロックの数を求めるステップと、

コンピュータが、B フレームに含まれるイントラ符号化マクロブロックの頻度的特徴を判別条件にして、この B フレーム又はこの B フレームを含むグループオブピクチャを、リプレイショットの先頭部分又は最後部分であるリプレイカット部と判別するステップと、を含むことを特徴とするカット部検出方法。

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は、動画像データから、ショットの分割位置であるカット部や、ショットの種類、複数の連続ショットからなるシーン等を検出するカット部検出システム及びショット検出システム並びにシーン検出システム、カット部検出方法に関するものである。

**【背景技術】****【0002】**

近年、インターネットのブロードバンド化に伴い映像配信が行われ、衛星放送やケーブルテレビによるチャンネル数の増加により、ユーザが見ることのできる映像が増加している。また、ストレージ技術の急速な発展にともない、テラバイトを超える大容量のハードディスクを搭載した HDD レコーダも登場してきており、膨大な番組（コンテンツ）を PC やレコーダのハードディスクに保存することが一般的になってきており、個人で映像データを利用することが頻繁に行われている。

そして、前記のように映像データが多くなるほど、映像データ内の見たい場面を素早く探し出すことが困難となる。このようなことから、膨大なビデオデータから、必要なシーンを効率良く、素早くアクセスする技術が重要となってきた。

**【0003】**

そこで、例えば、特許文献 1 に記載された発明では、画面内に占める芝の割合や、動きベクトル、歓声の特徴等から、特定のショットまたはシーンの一部を判定するようにしている。

また、特許文献 2 に記載された発明では、双方向予測符号化フレームの順方向予測符号化ブロック数、逆方向予測符号化ブロック数の大小関係を調べることにより、複数のショットの分割位置であるカット位置を検出するようにしている。

**【0004】**

しかしながら、先の映像の繰り返しを映しているリプレイショットについては、その分割位置（カット部）やその範囲を検出するのが困難であった。

また、リプレイショット以外のショットに関しても、そのショットの種類を明確に分類したり、複数の連続ショットからなるシーンを、ストーリー上意味のある連続した場面として抽出するのは困難であった。

**【特許文献 1】**特開 2005 - 252860 号公報

**【特許文献 2】**特開平 07 - 284071 号公報

**【発明の開示】****【発明が解決しようとする課題】****【0005】**

本発明は上記従来事情に鑑みてなされたものであり、その課題とする処は、リプレイショットの先頭部分又は最後部分であるリプレイカット部を検出可能なカット部検出システム、リプレイショットおよびその他の特定のショットを精度良く判別することが可能なカット部検出システム及びショット検出システム並びにシーン検出システム、カット部検出方法を提供することにある。

**【課題を解決するための手段】**

10

20

30

40

50

## 【0006】

上記課題を解決するために本発明に係る技術的手段は、マクロブロックタイプ情報を含む動画圧縮データから、少なくともBフレーム毎にマクロブロックタイプ情報を抽出する手段と、この抽出されたマクロブロックタイプ情報からフレーム毎のイントラ符号化マクロブロックの数を求める手段と、Bフレームに含まれるイントラ符号化マクロブロックの頻度的特徴を判別条件にして、このBフレーム又はこのBフレームを含むグループオブピクチャを、リプレイショットの先頭部分又は最後部分であるリプレイカット部と判別する手段と、を含むことを特徴とするカット部検出システムである。

## 【0007】

更なる技術的手段では、前記判別条件には、連続する二つのBフレームの内の少なくとも一方におけるイントラ符号化マクロブロックの数が所定数以上であるという条件が含まれていることを特徴とする。

10

## 【0008】

更なる技術的手段では、前記判別条件には、連続する二つのBフレームの内の少なくとも一方におけるイントラ符号化マクロブロックの数が所定範囲内であって、このパターンとなる前記二つのBフレームが、複数のグループオブピクチャ中で所定数以上あるという条件が含まれていることを特徴とする。

## 【0009】

更なる技術的手段では、上記マクロブロックタイプ情報からフレーム毎の順方向予測符号化マクロブロックの数およびフレーム毎の逆方向予測符号化マクロブロックの数を求める手段を備え、前記判別条件に加える条件として、順方向予測符号化マクロブロックの数が最頻度で、且つ逆方向予測符号化マクロブロックの数とイントラ符号化マクロブロックの数との合計が所定数以上であるという条件と、逆方向予測符号化マクロブロックの数が最頻度で、且つ順方向予測符号化マクロブロックの数とイントラ符号化マクロブロックの数との合計が所定数以上であるという条件と、イントラ符号化マクロブロックの数が最頻度であるという条件と、の三つ条件の内の何れかを含むことを特徴とする。

20

## 【0010】

更なる技術的手段は、前記カット部検出システムより検出される複数のリプレイカット部の内、連続する二つのリプレイカット部の間を、先の映像の繰り返しを映しているリプレイショットと判別することを特徴とする上記カット部検出システムを用いたショット検出システムである。

30

## 【0011】

更なる技術的手段では、前記カット部検出システムより検出される複数のリプレイカット部の内、連続する二つのリプレイカット部が、連続する二つの瞬時カット点の間に位置する場合に、これら連続する二つのリプレイカット部の間を、先の映像の繰り返しを映しているリプレイショットと判別することを特徴とする。

## 【0012】

更なる技術的手段では、前記連続する二つのリプレイカット部の内の一方と、この一方のリプレイカット部に連続する瞬時カット点との間に、所定数以上のグループオブピクチャを有するという条件を、前記リプレイショットの判別条件に加えたことを特徴とする。

40

## 【0013】

更なる技術的手段では、前記連続する二つのリプレイカット部の間に所定数以上のグループオブピクチャを有するという条件を、前記リプレイショットの判別条件に加えたことを特徴とする。

## 【0014】

更なる技術的手段では、動画圧縮データから抽出されたショットに対し、フレーム毎に動きベクトル情報を抽出する手段と、この動きベクトル情報からフレーム毎の動きベクトル量を求める手段と、前記抽出されたショットについて、その1ショット分の前記動きベクトル量をグループオブピクチャ数で割った平均の動きベクトル量を求める手段とを備え、前記平均の動きベクトル量が所定範囲内であることを判別条件にして、前記抽出された

50

ショットを特定のショットとして判別することを特徴とする。

【0015】

更なる技術的手段では、動画圧縮データから抽出されたショットに対し、フレーム毎に動きベクトル情報を抽出する手段と、各フレームのマクロブロック毎に動きベクトルの方向を所定数の方向に量子化する手段と、前記量子化により分類された動きベクトルの方向について、その分散度をフレーム毎に求める手段と、前記抽出されたショットについて、その1ショット分の前記分散度をグループオブピクチャ数で割った平均の分散度を求める手段とを備え、前記分散度が所定範囲内であることを判別条件にして、前記抽出されたショットを特定のショットとして判別することを特徴とする。

【0016】

更なる技術的手段では、動画圧縮データから、Iフレーム毎に輝度成分Y、色差成分Cb、色差成分Crを抽出する手段と、各Iフレームを平行な複数のラインに分割し、そのライン毎に、輝度成分Y、色差成分Cb、色差成分Cr、色差成分の比 $r = Cb / Cr$ が、それぞれ所定範囲内であるという色条件を満たすマクロブロックの数を求める手段と、Iフレーム毎に、前記色条件を満たすマクロブロック数が所定範囲以内であるラインの数の平均値を求める手段とを備え、前記平均値が所定範囲内であるIフレームを含むショットを、特定のショットとして判別することを特徴とする。

【0017】

更なる技術的手段では、Iフレーム毎に、前記ラインの変化に対する前記マクロブロック数の変化の度合を求め、その度合の最大値が所定範囲内であるIフレームを含むショットを、特定のショットとして判別することを特徴とする。

【0018】

更なる技術的手段は、前記ショット検出システムにより検出された各ショットに対し、その種類に応じたショット記号を付与することでショット記号列を作成する手段と、検出対象のシーンに対応するように複数の前記ショット記号を配列してなるモデルパターン記号列を記憶する手段と、前記ショット記号列中から、前記モデルパターン記憶手段のモデルパターン記号列に対応する記号列を検出する検出手段とを備え、前記検出手段により検出された記号列に対応する前記動画データの連続ショットを、特定のシーンとして認識するようにしたことを特徴とする上記ショット検出システムを用いたシーン検出システムである。

【0019】

更なる技術的手段は、前記検出手段は、前記ショット記号列に対し前記モデルパターン記号列を照合させて、前記ショット記号列中から前記モデルパターン記号列に略合致する記号列を検出するパターンマッチング処理を行うことを特徴とする。

【0020】

更なる技術的手段では、コンピュータが、マクロブロックタイプ情報を含む動画圧縮データから、少なくともBフレーム毎にマクロブロックタイプ情報を抽出するステップと、コンピュータが、この抽出されたマクロブロックタイプ情報からフレーム毎のイントラ符号化マクロブロックの数を求めるステップと、コンピュータが、Bフレームに含まれるイントラ符号化マクロブロックの頻度的特徴を判別条件にして、このBフレーム又はこのBフレームを含むグループオブピクチャを、リプレイショットの先頭部分又は最後部分であるリプレイカット部と判別するステップと、を含むことを特徴とするカット部検出方法である。

【発明の効果】

【0021】

本発明は、以上説明したように構成されているので、以下に記載されるような作用効果を奏する。

動画圧縮データからマクロブロックタイプ情報が抽出され、この抽出されたマクロブロックタイプ情報からフレーム毎のイントラ符号化マクロブロックの数が求められる。

イントラ符号化マクロブロックは、リプレイショットの先頭側または最後側のBフレー

10

20

30

40

50

ムにおいて比較的多くなる傾向があるため、Bフレームに含まれるイントラ符号化マクロブロックの数が所定範囲内であることを判別条件にして、このBフレーム又はこのBフレームを含むグループオブピクチャを、リプレイショットの先頭部分又は最後部分であるリプレイカット部と判別することができ、更には、連続する二つのリプレイカット部の間をリプレイショットとして判別することも可能である。

【発明を実施するための最良の形態】

【0022】

以下、本発明の実施の形態を図面に基づいて説明する。

図1は、本発明の一例を示すブロック図であり、本発明の特徴を明確にするために、一般的なMPEG復号器の構成を同ブロック図に併せて記載している。

10

【0023】

図1に示すように、DVDやハードディスク等の記録媒体に記憶された動画圧縮データは、バッファ1を経由して復号器2へ入力される。その一方で、本発明に係わるカット部検出システム及びショット検出システム並びにシーン検出システム3は、復号前の同動画圧縮データからMBT情報(マクロブロックタイプ情報)を抽出する。

これらの構成は、例えば、コンピュータやDVDレコーダ等の装置、および該装置を機能させるためのプログラムや電子回路等として実現される。

【0024】

復号器2は、可変長復号化、逆量子化、逆離散コサイン変換(IDCT)等の周知の復号処理により動画圧縮データを復号化する装置であり、この復号器2によって復号化された動画はディスプレイ等の表示装置に表示される。

20

【0025】

本実施の形態のカット部検出システム及びショット検出システム並びにシーン検出システム3は、動画圧縮データから直接的にマクロブロックタイプ情報(以降、MBT情報とも称する)を抽出し、そのMBT情報からマクロブロックタイプ記号列(以降、MBT記号列とも称する)を作成し、このMBT記号列を用いて、瞬時カット点やリプレイカット部等のカット部を検出する。

そして、カット部検出システム及びショット検出システム並びにシーン検出システム3は、前記カット部を区切り位置とした動画データ、リプレイショットを含む複数のショットに分割する。

30

【0026】

次に、このカット部検出システム及びショット検出システム並びにシーン検出システム3は、動きベクトルや、輝度成分Y、色差成分Cb、色差成分Cr、色差成分の比 $r = Cb / Cr$ 等の特徴量に基づき、前記ショットを複数の分類パターンに識別し、各ショットに対し種類毎に異なるショット記号を付与し、ショット順に並ぶショット記号列を作成する。

更に、カット部検出システム及びショット検出システム並びにシーン検出システム3は、ショット記号列から照合処理(パターンマッチング処理)により特定のパターンを探索することでシーンを検出する。

【0027】

40

前記動画圧縮データは、本実施の形態の一例によれば、サッカー試合動画データを圧縮符号化してなるMPEG2データを用いている。

MPEG2形式では、ビデオデータを圧縮する際に以下の3つフレームタイプを用いている。

- ・イントラフレーム(Intra-coded frame, Iフレーム)  
他のフレームの情報を使用せず、フレーム内圧縮のみ行うフレーム。
- ・前方向予測符号化フレーム(Predictive-coded frame, Pフレーム)  
フレーム内圧縮のほかに、過去のIフレーム及びPフレームを参照フレームとして、時間軸上で順方向動き予測符号化を行うフレーム。
- ・双方向予測符号化フレーム(Bidirectionally predictive-coded frame, Bフレーム)

50

フレーム内圧縮のほかに、過去と未来のIフレーム及びPフレームを参照フレームとして、時間軸上で双方向予測符号化を行うフレーム。

MPEG2形式で圧縮されたビデオデータはグループオブピクチャ(以後GOP(Group of Picture)と称する)単位でまとめられており、通常、1つのGOPは0.5秒分となる15フレーム分のデータで構成される。また、GOP内のフレームタイプの構成はビデオデータごとに決まっており、本実施の形態では、IBBPBBPBBPBBPBBという通常のパターン、すなわち連続する2つのBフレームを周期的に含むパターンを用いている。

#### 【0028】

GOPには、Closed GOPという方式と、Open GOPという方式が存在する。

Closed GOPとは、一つ前のGOPを参照しない方式であり、一つのGOP内だけで符号化を行う。このため、GOPの最初の2つのBフレームは、すべてのマクロブロックが必ず逆方向予測となる。一方、Open GOPとは、一つ前のGOPを参照する方式である。GOPの最初の2つのBフレームも、順方向予測が可能となるため、Closed GOPよりもOpen GOPの方が、より符号化による圧縮率が高くなる。

また、Open GOPでは、映像が切り替わる部分などにおいて、GOP内のフレーム数が少なくなることがある。例えば、GOPを構成するフレーム数が、3フレーム、6フレーム、9フレーム、12フレームといった3の倍数の数に減少する。この特徴を利用し、カット点検出の精度を上げることができる。

本実施の形態では、前記のような特徴を利用することで、比較的カット部の検出が行いやすく、また、圧縮効率もよいOpen GOP方式で符号化されたMPEGデータを用いることとする。

#### 【0029】

次に、MPEG2における符号化方式について説明する。

通常MPEG圧縮においては1枚の画像フレームを8×8ブロックに細分化した上で、各ブロック毎に離散コサイン変換を行う。輝度4ブロックと、色差(Cb, Cr)の2ブロックをマクロブロックという。動き補償のブロックサイズは16 pixel × 16 pixelであるのに対し、DCT(離散コサイン変換)のサイズは8pixel × 8pixelであるためにそれらを結合する仕組みがマクロブロックである。このマクロブロックが参照画像への予測を行う単位となる。また、MPEGでは、符号化を行う際に前後のI, Pフレームに対して参照を行うことで圧縮効果を向上させる工夫が施されている。このため、マクロブロック単位での符号化には、Intra符号化、順方向予測符号化、逆方向予測符号化、双方向予測符号化の4種類の方法があり、各符号化方法は以下の特徴を持つ。

#### 【0030】

##### ・Intra 符号化

他の画像を参照しないで画素値を復元する符号化方法。

##### ・順方向予測符号化

過去の画像に対して参照を行うことで画素値を復元する符号化方法。

##### ・逆方向予測符号化

未来の画像に対して参照を行うことで画素値を復元する符号化方法

##### ・双方向予測符号化

順方向と逆方向の二つの予測を対応画素間で平均することで画素値を復元する符号化方法。

#### 【0031】

各マクロブロックは、前記4つの方式で符号化される。MPEG2には3種類のフレーム(I, P, B)がある。

IフレームはIntra符号化のみを選ぶことができる。

PフレームはIntra符号化に加え、直前のIフレームもしくはPフレームを参照フレームとする順方向予測符号化を選ぶことができる。

Bフレームは、それらに加え、直後のI, Pフレームを参照フレームとする逆方向予測符号化と、直前、直後のI, Pフレームを参照フレームとする双方向予測符号化を選ぶこと

10

20

30

40

50



ができる。

【0032】

各フレームはマクロブロックと呼ばれる単位に分割されており、個々のマクロブロックごとにマクロブロックタイプと呼ばれる属性が割り当てられている。各フレーム毎に選ぶことのできる符号化方法を使ってマクロブロックの符号化が行われる。

前記4つの符号化方式のうちの方法を選ぶかは符号化の段階においてマクロブロック単位で決められ、本実施の形態が扱う動画圧縮データであるMPEG符号化データの中にマクロブロックタイプとして埋め込まれている。

【0033】

次にMPEG2のフォーマットについて説明する。

動画像には422と420というコンポーネントデジタル化された形式がある。

422formatでは、輝度画像が720pixel×480pixelの大きさの場合は、2つの色差画像は横方向にサブサンプル(例えば間引き)された各360pixel×480pixelの縦長な形をしている。

一方、420formatというものは、図2に示すように色差をさらに縦方向にもサブサンプルした画像形式をいう。色差画像の大きさは縦横とも輝度画像の1/2になる。これがもっとも一般的に動画像符号化で使われる形式である。420の一例では、輝度画像(Y)と縦横比の一致した2つの色差画像(Cb,Cr)がセットになって扱われる。また、各フレームは45×30のマクロブロックから構成されている。

本実施の形態で用いるMPEGデータは、この420でコンポーネントデジタル化された形式の動画像を用いる。

【0034】

次に、本実施の形態においてMPEG2データから抽出され、カット部の検出、及びショットの判別に用いる特徴量について説明する。

本実施の形態では、MPEG2データから、以下の3つの情報を抽出して利用する。

- ・マクロブロックタイプ
- ・動きベクトル
- ・マクロブロックの輝度、色差(DC成分)

【0035】

マクロブロックタイプとは、各マクロブロックが前後のどのフレームを参照しているかを表した情報である。上記したように、マクロブロック単位での符号化方式には、4種類の方式があり、この4つの方式が以下の4つのマクロブロックタイプに対応する。

- ・Intra符号化マクロブロックタイプ
- ・順方向予測符号化マクロブロックタイプ
- ・逆方向予測符号化マクロブロックタイプ
- ・双方向予測符号化マクロブロックタイプ

これら4種類のマクロブロックタイプは、マクロブロック毎に対応するコードとしてMPEG動画圧縮データのマクロブロック層に書き込まれている。

【0036】

Bフレームでは、4つのマクロブロックタイプを選択することができ、参照する画像によって、1フレーム内のそれぞれのマクロブロックタイプの数に大きな偏りが見られる。このような特徴を利用し、映像の切り替わる点を検出することができる。

また、カラー画像はRGBの3つの値または3つの画像で表される。これを線形変換して、輝度(Y)と2つの色差(Cb,Cr)で表わすと、色差画像の空間分解能を減らすことができる。輝度とは明るさを表し、色差とは色の度合いを表すものである。

【0037】

MPEGでは、符号化を行う方法として、冗長な情報を削減するために、離散コサイン変換(DCT)が行われる。DCTは画像信号を少ない低域係数に集中させる働きを持ち、画像の空間的方向の情報量削減に使われており、固定の変換係数の直交変換のなかでは画像符号化に最も有効な変換とされている。

DCT後の直流成分をDC成分と言い、8 pixel×8 pixelブロックの平均値を表している

10

20

30

40

50

ため、特徴量に用いる色情報として、輝度・色差のDC成分を用いる。マクロブロックでの輝度・色差成分を図3に示す。図3のように、420フォーマットでは、一つのマクロブロックに4つの輝度成分とそれぞれ1つのCb, Crの色差成分が含まれる。この輝度・色差のDC成分を用いることで、グラウンドエリアの色情報を抽出し、ショットの判別に利用する。また、ショット判別に用いる特徴量としては、一つのマクロブロックにおいて、Cb, Crの色差成分と、4つの輝度ブロックのうち左上の1つの輝度成分のみを利用する。

#### 【0038】

次に、本実施の形態で用いる動きベクトルについて説明する。

動きベクトルとは、現在の画像における予測画像の差分を求めるときの、現画像のマクロブロック位置と予測画像の参照を行う位置とのピクセル単位でのずれ量を表す。

この動きベクトルは、MPEG2の場合16pixel x 16pixel単位のマクロブロックで発生し、現マクロブロックから参照する位置までの距離をx成分とy成分で表す。このとき、x成分は右方向が正、y成分は下方向が正となる。

また、過去の画像との差分を表したものを順方向動きベクトルと呼び、未来の画像との差分を表したものを逆方向動きベクトルと呼ぶ。

Iフレームでは、フレーム間予測を行っていないので、動きベクトルが発生しない。Pフレームでは、過去のIまたはPフレームとフレーム間予測を行うため、順方向動きベクトルのみ発生する。Bフレームでは、過去と未来、両方のI, Pフレームを参照できるため、順方向動きベクトル、逆方向動きベクトルの2つの動きベクトルが発生する。動きベクトルは、フレーム間予測を行うため、映像における動きの特徴や、カメラの動きに対応した映像の流れの特徴を表している。

本実施の形態では、この動きベクトルをショット分類の判別を行うための特徴量として利用する。また、動きの特徴を求めるために十分な量と考えられるPフレームとBフレームにおける順方向動きベクトルのみを特徴量として用いる。

#### 【0039】

図4は、本実施の形態が、動画圧縮データのどのフレームからどのような特徴量を抽出するのかを示している。

すなわち、Iフレームからは、マクロブロックの輝度・色差のDC成分を取り出す。そして、P, Bフレームからは、順方向動きベクトルを取り出す。これら輝度・色差と動きベクトル等の情報は、ショット分類の判別に利用する。

また、Bフレームからは、マクロブロックタイプを取り出す。このマクロブロックタイプは、カット点およびリプレイカット部の検出に用いる。

#### 【0040】

ここで、サッカー試合映像を一例にして、シーン、ショット、カット点、リプレイカット部等の概念を、詳細に説明する。

シーンとは、ストーリー上意味のある連続した場面の動画であり、例えば、選手がシュートをしたシーン（ゴールはしていない）や、選手がシュートしたボールがゴールしたゴールシーン等である。このシーンは、図5に示すように、時系列順に複数連続して、例えばサッカー試合映像等の一つの動画を構成する。

#### 【0041】

そして、各シーンは、図5に示すように、連続する複数のショットから構成される。ショットとは、通常1台のカメラが捉えた時間的かつ空間的に連続した場面を意味し、各ショットは連続する複数のフレームからなる。

本実施の形態におけるショットには、前後のカット点により区分されたショットと、前後のリプレイカット部により区分されたリプレイショットとを含む。

#### 【0042】

本実施の形態では、ショットを以下のように分類し定義している。なお、以下の左側の括弧内の数字は、後述するショット検出で用いるショット記号を示している。

(1) センターロングショット (Center Long Shot)

10

20

30

40

50

グラウンド全体を映しているショットであって、最後の3つのグループオブピクチャ（以後GOPとも称する）がゴール付近の画面でないショット。

(2) ゴールロングショット (Goal Long Shot)

グラウンド全体を映しているショットであって、最後の3GOPがゴール付近の画面であるショット。

(3) クローズアップショット (Close-up Shot)

選手のアップを映しているショット。

(4) アウトフィールドショット (Out-field Shot)

観客などフィールド外を映しているショット。

(5) リプレイショット (Replay Shot)

先の映像の繰り返し（リプレイ）を映しているショット

【0043】

上記5つのショット中、(1)～(4)の各ショットは、その前後のカット点の間に位置する動画像の範囲である。カット点とは、カメラの切り替わりにより生じた映像の切り替わり点のことである。

一般的に、カット点には、映像が急激に変化する瞬時カット点（ハードカット点とも呼称される）や、ディゾルブカット点、ワイプカット点、その他の特殊なカット点等があるが、本実施の形態で扱うカット点は瞬時カット点である。

また、上記(5)のリプレイショットの最前部と最後部に位置するリプレイカット部は、本実施の形態では前記カット点と区別し、後に詳述する。

【0044】

各ショットを構成している複数のフレームは、図6に示す再生順のデータであるが、符号化されるときに順番が入れ替えられることで、同図6に示すビットストリーム順のデータとなる。

更に、このビットストリーム順のデータは、復号化されるときに順番が入れ替えられることで、元の順番（前記再生順と同じ順番）に戻され、同図6に示す出力順のデータとなる。

本実施の形態では、前記再生順または出力順のことを、フレーム順と称する。

【0045】

次に、上記リプレイショット及びリプレイカット部の概念について説明する。

スポーツ映像でのリプレイとは、シュートや、ゴール、ファウルなど視聴者が特に興味のあるシーンがあったときに、数秒後に再び同じシーンを、様々な角度のカメラからの映像がスローモーションなどの効果が加えられ、再生される場面のことを言う（図7参照）。

リプレイに加えられる特殊な編集効果に着目すると、スポーツ映像では、リプレイ部分の始めと終わりに特殊なショット切り替え操作が加えられる場合が多い。例えば、図8に示すような特殊な画像が入る場合や、ワイプカットなどの特殊なショット切り換え操作が入ることもある。このようなリプレイ部分の始めと終わりでの特殊な切り替え操作をDVE (Digital Video Effect) と呼ぶこととする。

本実施の形態で用いるサッカー試合動画像には、図8に示すように、サッカー試合映像に対し、該サッカー試合映像以外の映像（例えば、画面の略全体を覆うボールや旗等の画像）を重ね合わせてなるオーバーレイ画像を、前記DVEの一例として含んでいる。

【0046】

本実施の形態では、DVEが入るリプレイ部分の始めと終わりの部分をリプレイカット部と称し、連続する二つのリプレイカット部で挟まれる映像区間をリプレイショットと称する。なお、リプレイショットの映像区間では、映像の切り替わり点として、ディゾルブカット点を用いられる場合があるが、図9に示すように、このディゾルブカット点をカット点とせず、リプレイカット部間のリプレイショットを一つのショットとして判断する。

【0047】

次に、本実施の形態による瞬時カット点（リプレイカット部を含まない）の検出につい

10

20

30

40

50

て詳細に説明する。

瞬時カット点が発生する位置は、一对の連続したBフレーム ( $B_i, B_{i+1}$ ) に注目すると、以下の(i), (ii), (iii), (iv), (v)の場合に分類される。

- (i)  $B_i$ の直前にカット点が存在する場合
- (ii)  $B_i$ と $B_{i+1}$ の間にカット点が存在する場合
- (iii)  $B_{i+1}$ の直後にカット点が存在する場合
- (iv)  $B_i$ のフレーム上にカット点が存在する場合
- (v)  $B_{i+1}$ のフレーム上にカット点が存在する場合

以下に、それぞれの場合が生じたとき、Bフレームの符号化にどのような影響があるのかについて説明する。

なお、以下の(i)~(v)の場合の説明では、 $B_i$ の直前がIフレーム、 $B_{i+1}$ の直後がPフレームの場合を例にして説明しているが、 $B_i$ の直前のフレームと $B_{i+1}$ の直後のフレームの組み合わせとしては、(I, P) (P, P) (P, I)のいずれの場合も、数が多くなるマクロブロックタイプは同じである。

#### 【0048】

(i)の場合には、図10(a)に示すように、 $B_i, B_{i+1}$ は直前のIもしくはP(図ではIを例示)を参照しても類似した値を得ることができないため、類似した直後のIもしくはP(図ではPを例示)に対して参照を行う。そのため $B_i, B_{i+1}$ 共に逆方向予測を行うマクロブロックの数が多くなる。

#### 【0049】

(ii)の場合には、図10(b)に示すように、画像の類似性から $B_i$ は直前のIもしくはP(図ではIを例示)に対しての参照を行い、 $B_{i+1}$ は直後のIもしくはP(図ではPを例示)に対して参照を行う。そのため $B_i$ は直前のIもしくはP(図ではIを例示)に対しての順方向予測符号化を行ったマクロブロックの数が多く、 $B_{i+1}$ は直後のIもしくはP(図ではPを例示)に対しての逆方向予測符号化を行ったマクロブロックの数が多くなる。

#### 【0050】

(iii)の場合には、図10(c)に示すように、画像の類似性から $B_i, B_{i+1}$ 共に直前のIもしくはP(図ではIを例示)に対して参照を行う。そのため $B_i, B_{i+1}$ ともに、直前のIもしくはP(図ではIを例示)に対しての順方向予測化を行ったマクロブロックの数が多くなる。

#### 【0051】

(iv)の場合には、図11(d)に示すように、イントラ符号化されるマクロブロックが発生するという特徴がある。これは、1枚のフレームが縦方向に1列おきのピクセルで2枚のフィールドで構成されるフィールド構造の場合に顕著に生じる。また、 $B_{i+1}$ では、Pと画像が類似していることにより、逆方向予測マクロブロックが多くなる。

#### 【0052】

(v)の場合には、(iv)のときと同様に、 $B_{i+1}$ フレーム上に、図11(e)に示すように、イントラ符号化されるマクロブロックが発生するという特徴がある。また、 $B_i$ では、Iと画像が類似していることにより、順方向予測マクロブロックが多くなる。

#### 【0053】

以上の傾向はフレーム毎のMBT情報に反映されており、このことは、本願発明者が、一般的な市販のMP EG2エンコーダにより符号化された動画圧縮データを用いて、実験的に確認している。

以上のことに基づき、本実施の形態のカット部検出システム及びショット検出システム並びにシーン検出システム3では、フレーム毎のMBT情報から特定のマクロブロックタイプが含まれる頻度を求め、その頻度的特徴に応じてフレーム毎に所定のMBT記号を付与するとともに、このMBT記号をフレーム順に対応するように並べることで、MBT記号列を作成し、このMBT記号列中におけるMBT記号の配列的特徴から動画像のカット点を判別している。

#### 【0054】

10

20

30

40

50

より具体的に説明すれば、先ず、各フレームのMBT情報から、イントラ符号化マクロブロック（以降、mbIntraとも称する）の数、順方向予測符号化マクロブロック（以降、mbForwardとも称する）の数、逆方向予測符号化マクロブロック（以降、mbBackとも称する）の数、双方向予測符号化マクロブロック（以降、mbInterとも称する）の数、その他の符号化マクロブロック（以降、mbOtherとも称する）の数を、それぞれ求める。

【0055】

次に、図12の表に示すように、各条件（表中の右列内）を満たすBフレームに対し、対応するマクロブロックタイプ記号0, 1, 2, 3, 4, 5, 6, 7, 8の何れかが付与される。

なお、図12の表中では、前記マクロブロックタイプ記号の項目名を、Frame Typeと表現している。

【0056】

すなわち、図12の表中においてマクロブロックタイプ記号が「0」となる場合を一例にして、詳細に説明すれば、フレーム毎のMBT情報の頻度的特徴として、Bフレームであって、順方向予測符号化マクロブロック（mbForward）の頻度が最大であり、且つ、逆方向予測符号化マクロブロックとイントラ符号化マクロブロックの数の合計（mbBack + mbInter）が所定の閾値t（図示例によれば100）よりも小さい場合には、その頻度的特徴を、予め設定されたマクロブロックタイプ記号である数字の「0」という記号により表す。

すなわち、前記条件を満たすフレームに対し、マクロブロックタイプ記号である「0」という記号が付与される。

その他、マクロブロックタイプ記号が1, 2, 3, 4, 5, 6, 7, 8の各々となる場合についても、同様に処理される。

【0057】

また、マクロブロックタイプ記号が「7」となる場合の条件は、マクロブロックタイプ記号が、2, 5, 6の内の何れかとなる条件が満たされ、且つ、イントラ符号化マクロブロックの数が所定数（例えば100）以上であることを満たす。

より詳細に説明すれば、順方向予測符号化マクロブロックの数が最頻度で、且つ逆方向予測符号化マクロブロックの数とイントラ符号化マクロブロックの数との合計が所定数（図示例によれば350）以上であるという条件と、逆方向予測符号化マクロブロックの数が最頻度で、且つ順方向予測符号化マクロブロックの数とイントラ符号化マクロブロックの数との合計が所定数（図示例によれば350）以上であるという条件と、イントラ符号化マクロブロックの数が最頻度という条件との三つ条件の内の何れかが満たされ、且つ、イントラ符号化マクロブロックの数が所定の閾値（例えば100）以上であることを満たす。

【0058】

同様に、マクロブロックタイプ記号が「8」となる場合の条件は、マクロブロックタイプ記号が、2, 5, 6の内の何れかとなる条件が満たされ、且つ、イントラ符号化マクロブロックの数が所定範囲内（図示例によれば20～100の範囲内）であることを満たす。

【0059】

なお、図12の表中、マクロブロックタイプ記号が「2」、「5」、「7」となる場合の条件について、閾値の上限が記載されていないが、この上限値は、1フレーム中におけるマクロブロックの最大数（例えば1350）としてもよいし、他の適宜な値を設定するようにしてもよい。

また、表中の閾値（20, 100, 350等）は、実験的に予め設定されたものであり、動画の種類やその他の条件に応じて適宜変更してもよい。

【0060】

そして、上記のようにしてフレーム毎に付与されたマクロブロックタイプ記号は、フレーム順に並べられることで、一連のMBT記号列（例えば「023567118314・

10

20

30

40

50

・ ・ 」等)を構成する。

したがって、このMBT記号列において、各MBT記号は、対応するフレームにおける特定のマクロブロックの頻度的特徴を表していることになる。

【0061】

なお、上記態様によれば、Bフレームのみに上記MBT記号を付与することを説明しているが、Bフレームに対応するMBT記号のみからMBT記号列を作成する構成としてもよいし、他例としては、Bフレーム以外のフレーム(例えばIフレームやPフレーム)にも適当なMBT記号を付与し、Bフレームに対応するMBT記号と、Bフレーム以外の前記フレームに対応するMBT記号と、の双方のMBT記号の配列によりMBT記号列を作成するようにしてもよい。

10

【0062】

b 次に、カット部検出システム及びショット検出システム並びにシーン検出システム3は、以下に示すように、上記MBT記号列中におけるMBT記号の配列的特徴に応じて、瞬時カット点の位置を判断する。

【0063】

すなわち、図13に示すように、上記一連のMBT記号列について、連続するペアのBフレームに対応する二つのMBT記号b1, b2に注目し、その配列パターンに応じて、以下のように、二つのMBT記号b1, b2の直前、直後、間の何れかに瞬時カット点があると判定する。

【0064】

20

以下、左二桁の数字は、前記二つのMBT記号b1, b2の配列パターンを示し、これら数字の右側には、前記配列パターンに応じて判定される瞬時カット点の位置を示す。

00: 末尾の0の後にカット点があると判定する。

0\*: 0の直後にカット点があると判定する。なお、\*は0以外の数とする。

\*3: 3の直前にカット点があると判定する。なお、\*は3以外の数とする。

33: 先頭の3の直前にカット点があると判定する。

11: 末尾の1の直後にカット点があると判定する。

14: 1と4の間にカット点があると判定する。

44: 先頭の4の直前にカット点があると判定する。

17: 1と7の間にカット点があると判定する。

30

18: 8の直後にカット点があると判定する。

84: 8の直前にカット点があると判定する。

74: 7と4の間にカット点があると判定する。

【0065】

次に、リプレイカット部の検出について説明する。

リプレイショットの始めと終わりに入るDVEには、Bフレームにおいてイントラ符号化されるマクロブロックの割合が多くなるという傾向がある(図14参照)。これは、挿入される画像や、映像の切り替わりが激しく変化するためだと考えられる。本実施の形態では、このことを判別条件として利用することで、リプレイカット部を検出するようにしている。

40

【0066】

具体的に説明すれば、以下の二つの判別条件を用いる。

(1) 連続する二つのBフレームの内の少なくとも一方におけるイントラ符号化マクロブロックの数が所定数以上(図12の一例によれば100以上)であるという条件。

(2) 連続する二つのBフレームの内の少なくとも一方におけるイントラ符号化マクロブロックの数が所定範囲内(図12の一例によれば20以上100未満)であって、このパターンとなる前記二つのBフレームが、複数(例えば2つ)のグループオブピクチャ中で所定数以上(例えば2~3以上、好ましくは3以上)あるという条件。

そして、以上の(1)と(2)の内の何れか一方の判別条件を満たす場合に、前記Bフレームを含むグループオブピクチャを、リプレイショットの先頭部分又は最後部分である

50

リプレイカット部と判別する。なお、前記 B フレームをリプレイカット部とすることも可能である。

【 0 0 6 7 】

リプレイカット部の判別について、より詳細にすれば、上記 M B T 記号列中に二つの M B T 記号 b 1 , b 2 の内の何れかが「 7 」であるという条件を満たすパターンが存在した場合には、このパターンが存在するグループオブピクチャを、リプレイカット部とする。

また、上記 M B T 記号列中に、二つの M B T 記号 b 1 , b 2 の内の何れかが「 8 」であるペアが連続する二つのグループオブピクチャ中に 3 ペア以上あるという条件が満たされる場合も、このパターンが存在するグループオブピクチャを、リプレイカット部と判別する。

10

【 0 0 6 8 】

次に、リプレイショットの検出について説明する。

本実施の形態のカット部検出システム及びショット検出システム並びにシーン検出システム 3 では、上記 ( 1 ) 又は ( 2 ) の条件により判別したリプレイカット部の間をリプレイショットとして判別する。

特にその判別精度を向上するために、より好ましくは、図 1 5 に示すように、連続する 2 つの瞬時カット点の間に 2 つのリプレイカット部のパターン ( 上記 ( 1 ) 又は ( 2 ) の条件を満たすパターン ) が存在し、さらに、前記瞬時カット点とこの瞬時カット点に連続するリプレイカット点との間が 2GOP 以上あり、且つ、前記 2 つのリプレイカット点間においても 6GOP 以上あるとき、これら 2 つのリプレイカット部の間の区間をリプレイショットとする。

20

この構成によれば、特に動きの激しいシーンにおいて、イントラ符号化されるマクロブロックが多い場合に、このようなフレームをリプレイショットと誤判断するのを防ぐことができる。

【 0 0 6 9 】

よって、上記構成によれば、連続する 2 つの瞬時カット点間に 1 つだけリプレイカット点パターンが存在する場合や、瞬時カット点とリプレイカット点の間が 2GOP 以上ない場合、連続するリプレイカット点間が 6GOP 以上ない場合では、リプレイショットとして判断されない。

【 0 0 7 0 】

次に、ショットの分類について説明する。

サッカー中継での映像のそれぞれのショットを、基本的に図 1 6 のような (1)Center Long Shot ( センターロングショット ) , (2)Goal Long Shot ( ゴールロングショット ) , (3)Close-up Shot ( クローズアップショット ) , (4)Out-field Shot ( アウトフィールドショット ) , (5)Replay Shot ( リプレイショット ) の 5 つのパターンに分類する。

30

そして、動画像データから抽出された各ショットを、前記分類に応じて異なるショット記号により表す。このショット記号は、本実施の形態の一例によれば、前記各ショットの先頭に付された番号 ( 1 ~ 5 の何れか ) としている。

【 0 0 7 1 】

本実施の形態では、ショット分類の判別に用いる特徴量の一つとして、MPEGデータ内の動きベクトルを用いる。動きベクトルでは、前後の参照するフレームのマクロブロック位置を表すため、各ショットでは、映像のカメラワークや選手の動きの速さに応じて、動きベクトルの大きさも変化する。動きベクトルは、P, B フレームから動きベクトルを取り出し、各GOPにおいて、その動きベクトルの総量Mを数 1 ( 式(4.1) ) から求める。

40

【 0 0 7 2 】

【数 1】

$$M = \sum_{\Phi} \sqrt{v_x^2 + v_y^2} \quad (4.1)$$

$\bar{v} = \{v_x, v_y\}$  : 動きベクトル  $\Phi$  : 1 フレーム内のマクロブロック数

【0073】

本願発明者の実験によれば、(1)Center Long Shotでは、選手が大きく映っていないために動きベクトルがカメラワークに大きく影響される。カメラワークが左右に動くときと止まっているときがあるため、それに伴い、数 1 (式(4.1)) で表される動きベクトル量の変動も大きくなり、1 ショットでの平均動きベクトル量は小さくなる傾向にある。

10

また、(3)Close-up shot では、選手が大きく映っているため、動きベクトルが選手の動きに大きく影響される。大きく映されている選手が常に動いているために、前記動きベクトル量の変動も小さくなり、1 ショットでの平均動きベクトル量は大きくなる傾向にある。

【0074】

(1)Center Long Shot と(3)Close-up shot の1 ショットにおけるGOP 単位での動きベクトル量の推移を図 17 に示す。

図 17 のように動きベクトル量の推移に明確な違いがあることがわかる。1 ショットの動きベクトル量をGOP 数で割った平均の動きベクトル量バーM (バーMはMの平均を表すものとする) において、(1)Center Long Shot ではバーMは比較的小さい値となり、(3)Close-up shot では、バーMはLong Shot と比べて大きい値となる。このバーMをLong Shot とClose-up shot の判別に利用する。

20

また、(2)Goal Long Shot も(1)Center Long Shot と同じLong Shot であるので、同様の傾向を示す。

【0075】

次に、動きベクトルの方向による分散値について説明する。

各ショットでは、動きベクトルの参照するマクロブロック位置の方向にそれぞれ特徴があることから、動きベクトルの方向をショット分類の判別に利用する。

30

(1)Center Long Shotでは、選手が大きく映っていないために、動きベクトルの方向はカメラワークの動きの方向になり、フレーム内のそれぞれの動きベクトルの方向はほぼ同じ向きとなる。

(3)Close-up shotでは、選手が大きく映っているため、動きベクトルの参照位置も選手の部位の動きに合わせて、様々な方向となることから、それぞれの動きベクトルの方向も分散する。

【0076】

本実施の形態のカット部検出システム及びショット検出システム並びにシーン検出システム3は、動きベクトルの方向を図 18 のように上下左右の4つの方向に量子化し、各方向に分類されるP, Bフレームでのマクロブロック数を求める。

40

【0077】

図 19 は、(1)Center Long Shotと(3)Close-up shotの一例において、図 18 の方向の量子化による1 フレームでのマクロブロック数の値を示している。この図から、(1)Center Long Shotでは動きベクトルの方向が一方向に集中しており、それに比較して、(3)Close-up shotでは4方向へ分散していることがわかる。

【0078】

量子化されたマクロブロック数から、各方向の偏りを求めるために、以下の数 2 (式(4.2)) から分散度Dを算出する。分散度DをP, Bフレームから1 フレームずつ計算し、1GOPごとにその平均を求める。

【0079】

50



## 【数 2】

$$\begin{aligned}
 D &= \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2 \\
 &= \frac{1}{n} (n_1(1 - \bar{x})^2 + n_2(2 - \bar{x})^2 + n_3(3 - \bar{x})^2 + n_4(4 - \bar{x})^2)
 \end{aligned}
 \tag{4.2}$$

$x_k$  : 各方向の値 (1, 2, 3, 4)

$n_1, n_2, n_3, n_4$  : 方向の量子化により分類された各マクロブロック数

$n = n_1 + n_2 + n_3 + n_4$  : 1 フレームの全マクロブロック数

$$\bar{x} = \frac{1 \cdot n_1 + 2 \cdot n_2 + 3 \cdot n_3 + 4 \cdot n_4}{n}$$

10

## 【0080】

そして、(1)Center Long Shotと(3)Close-up shotの1ショットにおけるGOP単位での分散度Dの推移、および、1ショットでのDの平均値 $\bar{D}$ ( $\bar{D}$ はDの平均を示すものとする)をそれぞれ図20、図21に示す。これらの図より、1ショットでのDの平均値 $\bar{D}$ が、(1)Center Long Shotでは約0.2、(3)Close-up shotでは約0.65となり、分散度に大きな違いがあることがわかる。

20

本実施の形態によれば、前記 $\bar{D}$ を(1)Center Long Shotと(3)Close-up shotの判別に用いる。

## 【0081】

次に、色成分によるグラウンド領域の判別について説明する。

Long Shotの中において、ゴールエリア付近の画面では、グラウンドのセンター付近の画面に比べ、グラウンド領域の境界線の傾きが大きくなるという特徴がある。本実施の形態では、この特徴を利用することで、Long Shotの中で、(1)Center Long Shotと(2)Goal Long Shotを判別する。

## 【0082】

グラウンドの領域を抽出する特徴量としては、MPEGデータ内の輝度・色差成分の色情報を用いる。グラウンド、つまり、芝の色情報としての輝度・色差( $Y, C_b, C_r$ )の範囲をあらかじめ決めておき、その範囲内の輝度・色差を持ったマクロブロックをグラウンド上のブロックと判別する。

30

本実施の形態の一例では、輝度、色差、そしてこれらの比 $r = C_b/C_r$ の範囲を数3(式(4.3))のように定めた。芝の色には、色差成分の割合である $r$ に一定の値となることから、 $r$ と $C_b, C_r$ の色差の範囲と、明るさを表す輝度 $Y$ の範囲を定めることで芝領域に含まれるマクロブロックを抽出することができる。

## 【0083】

## 【数 3】

$$0.68 \leq r \leq 1.1, 85 \leq Y \leq 145, 80 \leq C_b \leq 125, 95 \leq C_r \leq 130
 \tag{4.3}$$

40

## 【0084】

より具体的に説明すれば、カット部検出システム及びショット検出システム並びにシーン検出システム3は、1フレームを平行な複数の水平状のラインに分割し、このライン毎に、輝度成分 $Y$ 、色差成分 $C_b$ 、色差成分 $C_r$ 、色差成分の比 $r = C_b/C_r$ が、それぞれ数3(式(4.3))で示される色条件を満たすマクロブロックの数を求める。

本願発明者の実験によれば、前記数3(式(4.3))を用いることで、グラウンド領域を正確に抽出ことができ、例えば、選手が着用している緑色のユニフォームを避けるこ

50

とができる上、スタジアムの影となっているグラウンド部分の芝を抽出することもできた。

【 0 0 8 5 】

グラウンド領域の境界線の傾きを表すために、抽出したグラウンド部分のマクロブロックの個数を横一列のライン毎に加算した水平投射グラフを図 2 2、図 2 3、図 2 4 に示す。

また、同図 2 2、図 2 3、図 2 4 の各々には、グラウンド領域の水平投射グラフから、一定の範囲内（例えば range=3）で、最小 2 乗法により、グラフの傾きを算出し、その範囲を 1 つずつずらしていくことで、前記水平投射グラフの傾きを表す。

すなわち、前記水平投射グラフの傾きとは、各フレームを平行な複数の水平ラインに分割した場合において、前記水平ラインを垂直方向へ変化させた場合の変化量に対する、芝を示すマクロブロック数の変化の割合を示すことになる。

【 0 0 8 6 】

なお、前記最小 2 乗法とは、 $n$  個のデータ  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  が得られたときに、最もフィットする直線を  $y=ax+b$  とし、この傾き  $a$  を数 4（式 (4.4)）から求める。

【 数 4 】

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (4.4)$$

図 2 2、図 2 3、図 2 4 によれば、垂直方向を  $y$  軸、水平方向を  $x$  軸とする。

【 0 0 8 7 】

図 2 2、図 2 3、図 2 4 は、代表的な (1) Center Long Shot、(2) Goal Long Shot、(3) Close-up Shot の一例について、左から順に、その代表フレームの画像、グラウンド領域の水平投射グラフ、水平投射グラフの傾きを示している。

【 0 0 8 8 】

本実施の形態の一例では、図 2 2、図 2 3、図 2 4 から、グラウンド領域のマクロブロック数が 40 以上の位置（ライン）の数  $X$  と、グラウンド領域のマクロブロック数が 5 以下の位置（ライン）の数  $Y$  と、傾きの大きさの最大値  $Z$  との 3 つの数値を求める。

そして、これらの 3 つの特徴量を、(1) Center Long Shot、(2) Goal Long Shot、(3) Close-up Shot、(4) Out-field Shot 等の判別に用いる。

【 0 0 8 9 】

すなわち、図 2 2 の (1) Center Long Shot では、グラウンド領域の境界線が水平方向の直線となるので、水平投射グラフの傾きの最大値  $Z$  は大きくなる。さらに、水平方向におけるグラウンド領域のマクロブロック数が 40 以上となる垂直方向の位置数  $X$  が 20 以上となり、40 以上の位置数  $X$  と 5 以下の位置数  $Y$  が垂直方向の位置での大部分を占めるため、 $X + Y$  は 25 以上となる。

【 0 0 9 0 】

一方、図 2 3 の (2) Goal Long Shot では、グラウンド領域の境界線が水平方向に対しての傾きが緩やかになるため、水平投射グラフの傾きの最大値  $Z$  は (1) Center Long Shot より小さくなる。また、 $X$  は、(1) Center Long Shot よりも領域が小さくなるので、 $X$  の範囲が 10 以上から 20 未満となり、 $X+Y$  は 10 以上から 25 未満となる。

【 0 0 9 1 】

図 2 4 の (3) Close-up Shot では、選手が映っているので、グラウンド領域と判断されるマクロブロックが少なくなる。このため、 $X$  の領域と  $Y$  の領域が小さくなることから、 $X+Y$  が 15 未満となる。

【 0 0 9 2 】

また、図 2 6 のように、1 フレーム内におけるグラウンド領域のマクロブロック数  $T$  にお

いて、(1)Center Long Shotでは、グラウンド領域のマクロブロックが多く、(3)Close-up Shotではグラウンド領域のマクロブロックが少ないという特徴がある。よって、GOPの1フレーム内のグラウンド領域と判断されたマクロブロック数Tをショット判別に利用する。

【0093】

そして、本実施の形態のカット部検出システム及びショット検出システム並びにシーン検出システム3は、リプレイショット以外のショットの判別に用いる特徴量として、以下の値をそれぞれ算出する。

パーM：1ショットでの動きベクトル量の平均。

パーD：1ショットでの動きベクトルの分散値の平均。

パーT：ショットの最後の3GOPの1フレームにおけるグラウンド領域のマクロブロック数の平均。

パーX：ショットの最後の3GOPの1フレームにおける横一列のグラウンド領域が40以上の位置数(ライン数)の平均

パーY：ショットの最後の3GOPの1フレームにおける横一列のグラウンド領域が5以下の位置数(ライン数)の平均。

パーZ：ショットの最後の3GOPの1フレームにおけるグラウンド領域の傾きの最大値の平均。

【0094】

上述した動きベクトル、及び色情報の特徴量を用いたショット判別の流れは、図27に示すようになる。

【0095】

詳細に説明すれば、先ず、カット部検出システム及びショット検出システム並びにシーン検出システム3は、上記瞬時カット点によって区切られたショットであって、上記リプレイショット以外のショットについて、パーY 27であって、かつ(パーD 0.8又はパーM 40000)の条件を満たす場合には、そのショットを、(4)Out-field Shotと判別する。そして、そうでない場合には、以下の処理を行う。

【0096】

次に、パーY 15、パーT 200、パーD 0.5、パーM 60000の内、何れか一つの条件を満たす場合には、前記ショットを、(3)Close-up Shotと判別する。そうでない場合には、以下の処理を行う。

【0097】

次に、4パーZ 11、8パーX 18、10(パーX + パーY) 25の三つの条件をすべて満たす場合には、前記ショットを、(2)Goal Long Shotと判別する。そうでない場合には、以下の処理を行う。

【0098】

次に、(パーX + パーY) 14という条件を満たす場合には、前記ショットを(3)Close-up Shotと判別し、そうでなければ、前記ショットを(1)Center Long Shotと判別する。

【0099】

なお、図27中の各閾値は、ショット判別の対象がサッカー試合映像である場合に特に好ましい一例を示すものである。

【0100】

また、図示例によれば閾値の上限又は下限を特定していない条件があるが、以下のように適宜な値に特定してもよい。

例えば、パーYの最上限値は、1フレームにおける垂直方向の最大マクロブロック数(例えば30)とすることが可能である。

また、パーDの最上限値は、例えば1とすることが可能である。

また、パーMの最下限値は、例えば0とすることが可能である。

また、パーMの最上限値は、例えば1ショット中の全ての動きベクトルの量とすること

10

20

30

40

50

が可能である。

また、P-Tの最下限値は、例えば0とすることが可能である。

また、(P-X + P-Y)の最下限値は、例えば0とすることが可能である。

#### 【0101】

次に、上記のようにして判別された5つのショットについて、その配列的な特徴からシーンを判別する手段について説明する。

本実施の形態では、検出するシーンとして、ゴールシーンと、シュートシーンを対象とする。ゴールシーンの定義として、シュートしたボールがゴール内に入り、得点が入る場面のこととする。

#### 【0102】

一方、シュートシーンの定義は、攻撃している側の選手が守備側のゴールに向かってシュートし、ボールがゴールの枠内に入らず、ゴールラインを割る場面のこととする。具体的には、攻撃側の選手が蹴ったボールが直接、または、キーパーが弾くなどして、守備側のペナルティエリア内のゴールラインを割り、さらに、シュートのショットの後に、リプレイショットが挿入される場面を対象にしている。すなわち、リプレイショットが入るシーンは、番組の制作者が注目している場面であり、視聴者が特に見たい重要な場面であると考えられるためである。

#### 【0103】

ゴールシーンなどの特定のシーンには、一連のショットの移り変わりに典型的なパターンが存在する。例として、ゴールシーンのショットの流れと、シュート(ゴールはしていない)シーンでのショットの流れをそれぞれ図28と図29に示す。

#### 【0104】

図28に示すように、一般的なサッカー試合映像のシュートシーンでは、(2)Goal Long Shotの後、(3)Close-up Shotが1~3ショット入り、その後(5)Replay Shotが1~2ショット入る傾向にある。

一方、図29のようなゴールシーンでは、(2)Goal Long Shotの後に、(3)Close-up Shotがシュートシーンより多く3~5ショット入り、間に(4)Out-field Shotも入ることが多い。また、リプレイショットの数も3~4ショットとシュートシーンよりも多い傾向にある。

本実施の形態は、このようなショットの遷移パターンの特徴を利用することで、ゴールシーンとシュートシーンを区別し、検出するようにしている。

#### 【0105】

すなわち、本実施の形態のカット部検出システム及びショット検出システム並びにシーン検出システム3は、各ショットを図16の分類によって、記号化、つまり、各ショットのラベル付けを行ない、映像データを対応するショットの記号列に置き換える。

図28や図29のように特定のシーンには特定のショットの移り変わりにパターンがあることから、ショットの記号列においてもシーン固有の典型的なパターンが存在する。このようなシーンの典型的な記号列パターンと、映像データの記号列とについて、パターンマッチング処理を行なうことでシーンを検出する。パターンマッチング処理には、後述するDPマッチングを用いる。

#### 【0106】

より具体的に説明すれば、カット部検出システム及びショット検出システム並びにシーン検出システム3は、図30に示すように、動画データから検出された各ショットに対しその種類に応じたショット記号(図示例によれば1~5の番号)を付与することで、ショット記号列(図示例ではショットタイプ列と称す)を作成し、このショット記号列を、記憶装置の所定の記憶領域(ショット記号列記憶手段)に記憶する。

#### 【0107】

なお、検出対象のシーンに対応するように複数のショット記号を配列してなるモデルパターン記号列(図示例によれば、2335や2333435等の数字列)は、予め、記憶装置の所定の記憶領域(モデルパターン記憶手段)に記憶しておく。

10

20

30

40

50

## 【0108】

このモデルパターン記号列の一例としては、図28（シュートシーンでのショットの流れ）に示すように、グラウンドを映しているショットであって且つ最後の三つのグループオブピクチャがゴール付近の映像であるゴールロングショットを示すショット記号（2）と、選手のアップを映しているクローズアップショットを示すショット記号（3）と、先の映像の繰り返しを映しているリプレイショットを示すショット記号（5）とを、この順番に配列してなる記号列「235」や、記号列「2335」等とすればよい。

## 【0109】

また、モデルパターン記号列の他例としては、図29（ゴールシーンでのショットの流れ）に示すように、グラウンドを映しているショットであって且つ最後の三つのグループオブピクチャがゴール付近の映像であるゴールロングショットを示すショット記号（2）と、選手のアップを映しているクローズアップショットを示す複数のショット記号（3）と、フィールド外を映しているアウトフィールドショットを示すショット記号（4）と、選手のアップを再度映しているクローズアップショットを示すショット記号（3）と、先の映像の繰り返しを映しているリプレイショットを示すショット記号（5）とを、この順番に配列してなる記号列「23435」や、記号列「233435」、記号列「2333435」等とすればよい。

10

## 【0110】

そして、本実施の形態のカット部検出システム及びショット検出システム並びにシーン検出システム3は、前記ショット記号列記憶手段のショット記号列中から、前記モデルパターン記憶手段のモデルパターン記号列に対応する記号列を検出する検出手段を備える。

20

この検出手段は、前記ショット記号列に対し前記モデルパターン記号列を照合させて、前記ショット記号列中から前記モデルパターン記号列に略合致する記号列を検出するパターンマッチング処理を行うものである。

## 【0111】

前記処理において、ショット記号列中の記号列と、モデルパターン記号列とは、予め設定された適宜な類似度で合致するようにすればよい。

例えば、図30によれば、ショット記号列中の「235」に対し、モデルパターン記号列「2335」は、完全一致しないが、略合致したものと判断されている。同様に、ショット記号列中の「233435」に対し、モデルパターン記号列「2333435」も略合致したものと判断されている。

30

## 【0112】

本実施の形態では、前記パターンマッチング処理の具体的手法として、DPマッチング処理を用いている。以下に、このDPマッチング処理について詳述する。

時系列パターンの類似度を求める手法としては、音声認識などによく用いられる動的計画法（DPマッチング）や、隠れマルコフモデル（HMM）などがある。

Dynamic Programming Matching（DPマッチング）は、異なる二つの時系列信号の類似度と最適な対応を求める手法であり、音声認識等のさまざまな分野で利用されている。DPマッチングは非線形の伸縮により、長さの異なるパターンの類似度を計算することができる。これらの利点により、本実施の形態では、時系列パターン間の類似度を求める手法として、DPマッチングを用いる。

40

DPマッチングの具体的な実現方法について、図31を用いて説明する。

2つのパターンAとBを特徴ベクトルの時系列として表現する。つまり、数5（5.1式）とする。

## 【0113】

【数 5】

$$\begin{aligned} A &= a_1 a_2 \cdots a_I \\ B &= b_1 b_2 \cdots b_J \end{aligned} \quad (5.1)$$

【0 1 1 4】

ここで、A,B 両パターンの時間の対応、すなわち時間伸縮関数は、図 3 1 の平面上の格子点  $c_k = (i, j)$  の系列 F で表現することができる。

【0 1 1 5】

10

【数 6】

$$F = c_1 c_2 \cdots c_K \quad (5.2)$$

【0 1 1 6】

2つの特徴ベクトル列  $a_i$  と  $b_j$  との(スペクトル)距離を  $d(c) = d(i, j)$  で表すと、F に沿った距離の総和は、数 7 (5.3式) で表すことができ、この値が小さいほど A と B の対応づけがよいことを示す。ここで、 $w_k$  は F に関連した正の重み関数である

20

ここで、上式を、次のような制限のもとで、F に関して最小化することを考える。

【0 1 1 7】

【数 7】

$$D(F) = \frac{\sum_{k=1}^K d(c_k) w_k}{\sum_{k=1}^K w_k} \quad (5.3)$$

## 【数 8】

(1) 単調性と連続性の条件

$$0 \leq i_k - i_{k-1} \leq 1, 0 \leq j_k - j_{k-1} \leq 1$$

(2) 境界条件

$$i_1 = j_1 = 1, i_K = I, j_K = J$$

(3) 整合窓の条件：極端な伸縮を防ぐため  $r$  を定数として、

$$|i_k - j_k| \leq r$$

部分点列  $c_1 c_2 \cdots c_k (c_k = (i, j))$  に対する部分和を考えると、

$$\begin{aligned} g(c_k) = g(i, j) &= \min_{c_1, \dots, c_{k-1}} \left[ \sum_{l=1}^k d(c_l) w_l \right] \\ &= \min_{c_1, \dots, c_{k-1}} \left[ \sum_{l=1}^{k-1} d(c_l) w_l + d(c_k) w_k \right] \\ &= \min_{c_1, \dots, c_{k-1}} \left[ \min_{c_1, \dots, c_{k-2}} \left\{ \sum_{l=1}^{k-1} d(c_l) w_l \right\} + d(c_k) w_k \right] \\ &= \min_{c_{k-1}} [g(c_{k-1}) + d(c_k) w_k] \end{aligned} \quad (5.4)$$

## 【0118】

点列  $F$  に沿ってとった距離の荷重平均を考え、その値を最小とるように  $F$  を変化させ、2つのパターン  $A$  と  $B$  のマッチングを行う。ここで  $F$  の変化には、あまり極端な時間軸の伸縮は許されないように制限をつけ、図 3 2 に示す 3 通りのみを考えることにする。2つの特徴ベクトル  $a_i$  と  $b_j$  との距離を  $d(c) = d(i, j)$  で表し、部分点列  $c_1 c_2 \cdots c_k$  に対する部分和を  $g(c_k) = g(i, j)$  とすると、最適性の原理(最適政策の部分系列はやはり最適政策になっている)により、数 9 (式(5.5)) のような漸化的な表現となる。

## 【0119】

## 【数 9】

$$g(i, j) = \min \left\{ \begin{array}{ll} g(i-1, j) & : (a) \\ g(i-1, j-1) & : (b) \\ g(i, j-1) & : (c) \end{array} \right\} + d(i, j) \quad (5.5)$$

## 【0120】

このようにして計算していき、最終的に求められた  $g(I, J)$  が 2つのパターン  $A$  と  $B$  との距離の最小値であり、その時の  $F$  が最適なマッチングを表す時間変換関数といえる。

格子点  $(i, j)$  までに選択された経路の長さを数 10 (式(5.6)) より求める。

## 【0121】

10

20

30

40

## 【数 1 0】

$$c(i, j) = \begin{cases} c(i-1, j) + 1 & : if (a) \\ c(i-1, j-1) + 2 & : if (b) \\ c(i, j-1) + 1 & : if (c) \end{cases} \quad (5.6)$$

## 【0 1 2 2】

$i = J$  まで、到達したときに、数 9 (式(5.5)) の累積距離を数 1 0 (式(5.6)) の経路の長さで割ることで距離の正規化をはかり、整合窓の中における  $G(i, J)$  が最小となる経路を求めることで、最適な経路を探し出すことができる。

## 【0 1 2 3】

## 【数 1 1】

$$G(i, J) = \frac{g(i, J)}{c(i, J)} \quad (5.7)$$

## 【0 1 2 4】

さらに、図 3 3 において、Input Pattern における  $i = 0$  の始点から  $G(i, J)$  の値が最小となる  $\min(G)$  の  $i$  までが、その始点において、最も Sample Pattern とマッチングされた系列となる。さらに、始点を  $i = 1, 2, \dots, l$  と 1 ずつ移動させながら、 $\min(G)$  までの Input Pattern の系列をそれぞれ求める。その中から、 $\min(G)$  が閾値以下となるパターンを最も Sample Pattern と類似した系列として求めることができる。

## 【0 1 2 5】

次に、上記構成のカット部検出システム及びショット検出システム並びにシーン検出システム 3 を用いて、実際のサッカー試合映像の MPEG2 データからカット点、ショット、リプレイショット、シーンを検出した結果について説明する。

ここで、前記サッカー試合映像の動画は、解像度  $640 \times 480$ 、ビットレートは平均 5.5 Mbps、最大 8 Mbps の VBR (Variable Bit Rate, 可変速度)、プロファイルとレベルは、MP@ML、GOP は IBBPBBPBBPBBPBB1BB $\dots$  という P フレームの後に B フレームが 2 枚続く  $N=2$  の構成をしている。1 つの GOP は基本的には I フレーム 1 枚、P フレーム 4 枚、B フレーム 10 枚の計 15 枚だが、それ以外のランダム GOP も存在し、その場合のランダム GOP は 15 枚未満の 3 の倍数で構成される。

結果は、図 3 4 ~ 3 7 に示す通りであり、カット点、ショット、リプレイショット、シーンの何れにおいても、比較的高確率の検出結果を得ることができた。

## 【0 1 2 6】

なお、上記実施の形態によれば、サッカー試合映像を対象に、カット点、ショット、リプレイショット、シーン等の検出を行ったが、他例としては、上述した色情報範囲 ( $0.68 \leq r \leq 1.1, 85 \leq Y \leq 145, 80 \leq Cb \leq 125, 95 \leq Cr \leq 130$ ) や特徴量 (バー M, バー D, バー T, バー X, バー Y, バー Z) の範囲等を適宜に変更することで、例えば、相撲の映像、野球映像、メロドラマ等、他の種類の動画を対象にすることも可能である。

## 【0 1 2 7】

また、上記実施の形態は、特に好ましい態様として MPEG 2 データを扱う態様としたが、上記したマクロブロック情報や、動きベクトル情報、輝度・色差情報等を含む他の動画圧縮データを扱うことも可能であり、例えば、MPEG 1 データや、MPEG 4 データ、H. 264 データ、その他の動画圧縮データを処理する可能性を有する。

## 【図面の簡単な説明】



【 0 1 2 8 】

【 図 1 】 本発明に係わるカット部検出システム及びショット検出システム並びにシーン検出システムの一例を示すブロック図である。

【 図 2 】 一般的な 4 2 0 フォーマットのマクロブロックを示す概念図である。

【 図 3 】 一般的なマクロブロックの輝度・色差成分について示す概念図である。

【 図 4 】 本発明の一例において、特徴量が抽出されるフレームを示す概念図である。

【 図 5 】 一般的な動画像の構造を示す概念図である。

【 図 6 】 動画像圧縮データについて、再生順、ビットストリーム順、出力順の関係を説明する概念図である。

【 図 7 】 リプレイショットについて説明する概念図である。

10

【 図 8 】 リプレイカット部の一例を示す画像である。

【 図 9 】 リプレイカットとリプレイショットの関係を示す概念図である。

【 図 1 0 】 カット点が発生する位置と B フレームが参照するフレームとの関係を示す概念図である。

【 図 1 1 】 カット点が発生する位置と B フレームが参照するフレームとの関係を示す概念図である。

【 図 1 2 】 M B T 記号と該 M B T 記号を与える条件との関係を示す表である。

【 図 1 3 】 瞬時カット点およびリプレイカット点のパターンを説明する概念図である。

【 図 1 4 】 リプレイカット部でのマクロブロックタイプを示す概念図である。

【 図 1 5 】 リプレイショットの判別について示す概念図である。

20

【 図 1 6 】 ショットの分類の一例を示す画像である。

【 図 1 7 】 1 ショットでの動きベクトル量の推移を示すグラフである。

【 図 1 8 】 方向の量子化を示す概念図である。

【 図 1 9 】 量子化された動きベクトルの方向とマクロブロック数の関係を示すグラフである。

【 図 2 0 】 センターロングショットでの動きベクトル方向の分散度の推移を示すグラフである。

【 図 2 1 】 クローズアップショットでの動きベクトル方向の分散度の推移を示すグラフである。

【 図 2 2 】 センターロングショットでのグラウンド領域の傾きについて説明するグラフである。

30

【 図 2 3 】 ゴールロングショットでのグラウンド領域の傾きについて説明するグラフである。

【 図 2 4 】 クローズアップショットでのグラウンド領域の傾きについて説明するグラフである。

【 図 2 5 】 水平投射グラフの傾きを説明するグラフである。

【 図 2 6 】 I フレーム内におけるグラウンド領域のマクロブロック数を示すグラフである。

【 図 2 7 】 ショット判別のフローを示す概念図である。

【 図 2 8 】 シュートシーンのパターンの一例を示す画像である。

40

【 図 2 9 】 ゴールシーンのパターンの一例を示す画像である。

【 図 3 0 】 シーン検出のフローを示す概念図である。

【 図 3 1 】 D P マッチング処理においてパターン A , B の時間の対応を示すグラフである。

【 図 3 2 】 D P マッチング処理において点 ( i , j ) における部分和  $g(i, j)$  の計算を示す概念図である。

【 図 3 3 】 D P マッチング処理におけるパターンマッチング方法を示す概念図である。

【 図 3 4 】 カット点の検出方法及び検出結果を示す説明図及び表である。

【 図 3 5 】 ショットの分類方法及び分類結果を示す説明図及び表である。

【 図 3 6 】 リプレイショットの検出方法及び検出結果を示す説明図及び表である。

50

【図37】シーンの検出方法及び検出結果を示す説明図及び表である。

【符号の説明】

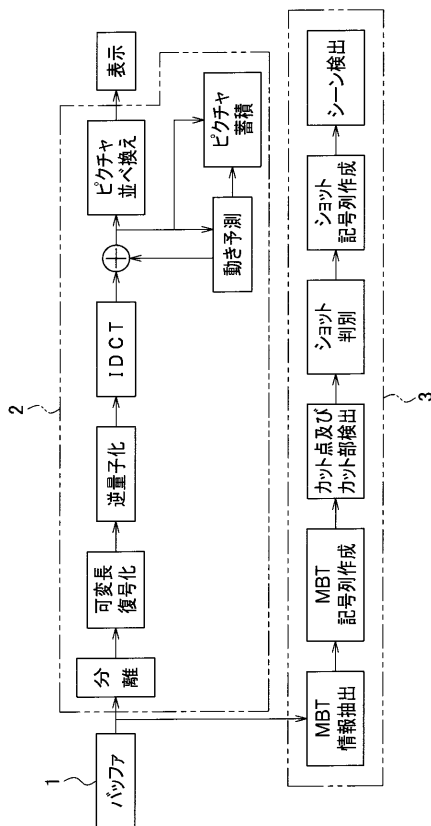
【0129】

1：バッファ

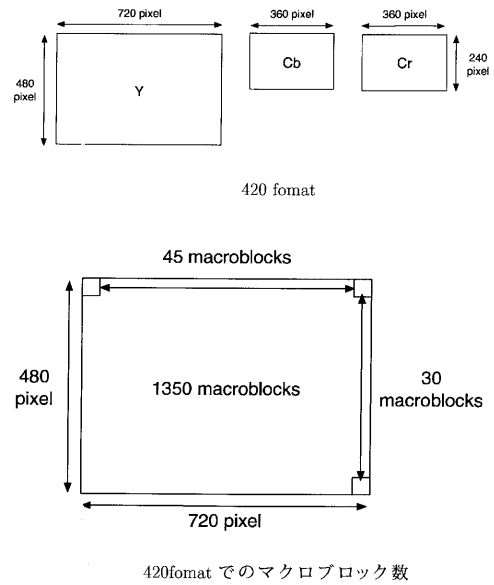
2：復号器

3：カット部検出システム及びショット検出システム並びにシーン検出システム

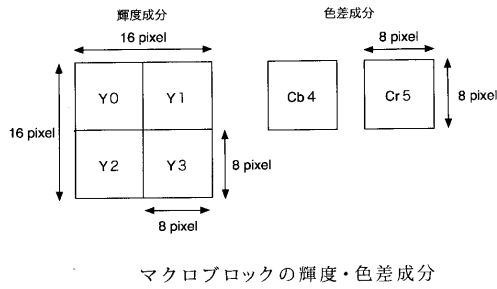
【図1】



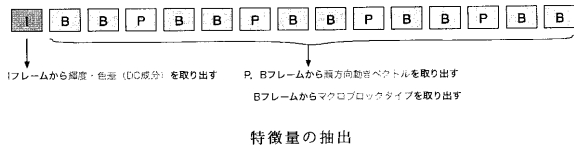
【図2】



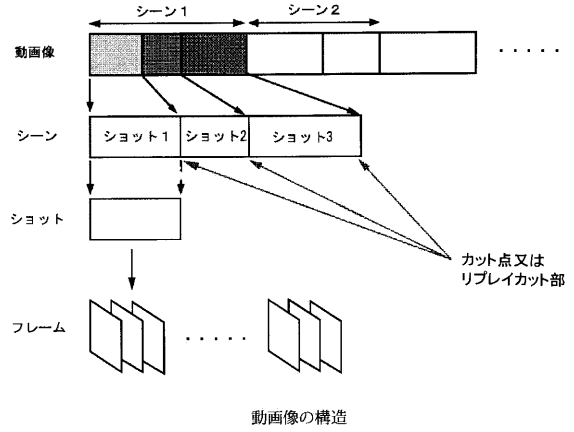
【 図 3 】



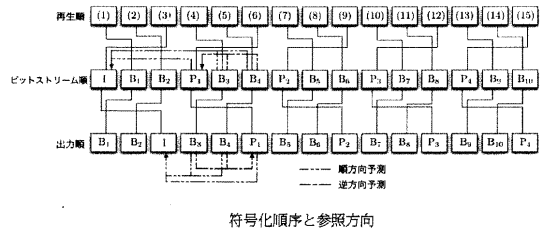
【 図 4 】



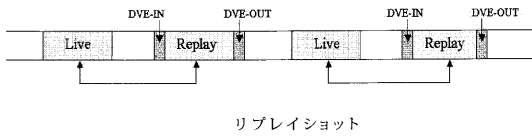
【 図 5 】



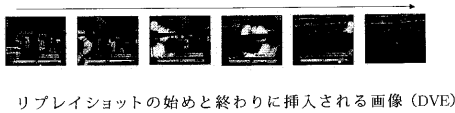
【 図 6 】



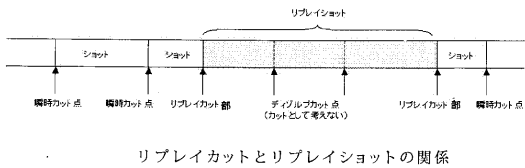
【 図 7 】



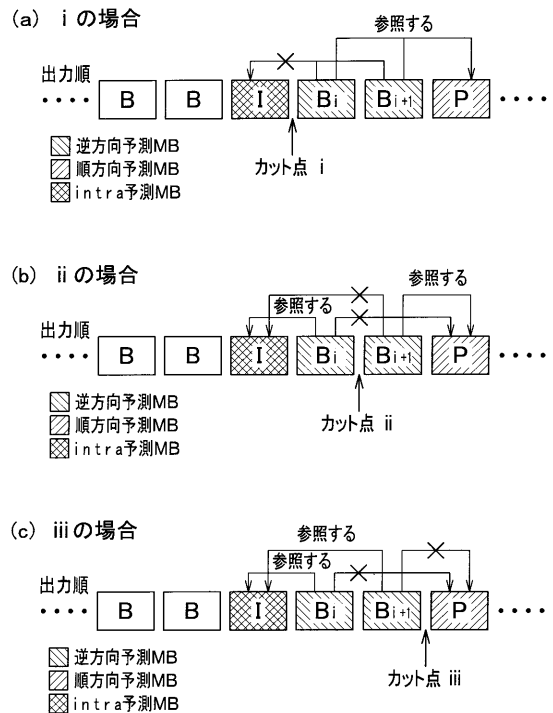
【 図 8 】



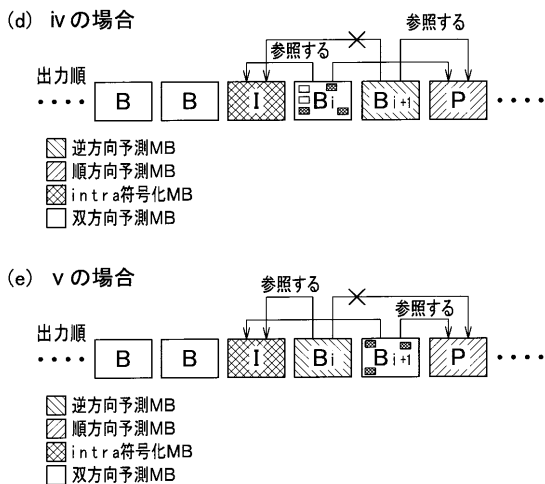
【 図 9 】



【 図 10 】



【 図 1 1 】

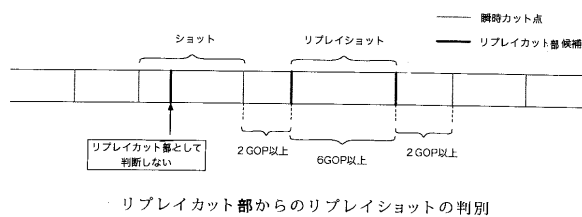


【 図 1 2 】

FrameType	条件
0	$mbForward$ が最頻度 かつ $100 > (mbBack + mbInter)$
1	$mbForward$ が最頻度 かつ $350 > (mbBack + mbInter) \geq 100$
2	$mbForward$ が最頻度 かつ $(mbBack + mbInter) \geq 350$
3	$mbBack$ が最頻度 かつ $100 > (mbForward + mbInter)$
4	$mbBack$ が最頻度 かつ $350 > (mbForward + mbInter) \geq 100$
5	$mbBack$ が最頻度 かつ $(mbForward + mbInter) \geq 350$
6	$mbInter$ が最頻度
7	$mbIntra \geq 100$ (FrameType が 2,5,6 のとき)
8	$100 > mbIntra \geq 20$ (FrameType が 2,5,6 のとき)

FrameType の条件

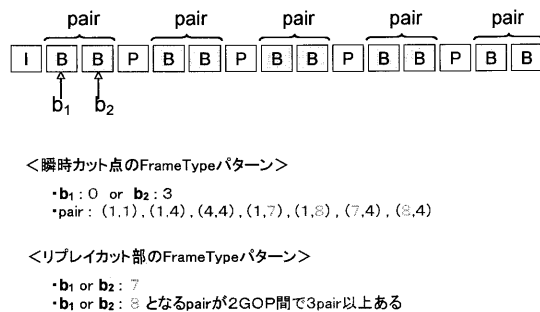
【 図 1 5 】



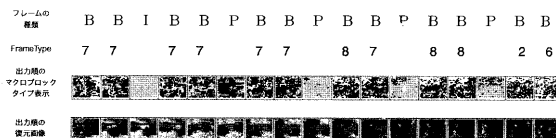
【 図 1 6 】



【 図 1 3 】



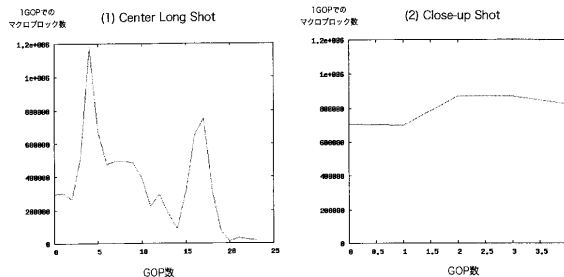
【 図 1 4 】



▨ 逆方向予測MB ▤ Intra符号化MB  
 ▩ 順方向予測MB □ 双方向予測MB

リプレイカット部でのマクロブロックタイプ

【 図 1 7 】



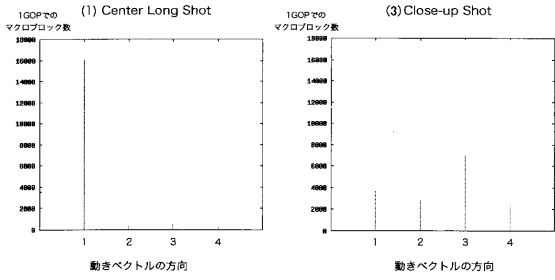
1 Shot での動きベクトル量の推移

【 図 1 8 】



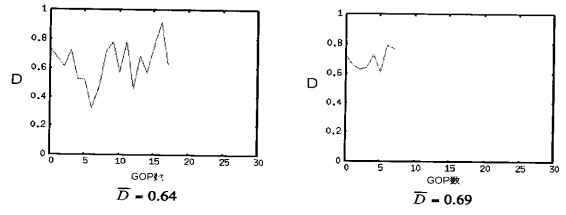
方向の量子化

【 図 1 9 】



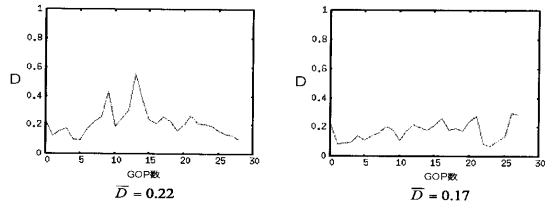
動きベクトルの方向による量子化

【 図 2 1 】



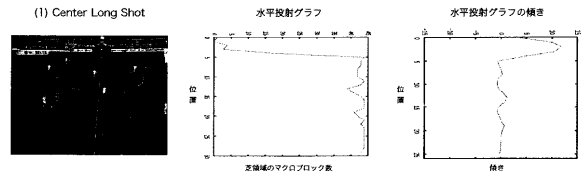
(3)Close-up Shot での動きベクトル方向の分散度の推移

【 図 2 0 】



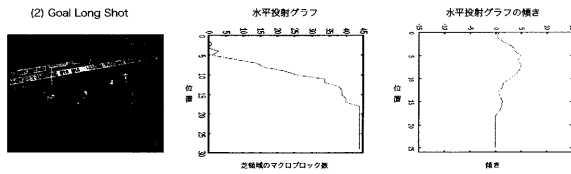
(1)Center Long Shot での動きベクトル方向の分散度の推移

【 図 2 2 】



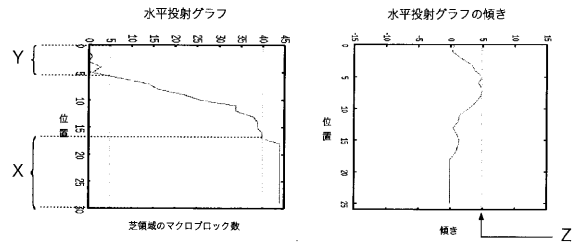
(1)Center Long Shot でのグラウンド領域の傾き

【 図 2 3 】

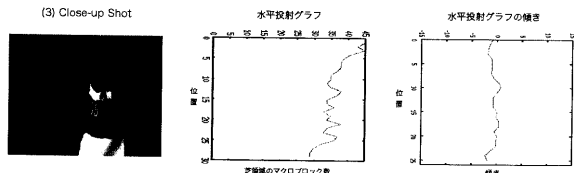


(2)Goal Long Shot でのグラウンド領域の傾き

【 図 2 5 】



【 図 2 4 】



(3)Close-up Shot でのグラウンド領域の傾き

【 図 2 6 】

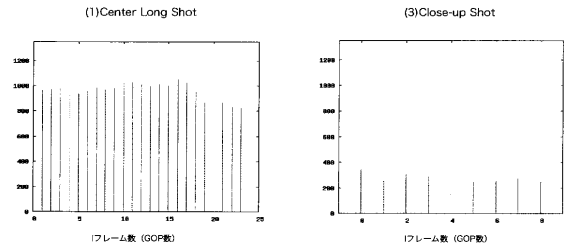
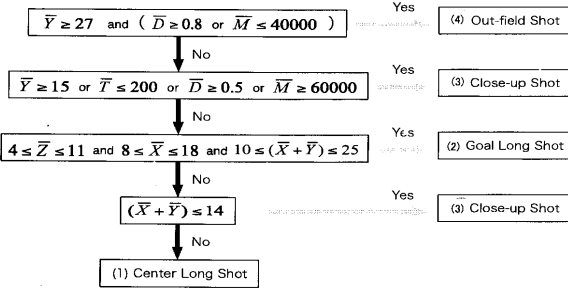
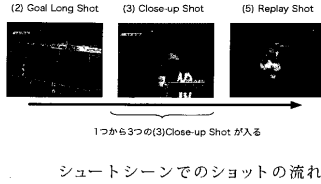


図 4.15: 1フレーム内におけるグラウンド領域のマクロブロック数

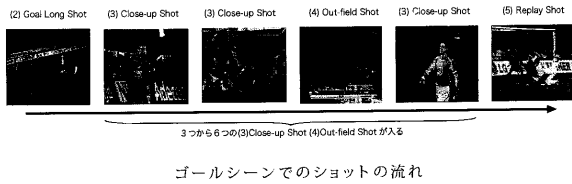
【 図 2 7 】



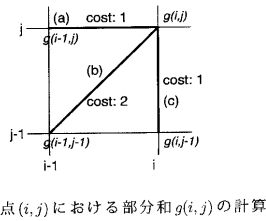
【 図 2 8 】



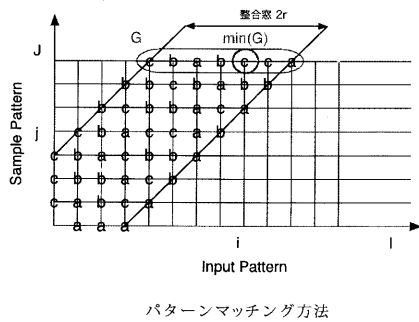
【 図 2 9 】



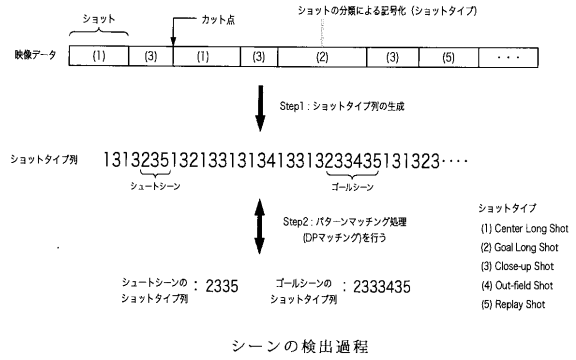
【 図 3 2 】



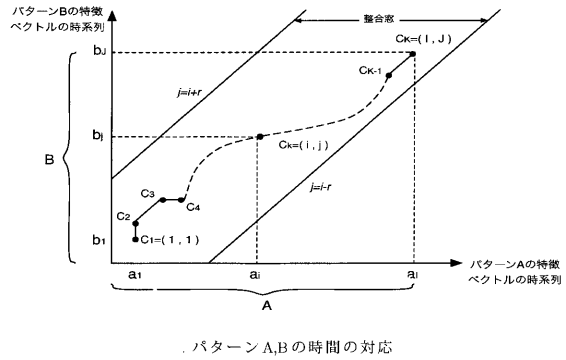
【 図 3 3 】



【 図 3 0 】



【 図 3 1 】



【 図 3 4 】

カット点の検出結果

- カット点検出の評価方法
  - CC(Correct Cut) : 既知の正解カット点数
  - CDC(Correct Detected Cut) : 正しく検出されたカット点数
  - ADC(All Detected Cut) : 検出した全カット点数
  - Recall: カット点検出の再現率  $Recall = CDC / CC \times 100$
  - Precision: カット点検出の適合率  $Precision = CDC / ADC \times 100$
- カット点の検出結果

	CC	CDC	ADC	Recall	Precision
Soccer1 (90分)	657	629	636	95.7%	98.9%
Soccer2 (90分)	669	635	649	94.9%	97.9%
6試合の合計	3298	3146	3298	95.3%	97.3%

【 図 3 5 】

ショットの分類結果

- ショット分類の評価方法
  - CSH(Correct SHot) : 既知の正解ショット数
  - CCSH Correct Classified SHot : 正しく分類されたショット数
  - CDSH Correct Detected SHot : ショットの区間が正しく検出されたショット数
  - RCC(Rate of Correct Classification) : ショットの分類率  $RCC = CCSH / CDSH \times 100$
  - RCR(Rate of Correct Recognition) : ショットの認識率  $RCR = CCSH / CSH \times 100$
- ショットの分類結果

	CSH	CCSH	CDSH	RCC	RCR
Soccer1 (90分)	657	557	597	93.3%	84.8%
Soccer2 (90分)	669	530	616	86.0%	79.2%
6試合の合計	3298	2680	2987	89.7%	81.3%

## 【 図 3 6 】

## リプレイショットの検出結果

## ■リプレイショット検出の評価方法

CRS (Correct Replay Shot) : 既知の正解リプレイショット数  
 CDRS (Correct Detected Replay Shot) : 正しく検出されたリプレイショット数  
 ADRS (All Detected Replay Shot) : 検出した全リプレイショット数  
 Recall : リプレイショット検出の再現率  $Recall = CDRS / CRS \times 100$   
 Precision : リプレイショット検出の適合率

## ■リプレイショットの検出結果

	CRS	CDRS	ADRS	Recall	Precision
Soccer1 (90分)	25	20	23	80.0%	87.0%
Soccer2 (90分)	36	29	32	80.6%	90.6%

## 【 図 3 7 】

## シーンの抽出結果

## ■シーン抽出の評価方法

CS (Correct Scene) : 既知の正解シーン数  
 CDS (Correct Detected Scene) : 正しく検出されたシーン数  
 ADS (All Detected Scene) : 検出した全シーン数  
 Recall : シーン検出の再現率  $Recall = CDS / CS \times 100$   
 Precision : シーン検出の適合率  $Precision = CDS / ADS \times 100$

## ■シーンの抽出結果

(6試合での合計)

	CS	CDS	ADS	Recall	Precision
シュートシーン	77	58	72	75.3%	80.6%
ゴールシーン	19	16	21	84.2%	76.2%

---

フロントページの続き

(72)発明者 政木 康生

大阪府大東市中垣内7丁目7番1号 船井電機株式会社内

(72)発明者 坪田 浩乃

大阪府大東市中垣内7丁目7番1号 船井電機株式会社内

Fターム(参考) 5C059 MA00 MA05 MA14 MA23 MC11 MC38 ME01 PP04 PP07 PP16  
SS02 SS08 SS19 TA00 TB04 TB07 TC12 TC27 TD03 TD07  
TD12 UA05  
5C159 MA00 MA05