

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2010-530702

(P2010-530702A)

(43) 公表日 平成22年9月9日(2010.9.9)

(51) Int.Cl. F I テーマコード (参考)
H04N 7/26 (2006.01) H04N 7/13 Z 5C159

審査請求 未請求 予備審査請求 未請求 (全 42 頁)

(21) 出願番号	特願2010-513120 (P2010-513120)	(71) 出願人	596180076
(86) (22) 出願日	平成20年6月19日 (2008.6.19)		韓国電子通信研究院
(85) 翻訳文提出日	平成22年2月22日 (2010.2.22)		Electronics and Telecommunications Research Institute
(86) 国際出願番号	PCT/KR2008/003481		大韓民国大田廣域市儒城區柯亭洞161
(87) 国際公開番号	W02008/156318		161 Kajong-dong, Yulsong-gu, Taejeon Korea
(87) 国際公開日	平成20年12月24日 (2008.12.24)		
(31) 優先権主張番号	10-2007-0059709	(74) 代理人	100077481
(32) 優先日	平成19年6月19日 (2007.6.19)		弁理士 谷 義一
(33) 優先権主張国	韓国 (KR)	(74) 代理人	100088915
(31) 優先権主張番号	10-2007-0061362		弁理士 阿部 和夫
(32) 優先日	平成19年6月22日 (2007.6.22)		
(33) 優先権主張国	韓国 (KR)		
(31) 優先権主張番号	10-2007-0096267		
(32) 優先日	平成19年9月21日 (2007.9.21)		
(33) 優先権主張国	韓国 (KR)		

最終頁に続く

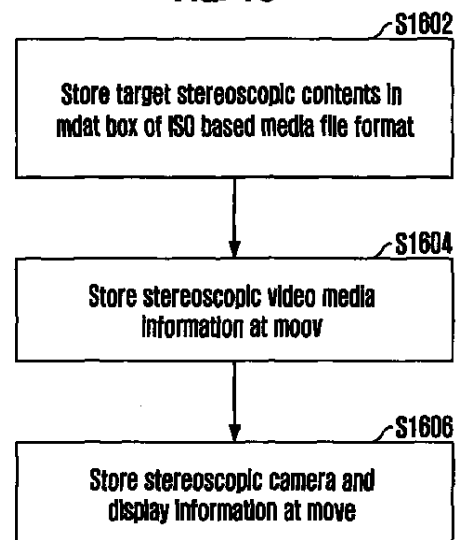
(54) 【発明の名称】 ステレオスコピックデータの保存および再生のためのメタデータ構造、およびこれを用いるステレオスコピックコンテンツファイルの保存方法

(57) 【要約】

本発明は、ステレオスコピックデータの保存および再生のためのメタデータ構造、およびこれを用いるステレオスコピックコンテンツファイルの保存方法に関するものである。

本発明による保存方法は、ステレオスコピックコンテンツの区間情報に関するステレオスコピックビデオメディア情報を、ステレオスコピックコンテンツを再生するために要求されるメタデータ情報として保存するステップと、ステレオスコピックコンテンツの符号化ストリームを保存するステップとを含む。ステレオスコピックデータは、モノスコピック（2次元）およびステレオスコピック（3次元）データが混用で構成される場合またはステレオスコピックデータだけ構成でされる場合のすべてを意味する。

FIG. 16



【特許請求の範囲】**【請求項 1】**

ステレオスコピックコンテンツを保存する方法であって、

前記ステレオスコピックコンテンツの区間情報に関するステレオスコピックビデオメディア情報を、前記ステレオスコピックコンテンツを再生するために要求されるメタデータ情報として保存するステップと、

前記ステレオスコピックコンテンツの符号化ストリームを保存するステップとを含むことを特徴とするステレオスコピックコンテンツの保存方法。

【請求項 2】

前記ステレオスコピックビデオメディア情報が、メディアファイルフォーマットの別のボックスで定義され保存されることを特徴とする請求項 1 に記載のステレオスコピックコンテンツの保存方法。

10

【請求項 3】

前記ステレオスコピックコンテンツに対するステレオスコピックカメラおよびディスプレイ情報を、前記ステレオスコピックコンテンツを再生するために要求されるメタデータ情報として保存するステップをさらに含み、

前記ステレオスコピックカメラおよびディスプレイ情報が、メディアファイルフォーマットの別のボックスで定義され保存されることを特徴とする請求項 1 に記載のステレオスコピックコンテンツの保存方法。

20

【請求項 4】

前記ステレオスコピックビデオメディア情報が、ステレオスコピック映像構成タイプ情報を含むことを特徴とする請求項 1 に記載のステレオスコピックコンテンツの保存方法。

【請求項 5】

前記ステレオスコピック映像構成タイプ情報が、サイドバイサイド (side by side)、フレームシーケンシャル (frame sequential)、および左右ビューシーケンス (left/right view sequence) のうち少なくとも 1 つを指し示す情報を含むことを特徴とする請求項 4 に記載のステレオスコピックコンテンツの保存方法。

【請求項 6】

前記ステレオスコピックビデオメディア情報が、前記ステレオスコピックコンテンツの基本符号化ストリームがステレオスコピック区間からモノスコピック区間に変更される場合、およびモノスコピック区間からステレオスコピック区間に変更される場合の、変更される区間の数を表示する情報を含むことを特徴とする請求項 1 に記載のステレオスコピックコンテンツの保存方法。

30

【請求項 7】

前記ステレオスコピックビデオメディア情報が、連続するモノスコピックまたはステレオスコピックサンプルの数をカウントする情報を含むことを特徴とする請求項 1 に記載のステレオスコピックコンテンツの保存方法。

【請求項 8】

前記ステレオスコピックビデオメディア情報が、連続するモノスコピックまたはステレオスコピック開始サンプルの数をカウントする情報を含むことを特徴とする請求項 1 に記載のステレオスコピックコンテンツの保存方法。

40

【請求項 9】

前記ステレオスコピックビデオメディア情報が、現在のサンプルがステレオスコピックであるか否かを表示する情報を含むことを特徴とする請求項 1 に記載のステレオスコピックコンテンツの保存方法。

【請求項 10】

前記ステレオスコピックカメラおよびディスプレイ情報が、カメラパラメータが存在するか否か、左右のカメラ間の距離、左右のカメラの配置、ディスプレイパラメータが存在するか否か、左右同一焦点距離 (focal length)、ベースラインから収束点 (convergence point) までの距離、および左右のカメラ移動 (translation) のうち、少なくとも 1 つを

50

表す情報を含むことを特徴とする請求項 3 に記載のステレオスコピックコンテンツの保存方法。

【請求項 1 1】

前記ステレオスコピックカメラおよびディスプレイ情報が、左右の映像間の最大視差および最小視差を含むことを特徴とする請求項 3 に記載のステレオスコピックコンテンツの保存方法。

【請求項 1 2】

前記ステレオスコピックコンテンツ保存方法が、ISO ベースメディアファイルを用いて行われることを特徴とする請求項 1 に記載のステレオスコピックコンテンツの保存方法。

10

【請求項 1 3】

前記ステレオスコピックカメラおよびディスプレイ情報のうち、ステレオスコピック区間を識別するためのアイテム ID は、アイテムロケーションボックス (item location box) のアイテム ID と連動して使用されることを特徴とする請求項 1 2 に記載のステレオスコピックコンテンツの保存方法。

【請求項 1 4】

ステレオスコピックコンテンツを保存するファイルフォーマット構造であって、

前記ステレオスコピックコンテンツの符号化ストリームを保存するメディアデータボックスと、

前記ステレオスコピックコンテンツの区間情報に関するステレオスコピックビデオメディア情報を保存するボックスと

20

を備えることを特徴とするファイルフォーマット構造。

【請求項 1 5】

前記ステレオスコピックコンテンツの基本符号化ストリームが 2 個以上の場合、前記基本符号化ストリームを主要トラックおよび付加トラックに区分する情報を保存するトラック参照ボックスをさらに含むことを特徴とする請求項 1 4 に記載のファイルフォーマット構造。

【発明の詳細な説明】

【技術分野】

【0001】

30

本発明は、ステレオスコピックデータの保存および再生のためのメタデータ構造、およびこれを用いるステレオスコピックコンテンツファイル保存方法に関し、特に 2 次元および 3 次元コンテンツがともに供給される場合に要求されるメタデータ構造 (制御情報) およびこれを用いるステレオスコピックコンテンツファイルの保存方法に関するものである。

【0002】

本発明は、M I C (Ministry of Information and Communication) / I I T A (Institute for Information Technology Advancement) の I T 研究開発プログラムから導出されたものである [課題管理番号 : 2 0 0 5 - S - 4 0 3 - 0 2 、課題名 : 知能型統合情報放送 (S m a r T V) 技術開発] 。

40

【背景技術】

【0003】

ステレオスコピックコンテンツを用いた応用サービスおよび関連機器に対する市場が国内外で携帯電話、デジタルカメラ、DVD、PDP などを中心に形成されている。これと関連して、ステレオスコピックコンテンツの獲得、保存および再生のために要求されるシステム情報または制御情報 (すなわち、メタデータ) と、これを含むファイルフォーマットに対する標準化が求められている。

【0004】

韓国公開特許第 2 0 0 6 - 0 0 5 6 0 7 0 号 (発明の名称 : M P E G - 4 オブジェクト記述子情報および構造を用いた 3 次元動画処理装置およびその方法、以下「特許文献 1」

50

という)では、3次元動画の種類、多様なディスプレイ方式および視点数などのような新しく要求される情報を追加した3次元動画オブジェクト記述子が開示されている。また、韓国特許出願第2006-0100258号(発明の名称:ステレオスコピック映像データの伝送方法、以下「特許文献2」という)では、ステレオスコピックコンテンツのデコードされた映像のデータに対するファイルフォーマットであって、ステレオスコピック映像情報を含む映像データ部、ステレオスコピック映像情報をデコードおよび再生するためのメタデータを含むヘッダ部を備えるファイルフォーマットが開示されている。

【0005】

しかし特許文献1および2は、2次元コンテンツおよび3次元コンテンツがともに構成され提供される場合(すなわち、2次元コンテンツおよび3次元コンテンツが混用される場合)にコンテンツを識別する方法、3次元コンテンツが異なるステレオスコピックカメラおよびディスプレイ情報を有する場合にステレオスコピックカメラおよびディスプレイ情報を提供する方法、2個の符号化ストリームで構成される3次元コンテンツの場合にステレオスコピックトラック参照情報などを提供することができない。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】韓国公開特許第2006-0056070号

【特許文献2】韓国特許出願第2006-0100258号

【発明の概要】

【発明が解決しようとする課題】

【0007】

本発明は、このような問題を解決するために提案されたもので、その目的は、ステレオスコピックコンテンツの保存方法を提供することにある。

【0008】

また、本発明の他の目的は、2次元(2D)コンテンツおよび3次元(3D)コンテンツが混用されて提供される場合に、これを多様な3次元端末環境でダウンロードして再生する時、2Dコンテンツおよび3Dコンテンツを識別して、自動でパララックスバリアをオン/オフし、識別した2Dコンテンツおよび3Dコンテンツをディスプレイすることができるステレオスコピックコンテンツの保存方法を提供することにある。

【0009】

また、本発明の他の目的は、ファイルに保存されるステレオスコピックコンテンツが区間(fragment)別に互いに異なるステレオスコピックカメラおよびディスプレイ情報を有する場合、ステレオスコピックカメラ/ディスプレイ情報を保存する方法を提供することにある。

【0010】

また、本発明の他の目的は、ステレオスコピックコンテンツが2個の基本符号化ストリームを有する場合、主要(メイン)/付加トラックを特定し、一般的な2D端末との互換性を維持しつつ、各トラックに含まれ使用され得るステレオスコピックカメラ/ディスプレイ情報の冗長性を除去することができるステレオスコピックコンテンツの保存方法を提供することにある。

【0011】

本発明の他の目的および長所は、以下の発明の詳細によって理解され、本発明の実施形態を参照することによって明らかになるだろう。また、本発明の目的および長所は、特許請求の範囲に記載した手段およびその組合せによって実現され得ることが、当業者には明らかである。

【課題を解決するための手段】

【0012】

本発明の態様に従い、ステレオスコピックコンテンツを保存する方法において、前記ステレオスコピックコンテンツの区間情報に関するステレオスコピックビデオメディア情報

10

20

30

40

50

を、前記ステレオスコピックコンテンツを再生するために要求されるメタデータ情報として保存するステップと、前記ステレオスコピックコンテンツの符号化ストリームを保存するステップとを含む方法が提供される。

【0013】

また、本発明の他の態様に従い、ステレオスコピックコンテンツを保存するファイルフォーマット構造において、前記ステレオスコピックコンテンツの符号化ストリームを保存するメディアデータボックスと、前記ステレオスコピックコンテンツの区間情報に関するステレオスコピックビデオメディア情報を保存するボックスとを備えるファイルフォーマット構造が提供される。

【発明の効果】

10

【0014】

本発明によれば、2次元コンテンツおよび3次元コンテンツから構成されるステレオスコピックコンテンツを容易に保存して提供することができる。2Dコンテンツおよび3Dコンテンツの区分情報によってパララックスバリアを自動で2Dまたは3Dディスプレイに変えることによって、ユーザはステレオスコピックコンテンツを容易に鑑賞することができる。

【図面の簡単な説明】

【0015】

【図1】ステレオスコピックコンテンツを提供するための多様なコンテンツ構成形態を示す図である。

20

【図2】本発明による、MPEG-4システム情報がある場合のステレオスコピックビデオアプリケーションフォーマットのための基本ファイルフォーマットの構造図である。

【図3】本発明による、MPEG-4システム情報がない場合のステレオスコピックビデオアプリケーションフォーマットのための基本ファイルフォーマットの構造図である。

【図4】本発明による、MPEG-4システム情報がある場合のステレオスコピックビデオアプリケーションフォーマットの第1の実施形態の構造図である。

【図5】本発明による、MPEG-4システム情報がない場合のステレオスコピックビデオアプリケーションフォーマットの第1の実施形態の構造図である。

【図6】(a)は、本発明による、1つのソースで構成された3次元コンテンツの場合の「ishd」ボックスを「moov」ボックスに含ませた保存フォーマットの構造図であり、(b)は、2つのソースで構成された3次元コンテンツの場合の「ishd」ボックスを「moov」ボックスに含ませた保存フォーマットの構造図である。

30

【図7】(a)は、本発明による、1つのソースで構成された3次元コンテンツの場合の「ishd」ボックスを「mdat」ボックスに含ませた保存フォーマットの構造図であり、(b)は、2つのソースで構成された3次元コンテンツの場合の「ishd」ボックスを「mdat」ボックスに含ませた保存フォーマットの構造図である。

【図8】(a)は、本発明による、1つのソースで構成された3次元コンテンツの場合の「ishd」を「meta」に含ませた保存フォーマットの構造図であり、(b)は、2つのソースで構成された3次元コンテンツの場合の「ishd」を「meta」に含ませた保存フォーマットの構造図である。

40

【図9】本発明による「ishd」ボックスとLASERを含めた保存フォーマットの構造図であって、(a)は、「ishd」ボックスとLASERが「moov」ボックスに含まれた場合の図であり、(b)は、「ishd」ボックスが「moov」ボックスに含まれ、LASERが「mdat」ボックスに含まれた場合の図であり、(c)は、「ishd」ボックスとLASERが「meta」ボックスに含まれた場合の図である。

【図10】本発明による「ishd」ボックスと「iloc」ボックスが含まれたSS-VAFの構造図であって、ES=1である場合の図である。

【図11】本発明による「ishd」ボックスと「iloc」ボックスが含まれたSS-VAFの構造図であって、ES=2である場合の図である。

【図12】(a)は、AVC内のSEIの基本構造のうち「stereo video i

50

n f o r m a t i o n S E I」、' r e s e r v e d _ s e i _ m e s s a g e」が含まれた部分を示す図であり、(b)は、A V Cストリーム内のS E Iの位置を示す図である。

【図13】本発明による's t e r e o v i d e o i n f o r m a t i o n S E I」と' r e s e r v e d _ s e i _ m e s s a g e」を用いたS S - V A Fの構造図である。

【図14】本発明による't r e f」、' i s h d」、' i l o c」ボックスを有するS S - V A Fの構造図であって、E S = 2である時、単一フォーマットのステレオスコピックストリームのみでコンテンツが構成される場合の図である。

【図15】本発明による't r e f」、' i l o c」、' s v m i」、' s c d i」ボックスを有するS S - V A Fの構造図である。

【図16】本発明によるステレオスコピックコンテンツ保存方法の一実施形態のフローチャートである。

【図17】本発明によるステレオスコピックコンテンツ保存方法の他の実施形態のフローチャートである。

【図18】本発明によるステレオスコピックコンテンツ保存方法のさらに他の実施形態のフローチャートである。

【発明を実施するための形態】

【0016】

図1は、ステレオスコピックコンテンツを提供するために構成することができる多様なコンテンツ構成形態を表す。図1Aは、基本符号化ストリーム (elementary stream: E S) が1個の場合 (すなわち、E S = 1の場合) であり、図1BはE S = 2の場合である。単一フォーマット (single format) は、ステレオスコピック映像構成方式が同一であり、1つのカメラパラメータおよび1つのディスプレイ情報のみを有するフォーマットである。多重フォーマット (multiple format) は、ステレオスコピック映像構成方式が異なる場合、またはステレオスコピック映像構成方式は同じであるが、カメラパラメータが異なる場合、または多数個のディスプレイ情報またはその他の情報を有する場合のフォーマットである。

【0017】

ステレオスコピックコンテンツのタイプには、i) 両眼式3次元ビデオサービスなどのステレオスコピックビデオコンテンツ (Stereoscopic Video Content)、ii) 両眼式3次元静止画像サービス (例: スライドショー)、2D (モノ) ビデオと3Dデータサービス (特定画面または部分) の結合などの、ステレオスコピックイメージコンテンツ (Stereoscopic Image Content)、iii) 2D (モノ) ビデオと3Dビデオ (特定画面または部分) の結合などの、ステレオスコピック混用コンテンツ (Monoscopic and Stereoscopic Mixed Content) がある。

【0018】

図2および図3は、本発明によるステレオスコピックビデオアプリケーションフォーマット (Stereoscopic video application format: S S - V A F) の基本構造図である。

【0019】

図2は、M P E G - 4システム情報があるファイルフォーマット構造であり、図3は、M P E G - 4システム情報がないファイルフォーマット構造である。図示したように、S S - V A Fは' f t y p」、' m o o v」、' m d a t」ボックスを含む。以下、S S - V A Fのボックス (b o x) のシンタックス (syntax) とセマンティックス (semantics) を説明する。本発明によるボックスはS S - V A Fの構造に含まれ、その位置は形態に応じて変わることができ、ボックスに含まれる情報は、別途に使用することができる。

【0020】

1. ' s c t y」 (Stereoscopic Content Type)

' s c t y」は、基本的に構成されるコンテンツのタイプを表したものであって、モノスコピックコンテンツ (一般2D画像) と、ステレオスコピックコンテンツを区分する。

表 1 は「s c t y」に対するシンタックスを表す。また「s c t y」を構成する「S t e r e o s c o p i c _ C o n t e n t _ t y p e」を、「f t y p」ボックスおよび他の既存のボックスに含めることができる。

【 0 0 2 1 】

【表 1】

```
Aligned(8) class StereoscopicContentTypeBox extend Box('sfty') {
    unsigned int(2) Stereoscopic_Content_Type:
    unsigned int(6) reserved;
}
```

10

【 0 0 2 2 】

表 1 において、S t e r e o s c o p i c _ C o n t e n t _ T y p e は、ステレオスコピックコンテンツタイプ (Stereoscopic content type) を表し、表 2 のような意味を有する。

【 0 0 2 3 】

【表 2】

Value	Description
00	Stereoscopic Content(3D)
01	Mono/stereoscopic Mixed Contents(2D+3D)
10～11	Reserved

20

【 0 0 2 4 】

2. 「s o v f」 (Stereoscopic Object Visual Format)

「s o v f」は、ステレオスコピックコンテンツの映像構成フォーマット (またはビジュアルフォーマット) を表す。表 3 は、「s o v f」のシンタックスを表す。また、「s o v f」内に含まれた「S t e r e o s c o i c _ O b j e c t _ V i s u a l F o r m a t」を、既存の他のボックス、またはステレオスコピックコンテンツの保存のために新たに定義されたボックス内に含めることができる。

30

【 0 0 2 5 】

【表 3】

```
Aligned(8) class StereoscopicObjectDataVisualFormat extend Box('sovf') {
    unsigned int(4) Stereoscopic_Object_VisualFormat:
    unsigned int(4) reserved;
}
```

【 0 0 2 6 】

表 3 において、「S t e r e o s c o p i c _ O b j e c t _ V i s u a l F o r m a t」は、ステレオスコピックコンテンツの映像構成情報を表し、表 4 のような意味を有する。

40

【 0 0 2 7 】

【表 4】

Value	Description
0000	Side by Side
0001	Top-Down
0010	Field Sequential based on horizontal
0011	Field Sequential based on vertical
0100	Frame Sequential
0101~0111	Reserved
1000	Main + additional (full size)
1001	Main + additional (half of vertical)
1010	Main + additional (half of horizontal)
1011	Main + additional (half of vertical/horizontal)
1100	Main + depth Map

10

【0028】

「full size」は、付加映像のサイズが主要映像のサイズと同一なことを意味し、「half of vertical」は、付加映像のサイズが主要映像の縦半分のサイズであることを意味し、「half of horizontal」は、付加映像のサイズが主要映像の横半分のサイズであることを意味し、「half of vertical/horizontal」は、付加映像のサイズが主要映像の縦／横半分のサイズであることを意味する。

20

【0029】

3. 「ssci」(Stereoscopic Content Information)

「ssci」は、ステレオスコピックコンテンツが含む最小／最大深さ(depth)または視差(disparity)情報であって、これによって3D端末は、3Dディスプレイに適合した立体映像を再生することができる。表5は「ssci」のシンタックスを表す。また、「ssci」内に含まれた最小／最大深さ情報を、既存の他のボックスまたはステレオスコピックコンテンツの保存のために新たに定義されたボックス内に含めることができる。表5において、Max_of_depth(disparity)は、最大深さ／視差情報(ピクセル単位)を表し、Min_of_depth(disparity)は最小深さ／視差情報(ピクセル単位)を表す。

30

【0030】

【表 5】

```
Aligned(8) class StereoscopicObjectContentInformation extend Box('ssci') {
    unsigned int(32) Max_of_depth or disparity;
    unsigned int(32) Min_of_depth or disparity;
}
```

40

【0031】

4. 「scpi」(Stereoscopic Camera Parameter Information)

「scpi」は、ステレオスコピックカメラによって獲得された、または関連ツールによって生成されたステレオスコピックコンテンツのカメラパラメータ情報を表す。表6は「scpi」のシンタックスを表す。また、「scpi」内に含まれた各フィールドを、既存の他のボックス、またはステレオスコピックコンテンツの保存のために新たに定義されたボックス内に含めることができる。

【0032】

【表 6】

```
Aligned(8) class StereoscopicCameraParameterInformation extend Box('scpi') {
    unsigned int(32) Baseline;
    unsigned int(32) Focal_Length;
    unsigned int(32) Convergence_point_distance;
    unsigned int(1) StereoscopicCamera_setting;
    unsigned int(7) Reserved;
}
```

【0033】

10

表 6 において、「Baseline」は、左右のカメラ間の距離を表し、「focal__Length」は、カメラ中心からイメージ面（image plane）（CCD センサ）までの距離を表し、「Convergence__point__distance」は、ベースライン（baseline）から収束点（convergence point）までの距離を表す。ここで、収束点は左右のカメラの視線が交差する地点を表す。そして「stereoscopicCamera__setting」は、ステレオスコピック撮影 / データのカメラ配置（camera arrangement）を表し、表 7 のような意味を有する。

【0034】

【表 7】

20

Value	Description
0	Parallel arrangement
1	Cross arrangement

【0035】

5. 「iods」（Object Descriptor Box）

「iods」は、BIFS または LASE R などのシーン（scene）情報がある場合、BIFS ストリームまたは OD ストリームの位置を表すためのIOD（Initial Object Descriptor）を表現するための情報を表す。表 8 は「iods」のシンタックスを表す。

30

【0036】

【表 8】

```
Aligned(8) class ObjectDescriptoratBox extend Box('iods') {
    ObjectDescriptor OD;
}
```

【0037】

6. 「soet」（Stereoscopic One ES Type Box）

「soet」は、エンコーダから出力される ES が 1 個である場合を表す。表 9 は、「soet」のシンタックスを表す。

40

【0038】

【表 9】

```
Aligned(8) class StereoscopicOnesTypeBox extend Box('soet') {
    unsigned int(3) Stereoscopic_OneES_Type;
    unsigned int(5) Reserved;
}
```

【0039】

表 9 において、「stereoscopic__OneES__Type」は、ES が 1 個で構成されるステレオスコピックデータの映像構成フォーマットの実施形態を表し、表 1

50

0 のような意味を有する。

【 0 0 4 0 】

【 表 1 0 】

Value	Description
000	Side by Side
001	Top-Down
010	Field Sequential based on horizontal
011	Field Sequential based on vertical
100	Frame Sequential
101	Mono/stereoscopic Mixed data
110	基準映像／付加データに基づいて構成されたステレオスコピックデータ
111	Reserved

10

【 0 0 4 1 】

7 . 「 s t e t 」 (Stereoscopic Two ES Type Box)

「 s t e t 」は、エンコーダから出力される E S が 2 個である場合の、各 E S のタイプを表す。表 1 1 は「 s t e t 」のシンタックスを表す。

【 0 0 4 2 】

20

【 表 1 1 】

```
Aligned(8) class StereoscopicTwoesTypeBox extend Box('stet') {
    unsigned int(2) Stereoscopic_TwoES_Type;
}
```

【 0 0 4 3 】

表 1 1 において、「 s t e r e o s c o p i c _ T w o E S _ T y p e 」は、2 個の E S で構成されるステレオスコピックデータの映像構成フォーマットの実施形態を表し、表 1 2 のような意味を有する。

【 0 0 4 4 】

30

【 表 1 2 】

Value	Description
000	基準映像
001	付加データ(基準映像と同一のサイズの映像)
010	付加データ(基準映像の縦半分のサイズの映像)
011	付加データ(基準映像の横半分のサイズの映像)
100	付加データ(基準映像の縦／横半分のサイズの映像)
101	付加データ(Depth map)
110	付加データ(Disparity map)
111	Reserved

40

【 0 0 4 5 】

8 . 「 s s t t 」 (Stereoscopic Time Table Box)

「 s s t t 」は、モノスコピックコンテンツおよびステレオスコピックコンテンツがともに使用される場合、シーン内でモノスコピックコンテンツおよびステレオスコピックコンテンツの開始および終了情報を表す。第 1 実施形態による「 s s t t 」のシンタックスが表 1 3 に表示される。また、下記に説明される「 s s t t 」内に含まれた各フィールドを、既存の他のボックスまたはステレオスコピックコンテンツの保存のために新たに定義

50

されたボックス内に含めることができる。

【 0 0 4 6 】

【 表 1 3 】

```
Aligned(8) class StereoscopicTimeTable Box extend Box('sstt') {
    int i;
    unsigned int(8) Mono/stereoscopic_Scene_count;
    for(i=0;i<=Mono/stereoscopic_Scene_count;i++)
    {
        unsigned int(4) Mono/stereoscopic_identifier;
        unsigned int(4) Reserved;
        unsigned int(32) Start_Time;
    }
}
```

10

【 0 0 4 7 】

表 1 3 において、「Mono/stereoscopic_Scene_count」は、ステレオスコピックコンテンツが 2 次元コンテンツおよび 3 次元コンテンツからともに構成される場合の、Mono/stereoscopic シーン変化の数を表す。すなわち、2D -> 3D -> 2D で構成されるステレオスコピックコンテンツの場合、「Mono/stereoscopic_Scene_count」は 2 と設定される。また、2 次元コンテンツを有せずに 3 次元コンテンツのみで構成される場合、「Mono/stereoscopic_Scene_count」は 1 と設定される。このような情報は、3 次元端末で 2D / 3D ディスプレイの自動変換に使用することができる。

20

【 0 0 4 8 】

「Mono/stereoscopic_identifier」は、時間別コンテンツタイプを表し、表 1 4 のような意味を有する。また、「Mono/stereoscopic_identifier」を用いて、2 次元コンテンツまたは 3 次元コンテンツを識別することができる。例えば、「Mono/stereoscopic_identifier」に 1 ビットが割当てられ、「Mono/stereoscopic_identifier」が「0」ならば 2D コンテンツを表し、「1」ならば 3D コンテンツを表す。「start_Time」は、時間別コンテンツの開始時間を表す。

30

【 0 0 4 9 】

【 表 1 4 】

Value	Description
0000	Monoscopic
0001	Side by Side
0010	Top-Down
0011	Field Sequential based on horizontal
0100	Field Sequential based on vertical
0101	Frame Sequential
0110~0111	Reserved
1000	Main + additional (full size)
1001	Main + additional (half of vertical)
1010	Main + additional (half of horizontal)
1011	Main + additional (half of vertical/horizontal)
1100	Main + Depth Map
1101	Main + Disparity Map

40

【 0 0 5 0 】

50

表 1 5 は、本発明の第 2 実施形態による「s s t t」のシンタックスを表示している。表 1 5 において、「s t a r t _ T i m e」は、ステレオスコピックコンテンツの開始時間を表し、「E n d _ T i m e」はステレオスコピックコンテンツの終了時間を表す。

【 0 0 5 1 】

【表 1 5】

```
Aligned(8) class StereoscopicTimeTable Box extend Box('sstt') {
    int i;
    unsigned int(8) Stereoscopic_Scene_count;
    for(i=0;i<=Stereoscopic_Scene_count;i++)
    {
        unsigned int(32) Start_Time;
        unsigned int(32) End_Time;
    }
```

10

【 0 0 5 2 】

表 1 6 は、第 3 実施形態による「s s t t」のシンタックスを表示している。「s t a r t _ S a m p l e _ n u m b e r」は、モノ/ステレオスコピックコンテンツの開始サンプル番号、またはサンプルの個数を表す。すなわち、サンプルの個数は、モノスコピックまたはステレオスコピックに対応するサンプル全体の個数を意味する。ここで、サンプル (sample) はビデオ、時間連続的な一連のビデオフレームの各々のフレーム (an individual frame of video and a time-contiguous series of video frames) を意味する。

20

【 0 0 5 3 】

【表 1 6】

```
Aligned(8) class StereoscopicTimeTable Box extend Box('sstt') {
    int i;
    unsigned int(16) Mono/stereoscopic_Scene_count;
    for(i=0;i<=Mono/stereoscopic_Scene_count;i++)
    {
        unsigned int(4) Mono/stereoscopic_identifier;
        unsigned int(32) Start_Sample_number;
    }
```

30

【 0 0 5 4 】

表 1 7 は、本発明の第 3 実施形態による「s s t t」のシンタックスを表示している。

【 0 0 5 5 】

【表 1 7】

```
Aligned(8) class StereoscopicTimeTable Box extend Box('sstt') {
    int i;
    unsigned int(8) Stereoscopic_Scene_count;
    for(i=0;i<=Stereoscopic_Scene_count;i++)
    {
        unsinged int(4) Stereoscopic_compositiontype;
        unsigned int(32) Start_sample_number;
        unsigned int(32) End_sample_number;
    }
```

40

【 0 0 5 6 】

表 1 7 において、「s t e r e o s c o p i c _ c o m p o s i t i o n t y p e」は、多様なステレオスコピック構成タイプで構成される場合の、時間別コンテンツタイプを表し、表 1 8 のような意味を有する。「E n d _ S a m p l e _ n u m b e r」は、ステ

50

レオスコピックコンテンツの終了サンプル番号またはサンプルの個数を表す。

【 0 0 5 7 】

【 表 1 8 】

Value	Description
0000	Side by Side
0001	Top-Down
0010	Field Sequential based on horizontal
0011	Field Sequential based on vertical
0100	Frame Sequential
0101~0111	Reserved
1000	Main + additional (full size)
1001	Main + additional (half of vertical)
1010	Main + additional (half of horizontal)
1011	Main + additional (half of vertical/horizontal)
1100	Main + Depth Map
1101	Main + Disparity Map

10

【 0 0 5 8 】

20

9 . 「 s e s n 」 (Stereoscopic ES Num Box)

「 s e s n 」は、エンコーダから出力される基本符号化ストリームの個数を表す。表 19 は、「 s e s n 」に対するシンタックスを表す。表 19 において、「 s t e r e o s c o p i c _ E S N u m 」は、ステレオスコピックコンテンツに対してエンコードから出力される E S の個数を表す。

【 0 0 5 9 】

【 表 1 9 】

```
Aligned(8) class StereoscopicESNumBox extend Box('seen') {
    unsigned int(16) stereoscopic_ESNum;
}
```

30

【 0 0 6 0 】

10 . 「 t r e f 」 (Track reference box)

「 t r e f 」は、1つのトラックから他のトラックを参照することができるよう情報を提供する I S O ベースファイルフォーマットで定義されたボックスである。「 t r e f 」は、「 t r a k 」 (Track Box) に含まれる。表 20 は、本発明の一実施形態による「 t r e f 」のシンタックスを表示している。ここで、「 t r a c k _ I D 」は、参照されるトラックの識別符号を表し、「 r e f e r e n c e _ t y p e 」は、表 21 のような意味を有する。

【 0 0 6 1 】

40

【 表 2 0 】

```
aligned(8) class TrackReferenceBox extends Box ('tref') {}
aligned(8) class TrackReferenceTypeBox (unsigned int(32) reference_type) extends Box(reference_type) {
    unsigned int(32) track_IDs[];
}
```

【 0 0 6 2 】

【表 2 1】

hint	参照されるトラック (複数) が本ヒントトラックのオリジナルのメディアを含む。
cdsc	本トラックが参照されるトラックを記述する。
svdp	本トラックが主要ビュートラックとしての参照されるトラックに対する依存を有し、参照されるトラックの「メタ」情報に関連するステレオスコピックを含むことができる。

【0063】

10

基本符号化ストリームが2個であるステレオスコピックビデオは、2個のトラックが存在し、従来のLASERのようなシーン記述 (scene description) としてLASER内のビデオ関連の2個のノードと連結される。すなわち、従来の技術ではESが2個であるステレオスコピックビデオは、2個のオブジェクトとして認識されるものである。しかし、ステレオスコピックビデオは、端末で最終的に1つの3Dビデオフォーマットに変換され再生されるため、1つのオブジェクトとして認識されなければならない。すなわち、ステレオスコピックビデオは2個のトラックを用いて構成されるが、これをシーン再生するためには1つの3Dビデオフォーマットに変換されるため、1つのノードのみを用いて連結しなければならない。よって、基本符号化ストリームが2個である場合、2個のトラックの関係性を表す情報が要求され、ステレオスコピックトラック参照情報は、表22に示すように「tref」内の「svdp」で定義され、使用される。LASERは使用されないが、ESが2個であるステレオスコピックコンテンツを保存するためには、「svdp」などのステレオスコピックトラック参照情報を使用しなければならない。

20

【0064】

ステレオスコピックコンテンツが2個のESを含む場合、「tref」により、2個のトラックを1つの主要トラック (main track) と1つの付加トラック (additional track) として識別することができる。また、「tref」により、付加トラックが主要トラックを参照し、ステレオスコピックビデオ関連ステレオスコピックカメラおよびディスプレイ情報を1つのトラックのみに保存することができる。その結果、情報の冗長性を除去することができる。また、ステレオスコピックコンテンツが2個のトラックを有するとしても、LASERのビデオノードの1個に連結することができるという長所を有する。

30

【0065】

ステレオスコピックコンテンツにより、多様なステレオスコピック映像構成フォーマットおよびカメラパラメータを支援することを可能とするために、本発明において初期ステレオスコピックヘッダ (initial stereoscopic header: ishd) の構造を提案する。本発明の実施形態によれば、「ishd」に含まれる情報を、別途に分離して使用することができる。

【0066】

ステレオスコピックストリーム内に多様なステレオスコピックフォーマットおよびカメラパラメータが存在する場合、各ステレオスコピックストリーム、各ステレオスコピックストリームの開始および長さの情報は、「iloc」によって識別され、各区間にitem_IDが割当てられる。よって、「ishd」は、item_IDに基づいて各々のステレオスコピックフォーマットまたはカメラパラメータ情報を提供しなければならない。ここで、アイテム (item) は、1つのシーケンスでステレオスコピック区間とモノスコピック区間が共に存在する場合、1つのステレオスコピック区間を意味する。

40

【0067】

また、ステレオスコピックストリーム内に3個のステレオスコピック区間があり、各区間に適用された「ishd」情報が異なる場合、ステレオスコピック区間はitem_IDにより識別され、各々の情報を記述する。一方、各区間に適用された「ishd」情報が同じである場合、第2および第3のステレオスコピック区間は、第1のステレオスコピ

50

ック区間と同じ「i s h d」情報を含む。このような構造は、「i s h d」内に含まれる情報の冗長性を除去することができるという長所がある。

【0068】

図4は、本発明によるSS-VAFの一実施形態の構図であって、MPEG-4システム情報がある場合である。図5は、本発明によるSS-VAFの他の実施形態の構図であって、MPEG-4システム情報がない場合である。

【0069】

一方、モノスコピックおよびステレオスコピックコンテンツが混用されている場合、モノスコピックまたはステレオスコピックコンテンツがいつ開始または終了するかを判定するために、区分情報が必要である。これはモノ/ステレオスコピックコンテンツの開始/終了情報、サンプルの2D/3Dの識別符号、およびそこに含まれるサンプル(AU)の個数に従って識別することができる。

【0070】

図6ないし図8は、図1のような多様なコンテンツ構成形態を支援するための保存フォーマットの概念的な構造を提供する。基本的に保存フォーマットの構造は大きく「f t y p」ボックス、「m o o v」ボックスおよび「m d a t」ボックスで構成される。「f t y p」ボックスは、ファイルタイプを定義するボックスであって、ステレオスコピックまたはモノスコピック/ステレオスコピック混用コンテンツファイルであることを表すフィールドを含むことによって、3次元コンテンツファイルであることを表す。「m o o v」ボックスは、メディアデータを再生するためのすべてのシステム(メタ)情報を含み、「m d a t」ボックスは、実際のメディアデータを含む。このような構造に基づきステレオスコピックコンテンツに対する新しい付加情報が要求され、このような付加情報の位置に応じて保存フォーマット構造は変化する。

【0071】

図6は、3次元コンテンツを構成するソースの個数および新しい付加情報を含む「i s h d」が「m o o v」ボックスに含まれる保存フォーマット構造を表す。3次元コンテンツに対してソースが1個である場合(図6A)は、あるフレームに、左右の映像情報が両方含まれる(例:side by side)。ソースが2個である場合(図6B)、左右の映像情報は各々1つのフレームに含まれる。「m o o v」ボックス内のトラックの個数は、含まれたメディアデータの個数に応じて変わる。「m o o v」ボックス内のトラックは、「m d a t」ボックスに含まれるメディアデータを再生するためのすべてのシステム(メタ)情報を含む。

【0072】

このような保存フォーマットは、新しい付加情報およびこれを支援するための構造が要求とされる。よって、本発明は「i s h d」を新たに定義し、これを「m o o v」ボックスのトラック内に含める。「i s h d」の位置は、「m o o v」ボックス内、または保存フォーマット内において変わることができる。

【0073】

図7は、先に定義された「i s h d」情報を「m d a t」ボックスに有する保存フォーマット構造を表す。図7Aは1つのソースで構成された3次元コンテンツの場合であり、図7Bは、2つのソースで構成された3次元コンテンツの場合である。これは既存の「m o o v」構造をそのまま維持しながら、「i s h d」ストリームが「m d a t」ボックス内に含まれているという情報を保存フォーマットに含めることによって可能である。

【0074】

図8は、先に定義された「i s h d」情報を「m e t a」ボックスに有する保存フォーマット構造を表す。図8Aは、1つのソースで構成された3次元コンテンツの場合であり、図8Bは、2つのソースで構成された3次元コンテンツの場合である。

【0075】

表22は、「i s h d」情報が「m d a t」ボックスに含まれているという情報を知らせるための構造を表し、このような構造は、「s t s d (sample description)」ボック

10

20

30

40

50

ス内に含まれる。

【 0 0 7 6 】

【 表 2 2 】

//ishd Sequences

Class ishdSampleEntry(codingname) extends SampleEntry ('ishd') { }

【 0 0 7 7 】

図 9 は、図 6 ないし 8 の構造に基づきシーン記述子情報が含まれる場合の保存フォーマット構造を表す。シーン記述子は、多様なマルチメディアの自由なシーン構成およびユーザとのインタラクションのためのものであって、本発明ではレーザー (L A S e R) を使用する。

10

【 0 0 7 8 】

図 9 A は、「 m o o v 」ボックス内にシーン記述子ストリームを保存するために別途のボックスを含む保存フォーマットを表す。図 9 B は、「 m d a t 」ボックス内にシーン記述子ストリームを有する保存フォーマットを表す。「 m o o v 」ボックスは、シーン記述子ストリームが「 m d a t 」ボックス内に含まれることを知らせるためのトラックを別途に有し、「 s t s d 」ボックスは、シーン記述子ストリームに関する情報を有する。すなわち、トラック内の「 s t s d 」ボックスを検索して、このトラックがどんな情報 (シーン記述子 / ビデオ / オーディオ) を表しているのかを分析し、この分析結果に基づいて「 m d a t 」ボックスに保存された情報を用いてデコードする方法である。図 9 C は、先に定義されたシーン記述子情報を「 m e t a 」ボックスに有する保存フォーマット構造を表す。

20

【 0 0 7 9 】

表 2 3 ないし表 2 5 は、図 1 の 3 次元コンテンツ構成形態をすべて支援する「 i s h d 」構造の実施形態を表す。

【 0 0 8 0 】

【 表 2 3 】

Class ishd {

unsigned int(16) num_MonoStereo_scene;

if(num_MonoStereo_scene){

for(i=0;i<num_MonoStereo_scene;i++)

unsigned int(16) start_sample_index;

unsigned int(3) Composition_type;

unsigned int(1) numofES;

unsigned int(1) LR_first;

unsigned int(3) Reserved;

}

} else

unsigned int(3) Composition_type;

unsigned int(1) numofES;

unsigned int(1) LR_first;

stereoscopicCameraInfo[0...1];

stereoscopicContentsInfo[0...1];

}

【 0 0 8 1 】

30

40

【表 2 4】

```

Class ishd {
    unsigned int(16) num_MonoStereo_scene;
    if(num_MonoStereo_scene){
        for(i=0;i<num_MonoStereo_scene;i++)

            unsigned int(16) start_sample_index;
            unsigned int(3) Composition_type;、
            unsigned int(1) numofES;
            unsigned int(1) LR_first;
            unsigned int(3) Reserved;
        }
    } else

        unsigned int(3) Composition_type;
        unsigned int(1) numofES;
        unsigned int(1) LR_first;
}

```

10

【 0 0 8 2】

20

【表 2 5】

```

Class ishd {
    unsigned int(16) num_MonoStereo_scene;
    if(num_MonoStereo_scene){
        for(i=0;i<num_MonoStereo_scene;i++)

            unsigned int(16) start_sample_index;
            unsigned int(3) Composition_type;
            unsigned int(1) numofES;
            unsigned int(1) LR_first;
            stereoscopicCameraInfo[0...1];
            stereoscopicContentsInfo[0...1];
            unsigned int(3) Reserved;
        }
    } else

        unsigned int(3) Composition_type;
        unsigned int(1) numofES;
        unsigned int(1) LR_first;
        stereoscopicCameraInfo[0...1];
        stereoscopicContentsInfo[0...1];
}

```

30

40

【 0 0 8 3】

表 2 3 ないし表 2 5 において、「num_MonoStereo_scene」は、ステレオスコピックコンテンツが 2 次元コンテンツおよび 3 次元コンテンツで構成される場合の、シーンの個数を表す。また、ステレオスコピックコンテンツが多様な 3 次元コンテンツで構成される場合の、シーンの個数を表す。例えば、ステレオスコピックコンテンツが (2D) (3D) (2D) で構成される場合、Num_MonoStereo_scene = 3 になる。また 3 ステレオスコピックコンテンツが side by side (fi

50

eld sequential)方式で構成される場合、num_MonoStereo_scene = 2となる。さらに、ステレオスコピックコンテンツが単一フォーマットの3次元コンテンツのみからなる場合、num_MonoStereo_scene = 1となる。

【0084】

「start_sample_index」は、各コンテンツの開始サンプル(すなわち、一般的なフレーム番号)番号または各コンテンツタイプに応じた含まれるサンプルの個数として使用されることができる。「num_of_ES」は「mdat」ボックスに含まれたビデオ符号化ストリームの個数を表す。

【0085】

「Composition_type」は、2次元コンテンツおよび3次元コンテンツのコンテンツフォーマットを識別するための情報を表す。「start_sample_index」および「Composition_type」を、2D/3Dディスプレイモードを支援する多様な3次元端末機において自動ディスプレイオン/オフするための基本情報として活用することができる。「Composition_type」は表26のような意味を有する。

【0086】

【表26】

Value		Description
3D	000	Side by Side
	001	Top-down
	010	Field sequential
	011	Frame sequential
	100	Vertical line interleaved format
	101	Left/right image
2D	110	Monoscopic left image
	111	Monoscopic right image

【0087】

「LR_first」は、左右の映像のうち優先視される映像を表したものであって、左の映像と右の映像のうち、先にエンコーディングされるイメージを知らせる。

【0088】

「stereoscopicCameraInfo」オブジェクトは、3次元コンテンツに対するカメラパラメータ情報を表したものであって、表27は、「stereoscopicCameraInfo」オブジェクトの一実施形態を表す。本実施形態によるカメラパラメータ情報は、既存の他のボックス内または新しく定義されたボックスに含まれ得る。表27で、「stereoscopicCamera_setting」は、3次元コンテンツを製作または撮影する時のカメラの配置形態を表したもので、「平行式(parallel)」または「交差式(cross)」を表す。

【0089】

「Baseline」は、ステレオスコピックカメラ間の距離を表し、「focal_Length」は、レンズからイメージ面までの距離を表す。また、「ConvergencePoint_distance」は、左右のカメラを連結するベースラインから収束点までの距離を表す。

【0090】

【表 2 7】

```

StereoscopicCameraInfo {
    unsigned int(1) StereoscopicCamera_setting;
    unsigned int(7) Reserved=1111;
    unsigned int(16) Baseline;
    unsigned int(16) Focal_Length;
    unsigned int(16) ConvergencePoint_distance;
}

```

【0091】

10

「stereoscopicContentsInfo」オブジェクトは、3次元コンテンツを表示するための最小限の情報を表す。表28は、「stereoscopicContentsInfo」オブジェクトの一実施形態を表し、「stereoscopicContentsInfo」に含まれる情報は、既存の他のボックス内または新たに定義されたボックス内に含むことができる。「Max_disparity」は、3次元コンテンツの最大視差(disparity)の大きさを表し、「Min_disparity」は3次元コンテンツの最小視差の大きさを表す。

【0092】

【表 2 8】

```

StereoscopicContentsinfo {
    unsigned int(16) Max_disparity;
    unsigned int(16) Min_disparity;
}

```

20

【0093】

「stereoscopicCameraInfo」および「stereoscopicContentsInfo」内の情報は、MPEG-7メタデータなどの別途の記述として表現し、保存することができる。

【0094】

図10は、ES = 1である場合のSS-VAFの一実施形態の構造図である。

30

【0095】

「ftyp」ボックスは、ステレオスコピックコンテンツを含むか否かを表す。符号化ストリームが全体3Dである場合、符号化ストリームが2D/3D混用のストリームで構成される場合、それらをステレオスコピックコンテンツとして見なす。

【0096】

ステレオスコピックコンテンツが2D/3Dのストリームで構成される場合、2D/3Dのストリームの開始情報および長さ情報が必要となる。開始情報および長さ情報のために、ISOベースファイルフォーマット(11496-12)の既存ボックスである「iloc」ボックスを利用する。「iloc」(item location)ボックスは、「meta」ボックスに含まれる。「iloc」ボックスは、ステレオスコピックコンテンツの場合、保存ファイル内のステレオスコピック区間の位置を提供する。

40

【0097】

また、2Dストリームおよび3Dストリームの区別に関連する情報は、「ishd」ボックスによって獲得する。2D/3D混用ストリームの場合、複数の3Dストリームが含まれても、複数の3Dストリームが同一の情報ならば(すなわち、単一フォーマットならば)、3Dストリームに関連する情報は、一つの「ishd」を参照することによって獲得する。

【0098】

単一フォーマットの3Dストリームのみで構成されたステレオスコピックコンテンツでは、「iloc」ボックスを使用せずに「ishd」ボックスのみを使用して、ステレオ

50

スコピックデータを表現することができる。また $ES = 1$ であるとき、3Dストリームが多重フォーマットで構成される場合は、「i l o c」ボックスを用いて各フォーマットのオフセット (offset) / 長さ (length) 値を知り、各フォーマット情報は「i s h d」ボックスによって獲得する。この場合「i s h d」ボックスは、多重フォーマットに関する情報を有している。

【0099】

図11は、 $ES = 2$ である時、 $SS \cdot VAF$ の一実施形態の構造図である。 $ES = 2$ である場合、左右のストリーム情報は対応する「t r a k」ボックスに含まれる。前述したように、ステレオスコピックデータは所定のフォーマットに変換されてディスプレイされるため、左右のストリーム情報は2個のトラックで構成されるが、1つのオブジェクトとして認識されるために相関関係を表す必要がある。例えば、左の映像がメイン映像であり、右の映像が付加映像ならば、右の映像ストリーム情報を含む「t r a k」ボックスが、左の映像ストリーム情報を含む「t r a k」ボックスとの関係性を表現することによって、「i s h d」ボックスのうちで重複したものを除去することができる。右の映像ストリーム情報に含まれる「i s h d」のカメラパラメータおよびディスプレイ情報が、左の映像ストリーム情報に含まれる「i s h d」と同一であるならば、別途に記述せず、左の映像ストリーム情報に含まれる「i s h d」を使用することができる。このような関係性を表現するために、本発明では「t r e f」ボックスおよび「s v d p」ボックスを提案する。

10

【0100】

左右ストリームで保存された3Dストリームに関して多重フォーマットの3Dストリームで構成される場合は、各フォーマットに対応する3Dストリームの区分、開始および長さ情報を知らなければならない。開始および長さ情報は、ISOベースファイルフォーマット(11496-12)の既存ボックスである「i l o c」によって獲得する。また、多重フォーマットの3Dストリームの区分に関連する情報は、「i s h d」ボックスによって獲得する。表29は、単一フォーマットの「i s h d」ボックスのシンタックス実施形態を表す。

20

【0101】

【表 2 9】

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

```

    unsigned int(8) Stereoscopic_Composition_type;
    unsigned int(1) LR_first;
    unsigned int(1) Is_camParams;
    unsigned int(1) Is_disInfo;
    unsigned int(5) Reserved;

```

//all the following are optional fields

```

if(Is_CamParams){
    unsigned int(32) Baseline;
    unsigned int(32) focalLength;
    unsigned int(32) convergence_distance;
    unsigned int(1) Is_camera_cross
}

```

10

```

if(Is_camera_cross){
    unsigned int(32) rotation[];
    unsigned int(32) translation[];
    unsigned int(7) Reserved;
}

```

20

```

if(Is_disInfo){
    Int(16) MinofDisparity;
    Int(16) MaxofDisparity;
}

```

【0 1 0 2】

「Is_camParams」は、カメラパラメータが存在するか否か、「Is_disInfo」は、ステレオスコピックコンテンツディスプレイ情報が存在するか否か、「Baseline」は、左右のカメラの距離、「focallength」は、レンズからイメージ面（フィルム）までの距離、「convergence_distance」は、左右のカメラを連結するベースラインの中心から左右のカメラの視線が交差する収束点までの距離を表す。「convergence_distance」は、平行軸カメラの場合、無限大の値を有し、これを表示するためにすべてのビットを1とする。

30

【0 1 0 3】

また、「Is_camera_cross」は、「1」のとき交差軸カメラ、「0」のとき平行軸カメラを表し、「rotation」は、オブジェクトまでのカメラ位置角度（camera position angle）を表す。「translation」は、ステレオカメラが移動したか否か（すべてのビットが0である場合、ステレオカメラの動きはない）、「MinofDisparity」は、左右の映像の最小視差の大きさ、「MaxofDisparity」は、左右の映像の最大視差の大きさを表す。

40

【0 1 0 4】

表 3 0 は、多重フォーマットの「ishd」ボックスのシンタックス実施形態を表す。「item_count」は、多重フォーマットの場合、フォーマットの情報を記述する個数を表し、「item_ID」は、各フォーマットの整数の名前（integer name）を表し、ステレオスコピックコンテンツ内の多様なステレオスコピックフォーマットを識別するために使用される。ここで、「item_ID」は、「iloc」ボックスのitem_IDと連動して使用される。

【0 1 0 5】

【表 3 0】

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

```

unsigned int(16) item_count;
for (i=0;i<item_count;i++) {
    unsigned int(16)    item_ID;
    unsigned int(8)     Stereoscopic_Composition_type;
    unsigned int(1)     LR_first;
    unsigned int(1)     Is_camParams;
    unsigned int(1)     Is_disInfo;
    unsigned int(5)     Reserved;
    //all the following are optional fields
    if(Is_CamParams){
        unsigned int(32) Baseline;
        unsigned int(32) focallLength;
        unsigned int(32) convergence_distance;
        unsigned int(1) Is_camera_cross
    }

    if(Is_camera_cross) {
        unsigned int(32) rotation[];
        unsigned int(32) translation[];
        unsigned int(7) Reserved;
    }

    if(Is_disInfo){
        Int(16) MinofDisparity;
        Int(16) MaxofDisparity;
    }

    //other additional information
    if(other_flag){}
}

```

【0 1 0 6】

以下、A V C (Advanced Video Coding)、S E I (Supplemental enhancement information) に対して説明する。S E I は、デコード、ディスプレイなどに関連したメッセージ情報を含む「stereo video information SEI」情報を含んでおり、S E I メッセージは、A V C ストリーム内に伝送される。

【0 1 0 7】

図 1 2 は、N A L ユニットを含む単一ビデオ E S のフローチャートであって、図 1 2 A は A V C 内の S E I の基本構造のうち「stereo video information SEI」、「reserved__sei__message」が含まれた部分を表し、図 1 2 B は A V C ストリーム内の S E I の位置を表す。表 3 1 は、「stereo video information SEI」メッセージを表す。

【0 1 0 8】

40

【表 3 1】

stereo_video_info(payloadSize){	C	Descriptor
field_views_flag	5	u(1)
if(field_views_flag)		
top_field_is_left_view_flag	5	u(1)
else{		
current_frame_is_left_view_flag	5	u(1)
next_frame_is_second_view_flag	5	u(1)
}		
left_view_self_contained_flag	5	u(1)
right_view_self_contained_flag	5	u(1)
}		

10

【0109】

「field_views_flag」は、フィールドベース (field based) ステレオスコピックストリームが存在するか否かを表す。「top_field_is_left_view_flag」が「1」の場合、垂直インターレース (vertical interlaced) フォーマットで構成されたステレオスコピックコンテンツ (左のビュー優先)、「0」の場合、垂直インターリーブ (vertical line interleaved) フォーマットで構成されたステレオスコピックコンテンツ (右のビュー優先) を表す。「Current_frame_is_left_view_flag」が「1」の場合、現在フレームが左のビューであることを表し、「0」の場合、現在フレームが右のビューであることを表す。「next_frame_is_second_view_flag」が「1」の場合、現在フレームと次のフレームでステレオスコピック映像を構成することを、「0」の場合、現在フレームと以前フレームでステレオスコピック映像を構成することを表す。「Left_view_self_contained_flag」が「1」の場合、右のビューとの相関を持たず、独立ストリームでコーディングしたものを、「0」の場合、右のビューとの相関に基づいてストリームをコーディングしたものを表す。「right_view_self_contained_flag」は「1」の場合、左のビューとの相関を持たず、独立ストリームでコーディングしたものを、「0」の場合、左のビューとの相関に基づいてストリームをコーディングしたものを表す。

20

30

【0110】

「stereo_video_information SEI」情報は、「stereoscopic_composition_type」のうちで表32のフォーマットを含んでいる。一方、「side_by_side」フォーマットは支援しない。

【0111】

【表 3 2】

1	Vertical interleaved format
2	Frame sequential format
3	Field sequential format
4	Monoscopic left image
5	Monoscopic right image

40

【0112】

以下、既存のAVC SEI情報を活用したステレオスコピックコンテンツをサービスするための方法および保存フォーマットを提案する。これは、ステレオスコピックコンテ

50

ンツがAVCによって符号化された場合のみ可能である。

【0113】

既存の「stereo_video_information SEI」に基づき各ステレオスコピックストリームに要求されるカメラパラメータおよびディスプレイ情報を追加する形態であって、「reserved_sei_message」を用いたSS-VAFを提案する。また、「stereo_video_information SEI」を表33のように拡張して、使用することができる。「side_by_side_flag」が「1」である場合、左のビューを優先し、1フレーム内の左右の映像として構成される。「0」である場合、右のビューを優先し、1フレーム内の左右の映像として構成される。表33で「C」は、シンタックスのカテゴリを意味し、「u(1)」は1ビットを使用する「符号なし整数 (unsigned integer)」を意味する。

10

【0114】

【表33】

stereo_video_info(payloadSize){	C	Descriptor
field_views_flag	5	u(1)
if(field_views_flag)		
top_field_is_left_view_flag	5	u(1)
else {		
current_frame_is_left_view_flag	5	u(1)
next_frame_is_second_view_flag	5	u(1)
}		
else {		
left_view_self_contained_flag	5	u(1)
right_view_self_contained_flag	5	u(1)
}		
side_by_side_flag	5	u(1)
}		

20

30

【0115】

表34は、AVCのSEI情報のうち「reserved_sei_message (payloadSize)」を用いて、ステレオスコピックカメラ情報を定義する。ここで、他のカメラ情報を追加することができ、追加した情報は別途に使用することができる。これに基づき、ステレオスコピックコンテンツストリームに対するカメラパラメータ情報を獲得する。

【0116】

【表34】

stereo_camera_info(payloadSize) {	C	Descriptor
if(Is_CamParams){		
Baseline;	5	U(32)
focalLength;	5	U(32)
convergence_distance;	5	U(32)
if(Is_camera_cross) {		
rotation[];	5	U(32)
translation[];	5	U(32)
}		
}		

40

50

【 0 1 1 7 】

表 3 5 は、A V C の S E I 情報のうち「r e s e r v e d _ s e i _ m e s s a g e (p a y l o a d S i z e)」を用いて、ステレオスコピックコンテンツを表示するために要求される情報を定義する。表 3 5 に定義される情報に基づき、ステレオスコピックコンテンツ視差の値を抽出する。

【 0 1 1 8 】

【表 3 5】

stereo_display_info(payloadSize) { if(Is_disInfo){ MinofDisparity; MaxofDisparity; }	C	Descriptor
MinofDisparity;	5	U(16)
MaxofDisparity;	5	U(16)

10

【 0 1 1 9 】

ただし、上記の情報を 1 つの S E I _ m e s s a g e に統合して、ステレオカメラおよびディスプレイ情報を提供することができる。

【 0 1 2 0 】

図 1 3 は、S t e r e o v i d e o i n f o r m a t i o n および R e s e r v e d S E I を用いた S S - V A F を表す。図示されたアプリケーションフォーマットにおいて、L A S e R は選択的に含まれる。

20

【 0 1 2 1 】

ステレオスコピックコンテンツが 2 D / 3 D 混用のストリームで構成される場合、「i l o c」ボックスによって区分された 3 D ストリームの区間の間は、A V C ストリーム S E I m e s s a g e 内で定義された 3 D ストリーム情報(「s t e r e o v i d e o i n f o r m a t i o n S E I」、「s t e r e o c a m e r a i n f o r m a t i o n S E I」、「s t e r e o d i s p l a y i n f o r m a t i o n S E I」)が獲得される。単一フォーマットの 3 D ストリームのみで構成される場合は、ステレオスコピックコンテンツを、A V C ストリーム S E I m e s s a g e 内で定義された 3 D ストリーム情報(「s t e r e o v i d e o i n f o r m a t i o n S E I」、「s t e r e o c a m e r a i n f o r m a t i o n S E I」、「s t e r e o d i s p l a y i n f o r m a t i o n S E I」)を使用して表現することができる。

30

【 0 1 2 2 】

図 1 4 は、E S = 2 であるとき、単一フォーマットのステレオスコピックストリームのみで構成される場合の S S - V A F の構造図である。E S = 2 である場合、左右のストリーム情報は各々の対応する「t r a k」に含まれる。ここで、左右のストリームは互いに相関関係を表す必要がある。例えば、左の映像がメイン映像であり、右の映像が付加映像ならば、右の映像ストリーム情報を含む「t r a k」ボックスが、左の映像ストリーム情報を含む「t r a k」ボックスとの関係性を表現することによって、「i s h d」ボックスのうちで重複したものを除去することができる。このような関係性は、I S O ベースファイルフォーマットに含まれた「t r e f」ボックスを利用する。これは再生に必要なすべての t r a k _ I D を記述することができる。すべての t r a k _ I D は、右映像(付加映像)ストリームの「t r a k」に、「t r e f」と共に記述される。

40

【 0 1 2 3 】

表 3 6 は、多様なステレオスコピック構成フォーマットおよびカメラパラメータを支援する場合の「i s h d」ボックスのシンタックス実施形態を表す。「i t e m _ I D」は、次の情報が定義されるアイテムの I D を表し、1 以上の値を有する。「c u r r e n t _ i n d i c a t o r」が「1」である場合、次に記述される情報の有効性を表し、「0

50

」である場合、次に記述される情報が先に記述された情報と同一であることを意味する。ただし、Item_ID = 1 である場合は、以後記述される情報がないことを意味する。すなわち、端末は「current_indicator」に基づいて、次に記述される情報の有効性を判定し、0 である場合は先に記述された「ishd」情報と同一であることを判定することになる。

【0124】

また、「LR_first」は左映像および右映像の基準位置選定、「Is_camParams」は、カメラパラメータが存在するか否か、「Is_displaySafeInfo」は、ステレオスコピックコンテンツディスプレイ情報が存在するか否かを表す。また「Baseline」は、左右のカメラの距離、「focallength」は、CCDからイメージ面（フィルム）までの距離を表す。「Is_camera_cross」は、「1」である場合、交差軸カメラ、「0」である場合、平行軸カメラを表す。

【0125】

【表36】

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

```
unsigned int(16) item_ID;
unsigned int(1) current_indicator;
unsigned int(8) Stereoscopic_Composition_type;
unsigned int(1) LR_first;
unsigned int(1) Is_camParams;
unsigned int(1) Is_displaySafeInfo;
unsigned int(4) Reserved;
```

//all the following are optional fields

```
if(Is_CamParams){
    unsigned int(32) Baseline;
    unsigned int(32) focalLength;
    unsigned int(1) Is_camera_cross
    if(Is_camera_cross) {
        unsigned int(32) convergence_distance;
        unsigned int(32) rotation[];
        unsigned int(32) translation[];
        unsigned int(7) eserved;
    }
}
```

```
if(Is_displaySafeInfo){
    unsigned int(16) ViewingDisatance;
    int(16) VerticalDisparity;
    int(16) MinofDisparity;
    Int(16) MaxofDisparity;
```

```
}
//other additional information
if(other_flag){}
```

```
aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0){
    unsigned int(16) item_count;
    ishdEntry[item_count] ishds;
}
```

10

20

30

40

50

【0126】

また、「convergence_distance」は、ベースラインの中心から収束点までの距離（平行軸カメラの場合、無限大の値を有する。すべてのビットが1である場合、無限大の距離を表す）、「rotation」は、オブジェクトに対するカメラ位置角度、「translation」は、ステレオカメラの移動を表す（すべてのビットが0である場合、ステレオカメラに動きはない）。また「VerticalDisparity」は、左右の映像の垂直視差の大きさ、「MinofDisparity」は、左右の映像の最小水平視差の大きさ、「MaxofDisparity」は、左右の映像の最大視差の大きさ、「item_count」は、次の配列内のエントリーの個数を表す。

10

【0127】

表37は、多様なカメラパラメータを支援する場合の「ishd」ボックスの第1実施形態のシンタックスを表す。ここで、ステレオスコピック構成フォーマットは同一であると仮定し、すぐ前の「ishd」情報が参照される。また「ishd」に含まれるカメラパラメータおよびディスプレイ情報は、別途のボックスで分離して保存され得る。

【0128】

【表 3 7】

```

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)
    unsigned int(16) item_ID;
    unsigned int(8) Stereoscopic_Composition_type;
    unsigned int(1) LR_first;
    unsigned int(1) current_indicator;
    unsigned int(1) Is_camParams;
    unsigned int(1) Is_displaySafeInfo;
    unsigned int(4) Reserved;
    //all the following are optional fields
    if(Is_CamParams){
        unsigned int(32) Baseline;
        unsigned int(32) focalLength;
        unsigned int(1) Is_camera_cross
        if(Is_camera_cross){
            unsigned int(32) convergence_distance;
            unsigned int(32) rotation[];
            unsigned int(32) translation[];
        }
        unsigned int(7) Reserved;
    }
    if(Is_displaySafeInfo){
        unsigned int(16) ViewingDisatance;
        int(16) VerticalDisparity;
        int(16) MinofDisparity;
        int(16) MaxofDisparity;
    }
    //other additional information
    if(other_flag){}

aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0){
    unsigned int(16) item_count;
    ishdEntry[item_count] ishds;
}

```

【0 1 2 9】

表 3 8 は、多様なカメラパラメータを支援する「i s h d」ボックスの第 2 実施形態のシンタックスを表す。ステレオスコピック構成フォーマットは同一であると仮定し、すぐ前の「i s h d」が参照される。

【0 1 3 0】

【表 3 8】

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

```

unsigned int(8) Stereoscopic_Composition_type;
unsigned int(1) LR_first;
unsigned int(7) Reserved;
unsigned int(16) item_count;
for(i=0;i<item_count;i++) {
    unsigned int(16) item_ID;
    unsigned int(1) current_indication;
    unsigned int(1) Is_camParams;
    unsigned int(1) Is_displaySafeInfo;
    unsigned int(5) Reserved;

    //all the following are optional fields
    if(Is_CamParams){
        unsigned int(32) Baseline;
        unsigned int(32) focalLength;
        unsigned int(1) Is_camera_cross
        unsigned int(7) Reserved;
        if(Is_camera_cross){
            unsigned int(32) convergence_distance;
            unsigned int(32) rotation[];
            unsigned int(32) translation[];
        }
    }
    if(Is_displaySafeInfo){
        unsigned int(16) ViewingDisatance;
        int(16) VerticalDisparity;
        int(16) MinofDisparity;
        int(16) MaxofDisparity;
    }
    //other additional information
    if(other_flag){

```

【0 1 3 1】

表 3 9 は、多様なカメラパラメータを支援する場合の「ishd」ボックスの第 3 実施形態のシンタックスを表す。ステレオスコピック構成フォーマットは同一であると仮定し、「cameParams」および「displaySafeInfo」が同じ所定の Item_ID を参照する。

【0 1 3 2】

表 3 9 において、「Is_ref」が「0」である場合、参照されるカメラパラメータおよびディスプレイ情報がないことを表し、「1」である場合、参照される Item_ID が有ることを表す。「current_refIndex」は、参照される Item_ID を表す。

【0 1 3 3】

【表 3 9】

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

```

unsigned int(8) Stereoscopic_Composition_type;
unsigned int(1) LR_first;
unsigned int(7) Reserved;
unsigned int(16) item_count;
for(i=0;i<item_count;i++) {
    unsigned int(16) item_ID;
    unsigned int(1) Is_camParams;
    unsigned int(1) Is_displaySafeInfo;
    unsigned int(1) Is_ref;
    unsigned int(5) Reserved;

    //all the following are optional fields
    If(Is_ref){
        unsigned int(16) current_refIndex;
    }
    else{
        if(Is_CamParams){
            unsigned int(32) Baseline;
            unsigned int(32) focalLength;
            unsigned int(1) Is_camera_cross
            unsigned int(7) Reserved;
        }
        if(Is_camera_cross) {
            unsigned int(32) convergence_distance;
            unsigned int(32) rotation[];
            unsigned int(32) translation[];
        }
        if(Is_displaySafeInfo){
            unsigned int(16) ViewingDistance;
            int(16) VerticalDisparity;
            int(16) MinofDisparity;
            int(16) MaxofDisparity;
        }
    }

    //other additional information
    if(other_flag){

```

【0 1 3 4】

表 4 0 は、多様なカメラパラメータを支援する場合の「ishd」ボックスの第 4 実施形態のシンタックスを表す。ステレオスコピック構成フォーマットは同一であると仮定し、「cameraParams」および「displaySafeInfo」が互いに異なる任意の Item_ID を参照する。

【0 1 3 5】

【表 4 0】

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

```

unsigned int(8) Stereoscopic_Composition_type;
unsigned int(1) LR_first;
unsigned int(7) Reserved;
unsigned int(16) item_count;
for(i=0;i<item_count;i++) {
    unsigned int(16) item_ID;
    unsigned int(1) Is_camParams;
    unsigned int(1) Is_displaySafeInfo;
    unsigned int(1) Is_camParamsref;
    unsigned int(1) Is_displaySafeInforef;
    unsigned int(4) Reserved;

    //all the following are optional fields
    If(Is_ camParamsref || Is_displaySafeInforef){
        unsigned int(16) current_camrefIndex;
        unsigned int(16) current_displayrefIndex;
    }
    else{
        if(Is_CamParams) {
            unsigned int(32) Baseline;
            unsigned int(32) focalLength;
            unsigned int(1) Is_camera_cross
            unsigned int(7) Reserved;
            if(Is_camera_cross) {
                unsigned int(32) convergence_distance;
                unsigned int(32) rotation[];
                unsigned int(32) translation[];
            }
            if(Is_displaySafeInfo){
                unsigned int(16) ViewingDisatance;
                int(16) VerticalDisparity;
                int(16) MinofDisparity;
                int(16) MaxofDisparity;
            }
        }
        //other additional information
        if(other_flag){
    }

```

【 0 1 3 6】

表 4 0 において、「Is__camParamsref」が「0」である場合、参照されるカメラパラメータ情報がないことを表し、「1」である場合、参照されるItem_IDが有ることを表す。「Is__displaySafeInforef」が「0」である場合、参照されるディスプレイsafe情報がないことを表し、「1」である場合、参照されるItem_IDが有ることを表す。「current__cameraIndex」は参照されるItem_IDを表し、「current__displayIndex」は

参照される I t e m __ I D を表す。

【 0 1 3 7 】

「 i s h d 」ボックスは、ステレオスコピックビデオメディア情報を記録する「 s v m i (stereoscopic video media information) 」ボックスと、カメラパラメータおよびディスプレイ情報を記録する「 s c d i (stereoscopic camera and display information) 」ボックスとに分離することができる。「 s v m i 」ボックスは、必須 (mandatory) であり、「 s c d i 」ボックスは必須でないため、「 i s h d 」ボックスを「 s v m i 」ボックスと「 s c d i 」ボックスとに分離するのは不必要な情報を除去するのに実益がある。

【 0 1 3 8 】

「 s v m i 」ボックスは、ステレオスコピックビジュアルタイプおよび区間情報を提供する。具体的に、ステレオスコピックビデオメディア情報は、ステレオスコピック映像構成タイプ情報と、左の映像および右の映像のうちで先にエンコーディングされる映像に関する情報と、ステレオスコピックコンテンツの基本符号化ストリームがステレオスコピック区間からモノスコピック区間に、またはモノスコピック区間からステレオスコピック区間に変更される場合の、区間の数に関する情報と、連続するサンプルの数または開始サンプル番号をカウントする情報と、現在のサンプルがステレオスコピックであるか否かに関する情報などを含む。

【 0 1 3 9 】

「 s c d i 」ボックスは、カメラパラメータが存在するか否か、左右のカメラ間の距離、左右のカメラの配置、メインビューカメラからサブビューカメラまでの相対角度、および左右の映像間の最大視差と最小視差などを含む。表 4 1 は、「 s c d i 」ボックスのシンタックスの一実施形態である。

【 0 1 4 0 】

10

20

【表 4 1】

```

unsigned int(16) item_count;
for(i=0;i<item_count;i++) {
    unsigned int(16) item_ID;
    unsigned int(1) Is_camParams;
    unsigned int(1) Is_displaySafeInfo;
    unsigned int(1) Is_ref;
    unsigned int(5) Reserved;
    //all the following are optional fields
    If(Is_ref){
        unsigned int(16) current_refIndex;
    }
    else{
        if(Is_CamParams){
            unsigned int(32) Baseline;
            unsigned int(32) focalLength;
            unsigned int(1) Is_camera_cross
            unsigned int(7) Reserved;
        }
        if(Is_camera_cross) {
            unsigned int(32) convergence_distance;
            unsigned int(32) rotation[];
            unsigned int(32) translation[];
        }
        if(Is_displaySafeInfo){
            unsigned int(16) ViewingDisatance;
            int(16) VerticalDisparity;
            int(16) MinofDisparity;
            int(16) MaxofDisparity;
        }
    }
}

```

【0 1 4 1】

示されるように「t r e f」ボックスによって、各トラックに含まれる「s c d i」情報の冗長性を除去することができる。E S = 2である場合、各トラックの「i l o c」ボックスにより、「s c d i」情報を提供するために、ステレオスコピック区間を区分する。ここで、各トラックのステレオスコピック区間は、同一のi t e m _ I Dを有し、同一のカメラおよびディスプレイ情報を有する。「t r e f」の「s v d p」によって、基本符号化ストリームは、主要 / 付加トラックに区分される。1つのトラックにのみ「i l o c」ボックスが含まれるが、3 Dディスプレイを行う場合、ステレオスコピック区間別に「i l o c」ボックスを同期することによって、再生が可能である。

【0 1 4 2】

また「t r e f」ボックスによって、各トラックに含まれる同一のステレオスコピックビデオメディア情報(「s v m i」)の冗長性を除去することもできる。「f t y p」ボックスによって、ステレオスコピックコンテンツを認識した後、「t r e f」ボックスの「s v d p」ボックスによって、主要 / 付加トラックに区分される。1つのトラックにのみ「s v m i」ボックスが存在すれば、他のトラックは自動的に認識されることができる。「s v m i」ボックスは必須であるため、主要 / 付加トラック内に存在することもでき、場合によって主要トラック内にのみ存在することができる。

【0 1 4 3】

図15はES = 2である場合、SS - VAFの一実施形態の構造図である。図15に図示されたSS - VAFは、「svmi」ボックスおよび「scdi」ボックスを備える。

【0144】

ステレオスコピックコンテンツを構成するESが2個の場合、2個のトラック(「trak」)が存在し、これは1つの主要トラックと付加トラックに区分することができる。これに伴い付加トラック内「tref」の「svdp」を用いて、主要トラックを参照して関連「scdi」情報に含まれる情報を提供する。このような構造により、各トラックに含まれる同一の「scdi」情報の冗長性を除去する長所を有する。ここで、track__IDは、参照されるトラックのIDを表す。また、reference__typeが「svdp」ならば、トラックが参照トラックに対してステレオスコピックカメラおよびディスプレイ情報を含むことを表す。

10

【0145】

ユーザが所定の映像を3Dディスプレイモードで視聴する途中、3Dディスプレイモードから2Dディスプレイモードに変更する場合、端末で主要トラックに対応する映像をディスプレイすることによって、2Dディスプレイモードを満足する。一方、ESが2個のトラックでは、基本符号化ストリームの中に同一のモノスコピックデータが存在することができる。モノスコピックデータでは、同一のコンテンツが2個のトラックに保存される。従って、モノスコピックデータは、3Dでディスプレイされ得ないデータである。このような場合、端末では2個のトラックのうち、どのトラックのモノスコピックデータをディスプレイするのが決定しなければならないが、本発明によって区分された主要トラックに対応する映像をディスプレイするようにする。

20

【0146】

また、ステレオスコピックビデオのESが2個の場合、2個のトラックが存在する。従来のLASERなどのシーン記述子によって、ステレオスコピックビデオが2個のオブジェクトとして認識され、LASER内のビデオ関連ノードの2個に連結される。しかし、ステレオスコピックビデオは、端末で最終的に1つの3Dビデオフォーマットに変換され再生されなければならないため、LASERに1つのオブジェクトとして認識されなければならない。すなわち、ステレオスコピックビデオをシーン再生するために、ステレオスコピックビデオを1つの3Dビデオフォーマットに変換する必要があるため、1つのノードのみを使用して連結される。よって本実施形態により、「tref」内の「svdp」を用いて、ステレオスコピックビデオは主要/付加トラックに区分され、LASER内のビデオ関連ノードは、主要トラックに対応する「trak」__IDまたはメディアストリームのみをリンクする。

30

【0147】

図16は、本発明によるISOベースメディアファイルフォーマットでステレオスコピックコンテンツを保存する方法の一実施形態のフローチャートである。まず、対象のステレオスコピックコンテンツをISOベースメディアファイルフォーマットの「mdat」ボックスに保存する(S1602)。次に、ステレオスコピックコンテンツのステレオスコピックビデオメディア情報と、ステレオスコピックカメラおよびディスプレイ情報を、ステレオスコピックコンテンツのメタデータとして「moov」ボックスに保存する(S1604、S1606)。

40

【0148】

図17は、本発明によるISOベースメディアファイルフォーマットでステレオスコピックコンテンツを保存する方法の他の実施形態のフローチャートである。

まず、対象のステレオスコピックコンテンツを、ISOベースメディアファイルフォーマットの「mdat」ボックスに保存する(S1702)。次に、ステレオスコピックコンテンツの基本符号化ストリームが2個以上の場合、基本符号化ストリームを主要トラックまたは付加トラックに区分する情報(「svdp」)をISOベースメディアファイルフォーマットの「tref」ボックスに保存する(S1704)。次に、主要トラックにのみリンクされるビデオ関連ノードを備えるステレオスコピックコンテンツに対するLAS

50

e Rを保存する (S 1 7 0 6)。ここで、L A S e Rにリンクされる部分は使用されず、「 t r e f」のみ使用される。

【 0 1 4 9 】

図 1 8 は、本発明によるステレオスコピックコンテンツ保存方法のまた他の実施形態のフローチャートである。ここでは前述した A V C S E Iを用いて、ステレオスコピックコンテンツを保存する。まず、A V Cによって、符号化されたステレオスコピックコンテンツを保存する (S 1 8 0 2)。次に「 r e s e r v e d _ s e i _ m e s s a g e」を用いて、ステレオスコピックコンテンツの各ストリームに要求されるカメラパラメータおよびディスプレイ情報を保存する (S 1 8 0 4)。

【 0 1 5 0 】

ここで「 s t e r e o v i d e o i n f o r m a t i o n S E I」は、ステレオスコピック映像構成が「 s i d e b y s i d e」タイプであることを示す情報を追加で含む。カメラパラメータは、左右のカメラ間の距離、左右同一焦点距離 (f o c a l _ l e n g t h)、ベースラインからコンバージェンスポイント (収束点)までの距離、左右のカメラが交差配列する場合の、左右のカメラの回転 (r o t a t i o n)と移動 (t r a n s l a t i o n)のうち、少なくとも一つを含む。ディスプレイ情報は、左右の映像間の最大視差と最小視差を含む。

【 0 1 5 1 】

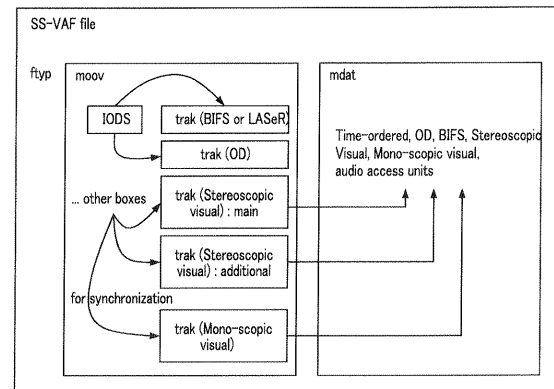
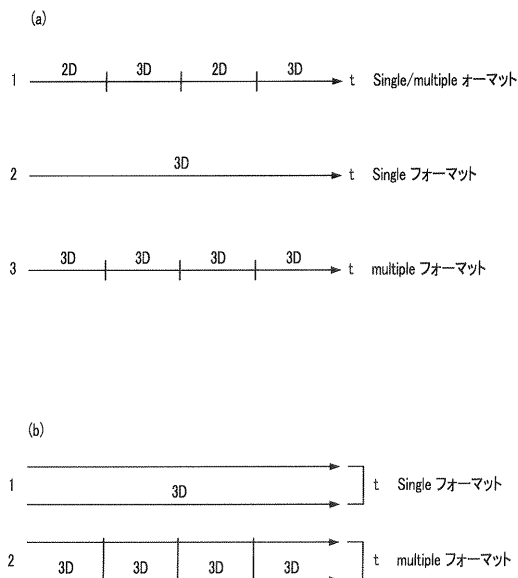
上述したように、本発明の方法は、プログラムで具現され、コンピュータ読み取り可能な記録媒体 (C D - R O M、R A M、R O M、フロッピー (登録商標) ディスク、ハードディスク、光磁気ディスクなど) に保存することができる。このような過程は当業者であれば容易に実施することができるので、ここでは詳述しない。

【 0 1 5 2 】

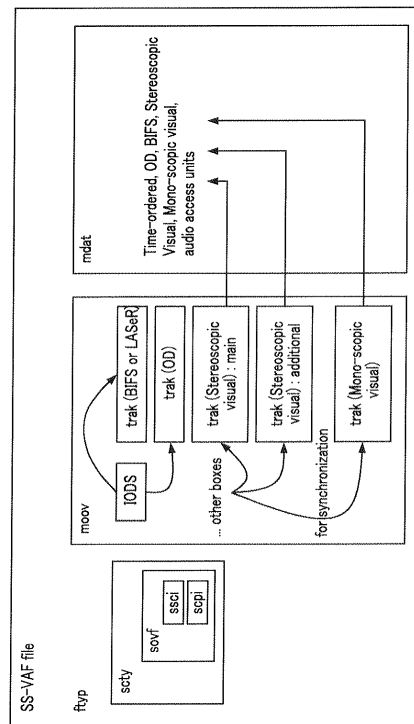
本発明を特定の実施形態に関連して説明してきたが、当業者であれば、添付の特許請求の範囲で定義される本発明の技術的思想および範囲から逸脱することなく、様々な置換、変形および変更を行うことが可能であることを理解されよう。

【 図 1 】

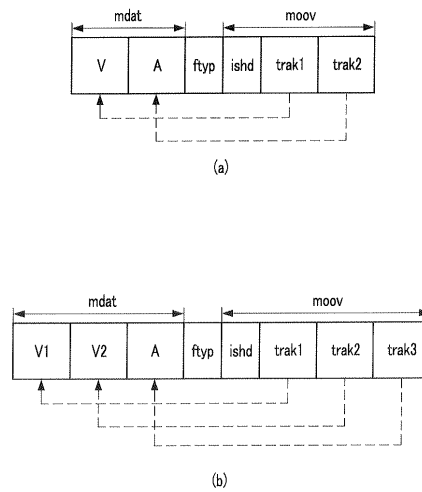
【 図 2 】



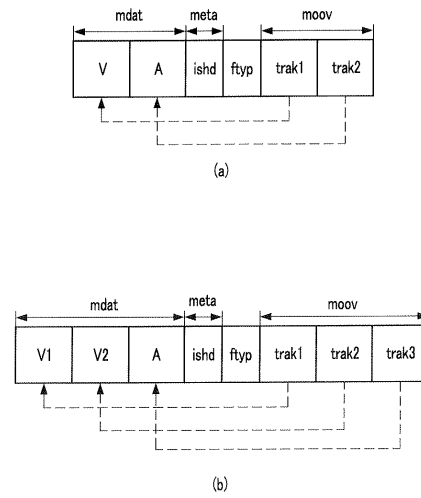
【 図 4 】



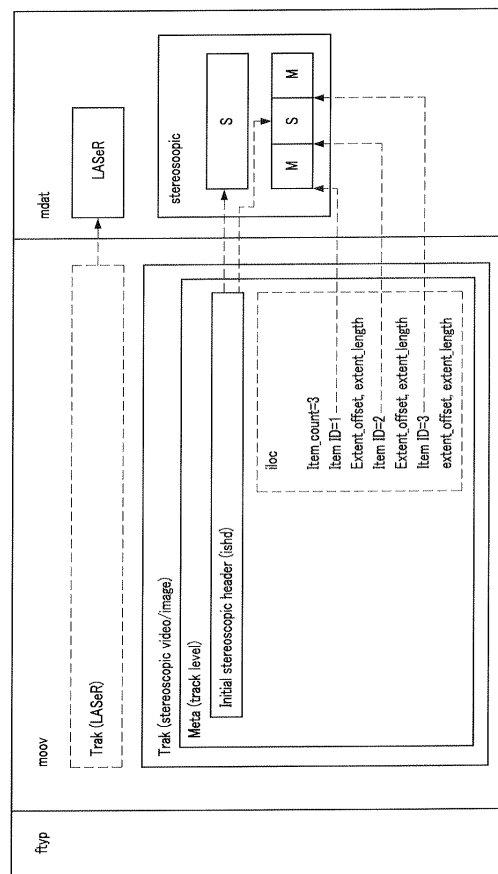
【 図 6 】



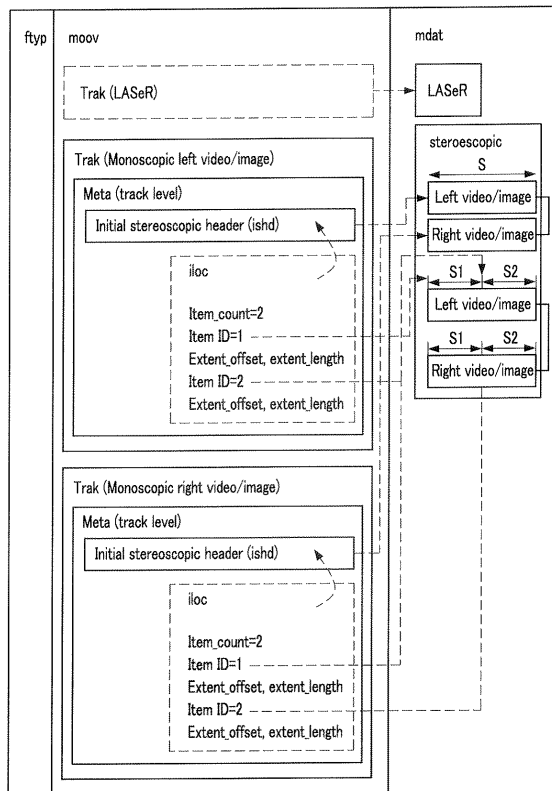
【 図 8 】



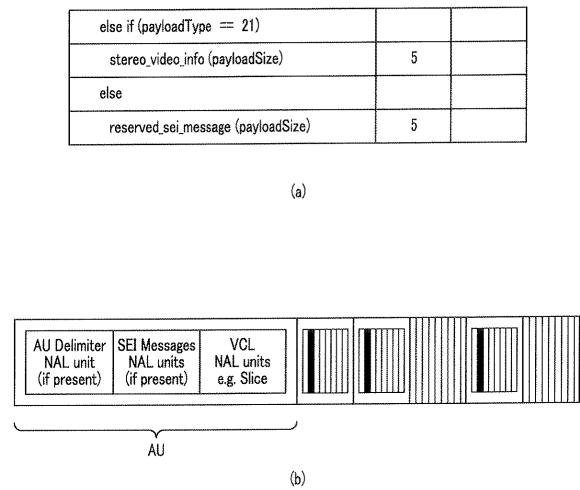
【 ㄨ 1 0 】



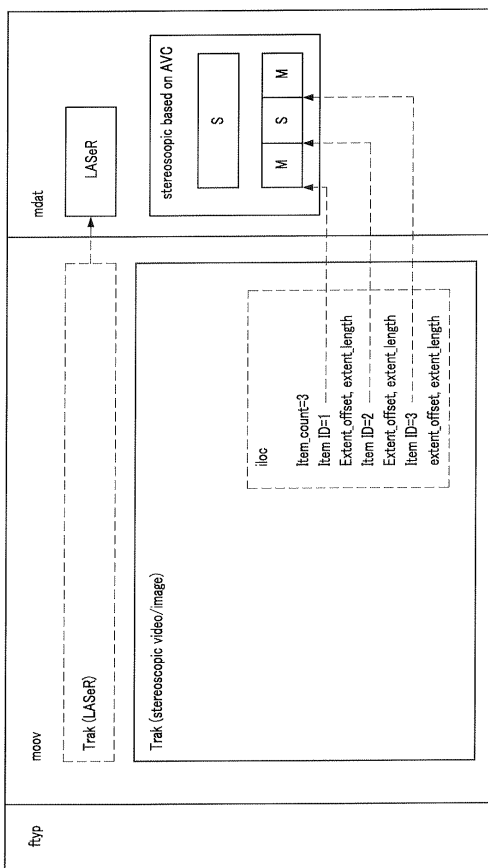
【図 1 1】



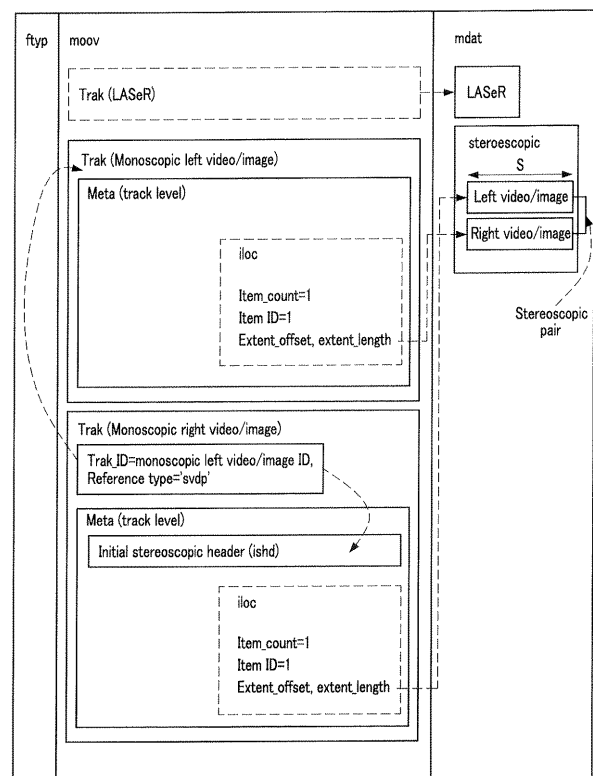
【図 1 2】



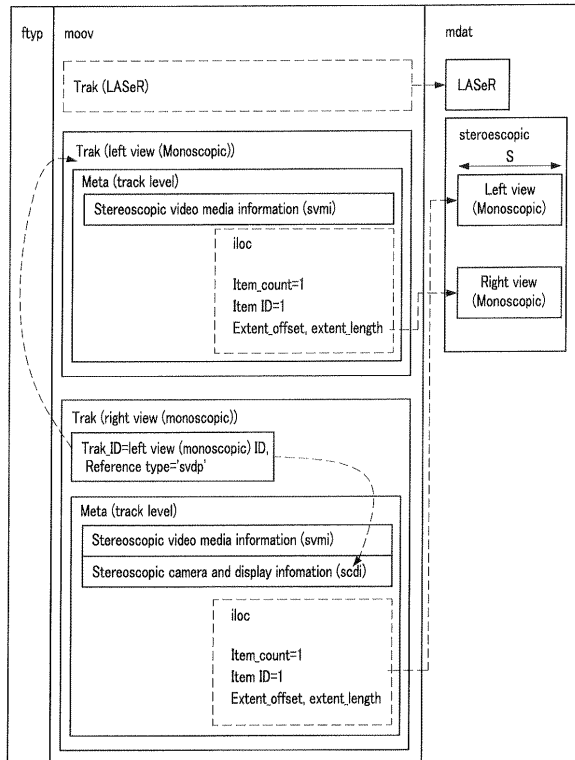
【図 1 3】



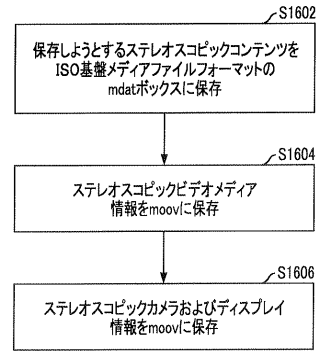
【図 1 4】



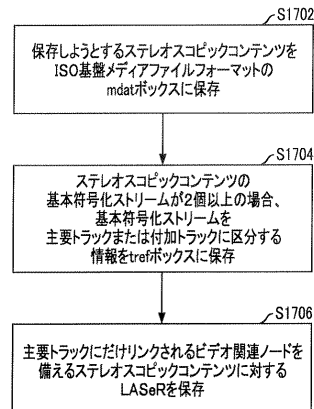
【図 15】



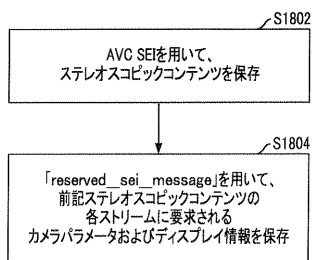
【図 16】





【図 17】



【図 18】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/KR2008/003481
A. CLASSIFICATION OF SUBJECT MATTER		
<i>H04N 13/00(2006.01)i, H04N 7/24(2006.01)i</i>		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) IPC 8 H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean Utility models and applications for Utility models since 1975 Japanese Utility models and applications for Utility models since 1975		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKIPASS (KIPO internal): "stereoscopic, content, metadata, information"		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-0716142 B1 (ECT INC. et al.) 11 May 2007 See the whole document.	1-15
A	JP 2005-026800 A (KONICA MINOLTA PHOTO IMAGING INC.) 27 January 2005 See abstract, paragraphs [0121]-[0126]; claims 2, 5, 9, 13, and 16.	1-15
A	JP 2006-013759 A (MATSUSHITA ELECTRIC IND. CO. LTD.) 12 January 2006 See abstract, paragraphs [0011]-[0012]; claims 1-12; figure 1.	1-15
A	US 2003-0095177A1 (KUG-JIN YUN et al.) 22 May 2003 See abstract, paragraphs [0011]-[0030].	1-15
A	KR 10-2007-0061227 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 13 June 2007 See the whole document.	1-15
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 27 NOVEMBER 2008 (27.11.2008)		Date of mailing of the international search report 27 NOVEMBER 2008 (27.11.2008)
Name and mailing address of the ISA/KR  Korean Intellectual Property Office Government Complex-Daejeon, 139 Seonsa-ro, Seo-gu, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140		Authorized officer Yoon, Yeomin Telephone No. 82-42-481-8515 

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2008/003481

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
KR 10-0716142 B1	11.05.2007	KR 10-2006-0100258 A	16.10.2006
JP 2005-026800 A	27.01.2005	None	
JP 2006-013759 A	12.01.2006	None	
US 2003-0095177 A1	22.05.2003	AU 2002-366143 A1	10.06.2003
		CN 1613263 A	04.05.2005
		EP 1457034 A2	15.09.2004
		JP 2005-510187 A	14.04.2005
		KR 10-2003-042090 A	28.05.2003
		US 2004-0120396 A1	24.06.2004
		WO 03-045046 A2	30.05.2003
KR 10-2007-0061227 A	13.06.2007	WO 2007-067020 A1	14.06.2007

フロントページの続き

(31)優先権主張番号 10-2007-0098665

(32)優先日 平成19年10月1日(2007.10.1)

(33)優先権主張国 韓国(KR)

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(72)発明者 ユン クグ - ジン

大韓民国 305 - 358 テジョン ユソング ハギドン ソンニム マウル ホバン リージ
ェンシー ヴィル ナンバー106 - 1504

(72)発明者 チョ スク - ヒ

大韓民国 305 - 804 テジョン ユソング シンソンドン 137 - 11 ナンバー301

(72)発明者 ホ ナムホ

大韓民国 305 - 325 テジョン ユソング ノウンドン (番地なし) ヨルメ マウル
アパートメント ナンバー801 - 1001

(72)発明者 キム ジン - ウン

大韓民国 305 - 761 テジョン ユソング チョンミンドン (番地なし) エクスポ アパ
ートメント ナンバー305 - 1603

(72)発明者 イ ス - イン

大韓民国 302 - 120 テジョン ソグ トウンサンドン (番地なし) クローバー アパ
ートメント ナンバー106 - 606

Fターム(参考) 5C159 LA05 MA00 PP01 PP13 SS11 TA00 TB12 TC00 UA31