

# (19)대한민국특허청(KR)

## (12) 공개특허공보(A)

(51) 。 Int. Cl.<sup>7</sup>  
H04N 7/24

(11) 공개번호 10-2005-0099488  
(43) 공개일자 2005년10월13일

(21) 출원번호 10-2005-0088630  
(22) 출원일자 2005년09월23일

(71) 출원인 한국정보통신대학교 산학협력단  
대전 유성구 문지동 103-6

(72) 발명자 노용만  
대전 유성구 노은동 새미래 8단지아파트 816-1102  
진성호  
경북 상주시 냉림동 358-3번지  
양승지  
강원 원주시 학성1동 954-13

(74) 대리인 장성구  
김원준

심사청구 : 있음

(54) 비디오 및 메타데이터의 통합을 위한 비디오 멀티미디어응용 파일 형식의 인코딩/디코딩 방법 및 시스템

### 요약

본 발명은 비디오 및 메타데이터의 통합을 위한 비디오 멀티미디어 응용 파일 형식의 인코딩/디코딩 방법 및 시스템에 관한 것이다. 즉, 본 발명은 비디오 데이터와 관계된 멀티미디어 콘텐츠 및 메타데이터(metadata)를 결합하는 새로운 비디오 멀티미디어 응용 포맷(Video Multimedia Application Format, Video MAF)을 구성하는 과정에 있어서, 효과적인 인덱싱(indexing) 및 검색을 위해 비디오 데이터 및 내용으로부터 획득한 오디오(audio)/이미지(image)/모션(motion) 특징정보를 비디오 MAF 파일의 메타데이터로 활용하며, 비디오 MAF의 구조에 따라 사용자가 MAF파일내에 포함된 비디오 및 메타데이터를 효과적으로 브라우징(browsing)할 수 있도록 한다.

### 대표도

도 1

### 색인어

비디오 멀티미디어 응용 파일 형식(Video MAF), 메타데이터, MPEG-4, MPEG-7, 오디오/이미지/비디오 특징정보

### 명세서

### 도면의 간단한 설명

도 1은 본 발명의 실시 예에 따른 전체 시스템 구성도,

도 2는 본 발명의 실시 예에 따른 비디오 프로세싱/저작부의 상세 구조도,

도 3은 본 발명의 실시 예에 따른 메타데이터의 구성 요소 및 요소들의 구조도,

도 4는 본 발명의 실시 예에 따른 비디오 MAF의 구성 요소 및 요소들의 구조도,

도 5는 본 발명의 실시 예에 따른 복수의 비디오 스트림을 포함하는 비디오 MAF의 구성 요소 및 요소들의 구조도.

## 발명의 상세한 설명

### 발명의 목적

#### 발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 비디오 멀티미디어 응용 포맷(Video Multimedia Application Format: Video MAF)에 관한 것으로, 특히, 비디오 데이터와 관계된 멀티미디어 콘텐츠 및 메타데이터(metadata)를 결합하는 새로운 비디오 멀티미디어 응용 포맷을 구성하는 과정에 있어서, 효과적인 인덱싱(indexing) 및 검색을 위해 비디오 데이터 및 내용으로부터 획득한 오디오(audio)/이미지(image)/모션(motion) 특징정보를 비디오 MAF 파일의 메타데이터로 활용하고, 비디오 MAF의 구조에 따라 사용자가 MAF파일 내에 포함된 비디오 및 메타데이터를 효과적으로 브라우징(browsing)할 수 있도록 하는 비디오 MAF 파일의 인코딩/디코딩 시스템 및 방법에 관한 것이다.

국제 표준화 기구인 MPEG(Moving Picture Experts Group)은 1988년 MPEG-1의 표준화를 시작으로 현재까지, MPEG-2, MPEG-4, MPEG-7, MPEG-21의 표준화를 수행 중에 있으며, 이와 같은 표준 기술들은 하나의 단일 표준 규격 내에서 필요한 요소들의 집합으로 표준 규격의 프로파일(profile)들로 구성하여, 특정한 응용 서비스를 지원하는 방식을 취하고 있다.

그러나, 위와 같은 제한된 표준 규격화는 산업의 다양한 기술적 요구 사항을 하나의 표준 기술로 만족시키기 어렵다는 점 때문에 그 동안 문제가 되어왔다. 즉, 다양한 멀티미디어 표준 기술들이 개발됨에 따라, 서로 다른 표준 기술들을 조합하여, 새로운 멀티미디어 응용 서비스들에 적용시킬 필요성이 대두되고 있다.

이를 위해, MPEG-A(MPEG Application: ISO/ICE 23000)에서는 이러한 필요성에 대한 표준화 활동을 진행하고 있다. MPEG-A에서 진행중인 멀티미디어 응용 형식(Multimedia Application Format, MAF)은 기존의 MPEG 표준들뿐만 아니라, JPEG과 같은 비 MPEG 표준들을 함께 조합함으로써, 산업계의 요구를 적극적으로 대응하여 표준의 활용 가치를 높이고자 하고 있다. 위와 같은 MAF 표준화의 주요 목적은 MPEG 기술을 쉽게 산업에서 이용할 수 있는 기회를 제공함으로써, 산업에서 필요한 응용 서비스를 위해 별도의 표준을 새로 만드는 노력 없이 이미 검증된 표준 기술을 쉽게 조합할 수 있도록 하는 것이다.

현재 음악 콘텐츠에 대한 응용 파일 형식 표준화가 MPEG Music Player Application Format(ISO/IEC 23000-2)이라는 이름으로 수행되어 음악 MAF는 FDIS(Final Draft International Standard) 상태로 표준화가 거의 완성 단계에 있다. 이로써 기존에 재생만 하던 MP3 재생기의 기능을 확장하여, 음악을 장르에 따라 자동으로 분류하여 재생하거나, 음악을 듣는 동시에 가사를 보여주거나, 음악과 관련된 앨범 재킷 사진 등을 브라우징 할 수 있게 되었다. 사용자들이 보다 향상된 음악 서비스를 받을 수 있는 파일 형식이 마련된 것이다. 특히, 최근 MP3 재생기가 휴대폰, 게임기(소니의 PSP), 또는 휴대형 멀티미디어 재생기(Portable Multimedia Player, PMP)에 탑재되는 등 소비자에게 인기가 높아지고 있다. 따라서 MAF를 이용한 향상된 기능의 음악 재생기를 상용화하는 것도 빠른 시일 내에 가능할 것이다.

또한 영상, 특히 사진 콘텐츠에 대한 응용 파일 형식 표준화가 MPEG Photo Player Application Format(ISO/IEC 23000-3)이라는 이름으로 시작되어, 사진 MAF의 형식 표준화 작업은 WD(Working Draft) 상태이다. MP3 음악과 마찬가지로, 디지털 카메라 시장이 꾸준히 증가하면서 디지털 카메라를 통하여 획득된 사진(일반적으로 JPEG) 데이터가 급속도로 증가하고 있다. 사진 데이터를 저장하는 매체(메모리 카드)가 소형화되고 집적화 됨에 따라, 하나의 메모리 카드에 수백 장의 사진을 저장할 수 있게 되었으나, 증가하는 사진의 양과 비례적으로 사용자들이 사진을 관리하면서 느끼는 어려움은 증가하고 있다. 따라서, 사진 MAF는 이러한 문제점을 해결하기 위하여 사진 콘텐츠와 관련된 메타데이터를 하나의 파일에 함께 묶는 새로운 파일 형식을 표준화가 시도되고 있다.

한편, 상기 음악 콘텐츠와 사진 콘텐츠에 대한 표준화 작업과 더불어 개인용 비디오나 방송용 비디오와 같은 비디오 MAF에 관해서도 표준화 연구에 대한 건의가 시작되고, 활용성이 증가할 것으로 예상된다. 따라서, 비디오 데이터와 관계된 멀티미디어 콘텐츠 및 메타데이터(metadata)를 결합하는 새로운 비디오 멀티미디어 응용 포맷을 구성하고, 사용자가 비디오 MAF 파일 내에 포함된 비디오 및 메타데이터를 효과적으로 브라우징(browsing)할 수 있도록 한다면 멀티미디어 콘텐츠의 보다 효과적인 활용이 기대된다.

### 발명이 이루고자 하는 기술적 과제

따라서, 본 발명의 목적은 기존의 비디오 데이터 및 메타데이터의 관리 방법의 문제점을 해결함으로써 사용자가 휴대용 단말에서 효과적으로 비디오를 브라우징할 수 있도록 하기 위하여, 비디오 데이터, 비디오의 구조 및 내용으로부터 얻는 오디오/이미지/모션 특징정보, 사진의 효과적인 인덱싱을 위한 다양한 특징정보를 메타데이터로 활용하여 비디오 MAF 파일로 인코딩하며, 상기 구조에 따라 사용자가 비디오 MAF 파일을 효과적으로 브라우징 및 소비할 수 있도록 비디오 데이터와 관계된 멀티미디어 콘텐츠 및 메타데이터를 결합하는 비디오 MAF 파일 인코딩/디코딩하는 방법 및 시스템을 제공함에 있다.

또한 본 발명의 다른 목적은 비디오 데이터와 이미지, 자막(subtitle) 그리고, 음악파일을 포함하는 그 외의 멀티미디어 콘텐츠들을 하나로 통합하여 비디오 MAF라는 응용 파일 형식으로 구성하는 과정에서, 비디오 데이터, 비디오의 구조 및 내용으로부터 얻는 오디오/이미지/모션 특징정보, 비디오의 효과적인 인덱싱 및 검색을 위한 요약 정보를 메타데이터로 포함하고, 메타데이터에 기반을 둔 콘텐츠 활용 방법적 도구들을 포함함으로써, 사용자에게 특별한 어플리케이션이 없거나, 메타데이터를 활용하는 기능을 포함하고 있지 않더라도, 멀티미디어 콘텐츠의 효과적인 소비가 가능하게 하는데 있다.

상술한 목적을 달성하기 위한 본 발명은, 개인 사용자가 PDA, 디지털 카메라, 휴대폰, 게임기와 같은 휴대용 단말에서 개인용 비디오를 획득/입력하거나 혹은 서비스 제공자나 콘텐츠 제작자가 방송용 비디오를 획득/입력하는 단계, 입력된 비디오로부터 특징정보를 획득하는 비디오 프로세싱 및 저작 단계, 획득된 특징정보들을 가지고 메타데이터를 생성하는 단계, 획득된 비디오 데이터와 메타데이터를 조합하여 비디오 MAF를 인코딩하는 단계, 비디오 MAF를 전송하는 단계, 전송된 비디오 MAF를 디코딩하는 단계, 디코딩된 비디오 데이터를 출력하는 단계, 디코딩된 메타데이터를 출력하는 단계를 포함하는 것을 특징으로 한다.

또한 본 발명은 발명의 핵심적인 요소로써, 비디오 MAF의 구조 및 비디오 MAF에 포함될 메타데이터의 구성 요소들을 정의하고, 각 구성 요소들의 관계 및 서술 구조를 포함하는 것을 특징으로 한다.

### 발명의 구성 및 작용

이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시 예의 동작을 상세하게 설명한다.

도 1을 본 발명의 실시 예에 따른 비디오 MAF를 효과적으로 구성하고 인코딩 및 디코딩하기 위한 시스템 구성을 도시한 블록도 이다.

이하 상기 도 1을 참조하여 각 블록에서의 동작을 상세히 설명하기로 한다. 먼저 비디오 획득/입력부(Video acquisition/input part)(1000)는 PDA, 휴대폰, 디지털 카메라, 게임기와 같은 휴대용 단말 장치를 이용하여 개인용 비디오를 획득하거나 입력받는다. 또한, 서비스 제공자나 콘텐츠 제작자에 의해서, 제공되는 양질의 방송용 비디오를 획득하거나 입력받는다.

비디오 프로세싱/저작부(Video processing/authoring part)(2000)는 입력된 비디오 데이터를 표현할 수 있는 특징정보를 프로세싱하고 저작한다. 즉, 비디오 프로세싱/저작부(2000)는 입력되는 비디오로부터 비디오의 샷(shot) 검출, MPEG-7의 오디오/이미지/모션 서술자(descriptor) 검출, 프리뷰(preview) 이미지 및 키프레임(key-frame) 이미지들을 검출한다.

비디오 메타데이터 생성부(Video metadata creation part)(3000)는 비디오 프로세싱/저작부(2000)에서 추출된 특징정보를 메타데이터로 생성한다. 이때 비디오 메타데이터 생성은 MPEG-7 MDS(Multimedia Description Scheme)를 기반으로 하여 XML 형태로 생성하게 된다.

비디오 MAF 인코딩부(Video MAF encoding part)(4000)는 입력된 비디오와 생성된 메타데이터를 조합하여 비디오 MAF를 인코딩 한다. 위와 같이 서비스 제공자나 콘텐츠 제작자에 의한 비디오 MAF는 기본적인 질의 응답 서비스 시스템을 통해 사용자의 요청에 의한 응답으로 사용자의 휴대 단말 장치로 전송된다.

비디오 MAF 디코딩부(Video MAF decoding part)(6000)는 전송된 비디오MAF를 휴대용 단말에서 비디오 데이터와 메타데이터로 디코딩 한다. 비디오 MAF 재생부(Video MAF player)(7000)는 디코딩된 비디오 데이터를 브라우징하고 디스플레이한다. 메타데이터 해석부(Metadata parser)(8000)는 메타데이터로부터 스토리(story)기반 비디오의 표현 및 비디오의 검색을 위한 특징정보를 해석한다.

본 발명에서는 입력된 비디오 데이터를 표현할 수 있는 특징정보를 프로세싱하고 저작하는 단계에 있어서, 프리뷰 이미지 추출, 샷 검출과 샷 단위의 모션 특징정보 및 오디오 특징정보의 추출, 키프레임의 추출, 키프레임들 단위의 이미지 특징정보의 추출하는 과정을 특징으로 한다.

도 2는 상기 도 1에 도시된 비디오 프로세싱/저작부(2000)의 상세 구조도를 도시한 것으로, 이하 상기 도 2를 참조하여 비디오 프로세싱/저작부 각 구성요소에서의 동작을 상세히 설명하기로 한다.

먼저, 프리뷰 이미지 선택부(Preview image selection part)(2110)는 입력된 비디오 데이터를 대표하는 프리뷰 이미지를 선택한다. 이때 프리뷰 이미지는 비디오 데이터를 대표하는 포스터 이미지나 광고 이미지들을 포함하며, 프리뷰 이미지의 형식은 JPEG, BMP, GIF등을 포함하나 이에 한정되지는 않는다.

샷 검출부(Shot detection)(2120)는 입력된 비디오 데이터를 샷(Shot) 단위로 세그먼트하며, 샷 검출의 결과로 샷 구간들과 각 샷을 대표하는 키프레임을 획득한다. 이때 키프레임의 형식은 JPEG, BMP, GIF등을 포함하나 이에 한정되지는 않는다.

모션 특징 획득부(Motion feature extraction part)(2130)는 샷 검출부(2120)에서 획득된 샷 구간들로부터, MPEG-7 모션 서술자들을 이용하여 모션 특징을 획득한다. 이때 모션 특징으로 카메라 모션(Camera motion)과 모션 액티비티(Motion activity)가 추출되는 것을 특징으로 한다.

오디오 특징 획득부(Audio feature extraction part)(2140)는 샷 검출부(2120)에서 획득된 샷 구간들로부터, MPEG-7 저레벨 오디오 서술자(Low-level audio descriptor)들을 이용하여 오디오 특징을 획득한다. 이때 오디오 특징으로 기본(Basic), 기본 스펙트럴(Basic spectral), 기본 신호(Basic signal), 시간적 음색(Temporal timbral), 스펙트럴 음색(Spectral timbral), 스펙트럴 기저(Spectral basis) 서술자들을 포함하는 것을 특징으로 한다.

키프레임 추출부(Key-frame extraction part)(2150)는 추출된 샷 구간내에서 키프레임을 추출한다. 이미지 특징 추출부(Image feature extraction part)(2160)는 추출된 키프레임으로부터 이미지 특징을 추출한다. 이때 이미지 특징은 칼라 스트럭처(Color structure), 칼라 레이아웃(Color layout), 스케일러블 칼라(Scalable color), 주요 칼라(Dominant color), 에지 히스토그램(Edge histogram), 균질 질감(Homogeneous texture) 서술자들을 포함한다.

본 발명에서는 메타데이터를 표준화된 형식 및 구조에 따라 서술하는 단계에 있어서, MPEG-7 이용하는 것을 특징으로 한다.

도 3은 상기 도 1의 비디오 메타데이터 생성부(3000)에서 생성되는 메타데이터의 구성 요소 및 요소들의 구조를 도시한 것이다.

상기 도 3을 참조하면, 비디오 MAF 메타데이터(Video MAF metadata, 3100)는 미디어 자체에 대한 서술인 미디어 정보 및 생성 서술부(3110)를 포함하고, 미디어 의미 서술부(3120)를 포함하고, 미디어의 디지털 권리를 서술하는 디지털 권리 서술부(3130)를 포함하는 것을 특징으로 한다.

상기 미디어 정보 및 생성 서술부(Media information/creation description)(3110)는 MPEG-7 MDS에 의한 생성 관련 메타데이터를 포함하고, 미디어의 시간적 구조를 서술하는 메타데이터를 포함하는 것을 특징으로 한다.

하기의 [표 1]은 상기 미디어의 정보 생성 서술을 위해 필요한 메타데이터를 표현하기 위한 서술 구조를 XML 형식으로 표현한 것이다.

표 1.

```
<complexType name="CreationInformationType">
  <complexContent>
    <extension base="mpeg7:DSType">
      <sequence>
        <element name="Creation" type="mpeg7:CreationType"/>
        <complexType>
          <sequence>
            <element name="Title" type="mpeg7:TitleType"
              minOccurs="1" maxOccurs="unbounded"/>
            <element name="PreviewImage" type="mpeg7:TitleMediaType"
              minOccurs="0"/>
            <element name="Abstract" type="mpeg7:TextAnnotationType"
              minOccurs="0" maxOccurs="unbounded"/>
            <element name="Creator" type="mpeg7:CreatorType"
              minOccurs="0" maxOccurs="unbounded"/>
            <element name="CreationCoordinates" minOccurs="0"
              maxOccurs="unbounded"/>
            <complexType>
              <sequence>
                <element name="Location" type="mpeg7:PlaceType" minOccurs="0"/>
                <element name="Date" type="mpeg7:TimeType" minOccurs="0"/>
              </sequence>
            </complexType>
          </sequence>
        </complexType>
        <element name="CreationTool" type="mpeg7:CreationToolType"
          minOccurs="0" maxOccurs="unbounded"/>
        <element name="CopyrightString" type="mpeg7:TextualType"
          minOccurs="0" maxOccurs="unbounded"/>
      </sequence>
    </complexType>
    <element name="Classification" type="mpeg7:ClassificationType" minOccurs="0"/>
    <element name="RelatedMaterial" type="mpeg7:RelatedMaterialType" minOccurs="0"
      maxOccurs="unbounded"/>
  </sequence>
</extension>
</complexContent>
</complexType>
```

미디어의 시간구조를 표현하는 메타데이터는 미디어의 시간적 구조를 세그먼트(segment)하는 정보를 포함하고, 각 세그먼트 당 요구되는 MPEG-7 오디오/모션 특징을 포함하고, 각 세그먼트당 추출되는 키프레임 이미지에서 추출되는 MPEG-7 이미지 특징을 포함하는 것을 특징으로 한다.

하기의 [표 2]는 상기 미디어의 시간적 구조를 표현하기 위한 서술 구조를 XML 형식으로 표현한 것이다.

표 2.

```
<complexType name="TemporalDecomposition">
  <complexContent>
    <sequence>
      <element name="AudioVisualSegment" type="mpeg7:AudioVisualSegmentType">
        <complexType>
          <sequence>
            <element name="TextAnnotation" type="mpeg7:TextAnnotationType"/>
          </sequence>
        </complexType>
      </element>
    </sequence>
  </complexType>
```

```

<element name="MediaTime" type="mpeg7:MediaTimeType" />
<element name="MediaSourceDecomposition"
type="mpeg7:VideoSegmentMediaSourceDecompositionType">
<sequence>
<element name="VideoSegment" type="mpeg7:VideoSegmentType" minOccurs="0"
maxOccurs="unbounded"/>
<element name="MediaTime" type="mpeg7:MediaTimeType"/>
<element name="Keyframe" type="mpeg7:ImageLocatorType"/>
<element name="VisualDescriptor" type="mpeg7:VisualDType" minOccurs="0"
maxOccurs="unbounded"/>
</Keyframe>
<element name="VisualDescriptor" type="mpeg7:VisualDType" minOccurs="0"
maxOccurs="unbounded"/>
</VideoSegment>
<element name="AudioSegment" type="mpeg7:AudioSegmentType" minOccurs="0"
maxOccurs="unbounded">
<element name="MediaTime" type="mpeg7:MediaTimeType"/>
<element name="AudioDescriptor" type="mpeg7:AudioDType" minOccurs="0"
maxOccurs="unbounded"/>
</AudioSegment>
<sequence>
</MediaSourceDecomposition>
</sequence>
</complexType>
</AudioVisualSegment>
</sequence>
</complexContent>
</complexType>

```

미디어 의미 서술부(Media semantic description)(3120)부 미디어의 주인공이나 요약정보와 같은 미디어 리소스의 의미 기반 특징의 서술을 포함하는 것을 특징으로 한다.

하기의 [표 3]은 상기 미디어 리소스의 의미를 표현하기 위한 서술 구조를 XML 형식으로 표현한 것이다.

**표 3.**

```

<complexType name="SemanticDescriptionType">
<complexContent>
<extension base="mpeg7:ContentAbstractionType">
<choice>
<element name="Semantics" type="mpeg7:SemanticType" minOccurs="1"
maxOccurs="unbounded"/>
<element name="ConceptCollection" type="mpeg7:ConceptCollectionType"
minOccurs="1" maxOccurs="unbounded"/>
</choice>
</extension>
</complexContent>
</complexType>

```

디지털 권리 서술부(Digital right description)(3130)는 비디오 MAF에 인코딩된 비디오의 접근 권한에 대해서 서술을 포함하는 것을 특징으로 한다.

하기의 [표 4]는 상기 미디어의 접근 권한에 대한 서술 구조를 XML 형식으로 표현한 것이다.

표 4.

```

<complexType name="DigitalRightType">
<complexContent>
<extension base="mpeg7: AvailabilityType">
<sequence>
<element name="EditPermission" type="boolean" use="required"/>
<element name="CopyPermission" type="boolean" use="required"/>
<element name="ReleasePermission" type="boolean" use="required"/>
<element name="ViewPermission" use="required"/>
<choice>
<element name="PulicView" type="boolean" />
<element name="LimitedView" type="boolean"/>
<element name="PrivateView" type="boolean"/>
<element name="FreeView" type="boolean"/>
</choice>
</ ViewPermission>
<sequence>
</extension>
</complexContent>
</complexType>

```

본 발명에서는 비디오와 관련된 멀티미디어 콘텐츠들 및 메타데이터를 조합하는 비디오 MAF 구조를 서술하는 단계에 있어서, MPEG-4 파일 포맷을 이용하는 것을 특징으로 한다.

도 4는 상기 도 1의 비디오 MAF 인코딩부로부터 생성되는 비디오 MAF 파일의 구성 요소 및 요소들의 구조를 도시한 것이다.

상기 도 4를 참조하면, 무브(moov)(5010)는 메타박스(meta)(5030)와 트랙박스(trak)(5040)를 포함하며, 무브는 멀티 트랙들을 포함하는 것을 특징으로 한다.

메타박스(meta)(5030)는 상기 도 3에서 서술하고 있는 비디오 MAF 메타데이터(3100)를 포함한다. 또한, 메타박스는 메타데이터를 이용하여, 복수의 비디오 데이터들이 인코딩된 비디오 MAF의 출력 방식을 포함하는 것을 특징으로 한다. 여기서, 출력방식은 각각의 비디오 데이터들의 시간 순서의 출력방식, 이벤트 단위의 출력방식, 스토리기반의 출력방식 등에 대해서 포함하나 이에 한정되지는 않는다.

트랙박스(trak)(5040)는 파일포맷내의 비디오 스트림, 오디오 스트림, 자막(subtitle) 스트림과 같은 미디어의 디코딩 정보를 포함하는 미디어박스(mdia)(5050)를 포함하는 것을 특징으로 한다. 미디어박스(5050)는 해당 미디어의 타임스탬프와 같은 디코딩 정보를 포함하는 구조를 특징으로 한다.

특히, 상기 도 4는 하나의 비디오 스트림에 대해서 관련된 복수의 멀티미디어 콘텐츠들 및 메타데이터를 인코딩하는 비디오 MAF의 경우를 도시한 것으로, 하나의 비디오 스트림 (video stream)(5070)에 대해서, N개의 오디오 스트림들(audio stream)(5080)과 N개의 자막들(subtitle stream)(5090)을 포함하는 것을 특징으로 한다. 이때 각각의 오디오 스트림들과 자막들은 관련된 언어를 각각 표현하는 콘텐츠이다. 예를 들면, 하나의 영화 콘텐츠에 대해서, 한국어, 영어, 일본어, 중국어 등등 다양한 언어들이 존재할 수 있다. 따라서, 하나의 비디오 스트림은 복수의 오디오 및 자막 스트림들을 가질 수 있다.

본 발명에서는 비디오 MAF 구조를 서술하는 단계에 있어서, 복수의 비디오 스트림들을 비디오 MAF에 포함하는 것을 특징으로 한다.

도 5는 복수의 비디오 스트림을 포함하는 비디오 MAF의 구성 요소 및 요소들의 구조를 도시한 것으로, 상기 도 4에서 설명된 하나의 비디오 스트림에 대한 비디오 MAF가 확장된 경우를 나타낸 것이다.

상기 도 5를 참조하면, 데이터 집합(5110) 및 (5120)은 각각 멀티미디어 데이터들과 메타데이터를 포함한다. 즉, 비디오 MAF는 복수의 비디오 데이터들의 집합들을 포함하는 것을 특징으로 한다. 예를 들어, 축구 경기의 슷장면에 대한 비디오



들은 다양하게 존재한다. 따라서, 각각의 숏장면 비디오에 대한 메타데이터와 프리뷰 이미지, 키프레임 이미지들을 조합하고, 다시 조합된 결과들을 비디오 MAF로 생성할 수 있다. 상기 도 5는 A개의 비디오들에 대해서 조합된 비디오 MAF를 나타낸 것이며, 이때 각각의 비디오들은 N개의 언어들과 M개의 키프레임 이미지들을 포함한다.

한편 상술한 본 발명의 설명에서는 구체적인 실시 예에 관해 설명하였으나, 여러 가지 변형이 본 발명의 범위에서 벗어나지 않고 실시될 수 있다. 따라서 발명의 범위는 설명된 실시 예에 의하여 정할 것이 아니고 특허청구범위에 의해 정하여져야 한다.

### 발명의 효과

이상에서 설명한 바와 같이, 본 발명은 비디오 데이터와 관계된 멀티미디어 콘텐츠 및 메타데이터(metadata)를 결합하는 새로운 비디오 멀티미디어 응용 포맷(Video Multimedia Application Format, Video MAF)을 구성하는 과정에 있어서, 효과적인 인덱싱(indexing) 및 검색을 위해 비디오 데이터 및 내용으로부터 획득한 오디오(audio)/이미지(image)/모션(motion) 특징정보를 비디오 MAF 파일의 메타데이터로 활용하며, 비디오 MAF의 구조에 따라 사용자가 MAF파일내에 포함된 비디오 및 메타데이터를 효과적으로 브라우징(browsing)할 수 있도록 하는 이점이 있다.

### (57) 청구의 범위

#### 청구항 1.

비디오와 관련된 멀티미디어 콘텐츠 및 메타데이터를 하나의 통합된 비디오 멀티미디어 응용 형식으로 인코딩/디코딩하는 방법으로서,

PDA, 디지털 카메라, 휴대폰, 게임기와 같은 휴대용 단말에서 비디오 단말을 획득/입력하거나 혹은 서비스 제공자나 콘텐츠 제작자로부터의 비디오 데이터를 획득/입력하는 단계와,

상기 입력된 비디오로부터 특징정보를 획득하는 비디오 프로세싱 및 저작 단계와,

상기 획득된 특징정보들을 이용하여 메타데이터를 생성하는 단계와,

상기 획득된 비디오 데이터와 메타데이터를 조합하여 비디오 MAF를 인코딩하는 단계와,

상기 인코딩된 비디오 MAF를 전송받아 이를 디코딩 하는 단계와,

상기 디코딩된 비디오 데이터와 메타데이터를 출력하는 단계

를 포함하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 2.

제 1 항에 있어서,

상기 입력된 비디오로부터 특징정보를 획득하는 비디오 프로세싱 및 저작 단계는,

프리뷰 이미지를 선택하는 과정과, 샷검출을 하는 과정과, 샷단위로 모션 특징을 추출하는 과정과, 샷단위로 오디오 특징을 추출하는 과정과, 샷단위로 키프레임을 추출하는 과정과, 키프레임 단위로 이미지 특징을 추출하는 과정을 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 3.

제 2 항에 있어서,



상기 모션 특징정보를 추출하는 과정에서 추출되는 모션 특징정보는,

MPEG-7 카메라 모션(camera motion)과 모션 액티비티(motion activity) 정보를 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 4.

제 2 항에 있어서,

상기 오디오 특징정보를 추출하는 과정에서 추출되는 오디오 특징정보는,

MPEG-7 저 레벨 오디오 서술자(low-level audio descriptor)들을 이용하여 추출되며, 오디오 특징으로 기본(basic), 기본 스펙트럴(basic spectral), 기본 신호(basic signal), 시간적 음색(temporal timbral), 스펙트럴 음색(spectral timbral), 스펙트럴 기저(spectral basis) 서술자들을 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 5.

제 2 항에 있어서,

상기 이미지 특징정보를 추출하는 과정에서, 추출되는 이미지 특징정보는,

MPEG-7 칼라 스트럭처(color structure), 칼라 레이아웃(color layout), 스케일러블 칼라(scalable color), 주요 칼라(dominant color), 에지 히스토그램(edge histogram), 균질 질감(homogeneous texture) 서술자들을 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 6.

제 1 항에 있어서,

상기 획득된 특징정보들을 가지고 메타데이터를 생성하는 단계에서의 비디오 MAF 메타데이터는,

미디어 자체에 대한 서술인 미디어 정보 및 생성 서술 정보와, 미디어 의미 서술 정보와, 미디어의 디지털 권리를 서술하는 디지털 권리 서술 정보를 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 7.

제 6 항에 있어서,

상기 미디어 정보 및 생성 서술 정보에서는,

미디어의 정보 생성 서술을 위해 필요한 메타데이터를 표현하기 위한 서술 구조를 XML 형식으로 구현하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 8.

제 6 항에 있어서,

상기 미디어 정보 및 생성 서술 정보에서는,

미디어의 시간적 구조를 표현하기 위한 서술 구조를 XML 형식으로 표현하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

## 청구항 9.

제 6 항에 있어서,

상기 미디어 의미 서술 정보에서는,

미디어 리소스의 의미를 표현하기 위한 서술 구조를 XML 형식으로 표현하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

## 청구항 10.

제 6 항에 있어서,

상기 디지털 권리 서술 정보에서는,

미디어의 접근 권한에 대한 서술 구조를 XML 형식으로 표현하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

## 청구항 11.

제 1 항에 있어서,

상기 획득된 비디오 데이터와 메타데이터를 조합하여 비디오 MAF를 인코딩하는 단계에서 인코딩된 비디오 MAF 데이터는,

하나의 비디오 데이터에 대해서 복수의 오디오 스트림과, 복수의 자막 스트림과, 복수의 키프레임과, 복수의 프리뷰 이미지를 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

## 청구항 12.

제 2 항에 있어서,

상기 비디오 MAF를 인코딩하는 단계에서 프리뷰 이미지와 키프레임에 대한 디코딩 정보는, 메타박스내의 메타데이터에 포함되고, 비디오 데이터, 오디오 스트림들, 자막 스트림들과 같은 미디어 데이터에 대한 디코딩 정보는, 미디어박스에 포함되는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

## 청구항 13.

제 2 항에 있어서,

상기 비디오 MAF를 인코딩하는 단계에서는, 각각 인코딩된 하나의 비디오 데이터와 관련된 메타데이터를 복수의 형태로 비디오 MAF를 인코딩하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 14.

제 1 항에 있어서,

상기 디코딩된 비디오 데이터와 메타데이터를 출력하는 단계에서, 복수의 비디오 데이터들의 출력 및 표현방식은, 각각의 비디오 데이터들의 시간 순서의 출력방식, 이벤트 단위의 출력방식, 스토리기반의 출력방식 중 어느 하나의 방식으로 구현되는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 방법.

#### 청구항 15.

비디오와 관련된 멀티미디어 콘텐츠 및 메타데이터를 하나의 통합된 비디오 멀티미디어 응용 형식으로 인코딩/디코딩하는 시스템으로서,

PDA, 디지털 카메라, 휴대폰, 게임기와 같은 휴대용 단말에서 비디오 단말을 획득/입력하거나 혹은 서비스 제공자나 콘텐츠 제작자로부터의 비디오 데이터를 획득/입력하는 비디오 획득/입력부와,

상기 비디오 획득/입력부를 통해 인가되는 비디오 데이터를 표현할 수 있는 특징 정보를 프로세싱하고 저장하는 비디오 프로세싱/저작부와,

상기 비디오 프로세싱/저작부를 통해 획득된 입력 비디오 데이터의 특징정보들을 이용하여 메타데이터를 생성하는 비디오 메타데이터 생성부와,

상기 입력된 비디오 데이터와 메타데이터를 조합하여 비디오 MAF를 인코딩하는 비디오 MAF 인코딩부와,

상기 인코딩된 비디오 MAF를 비디오 데이터와 메타데이터로 디코딩하는 비디오 MAF 디코딩부와,

상기 비디오 MAF 디코딩부로부터 디코딩된 비디오 데이터를 브라우징하고 디스플레이시키는 비디오 MAF 재생부와,

상기 메타데이터로부터 스토리 기반 비디오의 표현 및 비디오의 검색을 위한 특징정보를 해석하는 메타데이터 해석부

를 포함하는 비디오 MAF 파일 인코딩/디코딩 시스템.

#### 청구항 16.

제 15 항에 있어서,

상기 비디오 프로세싱/저작부는,

상기 입력된 비디오로부터 특징정보를 획득하는 단계에서, 프리뷰 이미지를 선택하는 과정과, 샷검출을 하는 과정과, 샷단위로 모션 특징을 추출하는 과정과, 샷단위로 오디오 특징을 추출하는 과정과, 샷단위로 키프레임을 추출하는 과정과, 키프레임 단위로 이미지 특징을 추출하는 과정을 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

#### 청구항 17.

제 16 항에 있어서,

상기 모션 특징정보를 추출하는 과정에서 추출되는 모션 특징정보는,

MPEG-7 카메라 모션(camera motion)과 모션 액티비티(motion activity) 정보를 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

## 청구항 18.

제 16 항에 있어서,

상기 오디오 특징정보를 추출하는 과정에서 추출되는 오디오 특징정보는,

MPEG-7 저 레벨 오디오 서술자(low-level audio descriptor)들을 이용하여 추출되며, 오디오 특징으로 기본(basic), 기본 스펙트럴(basic spectral), 기본 신호(basic signal), 시간적 음색(temporal timbral), 스펙트럴 음색(spectral timbral), 스펙트럴 기저(spectral basis) 서술자들을 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

## 청구항 19.

제 16 항에 있어서,

상기 이미지 특징정보를 추출하는 과정에서, 추출되는 이미지 특징정보는,

MPEG-7 칼라 스트럭처(color structure), 칼라 레이아웃(color layout), 스케일러블 칼라(scalable color), 주요 칼라(dominant color), 에지 히스토그램(edge histogram), 균질 질감(homogeneous texture) 서술자들을 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

## 청구항 20.

제 15 항에 있어서,

상기 비디오 메타데이터 생성부에서 생성되는 비디오 MAF 메타데이터는,

미디어 자체에 대한 서술인 미디어 정보 및 생성 서술 정보와, 미디어 의미 서술 정보와, 미디어의 디지털 권리를 서술하는 디지털 권리 서술 정보를 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

## 청구항 21.

제 20 항에 있어서,

상기 미디어 정보 및 생성 서술 정보에서는,

미디어의 정보 생성 서술을 위해 필요한 메타데이터를 표현하기 위한 서술 구조를 XML 형식으로 구현하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

## 청구항 22.

제 20 항에 있어서,

상기 미디어 정보 및 생성 서술 정보에서는,

미디어의 시간적 구조를 표현하기 위한 서술 구조를 XML 형식으로 표현하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

### 청구항 23.

제 20 항에 있어서,

상기 미디어 의미 서술 정보에서는,

미디어 리소스의 의미를 표현하기 위한 서술 구조를 XML 형식으로 표현하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

### 청구항 24.

제 20 항에 있어서,

상기 디지털 권리 서술 정보에서는,

미디어의 접근 권한에 대한 서술 구조를 XML 형식으로 표현하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

### 청구항 25.

제 15 항에 있어서,

상기 비디오 MAF 인코딩부에서 인코딩되는 비디오 MAF 데이터는,

하나의 비디오 데이터에 대해서 복수의 오디오 스트림과, 복수의 자막 스트림과, 복수의 키프레임과, 복수의 프리뷰 이미지를 포함하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

### 청구항 26.

제 16 항에 있어서,

상기 비디오 MAF를 인코딩하는 단계에서 프리뷰 이미지와 키프레임에 대한 디코딩 정보는, 메타박스내의 메타데이터에 포함되고, 비디오 데이터, 오디오 스트림들, 자막 스트림들과 같은 미디어 데이터에 대한 디코딩 정보는, 미디어박스에 포함되는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

### 청구항 27.

제 16 항에 있어서,

상기 비디오 MAF를 인코딩하는 단계에서는, 각각 인코딩된 하나의 비디오 데이터와 관련된 메타데이터를 복수의 형태로 비디오 MAF를 인코딩하는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

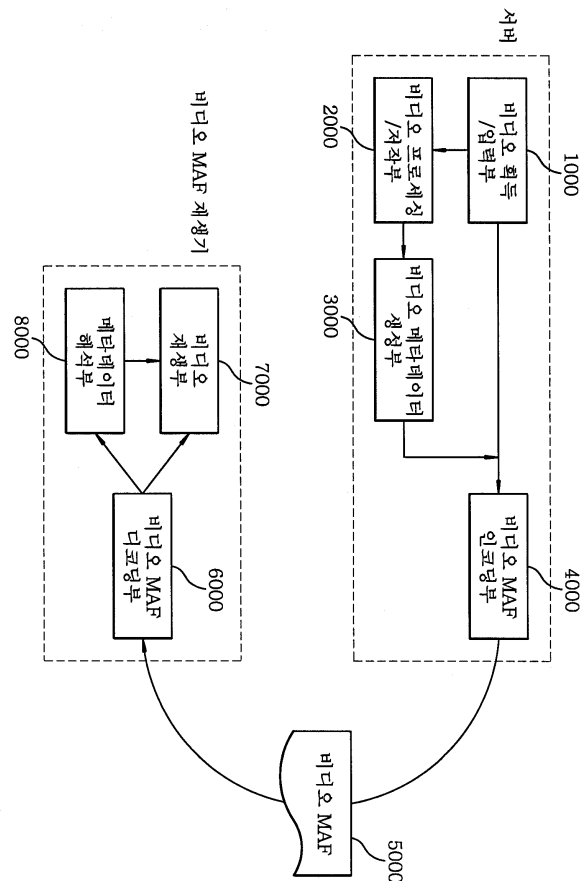
## 청구항 28.

제 15 항에 있어서,

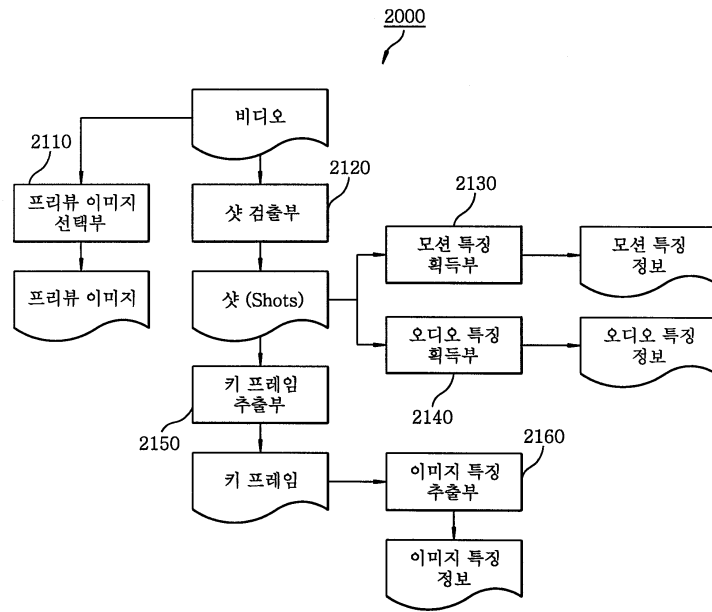
상기 비디오 MAF 디코딩부에서 비디오 데이터와 메타데이터를 출력하는 단계에서, 복수의 비디오 데이터들의 출력 및 표현방식은, 각각의 비디오 데이터들의 시간 순서의 출력방식, 이벤트 단위의 출력방식, 스토리기반의 출력방식 중 어느 하나의 방식으로 구현되는 것을 특징으로 하는 비디오 MAF 파일 인코딩/디코딩 시스템.

도면

도면1

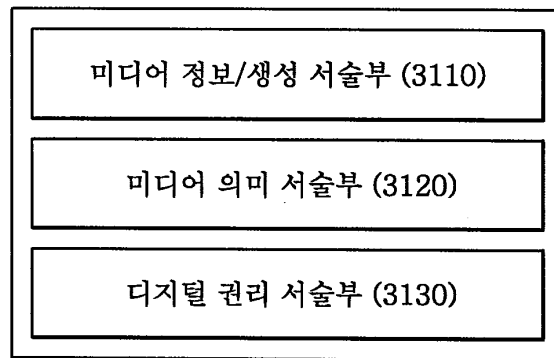


도면2



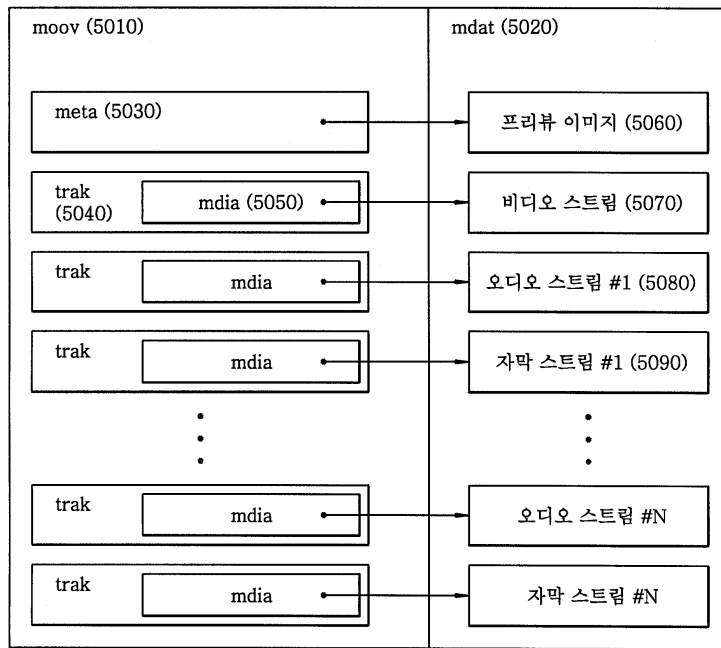
도면3

비디오 MAF 메타데이터 (3100)





도면4



도면5

