

메타데이터간 동기화에 대한 연구

전유성*, 최혁승*, 최명순*, 김두일*, 김영윤*
*삼성전자

e-mail : yuseong74@samsung.com

A study of synchronizing between metadata

YUSEONG JEON*, HyokSung Choi*, MYOUNGSOON CHOI*, Du-il Kim*, Youngyoon Kim*
*Samsung Electronics

요 약

본 논문은 자산의 메타데이터 편집 시, 자산의 백 워드 링크(backward-link) 파일을 이용하여 동일한 자산(asset)에 링크된 파일들의 메타데이터간 동기화 방법에 관한 것으로서, 본 논문에 따른 방법은, 편집된 메타데이터를 갖는 제 1 자산의 백 워드 링크 정보 파일에 상기 제 1 자산과 동일한 제 2 자산이 존재하는지 체크하고, 백 워드 링크 정보 파일에 제 2 자산이 적어도 하나 존재하면, 제 2 자산의 메타데이터를 제 1 자산의 편집된 메타데이터에 동기화 시키도록 함으로써, 동일한 자산에 링크되어 있는 파일들의 메타데이터를 용이하게 동기 시킬 수 있다.

1. 서론

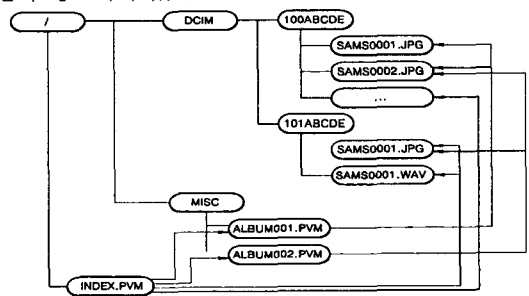
본 논문은 메타데이터(Metadata)간 동기화에 관한 것으로, 보다 구체적으로는, 자산(asset)에 링크된 파일들의 메타데이터간 동기화 방법에 관한 것이다.

디지털 데이터 콘텐츠를 다루는 정보기기 제품 또는 가전 제품들 간에 콘텐츠를 공유하고 교환하기 위한 다양한 기술들이 제시되고 있다. OSTA (Optical Storage Technology Association, 이하 OSTA 라 칭함)에 의해 표준화가 진행되고 있는 MPV (MusicPhotoVideo, 이하 MPV 라고 칭함)는 상기 다양한 기술들 중 하나이다. MPV 는 XML (eXtended Markup Language, 이하 XML 이라 칭함)을 기반으로 콘텐츠를 메타데이터화하여 제품들간에 콘텐츠의 공유 및 콘텐츠의 교환을 가능하게 하는 기술이다.

MPV 등 XML 기반 메타데이터 자산은 상위 인덱스 또는 앨범을 이용하여 링크된다. 이러한 링크는 XML 로 기술되며, 링크의 참조 방향은 부모 자산에서 자식 자산으로 향하는 단방향 참조 링크 (one-way referenced link)를 갖는다.

즉, 그림 1 에 도시된 MPV 링크 예를 참조하면, /DCIM/100ABCDE/SAMS001.JPG 자산은 album001.pvm 파일에 링크되어 있고, /DCIM/100ABCDE/SAMS002.JPG 자산은 album001.pvm 파일 및 album002.pvm 파일에 링크되어 있다. /DCIM/101ABCDE/SAMS0001.JPG 자산은

album002.pvm 파일과 index.pvm 파일에 링크되어 있고, /DCIM/101ABCDE/SAMS001.WAV 자산은 index.pvm 파일에 링크되어 있다.



[그림 1] MPV 링크 예시도

따라서, 그림 1 의 index.pvm 파일의 소스 코드는 그림 2 에 도시된 바와 같이 XML 로 기술될 수 있다. 그림 2 를 참조하면, index.pvm 파일에서 /DCIM/101ABCDE/SAMS001.JPG 자산의 mpv:id 는 "MariahPhoto"로 기록되고, /DCIM/101ABCDE/SAMS0001.WAV 자산의 mpv:id 는 "MariahSong00"으로 기록되고, StillWithAudio 자산의 mpv:id 는 "MariahAlbum00"으로 기록된다.

그림 1 의 album002.pvm 파일의 소스 코드는 그림 3 에 도시된 바와 같이 XML 로 기술될 수 있다. 그림

3 을 참조하면, album002.pvm 파일에서 /DCIM/101ABCDE/SAMS001.JPG 자산의 mpv:id 는 "MariahPhoto00"로 기록되고, /DCIM/100ABCDE/SAMS002.JPG 자산의 mpv:id 는 "MandyMoorePhoto00"으로 기록된다.

```
<!-- index.pvm -->
<file:Manifest>
  <mpv:AssetList>
    <mpv:Still mpv:id="MariahPhoto00">
      <LastURL>DCIM/101ABCDE/SAMS0001.JPG</LastURL>
      <nmf:Metadata>
        <Properties xmlns="http://purl.org/dc/elements/1.1/">
          <title>Beautiful Mariah</title>
          <creator>Pan of Mariah</creator>
        </Properties>
        <Properties xmlns="http://purl.org/dc/terms/">
          <created>2004-04-20T05:00:03Z</created>
        </Properties>
      </nmf:Metadata>
    </mpv:Still>
    <mpv:Audio mpv:id="MariahSong00">
      <LastURL>DCIM/101ABCDE/SAMS0001.WAV</LastURL>
    </mpv:Audio>
    <mpv:StillWithAudio mpv:id="MariahAlbum00">
      <mpv:StillRef mpv:idRef="MariahPhoto00"/>
      <mpv:AudioRef mpv:idRef="MariahSong00"/>
    </mpv:StillWithAudio>
    ...
  </mpv:AssetList>
</file:Manifest>
```

[그림 2] index.pvm 파일의 소스 코드 예

```
<!-- album002.pvm -->
<file:Manifest>
  <mpv:AssetList>
    <mpv:Still mpv:id="MariahPhoto00">
      <LastURL>DCIM/101ABCDE/SAMS0001.JPG</LastURL>
      <nmf:Metadata>
        <Properties xmlns="http://purl.org/dc/elements/1.1/">
          <title>Beautiful Mariah</title>
          <creator>Pan of Mariah</creator>
        </Properties>
        <Properties xmlns="http://purl.org/dc/terms/">
          <created>2004-04-20T05:00:03Z</created>
        </Properties>
      </nmf:Metadata>
    </mpv:Still>
    <mpv:Still mpv:id="MandyMoorePhoto00">
      <LastURL>DCIM/100ABCDE/SAMS0002.JPG</LastURL>
      <nmf:Metadata>
        <Properties xmlns="http://purl.org/dc/elements/1.1/">
          <title>Cute Mandy</title>
          <creator>Pan of Mandy</creator>
        </Properties>
        <Properties xmlns="http://purl.org/dc/terms/">
          <created>2004-04-21T05:00:03Z</created>
        </Properties>
      </nmf:Metadata>
    </mpv:Still>
    ...
  </mpv:AssetList>
</file:Manifest>
```

[그림 3] album002.pvm 파일의 소스 코드 예

그림 2 및 그림 3 의 소스 코드 예를 참조하면, DCIM/101ABCDE/SAMS001.JPG 자산의 mpv:id 를 index.pvm 파일과 album002.pvm 파일 모두 "MariahPhoto00"으로 기록하고 있다. 이 때, index.pvm 파일에서 DCIM/101ABCDE/SAMS001.JPG 자산의 메타데이터를 편집할 경우에, album002.pvm 파일의 DCIM/101ABCDE/SAMS001.JPG 자산의 메타데이터와 다른 내용을 갖게 된다.

즉, index.pvm 파일에서 mpv:id 가 "MariahPhoto00"으로 참조되고 있는 DCIM/101ABCDE/SAMS001.JPG 자산의 메타데이터인 <title>을 "Beautiful Mariah"에서 "Mariah in Korea"로 변경했을 때, album002.pvm 파일에서 mpv:id 가 MariahPhoto00"으로 참조되고 있는 DCIM/101ABCDE/SAMS001.JPG 자산의 메타데이터의 <title>은 여전히 "Beautiful Mariah"이다. 따라서,

DCIM/101ABCDE/SAMS001.JPG 자산에 링크되어 있는 index.pvm 파일과 album002.pvm 파일들의 DCIM/101ABCDE/SAMS001.JPG 자산에 대한 메타데이터간에 모순이 발생하게 된다.

이러한 모순은 링크의 참조 방향이 단방향 참조 링크이기 때문이다. 따라서, 메타데이터 편집 시, 동일한 자산을 참조하는 서로 다른 파일들의 메타데이터간의 모순을 해결하기 위하여, 단방향 링크들을 모두 검색하여야 한다.

2. 목적

본 논문이 이루고자 하는 기술적 과제는 자산의 메타데이터 편집 시, 자산의 백 워드 링크 파일을 이용하여 동일한 자산에 링크된 파일들의 메타데이터간 효율적인 동기화 방법을 제시한다.

3. 백워드 링크 정보

MPV 파일은 XML 또는 텍스트 형식으로 기술된 파일들로, *.pvm 파일, *.mum 파일로 표시된다. *.pvm, *.mum 파일은 인덱스와 앨범 파일이다.

본 논문에서 제안하고자 하는 *.mtd 파일(이하 mtd 파일이라 함)은 각 자산의 백워드 링크 정보를 정의한 백워드 링크 정보 파일이다. 즉, mtd 파일은 각 자산을 참조 또는 링크하는 MPV 파일의 정보를 정의한 파일로서, 그림 4 에 도시된 바와 XML 을 기반으로 기술될 수 있다.

```
<!-- DCIM/101ABCDE/SAMS0001.mtd -->
<?xml version="1.0">
<psfs:RevReference xmlns:psfs="http://psfs.org/">
  <psfs:LastURL>../MISC/album2.pvm?mpv_id=MariahPhoto00</psfs:LastURL>
  <psfs:LastURL>../index.pvm?mpv_id=MariahPhoto00</psfs:LastURL>
  <psfs:LastURL>../index.pvm?mpv_id=MariahSong00</psfs:LastURL>
  <psfs:LastURL>../index.pvm?mpv_id=MariahAlbum00</psfs:LastURL>
</psfs:RevReference>
```

[그림 4] mtd 파일의 소스 코드 예

그림 4 에 도시된 mtd 파일은 그림 1 의 MPV 링크 예에서 DCIM/101ABCDE/SAMS0001 자산을 참조하는 MPV 파일들의 정보를 정의한 경우이다. 따라서, 그림 4 는 파일 이름이 "DCIM/101ABCDE/SAMS0001"이고 확장자가 JPG, WAV, pvm 인 복수개의 자산에 링크된 MPV 파일 정보를 정의한 mtd 파일이라 할 수 있다.

그림 4 에 도시된 mtd 파일은 확장자가 "JPG"인 album.pvm 파일 정보와 index.pvm 파일 정보, 확장자가 "WAV"인 index.pvm 파일 정보, 및 확장자가 "pvm"인 index.pvm 파일 정보를 포함한다. 따라서 그림 4 에서 LastURL 이 ".../MISC/album2.pvm?mpv_id=MariahPhoto00" 파일과 LastURL 이 ".../index.pvm?mpv_id=Mariah Photo00" 파일은 동일한 자산에 링크된 파일이다.

4. 메타 데이터 편집

메타데이터 편집 모듈은 사용자 인터페이스 모듈을 통해 입력되는 사용자 요구에 따라 MPV 파일에 포함되어 있는 MPV 파일의 메타데이터의 생성 또는 복사, 제거, 변경, 및 이동 등의 편집을 수행한다.

예를 들어, index.pvm 파일이 그림 2에 도시된 바와 같이 XML을 기반으로 기술되고, 사용자가 mpv: id 가 "MariahPhoto00"으로 참조되고 있는 DCIM/101ABCDE/SAMS001.JPG 자산의 메타데이터인 타이틀을 "Beautiful Mariah"에서 "Mariah in Korea"로 변경할 것을 요구하면, 메타데이터 편집 모듈은 MPV 파일에서 index.pvm 파일을 검색하고, XML을 기반으로 기술되어 있는 index.pvm 파일의 자산 리스트에서 LastURL이 DCIM/101ABCDE/SAMS001.JPG인 자산을 검색하고, 검색된 자산의 타이틀을 "Beautiful Mariah"에서 "Mariah in Korea"로 변경시킨다.

사용자에 의해 요구된 편집이 완료되면, 메타데이터 편집 모듈은 메타데이터 동기화 장치로 편집된 자산에 대한 정보를 제공한다. 상기 편집된 자산에 대한 정보는 편집된 자산과 참조되는 mpv:id 정보 및 상기 편집된 자산이 링크된 파일 정보를 포함할 수 있다. 예를 들어, 편집된 자산이 DCIM/101ABCDE/SAMS001.JPG 자산이고, mpv:id는 MariahPhoto00이고, 상기 편집된 자산에 링크된 파일이 index.pvm 파일이라는 정보를 상기 편집된 자산에 대한 정보로 제공할 수 있다.

5. 메타데이터간 동기화

메타데이터간 동기화 모듈은 메타데이터 편집 모듈로부터 상기 편집된 자산에 대한 정보가 수신되면, MPV 파일에 편집이 수행된 자산의 mtd 파일을 검색하고, 분석하여 mtd 파일에서 상기 편집된 자산과 동일한 자산을 검출하고, 검출된 자산의 메타데이터를 상기 편집된 자산의 메타데이터에 동기시킨다.

상기 mtd 파일에서 검출되는 상기 편집된 자산과 동일한 자산은 실질적으로 동일한 자산이나 서로 다른 파일에 의해 참조 또는 링크되는 자산이다. 만약, 그림 4의 /MISC/album2.pvm 파일에 링크된 DCIM/101ABCDE/SAMS001.JPG 자산의 메타데이터가 편집되고, 편집된 자산의 mtd 파일이 그림 4에 도시된 바와 같이 XML로 기술되면, mtd 파일에서 검출되는 상기 편집된 자산과 동일한 자산은 LastURL이 /MISC/album2.pvm?mpv_id=MariahPhoto00인 자산과 LastURL이 /index.pvm?mpv_id=MariahPhoto00인 자산이다. 그러나, /MISC/album2.pvm에 링크된 DCIM/101ABCDE/SAMS001.JPG 자산의 메타데이터는 편집되었으므로, mtd 파일에서 검출되는 상기 편집된 자산과 동일한 자산은 /index.pvm에 링크된 DCIM/101ABCDE/SAMS001.JPG 자산이 된다. 이들 자산은 실질적으로 동일한 자산이나 상기 편집에 의해 상이한 메타데이터를 갖는다. 따라서, 메타데이터간 동기화 장치는 상기 mtd 파일에서 검출된 자산의 메타데이터를 상기 편집된 자산의 메타데이터에 동기시킨다. 이를 위하여 메타데이터간 동기화 모듈은

mtd 파일 검색 모듈, 자산 검출 모듈, 및 동기화 처리 모듈을 포함한다.

mtd 파일 검색 모듈은 메타데이터 편집 모듈로부터 편집된 자산에 대한 정보가 수신되면, 상기 수신된 편집된 자산에 대한 정보를 이용하여 MPV 파일에서 편집된 자산의 mtd 파일을 검색한다. mtd 파일은 대응되는 자산의 파일이름과 동일하고 확장자가 mtd인 파일이다. 따라서, 수신된 편집된 자산에 대한 정보 중에서 편집된 자산의 파일이름을 이용하여 mtd 파일을 검색한다. 예를 들어, 수신된 편집된 자산의 파일이름이 /DCIM/101ABCDE/SAMS001이면, /DCIM/101ABCDE/SAMS001.mtd 파일을 대응되는 mtd 파일로 검색한다. 검색된 mtd 파일은 자산 검출 모듈로 전송한다.

자산 검출 모듈은 입력된 mtd 파일에서 편집된 자산과 동일한 자산을 검출한다. 즉, 자산 검출 모듈은 입력된 mtd 파일을 분석하여, mtd 파일에 포함되어 있는 항목(그림 4의 경우에 LastURL 항목)이 편집된 자산과 동일한 자산인지, 만약 동일한 자산이면 편집된 자산과 동기화를 위해 이미 체크된 자산인지를 판단하고, 편집된 자산과 동일하면서 이미 체크되지 않은 자산을 검출한다. 상기 편집된 자산과 동일한 자산인지는 해당 항목의 mpv_id를 분석하여 판단할 수 있다. 이 때, 분석 기준은 메타데이터 편집 모듈로부터 제공되는 편집된 자산에 대한 정보에 포함되어 있는 mpv_id 정보이다. 상기 동기화를 위해 이미 체크된 자산인지는 동일한 자산으로 검출될 때마다 마킹을 수행하고, 상기 마킹된 정보를 이용하여 판단하거나 자산 검출 모듈 또는 미도시된 저장 수단에 이미 체크된 자산에 대한 정보를 관리하고, 이 관리 정보를 이용하여 판단하도록 구현할 수 있다. 검출된 자산이 복수 개이면, 자산 검출 모듈은 자산단위로 검출된 자산을 동기화 처리 모듈로 전송할 수 있다.

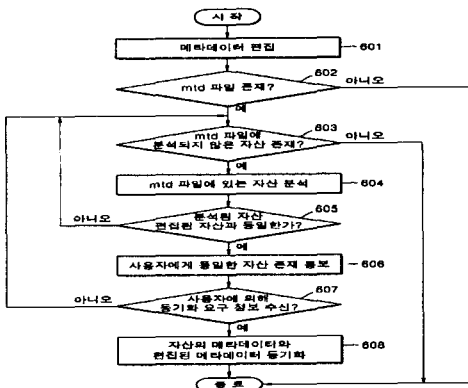
동기화 처리 모듈은 자산 검출 모듈로부터 전송되는 검출된 자산에 대한 정보를 사용자 인터페이스 모듈로 제공한다. 사용자 인터페이스 모듈로부터 사용자에 의한 동기화 요구 정보를 수신하면, 동기화 처리 모듈은 mtd 파일에 정의되어 있는 LastURL을 이용하여 MPV 파일에서 해당되는 자산을 읽고, 읽혀진 자산의 메타데이터를 편집된 메타데이터에 동기화시킨다. 또한, 편집된 메타데이터를 갖는 자산과 자산 검출 모듈에서 검출된 자산의 메타데이터를 MPV 파일로부터 모두 읽어 동기화 시킬 수 있다. 이를 위해 동기화 처리 모듈은 메타데이터 편집 모듈로부터 전송되는 편집된 자산에 대한 정보에 포함되어 있는 편집된 자산의 LastURL 정보를 이용할 수 있다. 그리고, 메타데이터 편집 모듈로부터 전송되는 편집된 자산에 대한 정보에 포함되어 있는 편집된 메타데이터 정보를 토대로 MPV 파일에서 읽혀진 자산의 메타데이터와 편집된 자산의 메타데이터를 동기화 시키도록 구현될 수도 있다.

동기화 처리 모듈은 자산 검출 모듈로부터 자산 단위로 검출된 자산 정보가 제공되면, 해당되는 자산에 대한 동기화가 완료된 후, 자산 검출 모듈로 그 다음

검출 자산에 대한 정보를 요구할 수 있고, 자산 검출 모듈로부터 더 이상 검출된 자산이 제공되지 않으면, 편집된 메타데이터를 갖는 자산에 의한 메타데이터간 동기화 작업을 종료한다.

사용자 인터페이스 모듈은 사용자에게 의해 편집 요구 정보가 입력되면, 메타데이터 편집 모듈로 상기 편집 요구 정보를 전송한다. 또한, 메타데이터간 동기화 장치로부터 검출된 자산에 대한 동기화를 사용자가 원하는지 확인하기 위한 정보가 수신되면, 사용자 인터페이스 모듈은 사용자에게 상기 확인 정보를 통보한다. 사용자로부터 동기화 요구 정보가 수신되면, 사용자 인터페이스 모듈은 이를 메타데이터간 동기화 장치의 동기화 처리 모듈로 제공한다.

6. 메타데이터 동기화 방법



[그림 5] 메타데이터간 동기화 방법에 대한 동작 흐름도

본 논문에 따른 메타데이터 간 동기화 방법에 대해서 그림 5 을 참조하여 설명하면 다음과 같다. 메타데이터가 편집되면, 메타데이터가 편집된 자산의 mtd 파일이 MPV 파일에 존재하는지 체크한다.

mtd 파일이 MPV 파일에 존재하면, mtd 파일에 메타데이터 동기화를 위해 분석되지 않은 자산이 존재하는지 체크한다. 즉, mtd 파일에 정의되어 있는 자산 중 편집된 자산과 동일한 자산인지 여부와 동일한 자산이면서 이미 동기화를 위해 체크된 자산인지 여부를 판단하기 위해 체크되지 않은 자산이 존재하는지 체크한다. 자산이 분석되지 않은 자산인지는 동기화를 위한 자산을 검출할 때 검출된 자산에 마킹된 정보를 이용하거나 동기화를 위한 자산을 검출할 때 저장된 관리 정보를 이용하여 판단될 수 있다.

mtd 파일에 분석되지 않은 자산이 존재하면, mtd 파일에 있는 분석되지 않은 자산을 분석한다. 예를 들어, 상기 분석되지 않은 자산의 mpv_id 를 분석한다. 분석되지 않은 자산이 복수 개이면, mtd 파일에서 정의된 순서에 기초하여 분석되지 않은 자산을 분석할 수 있다. 분석된 정보를 이용하여 mtd 파일에서 검출된 자산이 편집된 자산과 동일한 자산인지 판단한다. 분석된 자산의 mpv_id 와 편집된 자산이 mpv_id 가 동일하

면, 검출된 자산은 편집된 자산과 동일한 자산으로 판단된다. 그러나, 분석된 자산의 mpv_id 와 편집된 자산의 mpv_id 가 상이하면, 검출된 자산은 편집된 자산과 상이한 자산으로 판단된다.

만약, 검출된 자산이 편집된 자산과 동일한 자산이 아니면, mtd 파일에 메타데이터 동기화를 위해 분석되지 않은 자산이 존재하는지 체크하는 단계로 리턴 된다. 그러나, 검출된 자산이 편집된 자산과 동일하면, 사용자에게 동일한 자산이 존재하는 것을 통보한다.

사용자에게 의해 동기화 요구 정보가 수신되면, 검출된 자산의 메타데이터와 편집된 자산의 메타데이터를 동기화 한다. 그러나, 사용자에게 의해 동기화 요구 정보가 수신되지 않으면, 마찬가지로 mtd 파일에 메타데이터 동기화를 위해 분석되지 않은 자산이 존재하는지 체크하는 단계로 리턴 된다.

7. 결론

본 논문은, 메타데이터가 편집되면, 편집된 메타데이터를 갖는 자산의 백워드 링크 정보 파일을 참조하여 편집된 자산과 동일한 자산을 검출하고, 검출된 자산의 메타데이터를 편집된 자산의 메타데이터에 동기 시킴으로써, 동일한 자산에 링크되어 있는 파일들의 메타데이터를 고속으로 효율적으로 동기 시킬 수 있고, 동일한 자산에 링크되어 있는 파일들의 메타데이터간에 모순이 발생하는 것을 방지할 수 있다.

참고문헌

- [1] "Dublin Core Metadata Initiative", a Simple Content Description Model for Electronic Resources. Available at <http://purl.org/DC/>
- [2] "Dublin Core Normalized Metadata Format Profile Specification 1.0"; OSTA, 2002. Available at <http://www.osta.org/mpv/>
- [3] "XML Manifest Specification 1.0"; OSTA, 2002., Available at <http://www.osta.org/mpv/>
- [4] "MPV - Basic Profile Specification", OSTA, 2002, Available at <http://www.osta.org/mpv/>
- [5] "MPV Core Specification 1.0"; OSTA, 2002., Available at <http://www.osta.org/mpv/>
- [6] "MPV Presentation Profile Specification 1.0"; OSTA, 2002., Available at <http://www.osta.org/mpv/>
- [7] "Normalized Metadata Format Specification 1.0"; OSTA, 2002., Available at <http://www.osta.org/mpv/>
- [8] "Uniform Resource Identifiers (URI): Generic Syntax", T. Berners-Lee, R. Fielding, L. Masinter, August 1998. Note that RFC 2396 updates [RFC1738] and [RFC1808].
- [9] Yergeau, F., "UTF-8, a transformation format of ISO 10646", RFC 2279, January 1998.
- [10] "Extensible Markup Language (XML) 1.0" T. Bray, J. Paoli and C.M. Sperberg-McQueen. W3C Recommendation 10 February 1998, Available at <http://www.w3.org/TR/REC-xml>
- [11] "Namespaces in XML", Tim Bray, Dave Hollander, Andrew Layman. W3C Recommendation 14 January 1999, Available at <http://www.w3.org/TR/REC-xml-names>