Interoperable Summary Description Model Using Dublin Core

Do-Nyun Kim and Young-Won Song
LG Electronics, Information Technology Lab.,
16 Woomyeon-Dong, Secho-Gu, Seoul 137-724, Korea
dnkim@LG-Elite.com and ysong@LG-Elite.com

Abstract

This paper proposes an interoperable metadata model generating summary description for multimedia content using Dublin Core (DC). The motive is based on the fundamental concepts such as (1) Description information about the multimedia content is essential in multimedia content access, search and retrieval process (2) the existing metadata are too complicated to use in applications such as e-cataloguing and browsing of e-commerce. As an approach to solve the problem, summary description that may be optimally minimal descriptive elements set derived from existing metadata schemes (full descriptions) is described in this paper. The proposed summary description generator model is achieved using thesaurus approach built on the basis of DC and any existing various metadata schemes.
Keywords: summary description, interoperable, agent, thesaurus, Dublin Core, cataloguing

1 Introduction

Recently, the demand and appetite of multimedia contents consumer or end-users is increasing rapidly in virtue of the development and deployment of internet technologies. As this volume of demand increases, the importance and usefulness of metadata is emphasized and drives many metadata frameworks to be developed in order to build an infrastructure for the easy and interoperable delivery and consumption of multimedia content. However, unfortunately, there is no flexible, uniform, interoperable or integrated metadata framework that can be applied and reaching a worldwide consensus in the multimedia consumption value chain exists. In addition, some description metadata frameworks or schemes are too complex to be understood and implemented with reasonable cost. As a result, metadata users or sometimes, consumers tend to be overwhelmed by their complexities. As an approach to solve the current problems, summary description using DC that is considered as an essential and optimally minimal description set is proposed in this paper. Basically, this paper proposes to employ thesaurus approach in building the interoperable summary description model of DC element set derived from existing a variety of full multimedia description sets[1][2]. Section 2 addresses basic approaches to build the proposed model. Based on the basic approaches, section 3 describes the proposed interoperable summary description model in detail. Section 4 shows an example of summary description composed of DC element set derived from MPEG (Moving Picture Experts Group)-7 MDS (Multimedia Description Schemes) description set. In section 5, the validity of the proposed model is concluded regarding the practical e-commerce applications.

2 Basic approaches

In order to establish a summary description that is interoperable, optimally minimal, dynamic and maintainable, three fundamental approaches are employed and explained in this section as follow.

● Multiple-to-one mapping mechanism

The proposed model employs multiple-to-one mechanism that extracts DC elements set (one) from a variety of existing description schemes (multiple) containing full description information about multimedia contents. Since the purpose of this multiple-to-one mechanism is to obtain a DC element set from a variety of metadata element sets like a 'metadata filter', it is expected that the interoperability can be achieved. In a practical viewpoint, the 'DC metadata filter' is very useful in the e-commerce applications such as electronic cataloguing and browsing. One of the challenging tasks is to implement such mechanism that is able to generate *a* summary description automatically through summarization mechanism. One of the

reasonable methods is to use thesaurus approach.

● System-selected vs. User-selected elements

In the multimedia content search, cataloguing, browsing and retrieval, the optimized and minimal set of descriptive elements plays a key role in the process. However, since it is very difficult to define the optimally minimized set and to provide an interoperable metadata framework, the existing metadata frameworks may contain enormous number of elements as possible in order to overcome the practical applicability. As a fundamental approach, the DC is used. Also, in DC society, in order to provide more general metadata scheme, hierarchical qualifiers are recommended and used. However, still, there is an interoperability issue remained with the existing metadata frameworks such as MPEG-7 MDS, EPICS and etc[5][6][7]. This paper proposes to employ DC as a summary description. If end-users wish to investigate the more detail metadata, the full description can be accessed as a hierarchical manner. Since this DC summary description is extracted from the full description set, it is called as system-selected elements in this paper. In order to provide a flexible model with the end-user like user interaction, user defined elements (or user-selected elements) can be declared in addition to DC in the model such as event reporting that is defined in the MPEG-21 multimedia framework. In the viewpoint of interoperability, system-selected elements are efficient. User-selected elements are, however, more useful in the viewpoint of user interaction. Those two approaches are employed by summarization process described in section 3.

● Dynamic generation by users' demand

Summary description is a sort of filtered version of it's the original full description by nature. It is required to be generated by users' demand on time. This requirement is expected to prevent users from wasting efforts and cost. In terms of storage efficiency, summary description is not necessary to be stored permanently, because it is related to a specific application. Hence, the model is required that summary description interacts with the full description in a dynamic manner in response to end-user's demand. In such case, the summary description maintenance issue such as summary metadata consistency with the full description and version control in terms of metadata updatability is required to be solved. As a plausible approach to the dynamic summary

description generation, software agent technique is proposed. In this situation, summarization process takes advantage of agents such as automatic notification and automatic change tracking.

Figure 1 shows the relationship among resource, summary description, (full) description, metadata of summary description (or Admin Core, Administrative Container Metadata, A-Core, or DescriptionMetadata in MDS), and metadata of (full) description (or A-Core). The full description describes resource such as audio, video clip, image, textual document or software. The metadata of summary or full description such as A-Core describes simply the summary or full description (for example, who, where, when the description is prepared). Hence, the summary description is different from metadata of description such as A-Core [3][4].
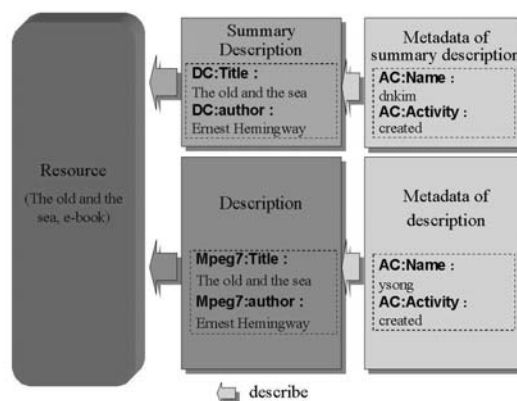
Figure 1. The relationship among resource, description, summary description, metadata of summary description and metadata of description

3 The proposed summary description model

This section describes the summary description model built based on the fundamental approaches addressed in section 2.

● Using Dublin Core

As a summary description, DC elements set is used. In the initial access, search and browsing procedure about the metadata, summary

DCPAPERS

Proc. Int'l. Conf. on Dublin Core and Metadata Applications 2001

description can play a major role in the perspective of end-users (In general they want simple, fast and automatic process). If users are interested in the corresponding content, users may request more detail description based on full description. Therefore, summary description needs not to be complicated and DC that has a wide consensus among e-commerce communities is sufficient to describe essential minimal description.

In case any search fails using DC or user-selected elements, or even if search results hit end-users' interests but they want to find detail description data, end-users may access the full description for more detail in a hierarchical manner as shown in Figure 2.
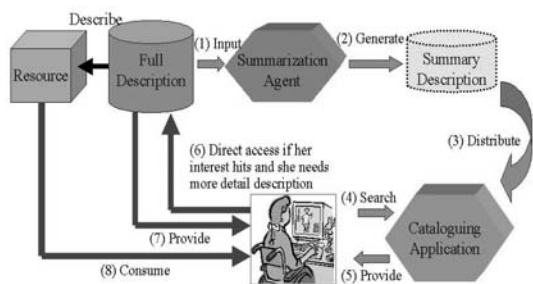

Figure 2. The sequence of actions using summary and full description in a hierarchical manner

● Using thesaurus

As described in section 2, the summarization process reflects many-to-one mechanism. In order to reduce ambiguous translation, the agent embedded in the summarization process utilizes thesaurus or dictionary – sometimes called controlled vocabulary. Thesaurus includes synonyms and antonyms. When elements of multiple and diverse descriptions are translated into those of summary description that is considered to be matched with DC elements set, the summarization process needs synonyms, antonyms and acronyms. Well-organized thesaurus has a significant influence on the performance of summarization agents.

Figure 3 describes the summarization process including thesaurus regarding to existing full description[5][6][7].
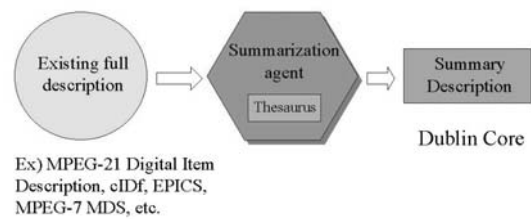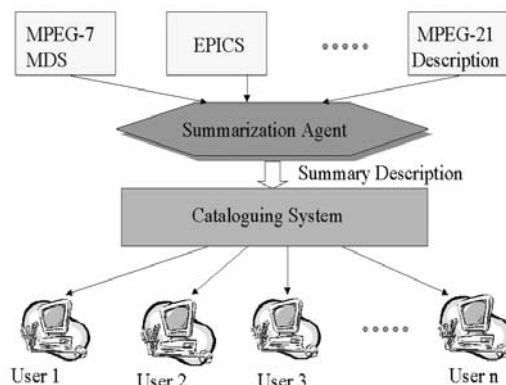

Figure 3. The input and output of the summarization process

Figure 4 shows a generic flow diagram of the summarization process.
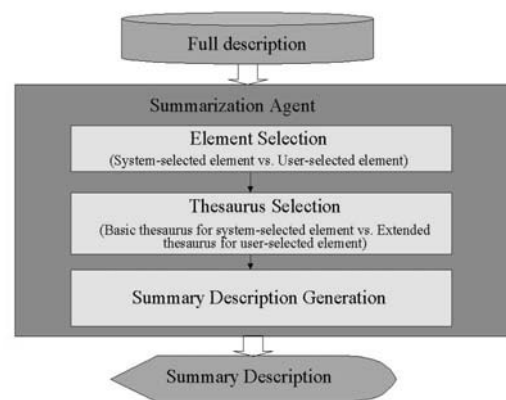
Figure 4. Flow diagram of summarization process

Figure 5 shows a block diagram that summary description is used in electronic cataloguing application.

Figure 5. Summary description usage in e-cataloguing application

4 Example of Thesaurus

Figure 6 shows a part of thesaurus example that is obtained using MPEG-7 MDS and DC that can be performed by summarization agent [1][2][4].

```
Title – [element][A name given to the resource]
    ⇨ <MPEG-7    MDS><Creation><Title
       type="main">
    ⇨ Alternative - [qualifier][Any form of the
       title used as a substitute or alternative
       to the formal title of the resource]
        ■  <MPEG-7 MDS> <Creation>
           <Title type="alternative">
        ■  <MPEG-7 MDS> <Creation>
           <Title type="secondary">
        ■  <MPEG-7 MDS> <Creation>
           <Title type="popular">
Description – [element][An account of the
content of the resource]
    ⇨ Abstract – [qualifier][The topic of the
       content of the resource
        ■  <MPEG-7 MDS> <Creation>
           <Abstract>
           <FreeTextAnnotation>
        ■  <MPEG-7 MDS> <Creation>
           <Abstract>
           <StructuredAnnotation>
Creator   – [element][An entity primarily
       responsible for making the content
       of the resource]
    ⇨ <MPEG-7    MDS>    <Creation>
       <Creator>
Subject – [element][The topic of the content of
the resource]
    ⇨ <MPEG-7   MDS>   <Classification>
       <Subject>
Language – [element][A language of the
intellectual content of the resource]
    ⇨ <MPEG-7   MDS>   <Classification>
       <Language>
Coverage – [element][The extent or scope of
the content of the resource]
    ⇨ <MPEG-7 MDS> <UsageInformation>
       <Availability>
    ⇨ Spatial  –   [qualifier]   [Spatial
       characteristics   of   the   intellectual
       content of the resource]
        ■  <MPEG-7 MDS>
           <UsageInformation>
           <Availability><OriginPlace>
    ⇨ Temporal  –  [qualifier][Temporal
       characteristics   of   the   intellectual
       content of the resource]
        ■  <MPEG-7 MDS>
           <UsageInformation>
           <Availability>
           <AvailabilityPeriod>
Publisher – [element][An entity responsible for
making the resource available]
        ■  <MPEG-7 MDS>
           <UsageInformation>
```

```
<Availability><Distributor>
```

Figure 6. An example of thesaurus for MPEG-7 MDS

Note. In case of Title, it is a DC element. The meaning is a titled name of the corresponding resource. The matching element in MPEG-7 MDS is <Title> element under <Creation>. If <Creation> is not specified before the <Title>, summarization process confuses with the <Title> of <PersonName> specified also in MPEG-7 MDS that indicates the titles of an individual, such as honorifics.

5 Conclusions

The conceptual model of interoperable summary description derived from the existing metadata frameworks or schemes is proposed. The validity and usefulness of the model is described. The summary description example obtained using MPEG-7 MDS shows the validity of the proposed model employing DC and thesaurus approaches. This summary description is expected to be meaningful in e-cataloguing and browsing applications of e-commerce.

6 References

[1] Dublin Core Metadata Initiative, *Dublin Core Metadata Element Set, Version 1.1: Reference Description,* July 2, 1999.

[2] Dublin Core Metadata Initiative, *Dublin Core Qualifiers,* July 11, 2000.

[3] R. Iannella and D. Campbell, *The A-Core: Metadata about Content Metadata,* June 30, 1999.

[4] ISO/IEC JTC1/SC29/WG11 N4040: *Information technology – Multimedia Content Description Interface– Part5: Multimedia Description Schemes, CD, Singapore*, March 2001.

[5] Godfrey Rust and Mark Bide, *The <indecs> metadata framework- Principles, model and data dictionary*, WP1a-006-2.0, June 2000.

[6] Carl Lagoze, Clifford A. Lynch and Ron Daniel, Jr.: *A Container Architecture for Aggregating Sets of Metadata*, June 28, 1996.

[7] R.V.Guha and Tim Bray, *Meta Content Framework Using XML,* http://www.w3.org/TR/ NOTE-MCF-XML - 970624, submitted to W3C 6 June, 2000.