

## MPEG-7 Visual Ontology

MPEG-7 [1] is an ISO/IEC standard developed by MPEG (Moving Picture Experts Group). MPEG-7, formally named "Multimedia Content Description Interface", is a standard for describing the multimedia content data that supports some degree of interpretation of the information meaning, which can be passed onto, or accessed by, a device or a computer code. MPEG-7 consists of several parts, such as the Audio part, which is the tool that deals with the description of Audio documents, the Visual part, which contains the specification for providing visual descriptions of multimedia documents and more. The standard language for encoding MPEG-7 descriptions is the XML language.

With the advent of the World Wide Web (WWW) and the Internet technologies, there is the opportunity for more and more companies and organizations that process and store multimedia documents to publish their archives on the WWW and make them accessible to a variety of users, such as other companies, researchers, artists, and the like. Hence, it is evident that for efficiently accessing this huge amount of information, meta-data descriptions, like the ones provided by the MPEG-7, become of high importance. Although, MPEG-7 is considered today one of the most successful metadata framework, its tight coupling on the XML language results in lack of formal semantics. Consequently, MPEG-7 descriptions are not interoperable, interchangeable and cannot be easily processed and accessed in a large scale by Web agents. For the purpose of describing and forming information and knowledge on the Web in a way that would consist information easily handled and accessed there exist Semantic Web technologies, like the Web Ontology Language (OWL) [2].

Today, there have been many attempts to move the MPEG-7 Schema from its XML definition to a more formally defined language like the OWL language, thus creating an MPEG-7 ontology. In order to correctly translate the MPEG-7 standard to the Semantic Web standards one has to take under consideration all the semantics of the MPEG-7 description elements that are described in the MPEG-7 documents. Moreover, other semantic information that is hidden inside the MPEG-7 XML Schema has to be externalized and encoded in the OWL MPEG-7 definition. Consider for example the XML representation of the ColorSpaceType Visual Descriptor given by,

```
<complexType name="ColorSpaceType" final="#all">
  <choice>
    <element name="ColorTransMat" minOccurs="0">
      <simpleType>
        <restriction>
          <simpleType>
            <list itemType="mpeg7:unsigned16">
            </simpleType>
            <length value="9"/>
          </restriction>
        </simpleType>
      </element>
    </choice>
    <attribute name="colorReferenceFlag" type="boolean"
      use="default" value="false"/>
    <attribute name="type">
      <simpleType>
        <restriction base="string">
          <enumeration value="RGB"/>
          <enumeration value="YCbCr"/>
          <enumeration value="HSV"/>
          <enumeration value="HMMD"/>
          <enumeration value="LinearMatrix"/>
          <enumeration value="Monochrome"/>
        </restriction>
      </simpleType>
    </attribute>
  </complexType>
```

```
</attribute>  
</complexType>
```

which specifies the colour model of a visual description. `ColorSpaceType` is the descriptor, which has three attributes, the `ColorTransMat`, the `ColorReferenceFlag` and the `Type` attributes. As we can see, the MPEG-7 specification is more programmatic centred rather than semantic centred, as it defines flags that would indicate the existence or non-existence of various elements. Obviously, `ColorSpaceType` can be translated to an OWL class, but it remains to figure out its properties and the range of them. Consider the “type” attribute. This could be transformed to an OWL property, but if we consider issues like interoperability, sharing and reuse in the WWW it would not be a good choice to translate it to an OWL property with the same name, as this word completely hides the intended meaning of the property. Thus, it would be difficult for a Semantic Web agent to automatically process such a property. Instead more descriptive names, such as “hasColorModel” can be used. Moreover, regarding the range of the attribute “type”, this is defined to be an enumeration of string values in the XML specification. Again the translation to the OWL language has to follow a more semantic centred approach and has to consider the hidden semantics that these string values represent. For example one can argue that all values, RGB, YCbCr, LinearMatrix, etc, represent concepts of our world, with internal description, e.g. the Red, Green and Blue components, in the case of RGB, the luminance and chrominance in the case of YCbCr, or the “ColorTransMat” in the case of the LinearMatrix. In that way we can also capture the fact that the `ColorTransMat` is an attribute of the concept `LinearMatrix`, rather than an attribute of the `ColorSpaceType` descriptor as it is falsely implied by the XML syntax of the descriptor. Hence, these entities should be defined as disjoint OWL Classes, while the type attribute should be an ObjectProperty with a range, the union of the disjoint classes defined earlier. This example shows that the naming followed in the MPEG-7 specification is not appropriate for the semantic web, as well as that the XML syntax many times implies false semantic conditions about the described elements.

Similar arguments can be made for most of the MPEG-7 Visual descriptors and their definitions. For example, the MPEG-7 standard defines a Visual descriptor called `ContourShape` with a property called `GlobalCurvatureVector`. Although, the XML Schema gives no semantics to `GlobalCurvatureVector`, as it is defined as a list of two strings, the MPEG-7 specification clearly specifies that the first string of `GlobalCurvatureVector` represent the *circularity* of the described object, while the second string its *eccentricity*. Such properties should be defined in an MPEG-7 owl ontology. Thus, we can see that the MPEG-7 standard describes the actual semantics of the elements inside its documents, and that such semantics have to be externalized and be represented in an MPEG-7 owl schema to enable system interoperability and intercommunication.

- [1]. MPEG-7 Overview, Jose M. Martinez, Available at: <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>
- [2]. *OWL Web Ontology Language Semantics and Abstract Syntax*, Peter F. Patel-Schneider, Patrick Hayes, and Ian Horrocks, Editors, W3C Recommendation 10 February 2004, <http://www.w3.org/TR/2004/REC-owl-semantics-20040210/>. [Latest version](#) available at <http://www.w3.org/TR/owl-semantics/>.