

# MPEG-7 E A RECUPERAÇÃO DA INFORMAÇÃO DE OBJETOS MULTIMÍDIA

Camila Monteiro de Barros\*  
Angel Freddy Godoy Viera\*\*

## RESUMO

A descrição de conteúdo de objetos multimídia é um grande desafio atual que incide diretamente no resultado da recuperação da informação. O MPEG-7, interface para descrição de conteúdos multimídia, é um complexo padrão baseado em modelos de metadados que atua na otimização da descrição desse tipo de conteúdo. Este artigo tem por objetivos descrever a estrutura do MPEG-7, estabelecer suas relações com a problemática da recuperação da informação de objetos multimídia e apresentar, brevemente, algumas iniciativas de desenvolvimento de sistemas de gerenciamento e recuperação da informação baseados no MPEG-7.

**Palavras-chave:** Recuperação da Informação. Metadados. MPEG-7. Objeto Multimídia.

\* Mestranda do Programa de Pós-Graduação em Ciência da Informação. Universidade Federal de Santa Catarina. Bolsista FAPESC. E-mail: camibarro2010@gmail.com

\*\* Professor do Programa de Pós-Graduação em Ciência da Informação. Departamento de Ciência da Informação da Universidade Federal de Santa Catarina. E-mail: godoy@cin.ufsc.br

## I INTRODUÇÃO

A partir da década de 1990, a *Web* no Brasil teve um grande desenvolvimento, possibilitando o acesso livre a inúmeros documentos, em diferentes formatos e suportes. Esses fatos contribuíram tanto para o aumento da produção quanto da demanda por informação, tornando a *Web* em um vasto repositório de documentos em variados formatos.

A organização da informação no ciberespaço é um tema bastante discutido por diversas áreas do conhecimento, que buscam desenvolver formas mais eficientes de se realizar a recuperação da informação- RI na *Web* de maneira satisfatória para o usuário. No que concerne ao ambiente virtual, a adoção de tecnologias que permitam a interoperabilidade dentre os diversos repositórios de documentos é imprescindível para uma recuperação eficiente da informação (CHANG, 2001).

Diferentemente dos documentos que trazem informação unicamente em formato de texto, nos documentos multimídia, o conteúdo

não se encontra explícito (como seriam o título, resumo ou a ocorrência de palavras no corpo de um artigo, por exemplo), os conteúdos ficam também correlacionados a diversos fatores externos ao próprio documento, como o propósito da sua produção, contexto histórico e geográfico do momento da concepção, e ainda características subjetivas que podem ser atribuídas ao documento pelo profissional/usuário que o descreve. Dessa forma, a aplicação unicamente da indexação automática não se torna suficiente para descrever a informação ali disponibilizada (SMEATON, 2004).

Nesse sentido, diversas ferramentas e aplicações têm sido desenvolvidas no intuito de possibilitar a identificação, recuperação ou filtragem da informação, como alguns modelos de metadados que permitem que informações adicionais sejam relacionadas ao objeto multimídia em questão. Alves e Souza (2007, p. 22) explicam que:

Os elementos de metadados têm o propósito primário de descrever,

identificar e definir um recurso de informação com o objetivo de modelar e filtrar o acesso. Os metadados são importantes na organização, gestão e recuperação da informação digital, principalmente [...] Portanto, são dados definidores que fornecem informação sobre ou documentação de outros dados dentro de uma aplicação ou de um ambiente.

Dentre as ferramentas desenvolvidas, destaca-se o MPEG-7 que se caracteriza como uma interface para gerenciamento de metadados, em que:

O objetivo final do MPEG-7 é promover interoperabilidade entre sistemas e aplicações usados na geração, gerenciamento, distribuição e consumo de descrições de conteúdos áudio-visuais. (CHANG et al, 2001, p. 688, tradução nossa<sup>1</sup>)

O MPEG-7 é uma ferramenta que desde sua criação vem se consolidando, sendo utilizado nas práticas de descrição de conteúdo de objetos multimídia em sistemas de gerenciamento e recuperação da informação que servem a diferentes grupos de usuários, e que têm apresentado resultados práticos satisfatórios quanto à eficiência do sistema. (DALLACOSTA et al, 2004; CAO et al, 2009; KANNAN et al, 2009; LUX, 2009; REY-LOPEZ et al, 2009; KAPELA et al, 2010)

Este artigo tem o intuito de descrever a estrutura do MPEG-7, estabelecer relações com a problemática da RI de objetos multimídia e apresentar, brevemente, algumas iniciativas de desenvolvimento de sistemas de recuperação da informação baseados no MPEG-7.

## 2 PROCEDIMENTOS METODOLÓGICOS

O levantamento das informações parte da exploração dos artigos publicados em periódicos eletrônicos que tratam diretamente das aplicações, conceitos e tendências relacionados com o MPEG-7.

As fontes utilizadas foram: periódicos indexados no portal CAPES e na base de dados *Scopus*, sendo que neste último se pode observar um maior retorno da informação buscada. No tópico que traz a discussão entre o MPEG-7 e recuperação da informação, inseriu-se a análise da obra de Ingwersen (1992), por este ser considerado um especialista no tema. Também foram levantados os conceitos de metadados e multimídia, inseridos de forma satisfatória para o contexto abrangido pelo presente artigo.

Em função de o tema estar relacionado a uma aplicação tecnológica, em que o aspecto da atualidade se torna relevante, optou-se por delimitar um corte cronológico de levantamento documental a partir do ano de 2001, excetuando-se a obra de Ingwersen (1992). Foram considerados para a pesquisa os documentos nos idiomas: português, espanhol e inglês.

Os artigos identificados que cumpriram com os critérios determinados para a pesquisa foram selecionados e utilizados para conceituar o padrão MPEG-7, identificar sua estrutura, suas aplicações na recuperação da informação e outras aplicações que estão sendo feitas sobre esse padrão e são apresentados a seguir neste artigo.

## 3 MPEG-7: ESTRUTURA

O MPEG-7 começou a ser desenvolvido em 1996 pelo *Moving Picture Experts Group-MPEG*, que é um grupo de trabalho ISO/IEC, com o objetivo de desenvolver padrões para representação codificada de dados digitais de áudio e vídeo. O MPEG é aprovado pela *International Organization for Standardization - ISO*, sendo que os padrões anteriores (MPEG-1, MPEG-2 e MPEG-4) tinham como foco a codificação e compactação dos conteúdos (VICENTE, 2005). Já o MPEG-7 é formalmente conhecido como *Multimedia Content Description Interface*, ou interface para descrição de conteúdos multimídia<sup>2</sup>, desenvolvido para que possa ser usado em grandes bases de material multimídia.

Vicente (2005) relata que o MPEG-1 e o MPEG-2 são padrões de compressão de áudio e vídeo utilizados para formatos como o MP3, sendo que o MPEG-2 é também utilizado para serviços de TV por satélite como Direct TV. O MPEG-4 teve sua função estendida para

<sup>1</sup> "The ultimate goal and objective of MPEG-7 is to provide interoperability among systems and applications used in generation, management, distribution, and consumption of audio-visual content descriptions" (CHANG et al, 2001, p. 688).

<sup>2</sup> Fonte: MPEG. Disponível em: <http://mpeg.chiariglione.org/>.

conteúdos 3D, suporte de objetos áudio e vídeo e aplicações para gestão de direitos de *copyright*.

O MPEG-7, que possui uma estrutura mais complexa, prevê ferramentas de descrição (*Description Tools*) representadas por elementos de metadados. Esses elementos são empregados “para criar descrições que serão utilizadas por ferramentas com funções para pesquisar, filtrar e navegar de forma eficiente em conteúdos multimídia” (CHELLA, 2004). É possível utilizar as ferramentas de descrição para arquivos digitais ou analógicos, imagem em movimento ou foto, som, vídeo, ou seja, elas independem dos formatos dos arquivos. Ainda segundo Chella (2004) as *Description Tools* podem criar descrições com informações sobre:

- Criação e processo de produção do conteúdo (diretor, título, resumo).
- Utilização do conteúdo (direito autoral, histórico de utilização, agendamento da transmissão).
- Estrutura da informação nos aspectos espacial ou espaço temporal (cortes na cena, segmentação em regiões, movimentos de elementos na cena).
- Conteúdo de baixo nível (cores, texturas, timbres de sons, descrição da melodia).
- Conceitos da realidade capturada (objetos e eventos, interações entre objetos).
- Como navegar pelo conteúdo de forma eficiente (sumários e variações).
- Coleções de objetos.
- Interação do usuário com o conteúdo (preferências do usuário, histórico de uso).
- Formato: indica o tipo de codificação utilizado para o arquivo (ex. JPEG, AVI). Esta informação auxilia na determinação de como o material será apresentado pelo terminal do usuário.
- Condições para acesso ao material: o que inclui links para registros com informações sobre propriedade intelectual, direito-autoral e preço.
- Classificação: qualifica o conteúdo em categorias pré-definidas.
- Links para outros materiais considerados relevantes: esta informação pode oferecer outros recursos relacionados ao tema pesquisado
- Contexto: No caso de material não ficcional, é importante se reconhecer a data da gravação.

As *Description Tools* são utilizadas segundo um conjunto de elementos normativos, a saber:

a) *Descriptor* - D: são descritores que definem características específicas relacionadas ao objeto multimídia. Em baixo nível de abstração, os descritores podem ser criados para descrever características como a cor, textura, movimento, tom do áudio, etc. Essas características podem ser extraídas automaticamente do documento (CHANG, 2001).

Quando a análise é realizada em nível de abstração mais alta, as características podem ser relacionadas com o evento (por exemplo, personagem X interage com elemento de cena Y) e conceitos mais abstratos como a subjetividade do conteúdo. Nesse caso é necessária a intervenção humana para anotações textuais.

b) *Description Schemes*- DS: são os esquemas de descrição que permitem a construção da complexa estrutura e relação semântica entre os elementos descritos. O conjunto de DS é denominado no padrão MPEG-7 de *Multimedia Description Schemes*- MDS.

Por exemplo, um esquema de descrição para um segmento de vídeo pode especificar a sintaxe e semântica dos elementos que o compõe [...], atributos individuais a cada segmento (por exemplo, comprimento do segmento, anotações textuais), e as relações entre os segmentos componentes da descrição. (CHANG et al, 2001, p. 689, tradução nossa<sup>3</sup>)

Ou seja, as DS permitem estabelecer relações tanto entre os Ds como entre DSs.

c) *Description Definition Language*- DDL: as informações construídas com os D, e as relações proporcionadas pelas DS, são feitas com base no XML *Schema*, que torna as descrições mais flexíveis. Assim, quando houver a necessidade de especializar as ferramentas de descrição, criar ferramentas adicionais, ou novos DS, é possível utilizando-se a linguagem XML em conformidade com os padrões da DDL.

Segundo Chella (2004), em razão do XML *Schema* não ter sido projetado especificamente para descrição de conteúdos audiovisuais, foram adicionadas algumas extensões. Deste modo, a DDL ficou dividida nos seguintes componentes: o XML *Schema* estrutura da linguagem, o XML *Schema* definição dos tipos de dados e extensões específicas para o padrão MPEG-7.

<sup>3</sup> “For example, the description scheme for a video segment may specify the syntax and semantics of the component elements [...], individual segment attributes (e.g., segment length, textual annotations), and relationships between component segments” (CHANG et al, 2001, p. 689).

A seguir apresentamos algumas aplicações identificadas na pesquisa sobre a aplicação do MPEG-7 na recuperação da informação.

#### 4 MPEG-7 E RECUPERAÇÃO DA INFORMAÇÃO

A recuperação da informação tem relação direta com diversos processos envolvidos no tratamento da informação: a representação (automática ou não) da informação por meio de descritores, textos explicativos, informações adicionais, reconhecimento automático de padrões gráficos e temporais; a estrutura do ambiente onde estas informações estão contidas (bases de dados fechadas, disponíveis na *web*); a aplicação de algoritmos específicos; os suportes das informações (texto, vídeo); dentre outros. Assim, é possível construir uma breve noção da complexidade do processo de recuperação da informação.

Ingwersen (1992) relata que, de forma geral, o problema central da RI é a noção de relevância das informações para atender precisamente a necessidade do usuário. Assim, “[...] relevância é definida como a medida ou grau de correspondência ou utilidade existente entre um texto ou documento e uma consulta ou requerimento de informação determinada por uma pessoa”. (INGWERSEN, 1992, p. 79, tradução nossa<sup>4</sup>)

Nos objetos em forma textual, é possível que os sistemas realizem uma representação objetiva da informação, por meio da extração automática de termos para a elaboração de um índice que servirá de referência no momento da busca. Essa extração pode ocorrer através do reconhecimento das palavras do título, resumo, ocorrência de palavras no corpo do texto, relações entre ocorrência de termos em diferentes textos. No entanto, a problemática da precisão dos sistemas de recuperação da informação está na construção de uma representação de conteúdo que englobe o sentido, o contexto da informação, ou seja, sua constituição semântica que, até o presente momento, só é possível com algum tipo de intervenção humana. A extração desse significado da informação é imprescindível para

o estabelecimento da relevância (INGWERSEN, 1992).

Essa intervenção humana é caracterizada de diferentes formas, dependendo do sistema que é utilizado na RI, dentre elas, a inserção de informações adicionais nos metadados, a respeito do conteúdo dos documentos ou na categorização desses conteúdos.

Como abordado anteriormente, os objetos multimídia não trazem seu conteúdo de forma explícita em termos e/ou expressões (ainda que estes também não esgotem as significações dos textos), tornando o processo de extração automática de informações mais complicada e de baixa eficiência descritiva quando se trata do conteúdo desses objetos. Essa configuração torna essencial a necessidade de anotações textuais para a posterior recuperação da informação (SMEATON, 2004). A tarefa de extração e geração das anotações textuais é complexa, onde as interpretações de alto nível de abstração são mais difíceis de serem realizadas por máquinas.

Seguindo esquemas preestabelecidos, que tornam as informações fornecidas pelo usuário passíveis de serem interpretadas por computador, é possível refinar o processo de descrição de conteúdo, agregando informações que otimizam a avaliação automática de relevância para a busca do usuário. Nesse sentido têm sido desenvolvidas diversas normas de especificações de metadados, com diferentes aplicações e desenvolvedores, como o ID3, iTunes, QT, Media RDF, MPEG-7<sup>5</sup>, entre outros.

Pretende-se, nesta discussão, estabelecer a relação entre a inferência do sentido de relevância de informações, com as características do MPEG-7 que contribuem para esse processo. Com essa ferramenta é possível descrever as características dos conteúdos multimídia para que os usuários possam pesquisar, recuperar e até mesmo navegar por estes conteúdos; as descrições podem estar fisicamente armazenadas em conjunto com o material áudio-visual, como também podem estar em outro lugar da rede (DALLACOSTA et al, 2004). Assim, tanto sistemas que atuam com usuários humanos como os automatizados podem utilizar o MPEG-7.

Chang et al (2001, p. 689, tradução nossa<sup>6</sup>) esclarece de forma sucinta algumas tarefas que

4 “[...] relevance, defined as the measure or degree of a correspondence or utility existing between a text or document and a query or information requirement as determined by a person” (INGWERSEN, 1992, p. 79).

5 W3C. Disponível em: <http://www.w3.org/standards/>

6 • Multimedia— Generate a customized program guide or summary of broadcast audio-visual content according to user's preference and use history.

usuários e aplicações podem executar partindo das descrições do MPEG-7:

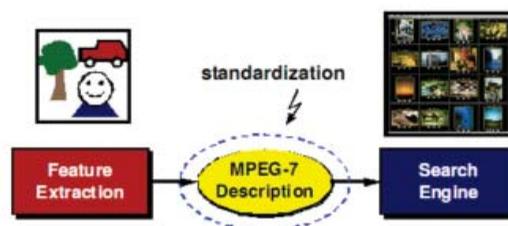
- Multimídia: pode gerar um guia personalizado de programas ou sumário de transmissão de conteúdo áudio-visual de acordo com as preferências do usuário e seu histórico de uso.
- Arquivo: pode gerar descrições de elementos individuais ou da coleção inteira de conteúdo áudio-visual e, de forma coerente, trocar conteúdos e suas descrições correspondentes entre os proprietários de conteúdos, agregadores, e consumidores.
- Adaptação: filtrar e transformar os fluxos de multimídia em ambientes com recursos limitados combinando a preferência do usuário, recursos disponíveis e descrição dos conteúdos.
  - Música/Áudio: tocar algumas notas no teclado e obter como resposta uma lista de peças musicais contendo a melodia que corresponde a essas notas.
  - Gráficos: desenhar algumas linhas na tela e obter como resposta, uma lista de imagens contendo gráficos, logos ou ideogramas semelhantes.
  - Movimento: com um determinado conjunto de objetos de vídeo, descrever movimentos e relações entre os objetos e obter, de volta, uma lista de animações ou vídeos que preencham as relações temporal e espacial descritas.
  - Cenário: em um determinado conteúdo áudio-visual, descrever ações e receber em contrapartida uma lista de cenários (segmentos áudio-visuais) onde ocorrem ações similares.

A eficiência da recuperação dos conteúdos multimídia proporcionada pelo MPEG-7 está relacionada com a gama de informações passíveis de serem inseridas através das *Description Tools*, que permite tratar os objetos de forma individual

- 
- Archive—Generate descriptions of individual elements or the entire collection of audio-visual content, and seamlessly exchange content and corresponding descriptions among content owners, aggregators, and consumers.
  - Adaptation—Filter and transform the multimedia streams in resource limited environments by matching the user preference, available resources, and content descriptions.
  - Music/Audio—Play a few notes on a keyboard and get in return a list of musical pieces containing (or close to) the tune somehow matching the notes.
  - Graphics—Draw a few lines on a screen and get, in return, a set of images containing similar graphics, logos, or ideograms.
  - Movement—With a given set of video objects, describe movements and relations between objects and get, in return, a list of animations or video clips fulfilling the described temporal and spatial relations.
  - Scenario—On a given audio-visual content, describe actions and get in return a list of scenarios (i.e., audio-visual segments) where similar actions take place.

ou coletiva, em diversos aspectos. O padrão permite a descrição de informações relacionadas aos aspectos espaciais e temporais (segmentação de vídeos, movimentos de elementos da cena) empregando a extração automática e manual de informações em cada trecho descrito. A partir da especificação do tipo de arquivo (JPEG, AVI...), é possível agendar a transmissão de determinado objeto ou trecho, relacionando-o ainda a determinada coleção que pode ser desenvolvida partindo da categorização dos conteúdos em características pré-definidas (REY-LÓPEZ, 2010). Esses aspectos permitem a especialização da descrição do conteúdo, aumentando o grau de correspondência e utilidade entre a solicitação do usuário e o objeto recuperado, ou seja, a relevância das informações recuperadas.

Na figura abaixo, Vicente (2005) ilustra, de forma sucinta, a relação do MPEG-7 com os sistemas de busca:



MPEG-7 e sistemas de busca de informações

Fonte: Vicente (2005, p. 4)

A figura mostra que, primeiramente, é realizada uma análise (manual ou automática) do objeto multimídia a ser descrito, suas características estruturais, semânticas, palavras-chave. A descrição das características é feita utilizando-se o MPEG-7 que disponibiliza os Ds e as DSs. Como a estrutura do MPEG-7 é projetada na linguagem XML, que é interoperável, as informações descritas podem ser “entendidas” e utilizadas por qualquer aplicação de busca.

Uma importante característica dos sistemas de recuperação da informação é o aprendizado do sistema a partir do histórico de uso e por meio do reconhecimento de tendências e perfis de usuários. Essa característica aparece no MPEG-7, pois o padrão reconhece e registra as preferências do usuário, a partir da interação deste com os conteúdos. O usuário pode ainda realizar as descrições em linguagem natural ou estruturada, contribuindo para a recuperação eficiente da informação.

Soluções como o MPEG-7 são importantes tanto para o usuário final quanto para os provedores de conteúdo, pois possibilita uma completa descrição do objeto, incluindo as características facilmente perceptíveis e também semânticas. Smeaton (2004, p. 379, tradução nossa<sup>7</sup>) conclui:

MPEG-7 pode ser usado como um veículo muito abrangente para descrever todos os aspectos do conteúdo do vídeo; a codificação da descrição de conteúdo de vídeos nesse formato padrão cria muitas oportunidades para exploração nas operações baseadas em conteúdo.

O MPEG-7 é uma ferramenta que vem se consolidando como interface de descrição de conteúdo multimídia, tendo sido adotado desde 2002 na descrição das coleções-teste de objetos multimídia da *Text Retrieval Conference- TREC*, conferência que tem a finalidade de “[...] apoiar a pesquisa junto à comunidade de recuperação da informação fornecendo a infraestrutura necessária para a avaliação em larga escala das metodologias de recuperação da informação” (TEXT..., tradução nossa<sup>8</sup>). É válido ressaltar que a adoção dos padrões MPEG nas TREC dá-se devido a sua característica que prevê uma completa interoperabilidade entre diferentes plataformas, o que facilita a operacionalização de um sistema colaborativo.

A utilização do formato acontece em sistemas com diferentes objetivos e estruturas, como acervos especializados e repositórios educacionais.

Pascual e Minguillón (2008) relatam a importância da utilização de padrões de metadados interoperáveis e que assegurem a qualidade no acesso à informação em cursos à distância. Os autores apresentam o projeto *Opera-Learning*, uma iniciativa do *Gran Teatre del Liceu* com a colaboração da *Facultat d'Informàtica de Barcelona* (UPC) e da *Techno Trends*, no oferecimento de um curso de ópera com uso de videoconferência voltado exclusivamente para estudantes universitários. O material oferecido

aos alunos do curso à distância foi baseado nos materiais anteriormente disponibilizados para cursos presenciais, assim, os materiais só puderam ser disponibilizados mediante sua gravação, digitalização e edição. As aulas são realizadas por videoconferência e o material disponibilizado é descrito e organizado com base no MPEG-7, que, por sua interoperabilidade, garante o acesso a todos os alunos.

A utilização deste padrão para descrever o assunto permite proporcionar um duplo sentido: por um lado, poder descrever sintática e semanticamente o assunto de forma completa, e por outro lado, poder reutilizar os materiais e conteúdos das edições anteriores. (PASCUAL E MINGUILLÓN, 2008, tradução nossa<sup>9</sup>)

Os autores expõem que, primeiramente, os conteúdos foram subdivididos em unidades elementares (áudio, vídeo e texto), cada unidade foi descrita no padrão MPEG-7, recebendo palavras-chave, anotações textuais, análise por trecho. Uma vez descritas as unidades elementares, foram estabelecidas as relações semânticas entre cada unidade. “Esta tarefa se realiza por meio dos esquemas de descrição (DS). O padrão MPEG-7 através da DDL facilita a criação e modificação de descritores e esquemas de descrição caso seja necessário” (PASCUAL; MINGUILLÓN, 2008, tradução nossa<sup>10</sup>), possibilitando a atualização e reutilização dos objetos descritos para outros cursos.

No artigo de Kapela et al (2010), é exposto o *Visual Content Description System- VCDS*, ou sistema de descrição de conteúdo visual. Esse sistema executa a descrição de vídeos em tempo real que, segundo os autores, acontece em duas partes: primeiro, a extração dos descritores utilizando o VCDS e, segundo, a utilização de valores para os descritores com a aplicação de um algoritmo particular. Esse sistema utiliza quatro formatos de descritores do MPEG-7 que se relacionam com diferentes objetos na cena

7 “MPEG-7 can be used as an all-encompassing vehicle to describe every aspect of (vídeo) content; the encoding of content description across videos in this standard format creates many opportunities for exploitation in content-based operations” (SMEATON, 2004, p. 379).

8 “[...] to support research within the information retrieval community by providing the infrastructure necessary for large-scale evaluation of information retrieval methodologies (TEXT...).

9 “La utilización de este estándar para describir la asignatura permite proporcionar un doble sentido: por una parte, poder describir sintáctica y semánticamente la asignatura de forma completa, y por otra parte, poder reutilizar los materiales y contenidos de las ediciones anteriores (PASCUAL E MINGUILLÓN, 2008).

10 “Este apartado se realiza mediante los diagramas de descripción (DS). El estándar MPEG-7 a través del DDL facilita la creación y modificación de descritores y diagramas de descripción en caso que sea necesario” (PASCUAL; MINGUILLÓN, 2008).

do vídeo, gerando um índice e permitindo uma rápida navegação pelo conteúdo do vídeo com base em medidas de similaridade.

Demonstrando a preocupação com a descrição semântica de vídeos de dança voltada especialmente para coreógrafos, dançarinos e estudantes, Kannan et al (2009) expõem a estrutura da aplicação desenvolvida pelos autores denominada *DanVideo*. O *DanVideo* oferece uma ferramenta MPEG-7 semi-automática de registro de autoria, gerando dados em formato MPEG-7. A ferramenta foi desenvolvida para ser utilizada por especialistas em dança na descrição dos vídeos, das complexas relações espaço-temporais que existe entre os dançarinos em termos dos movimentos dos corpos e das emoções expressadas na dança. O sistema tem um motor de busca que recupera informações semânticas conforme as consultas dos usuários e também dispõe de uma ontologia de domínio específico.

A abordagem de Dallacosta et al (2004) trata, de forma mais teórica, a importância da adoção de sistemas de descrição de vídeos especialmente para utilização na área da educação. Para os autores, com a utilização do MPEG-7, a recuperação de conteúdos em vídeo é otimizada, possibilitando aos

[...] professores de forma simples e rápida, localizar vídeos educacionais que melhor se encaixem para o apoio em sala de aula. Adicionalmente servirá como ferramenta para que o aluno navegue, de forma hipertextual, nos diversos vídeos armazenados neste repositório. (DALLACOSTA et al, 2004, p. 1)

Os autores ainda discutem a respeito das diferenças entre a educação linear e hipermídia e as vantagens no que concerne à qualidade do processo ensino-aprendizagem quando elementos áudio-visuais são implementados como apoio para o ensino. Quanto à aplicação prática do MPEG-7, é descrita a iniciativa de desenvolvimento do repositório de vídeos educacionais com descrições baseadas no MPEG-7, o Projeto CESTA.

O projeto CESTA (Coletânea de Entidades de Suporte ao uso de Tecnologia na Aprendizagem) foi idealizado com vistas a sistematizar e organizar a catalogação dos materiais educacionais (objetos educacionais). (DALLACOSTA et al, 2004, p. 9)

Neste projeto, um professor gera as palavras-chave e as indexa em formato MPEG-7. Essas palavras estão associadas a segmentos de vídeo, facilitando a recuperação de conteúdos específicos.

Já Lux (2009), relata a Caliph&Emir, uma aplicação baseada na linguagem Java para anotações de informações em imagens e posterior recuperação. O autor explica que a aplicação emprega o esquema de descritores do MPEG-7 e dispõe ao usuário uma interface para gerar arquivos XML do MPEG-7 para busca e recuperação de informações.

A recuperação da informação é um processo que visa o atendimento das necessidades informacionais do usuário. O desenvolvimento de ferramentas como o MPEG-7 é relevante para a organização e identificação dos conteúdos digitais especialmente dos objetos multimídia, oferecendo suporte para uma efetiva representação, de forma a garantir a relevância dos resultados de busca.

## 5 MPEG-7: OUTRAS APLICAÇÕES

Vicente (2005) ressalta que o MPEG-7 não é um padrão para busca e recuperação de conteúdos, mas sim uma ferramenta para descrição de conteúdo multimídia com a finalidade de proporcionar uma linguagem comum de forma a aperfeiçoar tais processos. Dessa forma, abordaremos brevemente algumas aplicações do MPEG-7 que não estão relacionadas diretamente com sistemas de recuperação de conteúdos.

Rey-Lopez et al (2009) relata o desenvolvimento de um sistema baseado no MPEG-7 para ser aplicado no gerenciamento de conteúdo da TV digital. A ideia é identificar os conteúdos de preferência do usuário, agrupando-os em categorias e permitindo que o sistema ofereça outros conteúdos similares adicionais para que o usuário acesse. Basicamente, o sistema foca duas áreas da TV digital: o oferecimento de conteúdos educativos relacionados com o programa que está sendo assistido (outros programas, *web sites*, etc) e o oferecimento ao espectador de anúncios de produtos relacionados com seus conteúdos de preferência.

Para a aplicação desse sistema, é necessário que seja possível estabelecer relações não apenas entre os conteúdos inteiros, mas também entre suas partes. A TV digital está

cada vez mais atrelando seus serviços à *web*, assim, principalmente por meio da característica que o MPEG-7 apresenta de agendamento da disponibilização de determinados conteúdos, torna-se possível que o oferecimento de conteúdos adicionais (incluindo conteúdos *web*) seja realizado de forma automática, por meio do reconhecimento do perfil do usuário (REY-LOPEZ et al, 2009).

Também referindo-se a serviços voltados à TV digital, Peláez (2008) aponta a importância de se associar a descrição de conteúdos (automática ou manual) a uma ontologia que delimite as relações semânticas em um domínio específico de conhecimento

[...] uma ontologia que descreva um domínio concreto, como por exemplo, o tênis, fragmentos de imagem que contém os conceitos ou elementos da ontologia, como por exemplo, raquetes. Desta maneira se une o plano conceitual (raquete) com o plano multimídia (seu aspecto). (PELÁEZ, 2008, p. 5)

O autor explica que nesse caso são utilizadas ontologias relacionadas com os descritores do MPEG-7 (padronização, redução de ambigüidade) e ontologias do domínio específico dos conteúdos a se tratar. A utilização dessa estrutura permite que o sistema realize, de forma automática por meio de um algoritmo específico, a análise das imagens englobadas pela ontologia e crie os descritores com base na ontologia relacionada ao MPEG-7. Assim, constitui-se um sistema de TV digital baseado no uso do MPEG-7, com a utilização de ontologias tanto para descrição de conteúdos quanto para recomendação de conteúdos; recomendações essas que são baseadas nos metadados de transmissão que o usuário está assistindo.

“Entre todos os formatos existentes o MPEG-7 parece ser o de maior potencial pela expressividade que permite nas descrições, por estar baseado em XML e por ser um padrão aberto” (PELÁEZ, 2008, p. 5). O autor ainda ressalta que as “tecnologias semânticas” têm potencial para melhorar tanto a geração de metadados quanto a sua aplicação em diferentes serviços.

Na pesquisa descrita por Cao et al (2009), os autores expõem o problema do alto custo de implantação e gerenciamento dos *middleware*, que

são programas de computador que transportam informações entre programas de diferentes protocolos de comunicação, plataformas e sistemas operacionais. Esses programas são utilizados principalmente para transmissão de informações em tecnologias móveis, onde o *middleware* torna possível a interoperabilidade entre diferentes sistemas. Assim, os autores visaram a interoperabilidade não mais a nível de *middleware*, mas a nível de gerenciamento de dados, sendo que o “MPEG-7 é um padrão bem estabelecido e amplamente utilizado no gerenciamento de dados multimídia” (CAO et al, 2009, p. 102, tradução nossa<sup>11</sup>). Os autores afirmam, como prova de conceito, que o MPEG-7 apresenta ótima acessibilidade e confiança na disponibilização de serviços multimídia em plataformas móveis.

## 6 CONSIDERAÇÕES

Devido a complexidade desse padrão, apenas alguns protótipos foram desenvolvidos em caráter científico e ainda existem poucas aplicações comerciais de criação, visualização e intercâmbio baseados no MPEG-7. A implantação do MPEG-7 exige que os responsáveis em gerenciar as informações recebam algum tipo de treinamento, e tenham algum conhecimento de linguagem XML. Esse processo gera um custo relacionado ao tempo de implantação e treinamento que precede a aplicação propriamente dita do modelo.

Em suma, o MPEG-7 oferece a vantagem de possibilitar a descrição de diversos aspectos dos objetos multimídia de forma bastante completa. A descrição independente de segmentos do objeto, bem como o estabelecimento de relações semânticas entre esses segmentos, também se configura em um importante aspecto para o atendimento das necessidades do usuário na busca da informação.

Nesse aspecto, acredita-se que as complexidades e custos associados de implantação e utilização do MPEG-7 estejam devidamente justificados e são superadas pela garantia de qualidade de gerenciamento, interoperabilidade, armazenamento e recuperação de informações que esse modelo proporciona.

<sup>11</sup> “MPEG-7 is a well-established and widely used standard in multimedia data management” (CAO et al, 2009, p. 102).

## MPEG-7 AND INFORMATION RETRIEVAL OF MULTIMEDIA OBJECTS

### Abstract

The description of the multimedia objects content is currently a great challenge which falls directly upon the result of information retrieval. MPEG-7, interface for description of multimedia contents, is a complex standard based on metadata templates that operates on the description optimization of this sort of content. This article is intended to describe the MPEG-7's framework, to determine its relations with the problem of multimedia objects' information retrieval, and briefly, to present some initiatives of information management and retrieval systems' development, based on MPEG-7.

### Keywords:

Information Retrieval. Metadata. MPEG-7. Multimedia Object.

---

Artigo recebido em 30/09/2010 e aceito para publicação em 04/12/2010

---

## REFERÊNCIAS

ALVES, M. D. R.; SOUZA, M. I. F. Estudo de correspondência de elementos metadados: Dublin Core e MARC 21. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 4, n. 2, p. 20-38, jan./jun. 2007. Disponível em: <<http://polaris.bc.unicamp.br/seer/ojs/viewarticle.php?id=94&layout=abstract>> Acesso em: 07 jun. 2009.

CAO, Y.; et al. Mobile access to MPEG-7 based multimedia Services. In: INTERNATIONAL CONFERENCE ON MOBILE DATA MANAGEMENT, 2009, Taipei, **Anais eletrônicos...** Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5088925>> Acesso em: 06 abr. 2010.

CHANG, S. et al. Overview of the MPEG-7 Standard. **IEEE transactions on circuits and systems for video technology**, v. 11, n. 6, jun. 2001. Disponível em: <[http://www.img.lx.it.pt/~fp/cav/Additional\\_material/MPEG7\\_overview\\_1.pdf](http://www.img.lx.it.pt/~fp/cav/Additional_material/MPEG7_overview_1.pdf)> Acesso: 06 abr. 2010.

CHELLA, M. T. **Sistema para classificação e recuperação de conteúdo multimídia baseado no padrão MPEG-7**. UNICAMP: São Paulo, 2004. Disponível em: <<http://www.nied.unicamp.br/~siros/doc/2232.pdf>> Acesso em: 06 abr. 2010.

DALLACOSTA, A; et al. A utilização da indexação de vídeos com MPEG-7 e sua aplicação

na educação. **Novas Tecnologias na Educação**, Porto Alegre, v. 2, n. 1, mar. 2004. Disponível em: <<http://www.cinted.ufrgs.br/ciclo3/af/35-autilizacao.pdf>> Acesso em: 30 mar. 2010.

INGWERSEN, P. **Information Retrieval Interaction**. London: Taylor Graham, 1992.

KANNAN, R.; et al. DanVideo: an MPEG-7 authoring and retrieval system for dance videos. **Multimedia Tools Appl**, out. 2009. Disponível em: <<http://www.springerlink.com/content/3858428730083007/fulltext.pdf>> Acesso em: 06 abr. 2010.

KAPELA, R; et al. Real-time visual content description system based on MPEG-7 descriptors. **Multimedia Tools Appl**, mar. 2010. Disponível em: <<http://www.springerlink.com/content/0x6j676817332463/fulltext.pdf>> Acesso: 06 abr. 2010.

LUX, M. Caliph & Emir: MPEG-7 photo annotation and retrieval. In: INTERNATIONAL MULTIMEDIA CONFERENCE, 2009, New York, **Anais eletrônicos...** Disponível em: <<http://www.springerlink.com/content/93032472212g4332/fulltext.pdf>> Acesso: 06 abr. 2010.

MPEG. Disponível em: <<http://mpeg.chiariglione.org>>. Acesso em: 07 abr. 2010.

PASCUAL, M.; MINGUILLÓN, J. Opera-Learning: Integración de estándares de distribución de

contenidos multimedia y learning objects. **Revista de Educación a Distancia**, Barcelona (Espanha), 2008. Disponível em: <[http://www.cc.uah.es/spdece/papers/Pascual\\_Final.pdf](http://www.cc.uah.es/spdece/papers/Pascual_Final.pdf)> Acesso em: 05 set. 2010.

PELÁEZ, V. M. Metadatos en contenidos multimedia para Televisión Digital. **Fundación CTIC**, Gijón (Espanha) 2008. Disponível em:<<http://www.di.uniovi.es/~labra/cursos/Doc08UniOvi/artiVictor.pdf>> Acesso em: 07 set. 2010.

REY-LÓPEZ, M.; et al. Enhancing TV programmes with additional contents using MPEG-7 segmentation information. **Science Direct**, v. 37, n. 2, p. 1124-1133, mar. 2010. Disponível em: <[http://www.sciencedirect.com/science?\\_ob=ArticleURL&\\_udi=B6V03-4WNXTWY-2&\\_user=687353&\\_coverDate=03%2F31%2F2010&\\_rdoc=1&\\_fmt=high&\\_orig=search&\\_sort=d&\\_docanchor=&view=c&\\_acct=C000037882&\\_](http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V03-4WNXTWY-2&_user=687353&_coverDate=03%2F31%2F2010&_rdoc=1&_fmt=high&_orig=search&_sort=d&_docanchor=&view=c&_acct=C000037882&_)

version=1&\_urlVersion=0&\_userid=687353&md5=3da7e221b47be61f83ca3cec412fa037>. Acesso em: 06 abr. 2010.

SMEATON, A. F.. Indexing, browsing, and searching of digital video. In: CRONIN, B. (Ed.). **Annual Review of Information Science and Technology**. Medford: Information Today, Inc., 2004, v. 38, p. 371-407.

TEXT Retrieval Conference. Disponível em:<<http://trec.nist.gov>> Acesso: 14 abr. 2010

VICENTE, P. J. V. El estándar MPEG-7. **Revista de Ingeniería Informática del CIIRM**, Murcia (Espanha), n.3, p. 1-5, 2005. Disponível em: <[http://www.cii-murcia.es/informas/jul05/articulos/El\\_estandar\\_MPEG-7.pdf](http://www.cii-murcia.es/informas/jul05/articulos/El_estandar_MPEG-7.pdf)> Acesso em: 03 set. 2010.

WORLD Wide Web Consortium. Disponível em: <<http://www.w3.org/>>. Acesso em: 07 abr. 2010.