

SACCA – Sistema Automático de Catalogação de Conteúdo Audiovisual

Renato Luís de Souza Dutra^{*}
Liane Margarida Rockenbach Tarouco^{**}
Mary Lucia Pedroso Konrath^{***}

Resumo: A crescente utilização de vídeos educacionais digitalizados no processo de ensino-aprendizagem ganhou importantes recursos com o surgimento de padrões que auxiliam a descrevê-los e indexá-los de forma padronizada, tais como o MPEG 7, possibilitando a busca e seleção de vídeos ou partes dele. Com base nessas tecnologias busca-se implementar um sistema que permita localizar vídeos educacionais no repositório de objetos educacionais já em operação, o CESTA, através de palavras-chave. Essas palavras-chave estarão associadas a uma descrição MPEG-7, que conterá os segmentos de vídeo que possuem aquela palavra-chave, o que possibilitará exibir o filme exatamente nos pontos onde ocorre cada palavra-chave.

Abstract: The increasing utilization of digitalized educational video in the teaching and learning process has gained important resources with the emerging of standards that helps to describe and index them in a standardized way, as mpeg 7, making possible the search and selection of videos or part of them. based on such technologies, we want to implement a system which allows the searching of educational videos in a learning objects repository – called CESTA - operated through key-word. this key-words will be associated with a mpeg-7 description, which will contain the videos segments that have that specific key-word. this will enable to play the movie from the exactly part where each key-word is associated to.

Palavras-chave: indexação de vídeos, MPEG7, vídeos na educação, Projeto SACCA

1. Introdução

Há muito tempo, vídeos são utilizados como ferramentas de apoio didático tanto na educação presencial como na educação à distância. Diversas iniciativas visam produzir vídeos educacionais, entre essas o Telecurso da Fundação Roberto Marinho, o Canal Futura e a TV Escola do MEC. Buscando oferecer esses recursos de forma organizada, um acervo de vídeos começa a ser produzido na UFRGS e disponibilizados pela TV Escola do MEC. Vídeos com conferências, aulas, experimentos, demonstração de procedimentos, visitas em campo, etc. estão se proliferando, mas ficam muitas vezes sem serem usados pelos professores pelo desconhecimento de seu conteúdo e pela potencial população de usuários. A utilização de vídeos educacionais no processo de ensino-aprendizagem ganhou importantes recursos com o surgimento de padrões que auxiliam a descrevê-los e indexá-los de forma padronizada, possibilitando a busca e seleção de vídeos ou partes dele.

Os sistemas de cadastramento existentes, tais como o do projeto CESTA (Coletânea de Entidades de Suporte ao uso de Tecnologia na Aprendizagem) para objetos de aprendizagem ou o ALEPH usado na biblioteca da UFRGS, usualmente se

^{*} Doutorando em Informática na Educação e Mestre em Ciência da Computação pela UFRGS

^{**} Doutora em Engenharia Elétrica pela USP, Mestre em Ciência da Computação pela UFRGS

^{***} Especialista em Informática na Educação pela UFRGS

limitam a registrar dados inerentes à catalogação e não necessariamente ao conteúdo em si. Perde-se nessas alternativas a possibilidade de indexação *full-text* que as máquinas de busca como Altavista e Google popularizaram e que são tão úteis e necessárias.

Esse artigo busca descrever o projeto SACCA, no qual pretende-se projetar e desenvolver um sistema que apoie o processo de catalogação *full content*, mas com a intermediação de um usuário que assiste ao vídeo e seleciona as palavras chave a serem incluídas na catalogação.

2. As iniciativas e acervos de vídeos

A utilização do vídeo na educação ocorre desde as primeiras televisões educativas surgidas nos EUA na década de 40 (MOORE, 1996). De lá para cá, a qualidade tem melhorado, bem como a diversificação de seu uso, podendo ser utilizado tanto como mídia de cursos a distância, assim como material de apoio no ensino presencial.

No Brasil, diversas iniciativas pioneiras merecem destaque, tanto pela qualidade do material produzido, quanto pelo alcance dos projetos. Entre esses, podemos citar o Projeto Telecurso da Fundação Roberto Marinho, o TV Escola do Mec e o Canal Futura, fruto de um consórcio de diversas empresas.

2.1. Projeto Telecurso

O Telecurso foi um projeto criado pela Fundação Roberto Marinho em 1978, visando inicialmente o 2º Grau, sendo transmitido na época por 60 emissoras de TV e 800 emissoras de rádio. Com grande sucesso, em 1995, o Telecurso 2º grau evoluiu para o Telecurso 2000, envolvendo o ensino fundamental, médio e profissionalizante. A qualidade do projeto é reconhecida internacionalmente, despertando a atenção de países como Estados Unidos, Argentina e Chile, e sendo também atualmente utilizado no Timor Leste. (Fundação Roberto Marinho, 2004)

No Telecurso, o aluno pode estudar assistindo aos programas de TV transmitidos por diversas emissoras, assistir a fitas de vídeo gravadas e estudar através do material impresso, podendo fazê-lo em casa ou nas telesalas, sendo que nas telesalas podem contar com orientação de profissionais especializados.

2.2. Canal Futura

O Futura é um canal de televisão com a missão de “contribuir para a formação educacional da sociedade, possibilitar o acesso ao conhecimento e incentivar a cidadania e a participação pessoal” (Fundação Roberto Marinho, 2004). O Futura é uma iniciativa de diversas empresas, entre elas a Rede Globo, a CNN, a Fundação Itaú, a Votorantim.

O Futura baseia-se em uma programação desenvolvida especialmente para servir de apoio a educação, onde os parceiros do projeto participam ativamente de sua gestão e diretrizes de programação, produzindo e divulgando programas e iniciativas na área social. O projeto conta também com uma mobilização comunitária que orienta as instituições sobre como utilizar os programas de maneira mais adequada.

2.3. TV Escola

A TV Escola é uma iniciativa do governo federal, a partir da Secretaria de Educação à Distância do MEC (Secretaria de Educação à Distância, 2004), que visa capacitar e atualizar professores do ensino médio e fundamental. Seu objetivo é propiciar a escolas da rede pública em todo o território nacional recursos de educação à distância para capacitar seus professores.

A TV escola disponibiliza programas produzidos por outras organizações e produzidos pela própria SEED, com programação diversificada. Entre os diversos programas podemos destacar o “Acervo” e “Sala do Professor”, que mostram determinados temas ou vídeos comentados por uma equipe multidisciplinar

2.4. Novas possibilidades do uso dessas produções na educação

Como podemos ver, os principais projetos que envolvem utilização da Televisão na educação possuem grande qualidade, despertando inclusive o interesse de outros países e ganhando prêmios internacionais. Porém, junto com o crescente ganho de qualidade dessas iniciativas brasileiras, ocorreram também importantes avanços no campo da tecnologia.

Com o surgimento da *Internet*, um novo leque se abriu diante desta tecnologia já bastante difundida, através da possibilidade de digitalização destes vídeos em formato *streaming*¹, possibilitando sua disponibilização através da rede e criando-se grandes bibliotecas digitais. Entretanto, para acompanhar essa crescente disponibilização de bibliotecas digitais, faz-se necessário a catalogação e indexação dos vídeos para disponibilizá-los de forma rápida e dinâmica.

3. Padrões de Indexação & Ferramentas

Neste contexto, surgiram diversas iniciativas visando não só a catalogação de vídeos digitais, como também sua sincronização, indexação e descrição. Podemos citar entre essas iniciativas, o SMIL, o Annodex e o MPEG 7.

3.1. SMIL

O SMIL (Synchronized Multimedia Integration Language), permite a criação de apresentações audiovisuais interativas (Synchronized Multimedia, 2004). O SMIL é tipicamente utilizado para apresentações multimídia do tipo “*rich media*”, que integram áudio e vídeo *streaming*, texto ou qualquer outro tipo de arquivo.

O SMIL é uma linguagem baseada no XML, trabalhando com *tags* semelhantes ao HTML, podendo ser editado por qualquer editor de texto comum, pois os elementos multimídia não são inseridos, somente referenciados. Com o SMIL, é possível gerenciar a transmissão de arquivos multimídia por *streaming*, sincronizando estes arquivos com animações flash, páginas html, figuras, etc.

¹ *Streaming* é um formato de transmissão de áudio e/ou vídeo via Internet (ou alguma outra rede) que é transmitido sob demanda.

Devido a esses recursos de sincronização, o SMIL também possibilita criar uma espécie de indexação destes arquivos de áudio e vídeo, através da sincronização das palavras chaves ou *thumbnails* com o vídeo desejado. Outra grande vantagem do SMIL é sua utilização pelos dois dos mais importantes *players* multimídia do mercado, o Real Player e o Quicktime, garantindo, portanto, grande portabilidade desses materiais.

3.2. ANNODEX – Annotation and Indexing

O Annodex, *anotating and indexing*, é uma especificação criada a partir do Continuous Media Project, iniciado na CSIRO (Australia's Commonwealth Scientific and Industrial Research Organisation), uma organização de pesquisa australiana envolvida nas mais diversas áreas. O Continuous Media Project foi iniciado por Silvia Pfeiffer em 2001, a partir das idéias contidas em um artigo de Bill Simpson-Young e Ken Yap em 1997, intitulado “*An open continuous media environment on the Web*”. (Open standards for annotating and indexing networked media, 2004)

Annodex possibilita a manipulação de arquivos de tempo-contínuo (áudio e vídeo) na Web. O formato Annodex, junto com o CMML (*Continuous Media Markup Language*), o *Temporal URI Reference*, servem de base para possibilitar a procura e navegação através de arquivos de tempo-contínuo.

O Annodex permite a especificação de âncoras em arquivos de tempo-contínuo, juntos com anotações textuais no formato de hiperlinks URI. Estas âncoras são anexadas sincronizadamente com os arquivos, quando criadas no formato Annodex. Este formato permite a encapsulação de qualquer arquivo do tipo *streaming*, independentemente de seu formato. Adicionalmente, as tags XML escolhidas pelo Annodex, são semelhantes ao do XHTML, facilitando assim a vida dos desenvolvedores de conteúdos.

O principal objetivo do Annodex é permitir a integração de arquivos de tempo-contínuo dentro das funcionalidades de procura e navegação existente na Web.

3.3. MPEG- 7

O MPEG-7 (CHIARIGLIONE, 2004) é uma especificação padronizada desenvolvida pelo Moving Picture Experts Group, um grupo de trabalho da ISO/IEC criado em 1988, encarregado pelo desenvolvimento de padrões para representação codificada de áudio e vídeo digitais.

O MPEG-7 descreve as características de conteúdos multimídia para que os usuários possam pesquisar e recuperar estes conteúdos de forma similar ao que ocorre nos mecanismos de busca na WWW. Para a criação destas descrições, o MPEG-7 disponibiliza um conjunto de elementos de metadados descritos em XML. Além das descrições textuais em arquivos XML, muitas vezes extensas, o MPEG-7 especifica descrições em formato binário, o BiM, permitindo o *streaming* e a compressão em até 98% das descrições.

```
<!-- ##### -->
<!-- Definition of ImageText DS -->
<!-- ##### -->
<complexType name="ImageTextType">
  <complexContent>
    <extension base="mpeg7:StillRegionType">
      <sequence>
        <element name="Text" type="mpeg7:TextualType" minOccurs="0"/>
      </sequence>
      <attribute name="textType" use="optional">
        <simpleType>
          <restriction base="string">
            <enumeration value="superimposed"/>
            <enumeration value="scene"/>
          </restriction>
        </simpleType>
      </attribute>
      <attribute name="fontSize" type="positiveInteger" use="optional"/>
      <attribute name="fontType" type="string" use="optional"/>
    </extension>
  </complexContent>
</complexType>
```

Figura 1 - Exemplo de Descrição MPEG 7

Diversos grupos de pesquisa e empresas do mercado já estão pesquisando ferramentas que utilizem o MPEG-7 como linguagem padrão de descrição de vídeo. Entre elas podemos destacar o *Ricoh MPEG-7 MovieTool*, utilizado para criar descrições, o *Canon MPEG-7 Spoken Content Transcription Service*, um serviço que transcreve em XML no padrão MPEG-7 um arquivo de áudio e o *IBM MPEG-7 Annotation Tool*.

3.3.1. IBM Annotation Tool

IBM Annotation Tool é uma ferramenta desenvolvida pela Alphaworks da IBM, um grupo criado pela IBM para pesquisar novas tecnologias. Esta ferramenta visa dar suporte para descrever sequências de vídeo gravadas no padrão de vídeo MPEG 1 e 2, através de metadados no padrão MPEG-7. Cada cena em uma sequência de vídeo pode ser descrita com descrições estáticas, descrições de objetos-chave, descrições de eventos e outras descrições léxicas. As descrições são associadas a cada cena e armazenadas como dados Mpeg 7 em um arquivo XML.

O IBM Annotation Tool requer também um arquivo de segmentação, no qual a sequência de vídeo é segmentada em pequenas unidades, chamadas de cenas. Este arquivo de segmentação pode ser carregado de outras fontes, como o "IBM CueVideo Shot Detection Tool Kit" ou segmentado automaticamente quando a ferramenta abre o arquivo de vídeo pela primeira vez.

3.4. Qual a melhor especificação para ser aplicada na Educação ?

As especificações descritas acima não são as únicas, nem as últimas voltadas para a descrição e anotação de vídeos digitais, mas com base nelas conseguimos ter um conjunto básico de opções disponíveis.

O MPEG 7 se destaca entre os padrões citados por ser um padrão desenvolvido dentro da ISO e, ao contrário do SMIL, por exemplo, ter sido criado especificamente com o propósito de descrever e indexar vídeos. Entretanto, o grande diferencial do MPEG-7 é sua abrangência, pois nele é possível fazer descrições que nos permitem indexar não somente palavras-chave em determinados intervalos do vídeo, como também permite a indexação por padrões de imagem, som e outros tipos de descrições. Desta maneira, podemos indexar um vídeo de tal forma que consigamos localizar quando determinada imagem, uma pessoa por exemplo, aparece no vídeo. O MPEG-7 pretende disponibilizar ferramentas para que a busca em imagens, vídeos e arquivos sonoros seja tão fácil quanto a busca em textos.

A aplicação do MPEG-7 na educação pode ocorrer de diversas formas, podendo ser utilizado para permitir a busca de determinadas palavras em um vídeo que sejam interessantes para uma aula, encapsulando informações adicionais de determinadas cenas e possibilitando a exploração pelos alunos, permitindo que alunos façam suas próprias anotações e observações em um determinado vídeo, etc. Um exemplo é o VEL (Virtual Entrepreneurship Lab), um ambiente virtual hipermídia para alunos de empreendedorismo, que permite a pesquisa de diferentes vídeos e informações relacionadas a eles através das descrições MPEG-7 (Virtual Entrepreneurship Lab, 2004).

4. Projeto SACCA

Como descrito anteriormente, existem diversas iniciativas no âmbito da criação de conteúdo audiovisual com vistas a sua aplicação na educação. A digitalização destes materiais e a criação de bibliotecas digitais é uma tendência cada vez mais apoiada no meio acadêmico. No sentido de catalogar conteúdos, foi desenvolvido por nosso grupo de trabalho o Projeto CESTA (FABRE, 2003), um repositório de objetos de aprendizagem que cataloga estes objetos seguindo o Learning Object Metadata (LOM), especificação da IEEE, visando descrever conteúdos educacionais. O SACCA (Sistema Automático de Catalogação de Conteúdo Audiovisual) é um projeto que visa complementar os esforços empreendidos no projeto CESTA, especificamente para a catalogação de vídeos digitais.

No âmbito da Internet2 já foi iniciado um experimento com catalogação automática *full content* em que o áudio do vídeo é analisado e palavras-chaves são identificadas e localizadas no tempo (em que momento são faladas). No projeto VALA, é usado o software comercial VIRAGE para fazer esta análise, mas este software somente funciona para a língua inglesa. Todavia, o resultado é muito relevante, pois possibilita ao aluno digitar uma ou mais palavras-chaves e receber uma coleção de URLs que apontam para vídeos armazenados no repositório (incluindo pontos específicos do vídeo onde a palavra é dita).

No presente projeto, pretende-se alcançar um resultado similar, mas operando inicialmente de forma não automática, porque não há no momento recursos (software de reconhecimento de áudio apropriado para o Português falado no Brasil). Assim, busca-se projetar e desenvolver um sistema que apoie o processo de catalogação *full content*, mas com a intermediação de um usuário que assiste ao vídeo e seleciona as palavras-chaves a serem incluídas na catalogação.

Será usado um software de domínio público que abre os tipos padrão de vídeo (MPEG), identifica e separa cenas. O IBM Annotation Tool pode ser utilizado para descrever arquivo de vídeo gravado no padrão MPEG (1 e 2), gerando uma descrição XML segundo o padrão MPEG 7. O sistema que está sendo proposto vai fazer o interfaceamento com este software e permitir ao usuário catalogador a digitação das palavras-chaves, agregando de forma semi-automática as informações de tempo de início em que as mesmas foram faladas.

As informações resultantes serão armazenadas em um banco de dados invertido, com todas as palavras-chaves catalogadas e ponteiros para os arquivos XML (MPEG 7) e para os registros do projeto CESTA, onde os objetos de aprendizagem tipo vídeo também estarão catalogados e onde estarão as informações referentes aos demais metadados de catalogação. O sistema CESTA será modificado de modo a permitir a interoperação com o sistema SACCA.

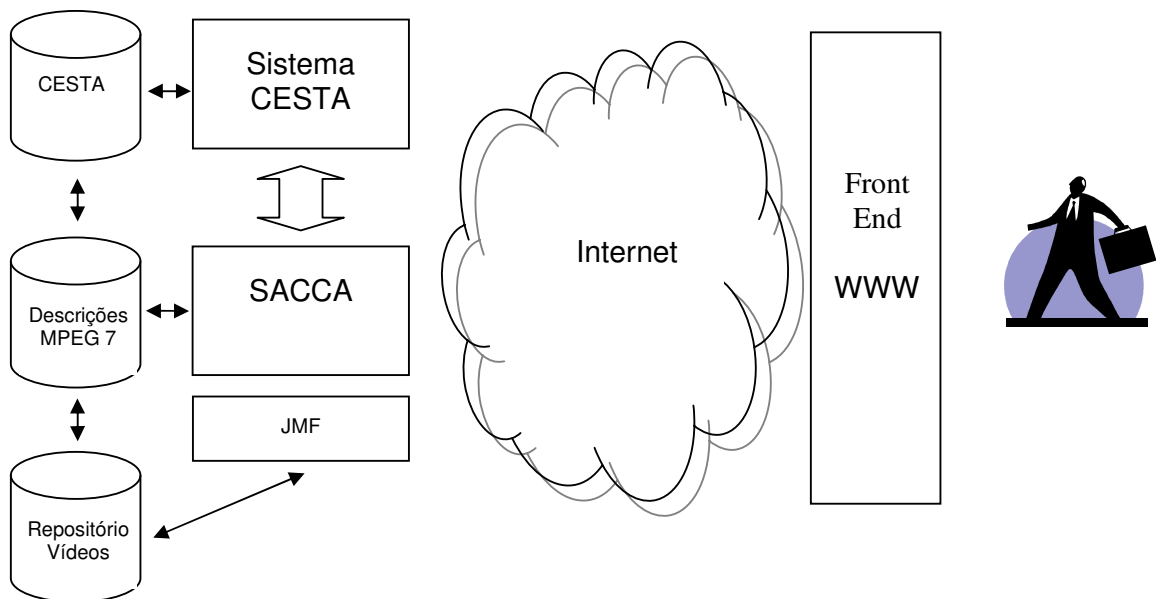


Figura 2 – Projeto SACCA – Esquema Geral

Um módulo de exibição será projetado e integrado a estes dois sistemas, com vistas a permitir a exibição dos vídeos, a partir de momentos referenciado por um link (uma URL associada a um instante em que uma dada palavra chave é falada). Este subsistema usará classes do Java Media Player, um subconjunto da API Java Media Framework (JMF), que permite que áudio, vídeo e outros tipos de mídia sejam adicionados em aplicações e applets desenvolvidos com tecnologia Java. Isso possibilitará a estas aplicações capturar, tocar, "stream" e transcodificar múltiplos formatos de mídia com total portabilidade.

O usuário utilizará o CESTA como ponto de partida, pesquisando e navegando nos diversos vídeos catalogados. Neste processo, o usuário poderá ver toda a descrição dos vídeos segundo as especificações do Learning Object Metadata do IEEE e também pesquisar por palavras-chaves do conteúdo dos vídeos. Isto irá ocorrer consultando o banco de dados invertido, uma extensão dos dados do CESTA, contendo todas as palavras-chaves gravadas no MPEG-7, o nome e a localização do arquivo de vídeo correspondente. A partir destes dados, o sistema abre o arquivo MPEG 7 e mapeia todas

as ocorrências da palavra-chave dentro do vídeo para mostrar ao usuário. Neste ponto, o usuário então poderá escolher qual o ponto exato que deseja assistir, ação que irá ocorrer através de um applet baseado no Java Media Player, alterado para tocar partes do vídeo a partir das descrições MPEG 7.

Conclusões

A utilização de vídeos educacionais digitalizados no processo de ensino-aprendizagem é um recurso multimídia importante diante crescente produção de vídeos educativos de qualidade. Acresce-se a este fato, o surgimento de padrões que auxiliam descrevê-los e indexá-los de forma padronizada, tais como o MPEG 7, possibilitando a busca e seleção de vídeos ou partes dele, integrados a conteúdos Web.

O Projeto SACCA apóia-se nestas tecnologias, integrando um repositório de conteúdos educacionais com conteúdos descritos em MPEG-7, possibilitando a busca, seleção e execução dos segmentos relacionados destes vídeos.

A utilização de tecnologias padronizadas como o MPEG-7, Html e Java, alia grande eficiência a uma total portabilidade e interoperabilidade, características estas muito importantes em qualquer aplicação acadêmica de grande alcance. Busca-se com este projeto oferecer à comunidade acadêmica uma ferramenta que auxilie alunos e educadores no enriquecimento de suas atividades acadêmicas.

Referências

CHIARIGLIONE, Leonardo. **The MPEG Home Page**. Tocantis: s.ed., 2004.
Disponível em: <<http://www.chiariglione.org/mpeg/>> Acesso em: 15 set. 2004.

FABRE, M. C. J. M.; TAROUCO, L. M. R.; TAMUSIUNAS, Fabrício R. **Reusabilidade de objetos educacionais**. RENOTE - Revista Novas Tecnologias na Educação. Porto Alegre: Centro Interdisciplinar de Novas Tecnologias na Educação (UFRGS), v. 1, n. 1, 2003. Disponível em: <http://www.cinted.ufrgs.br/renote/fev2003/artigos/marie_reusabilidade.pdf > Acesso em 22 set. 2004.

FUNDAÇÃO ROBERTO MARINHO. S. Local: Fundação Roberto Marinho, 2004.
Disponível em: <<http://www.frm.org.br>> Acesso em: 15 set. 2004.

MOORE, Michael G; KEARSLEY, Greg. **Distance Education: a systems view**. United States: Wadsworth Publishing, 1996.

Open standards for annotating and indexing networked media. S. local: ANNODEX, 2004. Disponível em: <<http://www.annodex.net/index.html>> Acesso em: 15 set. 2004.

Secretaria de Educação a Distância. Brasília: MEC, 2004. Disponível em: <<http://www.mec.gov.br/seed>> Acesso em: 15 set. 2004.



Synchronized Multimedia. S.local: W3C, 2004. Disponível em:
<<http://www.w3.org/AudioVideo/>> Acesso em: 15 set. 2004.

TV ESCOLA – O canal da Educação. Brasília: MEC, 2004. Disponível em:
<<http://www.mec.gov.br/seed/tvescola/default.shtm>> Acesso em: 15 set. 2004.

Virtual Entrepreneurship Lab. S. Local: VEL – 2004. Disponível em:
<http://www-i5.informatik.rwth-aachen.de/lehrstuhl/staff/klamma/Projects/virtual_entrepreneurship_lab.htm> Acesso em: 15 set. 2004.