

Desafios para a Localização e Recuperação de Informação em Conteúdos Educacionais Audiovisuais

Ismar F. Silveira, Paulo R. S. Borges

Programa de Pós-Graduação em Engenharia Elétrica (PPGEE)
Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

ismar.silveira@mackenzie.br, paulorenato.borges@mackenzista.com.br

***Abstract.** With the technology evolution, there was a meaningful increase of data volume in Web and of user's number, which occasioned in a big challenge the precision, efficiency and speed in the search and recovery of desired information. In the education context, the e-learning and the usage of audiovisual contents has been become popular and the size of data collections with these content types is growing. Hence it is necessary an efficient way to structure and organize these contents according to its meaning semantic. Therefore, the goal of this paper is to discuss about the challenges on semantic utilization to provide the process of search and retrieval information in the audiovisual educational contents.*

***Resumo.** Com a evolução tecnológica, houve um aumento significativo do volume de dados na Web e do número de usuários, acarretando em um grande desafio a rapidez, precisão e eficiência na localização e recuperação da informação desejada. No contexto da educação, o e-learning e o uso de conteúdos audiovisuais têm se popularizado e o tamanho das bases de dados com esses tipos de conteúdo está crescendo. Dessa forma, é necessário um modo eficiente de estruturar e organizar esses conteúdos de acordo com seu significado semântico. Assim, o objetivo deste trabalho é discutir os desafios da utilização da semântica para subsidiar o processo de localização e recuperação de informação em conteúdos educacionais audiovisuais.*

1. Introdução

Com a evolução tecnológica, houve um aumento significativo do volume de dados presentes na Web e do número de usuários, acarretando em um grande desafio, a rapidez, precisão e eficiência na localização e recuperação da informação desejada. Isso porque a busca e recuperação considerando essa grande quantidade e diversidade de dados e de usuários demandam uma forma eficaz de descrição e organização desses conteúdos Web, o que torna necessário um entendimento sobre as informações contidas nesses conteúdos.

A maior parte dos conteúdos digitais foi construída para a compreensão de humanos, não de softwares e isso faz com que a maioria das informações se torne inacessível pelos usuários. Assim, no início dos anos 1980, para melhorar a interação entre humanos e computadores surgiu a área de Interação Humano Computador (IHC), inicialmente como uma área de especialização em ciência da computação abrangendo ciência cognitiva e engenharia de fatores humanos (Carroll, 2012). Ainda nessa direção, em 2001, foi idealizada a Web Semântica ou Web dos Dados Conectados, como uma extensão da Web na qual a informação possui um significado bem definido e proporciona

um melhor modo de cooperação entre humanos e computadores (Berners-Lee, Hendler, Lassila, 2001).

De acordo com Berners-Lee, Hendler e Lassila (2001), na época, já existiam duas tecnologias importantes, que poderiam ser empregadas no desenvolvimento da Web Semântica: a *eXtensible markup Language (XML)* e o *Resource Description Framework (RDF)*, ambas criadas para facilitar o armazenamento e o compartilhamento de informações via Web. A XML é utilizada para criar *tags* e adicionar certa estrutura aos documentos e o RDF para atribuir significado a essa estrutura. Ainda, foi necessário um terceiro componente denominado ontologia, que consiste em um conjunto de especificações formais e possui uma taxonomia para definir classes de objetos e a relação entre esses objetos. Apesar dessas três tecnologias serem os pilares do funcionamento da Web Semântica até os dias atuais, com o passar dos anos, os conceitos da Web Semântica fundamentaram a criação de outras tecnologias que possibilitam a visão contemporânea do W3C sobre o termo Web Semântica como Web dos Dados Conectados.

Juntamente com a evolução da Web Semântica, a área de IHC passou por transformações que a fizeram crescer ao longo dos anos e atualmente agrega uma ampla quantidade de áreas de pesquisa e prática na informática centrada no ser humano.

Os conceitos e as ferramentas semânticas e de IHC têm sido utilizados em estudos nos mais variados segmentos como saúde, energia, economia, computação, etc., inclusive na área educacional, em que a aprendizagem *online* ou *e-learning* está sendo popularizada. Esse tipo de aprendizagem pode ser possibilitado por diversas tecnologias, mas necessita dos objetos digitais de aprendizagem (ODA), que são recursos digitais para auxiliar no processo educativo e passíveis de reutilização e disseminação pela internet (Borges, 2016). Os ODA podem ser de diferentes tipos como: textos; imagens; áudios; dentre outros, destacando-se os conteúdos audiovisuais pela sua capacidade de ampliar e dinamizar as habilidades cognitivas dos alunos, proporcionando-lhes uma experiência diferenciada por ser sensorial, visual e auditiva (Serafim e Sousa, 2011; Kate, Waghmare e Priyadarshi, 2015), o que justifica serem o foco desta investigação.

Assim como os demais conteúdos audiovisuais presentes na Web, os educacionais audiovisuais possuem uma heterogeneidade estrutural e semântica imensa que implica em diversos pontos a serem tratados, antes de se explorar o potencial da Web Semântica para a localização e recuperação de informações presentes nesses conteúdos. Essas questões, apesar de estarem ligadas essencialmente a Interação Humano Computador (IHC), englobam desafios em diferentes áreas como Educação, Semântica, Visão Computacional, Processamento de Imagem, Aprendizado de Máquina, dentre outras.

Considerando esses pressupostos, o objetivo deste estudo é discutir sobre os desafios no processo de localização e recuperação de informação em conteúdos educacionais audiovisuais, mais especificamente, apresentar sobre propostas de avanços na direção de solucionar esses desafios e mostrar a importância desses avanços para a educação.

Na seção 2, é apresentada uma revisão da literatura. Na seção 3, são evidenciadas as tecnologias existentes e os desafios em sua utilização. Na seção 4, são apontadas propostas para enfrentar os desafios nos processos de localização das informações presentes em conteúdos audiovisuais educacionais. E por fim, as conclusões do estudo.

2. Revisão da Literatura

O estudo de Wong e Reimann (2009) evidencia que há várias pesquisas com ênfase em experiências de utilização de conteúdos audiovisuais da Web, as quais apontam para a melhoria na interação dos usuários, reflexão colaborativa e aprendizagem. Mostra ainda que esses conteúdos estão sendo largamente utilizados na área educacional e, em decorrência, os bancos de dados desses vídeos crescem rapidamente (Wong e Reimann, 2009; Rawat, Bhatt e Kankanhalli, 2014). Localizar e recuperar informações em uma grande quantidade de dados, estejam eles estruturados ou não, tem sido desafiador. Muitos órgãos empresariais, governamentais e instituições de pesquisas têm investido no processamento de uma quantidade maciça de dados, porém, refletindo sobre os interesses dos usuários, a maior parte dos dados possuem informações que não lhes estão acessíveis. Dessa forma, é preciso tecnologias e ferramentas que sejam capazes de processar, transformar e analisar os dados, na direção de encontrar informações que possam ser recuperadas (Chakraborty, Padki e Bansal, 2017). Essa necessidade tem sido foco de estudos que abarcam sobre a utilização de semântica nesse processo, mais especificamente, para a incorporação de metadados nos conteúdos audiovisuais com o objetivo de facilitar o acesso às informações neles contidas.

Dong *et. al.* (2015) apresentam um algoritmo para a sumarização automática de vídeos com base no seu processamento, combinando a detecção das principais características visuais como faces humanas, cores e texturas com indicadores semânticos derivados dessas informações visuais e de texto. Em relação a anotação de segmentos de vídeos, Chamasemani *et. al.* (2015) e Yu, Cai e Liu (2016) trazem em seus estudos métodos para extrair as principais características das imagens contidas em um vídeo, selecionar os principais quadros e anotar esses quadros automaticamente utilizando conceitos semânticos e aprendizado de máquina. Já o estudo de Chou, Chen e Lee (2017) mostra um método que utiliza *Deep Learning* para efetuar a anotação automática de um vídeo, criando uma hierarquia de segmentação e utilizando características visuais similares, juntamente com palavras-chave semânticas, com a finalidade de agrupar os vídeos em *clusters* semânticos e, conseqüentemente, facilitar a recuperação de vídeos semelhantes.

Em relação aos trabalhos no âmbito educacional, a pesquisa de Rawat, Bhatt e Kankanhalli (2014) utiliza conceitos semânticos visuais e aprendizado de máquina para segmentar e anotar automaticamente vídeos de palestras com informações sobre o modo de ensinar do professor. Já o trabalho de Yang e Meinel (2014) concentra-se em uma abordagem que utiliza o áudio e textos presentes nos vídeos para extrair informações automaticamente. A investigação de Borges (2016), apesar de utilizar anotação semântica manual, mostra como a incorporação de metadados semânticos em conteúdos audiovisuais educacionais auxilia na melhoria dos processos de localização, recuperação e reutilização de informações neles contidas em um Portal Semântico Educacional. Belizário Júnior e Dorça (2016) apresentam um modelo que considera um programa de curso e mapeia suas seções em conceitos representados em ontologias para realizar a busca inteligente de objetos de aprendizagem e sua recomendação de acordo com os estilos cognitivos e de aprendizagem dos estudantes. Mendes *et. al.* (2016) anunciam uma proposta de agrupamento de conteúdos em grandes repositórios, visando facilitar e auxiliar o processo de recomendação de conteúdos referentes a estilos de aprendizagem de alunos.

Analisando os estudos apresentados nesta seção, observa-se que nenhum avalia a construção de modelos destinados à educação que englobam simultaneamente os aspectos educacionais, os conceitos de IHC, o processamento automatizado de conteúdos audiovisuais, organização semântica automática desses conteúdos com base nas informações neles contidas e recuperação de conteúdos de acordo com a informação buscada por determinado usuário, seja ele aluno ou professor. Dessa forma, a seção seguinte discute sobre os desafios que impactam na arquitetura desses modelos.

3. Tecnologias existentes e desafios em sua utilização

Quando se trata de conteúdos audiovisuais educacionais, a tarefa de extrair e recuperar informações é muito complexa, pois, com base na teoria da informação, a detecção e reconhecimento das informações presentes nesses conteúdos dependem de diversos fatores que vão desde os aspectos intrínsecos aos autores desses conteúdos, passando pela forma que são construídos e metodologias utilizadas, até as características dos usuários que buscam por uma informação. No que se relaciona aos autores dos conteúdos audiovisuais, cada professor tem uma forma de ensinar e apresentar os assuntos de interesse, além das preferências tecnológicas, área de atuação e nível de ensino.

Outro ponto a ser salientado é que os usuários que procuram por um conteúdo educacional audiovisual também possuem particularidades que influenciam na informação buscada. Nesse contexto, Pirolli e Card (1999) propuseram a teoria do Forrageamento de Informação, que descreve o comportamento para a recuperação de informação fazendo uma analogia com a teoria de forrageamento de alimentos. Os principais conceitos envolvidos nessa analogia são: a fonte de comida e o lugar no qual ela pode ser encontrada; as estratégias para encontrá-la; quais as ferramentas disponíveis para encontrá-la; e quais os benefícios no seu consumo. Esses conceitos podem ser traduzidos para o contexto da busca por informação como: a tomada de decisão para escolher uma fonte de informação sobre outra; como as informações estão agrupadas; o valor dessa informação para quem está buscando e o item de informação que é buscado ou encontrado em relação ao valor da informação desejada. Assim, de modo geral, o desafio de recuperar uma informação de um conteúdo encontra-se em considerar os conceitos presentes nessa analogia para arquitetar interfaces que realmente suportem esses conceitos com o objetivo de entender que fatores determinam as preferências de um usuário na busca de informações e suas estratégias de explorar as informações e consumi-las. Nessa direção, Pirolli e Card (1999) desenvolveram um modelo para ser utilizado para avaliar a tecnologia de recuperação de informação e entender o comportamento do usuário. Mas, criar esse modelo e utilizá-lo em uma aplicação real ainda é muito complexo, pois, apesar de toda tecnologia existente, depende de muitos fatores que não são inteiramente detalhados por esses autores.

Nesse contexto, além de entender o comportamento do usuário utilizando os conceitos de IHC, fica clara a necessidade de agrupar a informação de forma eficiente, sendo preciso uma organização dos conteúdos audiovisuais educacionais de acordo com as informações neles contidas, as quais podem incluir metadados como: títulos, sumários, data, autores, duração, tamanho do arquivo, formato, atores, legendas, dentre outros. Todas essas informações podem ser utilizadas para atribuir um significado semântico a esse vídeo, mas a maior parte dos vídeos presentes na Web não contém metadados suficientes para serem relevantes semanticamente (Hu *et. al.* 2011; Yang e Meinel, 2014). Apesar do significado semântico ser muito importante para localização e recuperação das

informações desses conteúdos, quando o processo de atribuição de semântica é manual, demanda muito tempo e é trabalhoso, o que sugere a necessidade da utilização de ferramentas para realizar a descrição e indexação automáticas (Yang e Meinel, 2014; Borges, 2016).

No que tange à incorporação automática de informações semânticas em conteúdos audiovisuais, a maior parte dos trabalhos existentes na atualidade efetuam o processamento do conteúdo visual e, geralmente, os vídeos são estruturados de acordo com uma hierarquia de clipes, cenas, fotos e quadros com o objetivo de segmentá-lo de acordo com o seu conteúdo semântico. Depois de segmentado o vídeo, os principais quadros, ou seja, aqueles que têm conteúdos mais relevantes, são analisados para realizar a extração de suas características visuais principais como cores, texturas, bordas, formas, textos e objetos, além de informações de movimentos. Em seguida, essas características principais são utilizadas para encontrar padrões de conteúdo; comportamento de objetos em movimentos; eventos e suas associações; e semântica conhecida de outros vídeos. Esses dados são então submetidos a regras que propiciam anotar os segmentos e classificar o vídeo. Mas, para o contexto educacional, foco desta pesquisa, nem sempre as técnicas baseadas no processamento do conteúdo visual podem ser aplicadas, já que vídeos de palestras ou aulas, por exemplo, são compostos por cenas homogêneas, tornando-se necessário processar também o áudio do vídeo. Para analisar o áudio do vídeo, primeiro é feita sua transcrição e posteriormente, organizada de acordo com os segmentos já estruturados segundo seu conteúdo visual, auxiliando na anotação e classificação semântica (Yang e Meinel, 2014). Entretanto, a diferença entre as informações extraídas automaticamente e a interpretação de um observador humano é que torna esse processo tão difícil (Hu *et. al.*, 2011; Yang e Meinel, 2014). Isso faz com que os resultados apresentados pelos estudos já existentes não sejam suficientes para serem utilizados como metodologia de extração automática de informações em conteúdos audiovisuais educacionais.

Uma vez que os conteúdos audiovisuais educacionais estão anotados e organizados semanticamente, a busca e recuperação de informações dos conteúdos em grandes repositórios ficam, de certa forma, facilitadas. Porém, ainda contam com entraves difíceis de resolver, pois as características e perfil do usuário que procura por uma informação também devem ser considerados para que os resultados lhe sejam relevantes. Existem ainda, inúmeras questões relacionadas à Educação, Pedagogia, Psicologia, Interação Humano Computador, dentre outras, as quais influenciam na acurácia dos resultados de acordo com o usuário que efetuou a busca, como por exemplo: palavras-chaves digitadas e o seu significado semântico; relação da busca atual com as buscas anteriores efetuadas por esse usuário; o tipo de vídeo educacional que esse usuário gosta; perfil do usuário e experiência como pesquisador; resultados acessados em uma pesquisa semelhante; estilo cognitivo de aprendizagem desse usuário; preferência desse usuário sobre a forma de visualizar os resultados, etc. Assim, a grande dificuldade consiste em encontrar as técnicas mais eficientes para a implementação de um modelo que seja capaz de englobar essas questões, dado que a linguagem natural utilizada pelos estudantes e professores não pode ser facilmente compreendida por algoritmos que automatizam esse processo na sua totalidade.

De maneira geral, pode-se afirmar que os principais desafios para a localização e recuperação de informação em conteúdos educacionais audiovisuais estão relacionados à IHC e semântica e isso requer que todo o processo seja subsidiado por um conjunto de

dados semânticos apropriado. Além disso, a acurácia nesse processo ainda depende da utilização de técnicas de processamento de imagem e aprendizado de máquina adequadas para que, não só ocorra a extração satisfatória de informações de um novo conteúdo, mas também uma atualização das informações dos conteúdos já existentes na base de dados fundamentada no aprendizado sobre os novos dados que estão sendo salvos.

Por fim, realizar todo esse processamento para localizar e recuperar informações em uma grande quantidade de conteúdos educacionais audiovisuais utilizando semântica pode demandar muito tempo, o que consiste em outro grande desafio, pois os recursos de hardware são limitados. Isto significa que muitas vezes a qualidade do processo fica prejudicada para não comprometer a experiência do usuário, principalmente na facilidade de acesso à informação.

Conforme mostrado nesta seção, apesar de existirem estudos sobre técnicas, metodologias e ferramentas para a localização e recuperação de informação em conteúdos audiovisuais, ainda são muitos os problemas em aberto que merecem a atenção de investigações futuras para que ocorram avanços nessa área. Dessa forma, a seção seguinte apresenta alguns desses problemas e possibilidades para reflexões sobre esses desafios.

4. Propostas para enfrentar os desafios

No que tange aos problemas relacionados à extração de informações, sugere-se pensar na construção de modelos e ontologias generalizáveis em larga escala utilizando técnicas multimodais para o processamento do áudio e do vídeo dos conteúdos educacionais audiovisuais, com o objetivo de subsidiar a extração automática de informações relevantes com maior acurácia. Considerando que, no contexto educacional, os conteúdos audiovisuais podem versar sobre variados temas, a acurácia dessa generalização está atrelada, em grande parte, a eliminação da dependência do domínio do conhecimento. Para tanto, é preciso que as características da percepção humana sejam melhor exploradas para que uma correspondência semântica dessas características seja empregada na extração automática de informações como forma de aumentar a acurácia das técnicas de processamento existentes. Esses modelos e ontologias generalizáveis podem ser utilizados também para a organização desses conteúdos em grandes repositórios de dados.

Em se tratando da busca e recuperação da informação desejada em grandes repositórios de dados, pode-se pensar na possibilidade de se estudar as maneiras de utilização de semânticas que sejam capazes de descrever comportamentos e sentimentos humanos com a finalidade de criar modelos cognitivos que representem o estilo e preferências do usuário.

É também oportuno que sejam investigadas metodologias para a combinação de técnicas de aprendizado de máquina de modo a obter mais acurácia em todas as etapas do processo de localização automática e recuperação de informações nos conteúdos educacionais audiovisuais.

Em suma, com base no princípio da Web Semântica de melhorar o trabalho em conjunto de humanos e máquinas (Berners-Lee, Hendler e Lassila, 2001), o sucesso das ações a serem tomadas para enfrentar os desafios evidenciados depende da utilização de conceitos de IHC, os quais são de grande importância para projetar sistemas interativos capazes de apoiar o modo com que os humanos se comunicam e interagem (Rogers, Sharp e Preece, 2013). Esses conceitos de IHC são essenciais para entender o comportamento do usuário e arquitetar soluções que o auxiliem na busca por informação.

Além de entender o estilo do usuário para trazer um conteúdo educacional audiovisual de acordo com a informação buscada, é preciso utilizar os princípios de usabilidade para construir as interfaces de usuário com o objetivo de facilitar a visualização das informações em uma grande quantidade de dados e compreender qual a relação existente entre essas informações. Em decorrência, o usuário poderá gastar mais tempo consumindo a informação do que buscando por ela. E essa deve ser a principal motivação para construir soluções para os desafios discutidos neste trabalho.

Para avaliar o sucesso e medir o impacto das ações sugeridas na direção de solucionar os desafios apresentados, é necessário a atuação de equipes multidisciplinares para que sejam aplicados os conceitos de IHC e reunidas habilidades de diversas áreas em prol do *design* iterativo (que envolve conceber uma solução, construir uma versão interativa e avaliá-la) de metodologias e ferramentas que auxiliem no processo de localização e recuperação de informações em conteúdos audiovisuais educacionais.

5. Conclusões

Este estudo evidenciou o cenário atual das pesquisas sobre a utilização de ferramentas e conceitos da semântica para localizar e recuperar informações em conteúdos educacionais audiovisuais. Além disso, contribuiu com propostas para avanços no emprego da semântica nos processos de reconhecimento das informações presentes em conteúdos educacionais audiovisuais, anotação semântica automática em seus segmentos e estabelecimento de modelos generalizáveis para a recuperação da informação de um conteúdo a partir de características e comportamentos do usuário. Nessa direção, espera-se que esta investigação possa servir de subsídio para reflexões e discussões sobre a indexação e recuperação semântica de informações em conteúdos educacionais audiovisuais.

Referências

- Belizário Júnior, C. F., e Dorça, F. A. (2016) “Uma Abordagem para a Recomendação Inteligente de Conteúdo Web na Criação de Materiais e Objetos de Aprendizagem”. ANAIS DO CSBC 2016, 5 DesafIE!, PUCRS.
- Berners-Lee, T.; Hendler, J.; Lassila, O. (2001). “The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities”. [S.l.: s.n.], 34+ p. ISSN 0036-8733.
- Borges, P. R. S. (2016). “Incorporação e recuperação de metadados semânticos em vídeos educacionais”. 110 f. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Presbiteriana Mackenzie, São Paulo, 2016.
- Carrol, J. M. “Human Computer Interaction - brief intro”. In: Soegaard, Mads, and Rikke Friis Dam. "The Encyclopedia of Human-Computer Interaction." The Encyclopedia of Human-Computer Interaction (2012).
- Chakraborty, J. Padki, A. e Bansal, S. K. (2017) "Semantic ETL — State-of-the-Art and Open Research Challenges," 2017 IEEE 11th International Conference on Semantic Computing (ICSC), San Diego, CA, USA, pp. 413-418.
- Chamasemani, F. F. , Affendey, L. S., Mustapha, N. e Khalid, F. (2015). “Automatic Video Annotation Framework Using Concept Detectors”. Journal of Applied Sciences, 15: 256-263.

- Chou, C. L. , Chen, H. T. e Lee, S. Y. (2017). "Multimodal Video-to-Near-Scene Annotation," in IEEE Transactions on Multimedia, vol. 19, no. 2, pp. 354-366, Feb. 2017.
- Dong, P., Xia, Y., Wang, S., Zhuo, L. Feng, D. D. (2015) "An iteratively reweighting algorithm for dynamic video summarization. Multimedia Tools and Applications", p. 1-25.
- Kate, L. S. Waghmare, M. M. e Priyadarshi, A. (2015). "An approach for automated video indexing and video search in large lecture video archives," 2015 International Conference on Pervasive Computing (ICPC), Pune, 2015, pp. 1-5.
- Mendes, M. M., de Carvalho, V. C., Dorça, F. A., & Cattelan, R. G. (2016) "Uma Abordagem para a Organização Automática de Objetos de Aprendizagem em Ambientes Educacionais Ubíquos com Base em Estilos de Aprendizagem". ANAIS DO CSBC 2016, 5 Desafie!, PUCRS.
- Pirolli, Peter, Card, Stuart K. (1999): Information foraging. In Psychological Review, 106 (4) pp. 643-675.
- Rawat, Y. S., Bhatt, C. e Kankanhalli, M. S. (2014) "Mode of teaching based segmentation and annotation of video lectures," 2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI), Klagenfurt, 2014, pp. 1-4.
- Rogers, Y., Sharp, H. e Preece, J. (2013) "Design de interação: além da interação humano-computador. Bookman".
- S. Yu, H. Cai and A. Liu. (2016) "Multi-semantic Video Annotation with Semantic Network," 2016 International Conference on Cyberworlds (CW), Chongqing, 2016, pp. 239-242.
- Serafim, M. L.; Sousa, R. P. d. Multimídia na educação: o vídeo digital integrado ao contexto escolar. [S.l.: s.n.], 2011. www.teses.usp.br/teses/disponiveis/55/55134/tde.../tesePrazeres.pdf. [Online; Acesso em: 17-out-2014].
- Hu, W. Xie, N., Li, L., Zeng, X. e Maybank, S. (2011)"A Survey on Visual ContentBased Video Indexing and Retrieval," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 41, no. 6, pp. 797-819, Nov. 2011.
- Wong, W. Y. e Reimann, P. (2009) "Web Based Educational Video Teaching and Learning Platform with Collaborative Annotation," 2009 Ninth IEEE International Conference on Advanced Learning Technologies, Riga, 2009, pp. 696-700.
- Yang, H. e Meinel, C. (2014) "Content Based Lecture Video Retrieval Using Speech and Video Text Information," in IEEE Transactions on Learning Technologies, vol. 7, no. 2, pp. 142-154, April-June 2014.