

Personalização de Conteúdos Multimédia.

Análise aos atributos relevantes para a sua anotação

Inês Oliveira¹

Introdução

Com o aparecimento dos acessos de banda larga, a globalização da informação multimédia tornou-se uma realidade incontestável. No entanto, embora se verifiquem avanços tecnológicos significativos, nomeadamente nas taxas de compressão, na velocidade e capacidade das redes, persistem dificuldades na recuperação da informação audiovisual [Qun 2001, Dimitrova 2002].

A complexidade é introduzida por algumas das qualidades intrínsecas à informação audiovisual, nas quais se incluem a elevada quantidade de dados, a diversidade de estruturas e heterogeneidade de tipos de media, mas sobretudo pela própria subjectividade da indexação. Na verdade, embora exista um conjunto diversificado de convenções e formatos com meta-dados [Koenen 2001, AAF 2002, Pro-Mpeg 2002] que abrangem um leque alargado de atributos, a indexação mantém-se subjectiva, quer seja manual ou automática:

- quando é manual, reflecte frequentemente a subjectividade de quem anota;
- quando é automática, espelha as propriedades da informação consideradas relevantes por quem programa.

Por esta razão, a probabilidade da indexação reflectir os critérios pelos quais os utilizadores pesquisam e personalizam os seus conteúdos torna-se reduzida.

Uma alternativa para aumentar a relevância dos atributos utilizados para indexar vídeo passa pela sua contextualização e personalização. A questão da contextualização dos conteúdos tem vindo a ser estudada com algum detalhe no âmbito das técnicas de análise de conteúdo, sendo considerada uma componente importantíssima da análise. Partindo deste conceito, será discutida neste artigo a sua aplicabilidade no âmbito da anotação e recuperação de conteúdos multimédia.

Organização do Documento

Na **secção 2** são apresentadas as definições de conteúdo e contexto, segundo a perspectiva da análise de conteúdos. Na **secção 3** é discutida a sua aplicação actual no âmbito da multimédia, e no seu seguimento, a **secção 4** apresenta algumas sistemas e abordagens actualmente existentes. Na **secção 5** é analisada a aplicabilidade da contextualização à recuperação de conteúdos multimédia. Finalmente, nas **secções 6 e 7**, encontram-se as conclusões e as referências.

Conteúdos e contextos – Os conceitos da análise de conteúdos

A necessidade de contextualização dos conteúdos é apresentada por Krippendorff [Krippendorff 2004] a partir da sua definição de conteúdo como algo que emerge do processo de análise de um “texto”² relativamente a um contexto particular. Esta abordagem fundamenta-se essencialmente nos seguintes aspectos que caracterizam os “textos”:

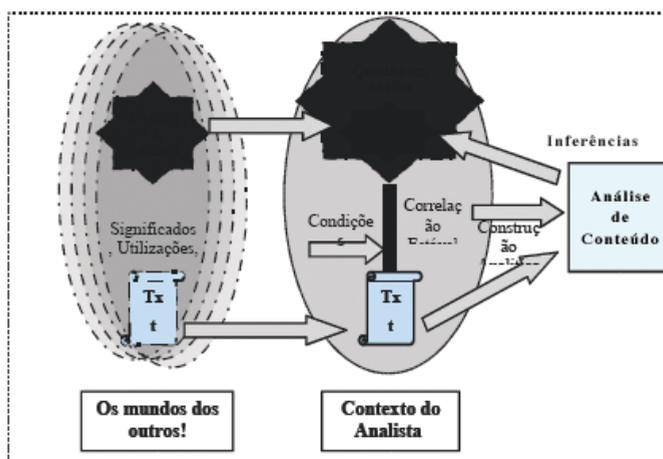
- os seus significados são sempre relativos a contextos, discursos e objectivos particulares;
- não têm qualidades independentes dos “leitores”, e portanto não têm um significado único que possa ser descoberto, identificado ou descrito;
- não “contêm” ou “possuem” os significados, uma vez que estes informam os leitores, invocam sentimentos e provocam mudanças comportamentais.

O contexto é aliás parte integrante da *framework* para a análise de conteúdos sugerida por Krippendorff (Figura 1).

Conteúdos e contextos em multimédia

Na área de multimédia o entendimento de conteúdo é sobretudo que este é inerente

Figura 1 - A framework da análise de conteúdo por Kripendorff



ao formato, daí ter-se apostado no desenvolvimento de algoritmos de processamento e extração de características da informação audiovisual [Oliveira 1997, Zhao 2002]. O esforço efectuado tem-se centrado em conseguir uma representação objectiva e consensual do conteúdo do vídeo [Davis 1995].

No que se refere à questão propriamente dita da contextualização, esta tem sido abordada por diversas vezes na área de multimédia, mas sobretudo relativamente à temporalidade do vídeo. Davenport et al. [Davenport 1991], por exemplo, apresentam já o contexto como o significado adicional representado no *shot*³ com base na sua adjacência com outros *shots* ou relacionado com o domínio de conhecimento. Posteriormente, Davis [Davis 1995] fala do distanciamento do contexto de utilização relativamente à anotação utilizada pelos editores de vídeo. Este autor sublinha a influência da montagem na semântica do vídeo, e constrói uma linguagem para anotação icónica, mas que se destina unicamente a representar os aspectos semanticamente invariantes da informação, isto é, independentes de contexto.

Butler et al., [Butler 1996] definem também contexto, ordem e significado no sentido temporal do termo, apresentando o sistema LIVE. Este sistema destina-se à construção de sequências de vídeo com significado a partir de uma base de dados

contendo segmentos de vídeo textualmente anotados.

Abordagens e sistemas existentes

Nesta secção descrevem-se algumas aproximações que de formas diversas permitem reduzir a distância entre os contextos da anotação e pesquisa.

Ontologias

A criação de ontologias (<http://www.w3.org/2001/sw/WebOnt>) procura no âmbito da anotação garantir que conteúdos semanticamente semelhantes sejam cruzados, mesmo quando descritos por sujeitos distintos com vocabulários distintos. Isto permitiria por exemplo que uma anotação fosse pesquisada por critérios semanticamente semelhantes, mas não iguais.

No entanto, as ontologias não resolvem a questão dos diversos grupos de interesse que podem criar e pesquisar meta-dados segundo conceitos e objectivos distintos. Neste âmbito Shabajee et al. [Shabajee 2002] apresentam uma abordagem com base em comunidades de interesse, descrita no ponto seguinte.

Anotação comunitária

A identificação das várias comunidades de interesse, que poderão aceder ao um

determinado repositório de informação irá permitir ajustá-lo melhor às suas necessidades específicas. Shabajee et al. [Shabajee 2002] propõem ainda que essas comunidades efectuem as suas próprias anotações. Para resolver o problema da qualidade, fiabilidade e relevância permitem quatro modos de anotação, relacionados com diferentes tipos de acesso:

- **Comunidades de Confiança (Trusted):** organizações que fornecem e validam informação e que se consideram seguras sob o ponto de vista da validade e relevância.

- **Comunidades Moderadas:** organizações que fornecem e validam informação, sendo responsáveis pela sua própria moderação e administração, por exemplo através de um moderador de forma similar a um fórum moderado.

- **Anotações Abertas:** Podem ser feitas por qualquer utilizador, e ser moderadas ou não.

- **Anotações de Terceiros:** Possibilidade dos utilizadores produzirem sites externos com as suas próprias anotações, sobre os conteúdos do sistema, não existindo neste caso um controlo efectivo sobre essa informação.

A qualidade da informação a que os utilizadores terão acesso, estará dependente do tipo de anotação.

Normas e formatos com meta-dados

Apresentam-se de seguida algumas normas e formatos com meta-dados, que contribuem de forma importante para a uniformização da anotação.

DCMI - Dublin Core Metadata Initiative

O DCMI (<http://dublincore.org>) é um fórum aberto para o desenvolvimento de normas interoperáveis de meta-dados que suportem um conjunto alargado de objectivos e modelos de negócio.

Um dos princípios pelos quais esta norma se rege faz alusão à problemática dos contextos de utilização e anotação: *“As boas práticas para um dado elemento ou qualificador podem variar com o contexto, [...] O requisito de utilidade para a sua pesquisa não deve por isso ser esquecido”*. Os termos principais que são utilizados pelo DMCI são os seguintes:

- Título
- Contribuidor
- Cobertura
- Assunto
- Identificador
- Data
- Autor
- Língua
- Formato
- Editor
- Relação
- Tipo
- Direitos
- Descrição
- Fonte

MPEG - Moving Picture Experts Group

O MPEG é uma família de normas não proprietárias de compressão de vídeo. Dentro destas, o MPEG-7 convencionou mecanismos para descrever a estrutura e a semântica de conteúdos multimédia. O objectivo desde formato é aumentar a eficiência do acesso à informação audiovisual e tornar possível a sua pesquisa e filtragem. A informação que pode ser guardada pelo MPEG-7 é a seguinte [Salembier 2001] (Figura 2):

O contexto de utilização é suportado por este formato, através do descritor para Interacção do utilizador com o conteúdo. As preferências podem ser descritas para dife-

rentes tipos de conteúdo e formas de navegação, permitindo dependências do contexto em termos de temporais e espaciais.

O formato MPEG-21 tem como objectivo principal permitir o acesso universal aos conteúdos multimédia. Esta norma unifica a descrição dos ambientes de utilização, englobando-se aqui redes, terminais e condições de acesso, permitindo que um dado conteúdo se adapte dinamicamente face a determinadas circunstâncias de consumo [Koenen 2001].

O MPEG-21 permite ainda expressões sobre os direitos relativamente à propriedade intelectual, completando o MPEG-7, razão

Figura 2 - Informação que pode ser guardada nas descrições MPEG-7

Tipo de Descritor	Exemplos de Informação
Processo de criação e produção Utilização do conteúdo	director, título, etc. copyright, história de utilização, horário de transmissão,...
Armazenamento do conteúdo	Formato e codificação
Componentes espacio-temporais	Cortes, segmentação espacial, detecção do movimento,...
Propriedades de baixo nível Realidade capturada	Cor, texturas, timbres, descrição da melodia Objectos, eventos e a interacção entre eles
Formas de navegação alternativas Colecções de objectos	Sumários, variações,...
Interacção do utilizador com o conteúdo	Preferências e história de utilização

pela qual existem já sistemas que utilizam ambas as normas [Steiger 2003, Tseng 2004].

AAF - Advanced Authoring Format

O AAF [AAF 2002] foi lançado em 2000 e é uma norma para a pós-produção e autoria de conteúdos multimédia. Este formato permite que os criadores dos conteúdos troquem facilmente informação audiovisual e meta-dados entre aplicações e plataformas. O modelo do AAF suporta as seguintes categorias de meta-dados (Figura 3).

MXF - Material eXchange Format

O MXF [Pro-Mpeg 2002] é um formato não proprietário muito recente, fundamentalmente direccionado para a troca de conteúdos audiovisuais associados com dados e meta-dados.

Como a informação que pode ser guardada sob a forma de meta-dados é infundável, este formato permite filtrar o que é relevante para um determinado contexto operacional. Inclui os seguintes tipos de meta-dados: a estrutura de ficheiros, os próprios conteúdos, palavras-chave ou títulos, notas de edição, localização, tempo, data, versão, etc.

Modelos de domínio

Os modelos de domínio procuram representar a informação multimédia com conhecimento acerca do seu domínio. Estes modelos restringem o contexto de utilização, sendo sobretudo utilizados para a segmentação e/ou classificação.

Zhang et al. [Zhang 1994], por exemplo, utilizaram o modelo do domínio dos noti-

ciários televisivos para segmentar e identificar os vários segmentos vídeo que os compõem. Isto foi feito com base em conhecimento sobre a estrutura espacial e temporal típica deste tipo de informação.

Fisher et al. [Fisher 1995] procuraram classificar programas televisivos com base nos chamados perfis de estilo, um género de assinatura contendo aspectos característicos de uma determinada classe de programas.

Mais recentemente, Xie et al. [Xie 2003] apresentam algoritmos para a análise da estrutura de vídeos de jogos de futebol utilizando conhecimento do domínio. É com base nestes algoritmos que estes autores efectuam a segmentação temporal da informação e a classificação automática dos segmentos obtidos.

Pesquisas e anotações em formato audiovisual

Vários autores referem as limitações das anotações textuais quando utilizadas para representar uma série de aspectos que existem nos conteúdos multimédia [Davis 1995, Elmagarmid 1997], pelo que existem alguns sistemas que optaram por permitir anotações e pesquisas em formato audiovisual.

Davis et al. [Davis 1995], por exemplo, criaram uma linguagem de anotação icónica para o sistema *MediaStreams*, para descrever os aspectos objectivos do conteúdo do vídeo. Na área do áudio, Ghias et al. [Ghias 1995] permitem a recuperação de uma dada melodia simplesmente cantarolando-a (*by humming*).

Na área da imagem, o sistema QBIC da IBM (*Query By Image Content*) permite

Figura 3 – As categorias suportadas pelo modelo do AAF

Categoria	Descrição
Identificação e Localização	Permite identificar e localizar um dado elemento seja ele dados ou meta-dados.
Administração	Permite definir os direitos, o nível de acesso, classificações de segurança, etc.
Interpretação	Permite anotação manual, incluindo por exemplo artistas, organizações, etc.
Processual	Permite descrever a forma como os dados são "montados".
Relacional	Permite descrever relações entre dados e meta-dados.
Espacio-temporal	Permite descrever lugares e tempo, incluindo ângulos, coordenadas, datas, etc.

pesquisar bases de dados de imagens (<http://www.qbic.almaden.ibm.com/>). Esta pesquisa é efectuada quer por objectos, quer por imagens, utilizando propriedades como: cor média, histogramas, textura, forma, esboço, localização e desenho.

Contextos na recuperação de conteúdos

Aplicando o conceito de contextualização à recuperação de conteúdos observa-se logo à partida a existência de três contextos distintos:

- o do **criador** dos conteúdos, que os cria num dado contexto e com determinado objectivo;
- o do **anotador** dos conteúdos, que os anota num dado contexto e com determinado objectivo;
- e o do **utilizador** final que os pesquisa num dado contexto e com determinado objectivo.

Para a recuperação ser bem sucedida é preciso que exista pelo menos um contexto de anotação que coincida com o contexto de utilização, podendo considerar-se irrelevante o contexto do criador dos conteúdos. Este processo é contudo complexo, veja-se por exemplo um caso simples ilustrativo: digamos que o nosso utilizador, que por acaso é arquitecto, pesquisa a palavra “planta”. A primeira interrogação que surge é: Será que é uma planta de um edifício ou uma planta no sentido biológico do termo?

Mas as dificuldades não terminam por aqui. Será que o anotador teve em conta a possibilidade de os utilizadores do sistema pesquisarem as plantas dos edifícios? Qual

é que terá sido o contexto da sua anotação? Arquitectura? Ciências? O ideal seriam as duas, mas qual será o número de contextos previsível para os utilizadores do nosso sistema?

Abordagens para a aproximação dos contextos

Existem algumas abordagens que se podem conjugar para aproximar os contextos de anotação e utilização, nomeadamente:

- contextualizar o utilizador no contexto da anotação,
- contextualizar a anotação no contexto do utilizador,
- e de acordo com a *framework* da análise de conteúdo, validar aos resultados obtidos

Técnicas para Contextualizar o Utilizador no Contexto da Anotação

A contextualização do utilizador na anotação pode acontecer em momentos distintos:

- **antes da pesquisa**, reforçando, por exemplo, as mensagens visuais do nosso sistema e disponibilizando manuais *on-line*.

Partindo novamente do caso da pesquisa por “planta”, uma das alternativas seria, por exemplo, acrescentar uma breve introdução sobre os objectivos da anotação (biologia ou arquitectura), para o utilizador perceber logo à partida o seu contexto e aquilo que pode esperar como resultado.

- **durante a pesquisa**, por exemplo, encaminhando o utilizador com base num dicionário de sinónimos e sugerindo-lhe

alternativas, ou mesmo utilizando mecanismos de personalização.

No caso da pesquisa por “planta”, por exemplo, se a anotação do sistema incluir de facto os contextos de Arquitectura e Biologia, podem-se sugerir ao nosso utilizador as alternativas de pesquisa “vegetal” e “planta de edifício”, direccionando-o e permitindo-lhe compreender melhor o contexto da anotação.

- **depois da pesquisa** (Ver secção).

Contextualizar a anotação no contexto do utilizador

A contextualização da anotação pode também processar-se em três momentos:

- **antes da anotação:** Isto requer, por exemplo, a prévia identificação dos tipos de comunidades existentes e das suas necessidades de anotação, e construir a anotação com base nesta informação. A identificação das comunidades e a classificação dos utilizadores numa dessas comunidades, pode ser conseguida por exemplo através de inquéritos prévios e fazendo com que todos os utilizadores sejam registados.

Neste âmbito, podem-se ainda fazer uso de mecanismos de personalização e também permitir que os próprios utilizadores colaborem na anotação.

- **durante a pesquisa**, por exemplo, solicitando de forma explícita ao utilizador os objectivos da sua pesquisa. Neste âmbito, podem-se ainda utilizar dicionários para que, com base no perfil de utilizador, se tenha acesso a um conjunto termos relacionados.

- **depois da pesquisa** (Ver secção).

Validar os resultados da recuperação

A validação dos resultados obtidos pelo sistema permite não só aferir o sucesso da pesquisa, mas também afinar e enriquecer a anotação, possibilitando a aprendizagem do

sistema. Neste âmbito, podem-se por exemplo:

- realizar inquéritos para avaliação dos resultados periódicos ou mesmo *on-line*;
- solicitar a colaboração dos utilizadores no enriquecimento da anotação;
- e analisar a historia de utilização do sistema, verificando, por exemplo, as tentativas dos utilizadores e os caminhos por estes adoptados.

Conclusões e trabalho futuro

A definição de contexto e conteúdo apresentada no ponto é aplicável à recuperação de conteúdos multimédia. O contexto é aflorado no âmbito da multimédia, mas sobretudo no que se refere à sua temporalidade. Actualmente existem várias abordagens que permitem de formas diversas aproximar os contextos de utilização e anotação, mas em geral nesta área o conteúdo continua a ser tido como uma coisa inerente ao formato.

A combinação destas abordagens com alguns dos mecanismos referidos no ponto anterior pode efectivamente aproximar os contextos de utilização e anotação, e deste modo melhor os resultados das pesquisas. A validação dos resultados obtidos e a utilização desta informação para refinar e afinar o sistema é, neste âmbito, um aspecto importante a considerar.

Com base nisto, e como trabalho futuro, pretende-se desenvolver um protótipo para o arquivo de vídeo da faculdade utilizando algumas destas técnicas. A comunidade será fechada consistindo numa primeira fase nos professores e alunos da licenciatura de Cinema, Vídeo e Multimédia. Pretende-se deste modo elaborar inquéritos prévios para identificar as suas necessidades de anotação, verificar os resultados obtidos e solicitar a sua colaboração para a anotação do sistema.

Bibliografia

[AAF 2002] AAF Association, *Advanced Authoring Format Specification, Version V1.0.1*, 2002.

[Butler 1996] S. Butler, A. P. Parkes. "Automatic Vídeo Editing by Filmic Decomposition of Non-Filmic Queries". *Multimedia Technology and Applications*, Springer, 1996.

[Davenport 1991] Gloriaana Davenport, Thonas A. Smith, Natalio Pincever. "Cinematic Primitives for Multimedia". *Proceedings of IEEE Computer Graphics & Applications*, 1991.

[Davis 1995] Marc Davis. "Media Streams: An Iconic Visual Language for Vídeo Representation". *Reedings in Human-Computer Interaction: Toward the Year 2000*. Morgan Kaufman Publishers Inc., 1995.

[Dimitrova 2002] Nevenka Dimitrova, Hong-Jiang Zhang, Behzad Shahraray, Ibrahim Sezan, Thomas Huang e Avideh Zakhor. "Applications of Video-Content Analysis and Retrieval", *IEEE Multimedia*, 2002.

[Elmagarmi d 1997] Ahmed K. Elmagarmid, hatao Jiang, Abdelsalam Helal, Anupam Joshi e Magdy Ahmed. *Video Database Systems: Issues, products and Applications*, Kluwer Academic Publications, 1997.

[Fisher 1995] Stephan Fisher, Rainer Lienhart, Wofgag Effelsberg. *Automatic Recognition of Film Genres*. Technical Report TR-95-006, University of Manhnheim, 1995.

[Ghias 1995] Asif Ghias, Jonathan Logan, David Chaberlin, Brian S. Smith. "Query By Humming". *Proceedings of ACM Multimedia' 95*, 1995.

[Koenen 2001] Rob Koenen. "From MPEG-1 to MPEG-21: Creating an Interoperable Multimedia Infrastructure", *ISO/IEC JTC1/SC29/WG11 N4518*, Dezembro 2001.

[Krippendorff 2004] Klaus Krippendorff. *Content Analysis: An Introduction to Its Methodology – Second Edition*. Sage Publications, 2004.

[Oliveira 1997] Inês Oliveira, Nuno Correia, Nuno Guimarães. "Image processing

Techniques for Video Content Extraction". *Proceedings of 4th Dellos Workshop*, San Mineato, Itália, Agosto 1997.

[Pro-Mpeg 2002] Pro-Mpeg Fórum, "Working Together with MFX", 2002.

[Qun 2001] Li-Qun Xua, Jian Zhu e Fred Stentiford. "Video summarization and semantics editing tools." *Proceedings of Storage and Retrieval for Media Databases 2001*, vol. SPIE 4315, San Jose, EUA, 2001.

[Salembier 2001] Philippe Salembier, "MPEG-7 Multimedia Description Schemes". 748 *IEEE Transactions on Circuits and Systems for Video Technology*, 2001.

[Shabajee 2002] Paul Shabajee, Libby Miller. "Adding Value To Large Multimedia Collections Through Annotation Technologies And Tools: serving Communities Of Interest", *Papers Museums and the Web*, 2002.

[Steiger 2003] Olivier Steiger, Touradj Ebrahimi e David Marim'on Sanju'an. "MPEG-Based Personalized Content Delivery". *Proceedings IEEE Int. Conf. on Image Processing, ICIP2003*, 2003.

[Tseng 2004] Belle L. Tseng, ching-yung lin e John R. Smith. "Using MPEG-21 and MPEG-7 for Personalizing Video", *IEEE Computer Society*, 2004.

[Xie 2003] Lexing Xie, Peng Xu e Shih-Fu Chang. "Structure Analysis of Soccer Video with Domain Knowledge and Hidden Markov Models" *Elsevier Science*, 2003

[Zhang 1994] H. ZHANG, g. Yihong, S. Smoliar, T. Yong. "Automatic Parsing of News Video". *Proceedings of IEE ICMS'94 Conference*, 1994.

[Zhao 2002] Rong Zhao, William I. Grosky. "Content-Based Retrieval and Image Database Techniques", *Bridging the Semantic Gap- Part II*, Wayne State University, USA, 2002.

¹ Universidade Lusófona, Departamento de Ciências da Comunicação, Artes e Tecnologias da Informação.

² Como *texto* entende-se um conteúdo qualquer, texto, imagem, vídeo, etc.

³ Como *shot* entende-se um segmento de vídeo contínuo, sem cortes ou quaisquer outras transições.