

**FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO**

**Vídeo Escalável e Interoperabilidade  
em Redes Multimédia**

Pedro Miguel Soares Vieira  
Licenciado em Engenharia Publicitária pela  
Universidade Fernando Pessoa

**Dissertação submetida para satisfação  
parcial dos requisitos do grau de mestre em  
Tecnologia Multimédia**

Dissertação realizada sob a supervisão de  
Professor Doutor Luís Corte-Real  
do Departamento de Engenharia Electrotécnica e de computadores  
da Faculdade de Engenharia da Universidade do Porto

Porto, Novembro de 2006

## Resumo

A diversidade de dispositivos e de redes multimédia com capacidade de reprodução de conteúdos audiovisuais tem registado um aumento significativo nos últimos anos. Torna-se por isso importante criar conteúdos capazes de responderem às solicitações provenientes de diferentes contextos de utilização.

O processo de adaptação destes conteúdos audiovisuais é o núcleo desta dissertação. É desenvolvido um trabalho de análise de conteúdos com vista à classificação dos diversos conteúdos multimédia. Com base nessa classificação, e tomando como base de trabalho a *framework* edVO[1], é depois proposto um modelo de adaptação de conteúdos. Este modelo de adaptação deverá realizar a adaptação em função das condições de utilização como a própria natureza do conteúdo adaptado.

O resultado das propostas no âmbito da classificação dos conteúdos e sua adaptação é reunido num módulo de adaptação que permite acrescentar ao edVO a capacidade de criar e distribuir conteúdos adaptáveis.

# Abstract

The great number of multimedia devices and networks capable of transmitting audiovisual content has been growing increasingly in the last years. Taking this phenomenon under consideration it becomes important to create audiovisual contents capable of meeting the demands of various contexts of use.

The adaptation process of those audiovisual contents is the core of this thesis. Taking as starting point and work base the edVO[1] framework, this thesis develops an analysis of the different types of content regarding its classification. Based on that classification is proposed a content adaptation model that allows that the adaptation to be executed considering both the context of use and the nature of the content itself.

The result of the proposal concerning the content classification and its adaptation is then assembled in an adaptation module which allows to add to edVO the ability of create and distribute scalable contents.

## Prefácio

A sociedade de informação em que actualmente vivemos origina a quase constante emergência de novas formas de comunicarmos e de nos relacionarmos com o mundo em nosso redor. A procura incessante por um meio que garanta a ubiquidade da sua mensagem levou o homem a criar diferentes e cada vez mais engenhosos meios de este se expressar. As pinturas rupestres na pré-história, a impressão com recurso a caracteres móveis e o telégrafo são exemplos considerados primitivos, numa época em que redes de satélites permitem transmitir simultaneamente para todo o globo e em tempo real vídeo, som e dados. A digitalização da informação, ao submete-la de igual forma ao paradigma dos bits, torna possível a sua transmissão sob a mesma forma e recorrendo às diversas redes que se foram disseminando pelo planeta.

O multimédia não é estranho a esta realidade e paralelamente ao surgimento de novas redes de comunicação vão aparecendo novos e diversos meios de criar, armazenar, transportar e visualizar conteúdo multimédia. Esta abundância de redes e de dispositivos de recepção de conteúdos (cada um com diferentes características e capacidades) torna necessária a criação de mecanismos que permitam a transmissão da mensagem para que possa ser transportada através das redes disponíveis e que possua as características indispensáveis para que esta possa ser recebida com a melhor qualidade possível no dispositivo que efectuou o pedido.

Neste contexto o trabalho desenvolvido no âmbito da plataforma edVO[1] tem vindo a proporcionar ferramentas que dotam os seus utilizadores de capacidade de codificação e autoria de conteúdos multimédia MPEG-4, bem como das ferramentas necessárias para a sua indexação (MPEG-7) e disponibilização.

Para a realização deste trabalho foi preciosa a colaboração e apoio do meu orientador Eng. Luís António Corte-Real e do Eng. Luís F. Teixeira, a quem gostaria aqui de deixar o meu grande agradecimento. Seria também imperdoável não referir o apoio constante dos meus avós que por meio do seu incentivo me foram ajudando nesta caminhada. Ainda que nem todos eles tivessem estado presentes, trazia-os sempre

comigo no meu pensamento. Por último um sincero agradecimento a todos os meus amigos, a todos muito obrigado pelo apoio sempre presente.

# Índice

1. Introdução .....	1
2. Tecnologias e Potencialidades.....	3
2 .1 Codificação e classificação.....	3
2.1.1 O MPEG-4.....	4
2.1.2 O MPEG-7.....	12
Ferramentas de descrição.....	14
Combinação de tecnologias MPEG-4 e MPEG-7.....	15
2.1.3 O edVO.....	15
Ferramentas Externas.....	16
2.2 Tecnologias vídeo .....	20
2.2.1 Vídeo escalável.....	23
Tipos de escalabilidade.....	24
2.2.2 Controlo de Streams.....	25
2.2.3 Aplicações.....	26
3 Adaptação e escalabilidade.....	29
3.1 Módulo de adaptação.....	29
3.1.1 Tecnologias.....	30
3.1.2 Funcionamento.....	31
3.2 Classificação de conteúdos.....	32
3.2.1 Descrição de cenas e objectos.....	32
3.2.2 Descritores usados.....	37
4 Solução proposta.....	43
4.1 Adaptação de conteúdos.....	43
4.1.1 O cliente.....	44
4.1.2 O Servidor.....	46

4.2 Modelo conceptual .....	56
4.2.1 Análise de desempenho e trabalho futuro .....	58
5. Conclusão.....	59
Lista de Referências.....	62
Bibliografia.....	64

## Índice de Figuras

Figura 1 - Os diferentes objectos presentes numa cena MPEG-4.....	9
Figura 2 - A estrutura de uma cena MPEG-4.....	9
Figura 3 - Exemplo de uma cena de conteúdo desportivo.....	36
Figura 4 - Descrição parcial de uma cena em MPEG-7.....	38
Figura 5- DDL MPEG-7.....	39
Figura 6 - Descrição em MPEG- 7 de um objecto.....	40
Figura 7 - Classificação de objecto.....	41
Figura 8 - Modelo desenvolvido para o OSMO4.....	44
Figura 9 - - integração do módulo de adaptação do edVO.Server.....	47
Figura 10 - Talking head sem grafismo.....	48
Figura 11- Descrição parcial em XMT de uma cena do tipo "talking head".....	48
Figura 12- Cena após inserção de grafismo.....	49
Figura 13 - Descrição corresponde ao objecto "grafismo".....	50
Figura 14 - Processo de adaptação do conteúdo do tipo cinema /documentário.....	52
Figura 15 - Processo de adaptação de conteúdos do tipo "música".....	53
Figura 16 - Diagrama de casos de uso.....	54
Figura 17 - Estrutura do edVO já contemplando o modulo de adaptação.....	55
Figura 18 A mesma cena em diferentes streams MPEG-4, com conteúdos e codificações diferentes.....	57
Figura 19 – Diferentes streams de notícias .....	61

# 1- Introdução

A presente dissertação irá incidir sobre as potencialidades do MPEG-4 e MPEG-7 quando combinados para criar *streams* de vídeo escalável. As potencialidades do MPEG-4, inerentes à sua própria estrutura de objectos, podem ser potenciadas ao nível da escalabilidade se a estas for adicionada uma descrição MPEG-7. O trabalho desenvolvido neste âmbito incidiu sobre a criação de um modelo de classificação de conteúdos audiovisuais e a sua tradução num módulo de software capaz de interpretar as descrições MPEG-7 para efectuar a adaptação dos conteúdos MPEG-4.

Todo o trabalho de desenvolvimento foi construído sobre a *framework* edVO, por esta já fornecer algumas das ferramentas necessárias para classificar e distribuir conteúdos.

O edVO[1] é uma *framework* desenvolvida por Luís Filipe Teixeira que tira partido das potencialidades do MPEG-4 para compor e disponibilizar conteúdos aos utilizadores. Posteriormente foi acrescentado ao edVO a capacidade de lidar com informação MPEG-7, permitindo deste modo um acesso mais fácil aos conteúdos armazenados no módulo de servidor.

É num contexto de explosão de dispositivos multimédia que tomamos contacto com esta tecnologia e nos apercebemos das suas potencialidades ao nível da personalização de conteúdos. À medida que vão proliferando no mercado diferentes dispositivos com capacidades multimédia como são os casos dos telefones móveis, PDAS, dispositivos portáteis de áudio e vídeo, é construída uma realidade marcada pela pluralidade de soluções tecnológicas e contextos de utilização.

Este trabalho incide sobre as potencialidades da *framework* desenvolvida por Luís Filipe Teixeira, procurando identificar os mecanismos para a adaptação de conteúdos, para que estes vão de encontro às especificidades dos terminais e dos cada vez mais variados contextos de utilização. Pretende-se desenvolver um modelo de indexação de conteúdos que permita efectuar uma adaptação que leve em conta as suas características específicas, garantindo assim o melhor serviço possível ao utilizador.

A presente dissertação irá por isso apoiar-se nas principais tecnologias presentes na incidir sobre a *framework* (o MPEG-4 e o MPEG-7) procurando combinar as suas características para construir um modelo de adaptação de conteúdos.

No capítulo dois são analisadas as diferentes tecnologias envolvidas no processo de adaptação e escalabilidade de conteúdos, procurando-se definir as ferramentas disponíveis para efectuar a sua adaptação, bem como as diferentes formas de a concretizar.

O terceiro capítulo procura estabelecer as linhas gerais do processo de adaptação de conteúdos definindo o modelo de classificação e a forma como esse modelo irá interagir com o módulo de adaptação.

No capítulo quatro o trabalho até então desenvolvido é vertido num modelo conceptual do módulo de adaptação do edVO, sendo dados exemplos concretos de como esta poderá ser realizada. Este capítulo termina tirando conclusões do trabalho desenvolvido e traçando um caminho para trabalho futuro.

## 2- Tecnologias e potencialidades

De acordo com o que foi abordado no início desta dissertação são várias as tecnologias que estão implicadas na adaptação e escalabilidade de conteúdos multimédia como também na própria *framework*.

Este capítulo aborda cada uma dessas tecnologias, procurando fornecer uma explicação sucinta sobre o seu funcionamento e a forma como estas se integram para constituir o edVO, a plataforma que serviu de base ao trabalho desenvolvido.

### 2.1 - Tecnologias multimédia

O trabalho desenvolvido assenta no módulo que suporta a adaptação de conteúdos, o edVO Server, acrescentando à aplicação funcionalidades que lhe permitem associar a um mesmo objecto várias representações alternativas. Esta característica é determinante para o trabalho que irá ser desenvolvido seguidamente, uma vez que dota a plataforma edVO da capacidade de adequar os conteúdos ao meio (ou outra qualquer característica que possa eventualmente vir a ser definida quer a nível de servidor, quer mesmo a nível de cliente).

Para começarmos a analisar o objecto sobre o qual o trabalho irá ser desenvolvido importa entender quais as tecnologias que foram empregues na sua elaboração, bem como as alternativas que possam eventualmente existir, quer sejam estas relativas ao trabalho já desenvolvido, quer sejam referentes ao trabalho a desenvolver.

As normas definidas pelo comité MPEG, em particular MPEG-4 e MPEG-7, são talvez o ponto fulcral desta análise uma vez que são estas duas tecnologias que suportam as principais funcionalidades do edVO. Num primeiro momento a nossa análise irá incidir sobre o MPEG-4 para posteriormente se focar no MPEG-7, ao que se seguirá uma análise mais detalhada sobre a estrutura e o funcionamento do edVo.

Para além das duas normas já referidas existem outros *standards* desenvolvidos pelo comité MPEG e os quais iremos aqui procurar explicar sucintamente. O Comité foi

formado em 1988 tendo começado a trabalhar na primeira norma MPEG, o MPEG-1, que visava criar um standard de compressão que permitisse a mais fácil difusão de conteúdos audiovisuais em formato digital, de que são o exemplo *video-cd* e o áudio MP3. Mais tarde surgiu o MPEG-2 que fez incidir o seu estudo sobre uma forma de distribuir conteúdos audiovisuais com qualidade *broadcast*, e nos quais se basearam as tecnologias DVD e a televisão digital. Mais recentemente surgiu o MPEG-21, também conhecido por multimédia *framework*, visto o propósito desta norma ser o de construir uma *framework* universal para a difusão de todos os conteúdos multimédia. No período de transição do MPEG-2 para o MPEG-21 surgiram outros dois standards MPEG: o MPEG-4 e o MPEG-7. O MPEG-4 destina-se à criação e difusão de conteúdos multimédia a *bitrates* muito baixas e o MPEG-7 estabelece ferramentas para a descrição de conteúdos. São estas duas normas que iremos analisar mais detalhadamente.

### 2.1.1 - O MPEG-4

O MPEG-4 é um standard ISO/IEC para a composição de conteúdos audiovisuais, naturais ou sintéticos, organizados sobre a forma de cenas. Este conjunto de normas foi pensado de forma a aproveitar o surgimento de novos dispositivos para a reprodução de conteúdos multimédia a *bitrates* bastante baixas (64Kb/s), muito embora a norma permita que, dependendo dos recursos disponíveis, o MPEG-4 possa atingir débitos até 4Mb/s. Estas características fazem desta norma uma solução ideal para operar em ambientes diversos que requerem uma grande flexibilidade. Tirando partido das suas características é possível adequar o conteúdo transmitido às mais diversas condições de transporte, processamento e visualização [1].

De uma maneira sucinta, este standard procura cumprir os seguintes requisitos: uma codificação eficiente de objectos audiovisuais, naturais ou sintéticos; composição de cenas audiovisuais através da descrição de grupos de objectos audiovisuais; acesso universal através da multiplexagem e sincronização dos dados associados aos objectos, para que estes possam ser transportados sobre uma rede com a Qualidade de Serviço

(QoS) apropriada; interacção do utilizador com a cena audiovisual. Para atingir estes objectivos o MPEG-4 disponibiliza ferramentas standard que suportam:

- Codificação: representa unidades de conteúdo áudio, vídeo ou audiovisuais, denominados *media objects*, que podem ser de origem natural ou sintética.
- Composição: descreve a composição dos *media objects* para criar *media objects* compostos que formam uma cena multimédia.
- Multiplexagem: efectua a multiplexagem e sincronização dos dados associados aos *media objects* de forma a transportá-los através de redes proporcionando uma Qualidade de Serviço apropriada à natureza de cada *media object*.
- Interacção: permite a interacção local ou remota com a cena, através de um canal de retorno.

A estrutura do MPEG-4 é composta por diversas partes, no entanto nem todas elas já são considerados standards internacionais (IS). Actualmente são considerados IS as seguintes partes: Systems, Visual, Audio, Conformance Testing, Reference Software, e Delivery Multimédia Integration Framework (DMIF).

## Sistema Típico

Tal como já foi explicado anteriormente, uma das características do MPEG-4 é a sua capacidade de se adaptar a diferentes ambientes. No entanto a estrutura base um sistema MPEG-4 normalmente permanece inalterada na maior parte dos cenários. O que iremos fazer de seguida é fornecer uma visão geral sobre o funcionamento de um sistema MPEG-4 padrão.

No que diz respeito à criação de conteúdos podemos estar a falar essencialmente de dois tipos diferentes: conteúdos editados digitalmente ou então de conteúdos captados ao vivo (*streaming*), um pouco como acontece com uma transmissão televisiva. Em relação ao material editado, este é produzido em duas etapas: a primeira designada por autoria e da qual fazem parte tarefas como a produção dos conteúdos, incluindo o material em bruto, a sua edição e organização em cenas bem como a interacção do receptor com o media; A segunda parte é designada por publicação e envolve adaptar o conteúdo

produzido para que este possa ser exibido correctamente no destinatário. Esta última parte implica ter em conta as limitações impostas, por exemplo, por redes e terminais. Enquanto que os eventos transmitidos em *streaming* são produzidos utilizando um dos vários formatos de *streaming* MPEG-4, o material produzido em *off-line* pode ser publicado nos servidores utilizando um formato que assegura a interoperabilidade através de XMT (eXtensible MPEG-4 Textual format), ou então através do formato de ficheiro MP4. A opção entre um destes formatos depende sempre da liberdade que o autor que oferecer aos utilizadores no processo de recepção dos conteúdos, ou mesmo a outros autores em processos subsequentes de edição.

O XMT fornece uma representação estruturada e extensível da descrição das cenas, e como tal fornece mais liberdade a eventuais processos de adaptação ou personalização dos dados. Já o MP4 é bastante mais rígido uma vez que as cenas estão codificadas dentro de forma binária, tornando-a mais impermeável a futuras alterações. Este aspecto, tal como anterior, acarreta vantagens e desvantagens. A opção pelo formato MP4 traz vantagens não só porque se trata de uma forma mais compacta de representação de cenas, mas também porque sendo mais difícil de alterar, assegura que as intenções do criador da cena são melhor preservadas.

Não é no entanto necessário que os dados sejam armazenados no servidor em qualquer uma destas formas. Futuras implementações podem encontrar uma maneira mais eficiente para armazenar os ficheiros MPEG-4, optimizando o formato em função das especificações de *hardware* e *software*. O que importa assegurar é que os dados saem do servidor num formato binário. O que na verdade é enviado pelo servidor são cadeias de dados MPEG-4 denominadas *elementary streams* (ESs). Uma apresentação MPEG-4 contém várias dessas ESs enviadas pelo servidor, se necessário recorrendo a diferentes redes para cada uma das *streams* que são mais tarde sincronizadas de forma a delas se obter um conteúdo coerente. No caso de se tratar de um conteúdo produzido em directo este pode ser submetido no servidor sobre a forma de ESs dotadas de interoperabilidade. Neste ponto importa falar novamente no objectivo da norma MPEG-4 que visa garantir que o mesmo conteúdo possa ser transmitido de diversas formas de maneira a ser recepcionado em diferentes ambientes. Para tal, foi especificado que o conteúdo seria abstraído da *delivery layer* utilizando um interface comum e detalhadamente definido.

Deste modo, quando o conteúdo é entregue no destino os pontos necessários para garantir a interoperabilidade são definidos pelos formatos de cada ES e pelos passos estabelecidos na interface para aceder a cada uma das ESs. Tudo aquilo que se encontra abaixo desta interface fica fora do âmbito da norma MPEG-4. Deste modo assegura-se que a representação do conteúdo multimédia é completamente independente da forma como este é depois distribuído.

## Media Objects

Uma cena é composta por diversos tipos de *media objects*, organizados de forma hierárquica numa estrutura tipo árvore em que cada folha corresponde a um *media object* primário [2]. Alguns exemplos de *media objects* primários são imagens estáticas, vídeo e áudio. Imaginando uma cena onde estes tipos de *media objects* poderiam estar presentes, teríamos um ambiente de documentário em que uma pessoa falava em primeiro plano, sob um fundo de natureza, ouvindo-se a voz do apresentador à qual se sobrepunha uma melodia. A norma MPEG-4 estabelece um conjunto de *media objects* capazes de representar conteúdos naturais ou sintéticos. Além daqueles que já foram mencionados previamente são também considerados *media objects*:

- Texto e gráficos;
- Rostos humanos sintetizados e animados e texto associado, utilizado para sintetizar a voz e animar o rosto;
- Corpos humanos sintetizados e animados de uma forma idêntica à dos rostos;
- Sons sintéticos.

Todos estes elementos podem ser posicionados num sistema de coordenadas, 2D ou 3D, dependendo do tipo de *media object* a que nos estejamos a referir.

Entende-se que um *media object*, na sua forma codificada, consiste numa série de elementos que permitem que esse mesmo objecto seja manipulado numa determinada cena multimédia, assim como um conjunto de informações destinadas a auxiliar no processo de *streaming*, caso tal seja necessário. Uma das características fundamentais

dos *media objects* é que todos eles podem ser representados independentemente daquilo que os rodeia ou do meio no qual estão inseridos [3].

A nível da estrutura hierárquica de representação dos *media objects* é ainda de referir que existem ainda *group objects*, destinados a agrupar vários *media objects*. Utilizando o exemplo do comentário mencionado anteriormente, a imagem e a voz do apresentador seriam parte de um *group object* denominado “pessoa” ou “anfitrião”. Isto permite que os utilizadores possam elaborar cenas complexas ao mesmo tempo que fornece aos consumidores a hipótese de estes manipularem um conjunto coerente de objectos.

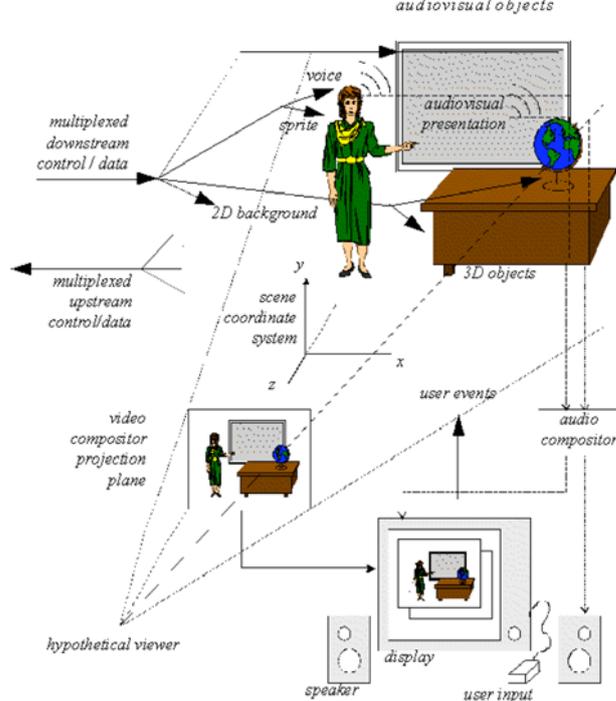
## Composição e interacção de media objects

O acto de se compor uma cena a partir de vários *media objects* implica que exista uma forma de descrever as propriedades de cada objecto bem como de definir as relações espaciais e temporais entre cada um deles [1]. Outras operações que implicam a existência de um formato capaz de as descrever são:

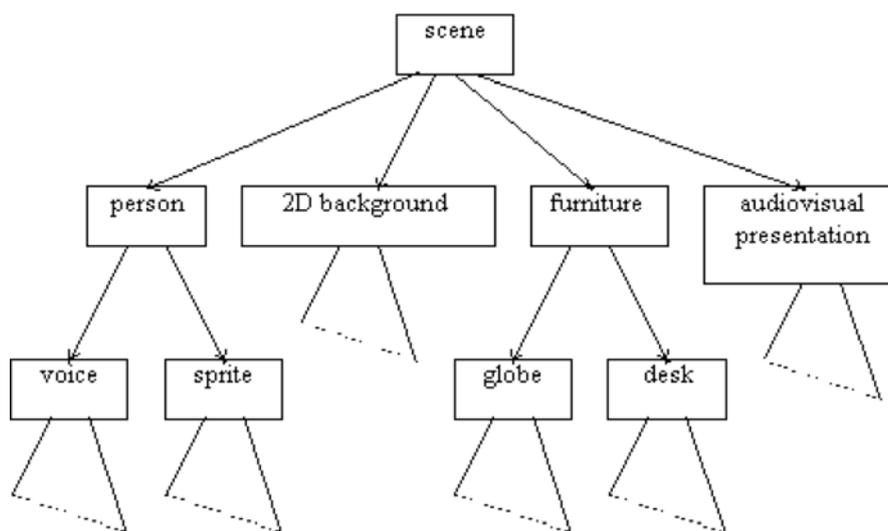
- Dispor objectos num sistema de coordenadas;
- Aplicar transformações a um objecto de forma a alterar as suas propriedades;
- Agrupar objectos primários;
- Aplicar transformações em *streaming* de forma a alterar as propriedades dos objectos (por exemplo, animando um rosto ou textura);
- Alterar de modo interactivo a posição do utilizador na cena.

Esta última característica não é a única forma que um utilizador tem para interagir com uma cena, existem outras formas de o fazer como por exemplo:

- Alterar a posição de determinados objectos;
- Iniciar eventos através de cliques em objectos
- Seleccionar um determinado idioma.



Para conseguir disponibilizar estas funcionalidades a norma MPEG-4 criou uma forma standard para descrever uma cena e chamou-lhe *Binary Format for Scenes* (BIFS). Este formato define a forma como os objectos são combinados de forma a constituírem uma apresentação coerente. Na base do BIFS está um formato de modelação 3D, o VRML. O grupo encarregue de trabalhar no BIFS aproveitou a estrutura já definida pelo VRML, adaptando-a e acrescentando-lhe algumas funcionalidades necessárias para suportar as inovações introduzidas pelo MPEG-4.



**Figura 2 - A estrutura de uma cena MPEG-4[3]**

De forma a poder organizar uma cena e actuar sobre os seus objectos é também necessário possuir uma forma de descrever as propriedades de cada *stream* que compõe a cena, para tal foram definidos *object descriptors*. Estes ODs servem para descrever informação sobre *codecs*, bit rates bem como as relações entre as *streams* e esquemas de gestão de direitos de autor. Uma outra funcionalidade importante do BIFS é dar a hipótese de ir gradualmente fazendo *upload* das diferentes *streams* que compõem as cenas, colmatando as dificuldades colocadas pela operação em redes de largura de banda reduzida.

Para facilitar a troca de conteúdos entre autores ao mesmo tempo que é garantida a preservação do conteúdo semântico foi desenvolvido o formato XMT. O formato XMT permite criar abstracções a alto-nível facilitando, para além das operações sobre os conteúdos, a interoperabilidade com as especificações X3D e SMIL.

A plataforma XMT é composta por dois níveis de sintaxe e semântica: o XMT-A e o XMT-O.

O formato XMT-A é uma forma baseada em XML de descrever conteúdo MPEG-4 e que é bastante semelhante à representação sob a forma binária. Este formato foi criado para fornecer uma representação em modo de texto da codificação MPEG-4 em modo binário, para além de suportar uma abordagem determinística da representação binária e para assegurar a compatibilidade com o formato X3D.

Já o formato XMT-O foi criado baseado na SMIL, destina-se a permitir aos autores criarem conteúdos multimédia dinâmicos e interactivos simples e estruturadamente, utilizando uma linguagem baseada em XML.

O XMT possui ainda um mecanismo que permite criar conteúdos que sejam conformes com os dois formatos (XMT-A e XMT-O) de forma a tirar melhor proveito das características de cada um deles.

## Descrição e sincronização de objectos

Por vezes os *media objects* podem necessitar que lhes seja transmitida informação adicional sob a forma de *streams*, como é o caso de um objecto que pretendemos animar. Para tal o MPEG-4 utiliza os ODs de forma a associar o objecto às suas *streams*.

Cada *stream* OD contém um variado conjunto de informações indispensáveis para a compreensão do papel do objecto na cena. A forma como o objecto foi codificado, a sua sincronização com os restantes objectos e os recursos necessários para adaptar o objecto às condições da rede são algumas das situações que se encontram descritas numa *stream* OD. Estas informações podem ser usadas no destinatário, por exemplo para determinar as capacidades de reprodução do software do cliente ou então para auxiliar garantir algum nível de QoS.

Além destas informações os ODs podem ainda transportar um conjunto adicional de informações referentes ao conteúdo do objecto, denominadas *Object Content Information*. Um tipo adicional de informação que pode estar contida num OD é a informação respeitante à gestão de direitos de autor.

A sincronização de cada uma das ES é feita pela camada de sincronização (SL). Esta camada utiliza unidades de tempo para identificar “marcadores” contidos em cada ES. Estes “marcadores” são denominados *Access Units* (AUs) e a SL consegue identifica-los bem como determinar a natureza de cada *stream* (vídeo, áudio, texto, etc.) e ordena-los para que as várias ES que compõem a cena estejam sincronizadas entre si.

## Conteúdo adaptável

Conforme vimos anteriormente é possível tirar partido das características estruturais do MPEG-4 para manipular uma cena, utilizando para o efeito comandos enviados através de *streams*. Esta característica permite a possibilidade de adaptar também vídeo se tal for necessário.

Apesar de uma cena ser visionada como o todo, não nos podemos esquecer que num ficheiro MPEG-4 uma cena é composta por objectos independentes que são compostos numa mesma cena utilizando o BIFS. Como tal é possível executar comandos que actuem sobre a composição da cena, acrescentado ou removendo objectos e até mesmo alterando as suas propriedades para que este se adapte a diferentes condições, isto é, a diferentes redes e dispositivos, demonstrando mais uma vez as sua potencialidades

multi-plataforma. Esta flexibilidade pode ainda ser aproveitada para permitir ao próprio utilizador modificar a cena, interagindo com esta [1].

### 2.1.2 - O MPEG-7

Tão importante como ter um acesso fácil aos conteúdos é encontrar uma forma de os descrever para que estes possam ser pesquisados e a informação relevante localizada facilmente. É para cumprir este objectivo que se criou o standard MPEG-7 (ISSO/IEC 15938), inicialmente conhecido como *Multimedia Content Description Interface*. Um ficheiro MPEG-7 pode descrever qualquer tipo de conteúdo multimédia, seja ele uma imagem estática, um vídeo, áudio ou imagens e animações 3D. Para além de descrever cada conteúdo individualmente o MPEG-7 pode ainda ser usado para descrever a forma como os conteúdos se combinam numa apresentação [1].

Uma característica importante do MPEG-7 é que este não tem qualquer relação com a forma como o conteúdo é produzido, distribuído ou armazenado. É possível criar uma descrição MPEG-7 de qualquer conteúdo em qualquer suporte, incluindo papel e caneta [4].

É no entanto importante observar as sinergias que o MPEG-7 pode criar com alguns formatos, nomeadamente com o MPEG-4. Quando um ficheiro MPEG-7 é criado a partir de um ficheiro codificado em MPEG-4, o MPEG-7 consegue ler a descrição nativa da norma MPEG-4 e utilizar essa informação para criar uma descrição MPEG-7 do ficheiro.

Uma outra característica importante do MPEG-7 é a sua capacidade de construir uma hierarquia em torno dos metadados, isto é, com o MPEG-7 é possível estruturar a informação em diferentes níveis consoante a característica que se está a descrever. Existem dois tipos de características, a saber: características de alto-nível e características de baixo-nível. Em relação a este último grupo, são consideradas características de baixo nível os *codecs* usados, a *bitrate*, a cor, a textura, etc. O nome de um participante, a classificação etária do filme, a identificação de um local ou de um objecto presente são consideradas características de alto-nível.

Por aquilo já foi dito é possível perceber que as características de baixo nível possuem três propriedades fundamentais:

- Podem ser extraídas automaticamente;
- São objectivas; e
- São nativas do conteúdo audiovisual.

Sendo assim, qualquer uma destas características pode ser extraída com relativa simplicidade por uma máquina que tenha a capacidade de distinguir entre diferentes algoritmos e ferramentas. Esta importante característica simplifica bastante a tarefa de analisar e descrever o conteúdo uma vez que permite a automatização do processo de extracção das propriedades de baixo-nível. O valor de possuir estas informações é bastante elevado, aumentando em função do “peso” da organização, do volume da informação e do número de pessoas que dependem dela. A natureza objectiva dos conteúdos descritos permite aumentar a eficiência nas pesquisas.

Quanto a implementar um processo idêntico para extrair informação das características de alto-nível, tal já se revela uma tarefa mais complexa ainda que este seja uma característica que ofereça grandes atractivos para os utilizadores. É no entanto importante reter que o MPEG-7 permite combinar características de alto e baixo-nível numa única descrição MPEG-7. Seguem-se alguns exemplos de características que podem ser descritas em MPEG-7:

- Informação referente ao processo de criação e produção do conteúdo (realizador, editor, som, produção, título, etc);
- Informações respeitantes à forma como o conteúdo deve ser utilizado (direitos de autor e cópia, horários de emissão, etc);
- Informação estrutural referente a características espaciais, temporais ou espaço-temporais (cortes, segmentações, etc...);
- Informações sobre as características de armazenamento do conteúdo (formato de arquivo, codificação);
- Informações sobre características de baixo-nível do conteúdo (cores, texturas, descrição melódica, etc);
- Informação conceptual sobre a realidade capturada no conteúdo (objectos, acontecimentos e interacções entre objectos);
- Informações sobre como aceder ao conteúdo de uma forma eficiente (sumários, variações, sub-bandas espaciais e de frequência);
- Informações sobre grupos de objectos; e

- Informações sobre a interacção do utilizador com o conteúdo (preferências e histórico de utilização).

O standard que temos estado a descrever é composto por várias partes, das quais podemos nomear as seguintes: Sistemas, DDL, Conteúdos visuais, Áudio, e MDS. Para efeitos do tema em estudo algumas destas partes (as que se referem a conteúdos de alto nível) irão ser estudadas mais em detalhe.

## Ferramentas de descrição

O MPEG-7 descreve o conteúdo multimédia com uma abordagem que parte de diversos pontos de vista. A cada forma de descrever o conteúdo corresponde uma ferramenta específica. Muito embora essas ferramentas e descrições sejam independentes entre si, estas podem ser combinadas numa mesma descrição[1]. Em uma descrição MPEG-7 uma entidade pode ser incluída total ou parcialmente, flexibilizando-se as ferramentas em função do conteúdo que se pretende descrever.

Os principais elementos do MPEG-7 são as ferramentas de descrição, a linguagem de definição da descrição (DDL) e as ferramentas de sistema[5]. Das ferramentas de descrição fazem parte os descritores (Ds), responsáveis pela definição da sintaxe e semântica de cada característica, e os esquemas de descrição (DSs) que por sua vez especificam a estrutura e a semântica das relações entre os componentes destas.

A DDL é a base para a sintaxe das ferramentas de descrição MPEG-7, permitindo a criação de novos DSs, ao mesmo tempo que permite a extensão e modificação daqueles já existentes. Quanto às ferramentas de sistema, estas fornecem suporte para um elevado número de funções, como sejam:

- Representação de informação codificada de forma binária;
- Mecanismos de transmissão;
- Multiplexagem de descrições;
- Sincronização das descrições com o conteúdo;
- Protecção da propriedade intelectual nas descrições MPEG-7;
- Etc.

Deste modo, o processo de criar uma descrição em MPEG-7 passa simplesmente por instanciar um ou mais DSs e os correspondentes Ds, de acordo com aquilo que o utilizador entenda que é mais indicado para descrever o conteúdo. Caso se verifique ser necessário um descritor não existente, é sempre possível recorrer à DDL de forma a criar um que se adequa às necessidades do utilizador.

## Combinação de tecnologias MPEG-4 e MPEG-7

A combinação das características dos formatos MPEG-4 e MPEG-7 criam uma combinação poderosa mas também flexível. Não obstante o facto de serem ambas normas bem definidas (as que já estão estandardizadas) a sua estrutura permite que estas se possam adaptar às mais diversas condições de forma a preencher os requisitos necessários. Se por um lado o tratamento dos conteúdos multimédia como *media objects* permite manipula-los melhor de acordo com o meio, a sua combinação com uma descrição destes potencia também uma melhor classificação, procura e manuseamento dos conteúdos.

### 2.1.3 - O edVO

O edVO é uma ferramenta modular que fornece um conjunto de serviços ao nível da autoria, armazenamento, gestão e distribuição de conteúdo multimédia. São 4 os módulos que compõem o edVO: o edVO.Composer, o edVO.Admin, o edVO.Server e o edVO.Client [4].

O edVO.Composer é o módulo que corresponde à ferramenta de autoria multimédia, fornecendo aos utilizadores a capacidade de criarem conteúdos multimédia MPEG-4 e MPEG-7. Toda a comunicação com o edVO.Server é feita através de XML-RPC e HTML sobre HTTP, sendo que este último envia comandos de controlo enquanto que o primeiro é responsável pelas transacções das interfaces de utilizador e comandos [1].

O edVO.Server é o núcleo de todo o sistema e actua como sistema de armazenamento e fornecimento de conteúdos multimédia, para além de implementar funcionalidades das

ferramentas de administração e publicação de conteúdos. Este módulo comunica com as restantes componentes utilizando XML-RPC e HTML sobre HTTP para enviar comandos de controlo e operações relacionadas com a interface do utilizador, já para a transferência de conteúdos suporta FTP, http e RTP/RTSP.

O módulo edVO.Admin não é mais do que uma interface para a administração e publicação de conteúdos no edVO.Server. Todas as interfaces estão acessíveis via HTML e como tal podem ser acedidas a partir de um browser.

Por último, o EdVO.Client é o módulo que corresponde à parte da plataforma que acede aos conteúdos através do edVO.Server. O cliente comunica com o servidor através de uma interface HTML, este depois envia para o *player* o conteúdo pedido utilizando FTP, HTTP ou RTP/RTSP. O cliente é actualmente uma aplicação *stand-alone*, mas pode também funcionar como um plug-in de um browser.

## Ferramentas externas

No desenvolvimento do edVO foram utilizadas algumas ferramentas externas como forma de acelerar o processo de criação da plataforma. De seguida iremos nomear cada uma dessas ferramentas, procurando explicar as suas características e a forma como estas se integram no projecto.

## GPAC

O GPAC (GPAC Project on Advanced Content) é uma plataforma baseada na norma MPEG-4. Toda ela foi desenvolvida em ANSI C e desde a versão 0.40 que se encontra disponível sob a GNU Lesser General Public Licence. Versões anteriores estão também disponíveis sob GNU General Public Licence. Actualmente existem versões desenvolvidas para Windows, Linux e WindowsCE/Pocket PC 2002, embora esta última ainda ofereça funcionalidades limitadas.

As ferramentas disponibilizadas pelo GPAC permitem ao edVO construir cenas MPEG-4 a partir de descrições em XMT.

Ainda que o GPAC seja uma versão mais simplificada do extenso conjunto de ferramentas disponibilizado pelo MPEG-4 Systems Reference Software, as bibliotecas do GPAC disponibilizam um importante conjunto de funcionalidades que podem ser agrupadas em diversos módulos. Seguidamente enumeram-se cada um desses módulos e descrevem-se as suas características [6].

- ISO Media File Format
  - Suporta edição, leitura, escrita/captura;
  - Interlaçamento flexível dos conteúdos media.
  - Suporte para RTP hint tracks; e
  - Suporte para fragmentos de clips- permite a escrita de um ficheiro como um conjunto de pequenos blocos de metadados e media.
- Object Descriptor (OD) Framework
  - Inclui o codec do OD; e
  - Inclui o codec do OCI.
- Binary Format for Scenes (BIFS) e gráfico de cena
  - Suporte para gráficos de cena MPEG-4 com nós configuráveis;
  - Suporte para interactividade (eventos, caminhos, interpoladores, etc);
  - Inclui um gestor de cenas para autoria de conteúdos em modo de texto via BT/XMT; e
  - Inclui codificação BIFS.
- Gestão de *Elementary Streams*
  - Suporte para codecs escaláveis (áudio, vídeo e sistemas)
  - Efectua a sincronização, gestão de media e temporização de clips; e
  - Efectua o controlo de tempo e dos descritores de segmentos.
- Rendering
  - Suporta *rendering* directo e indirecto;
  - Implementa a maioria dos nós definidos no BIFS e a maioria das emendas entretanto introduzidas no standard (texto e gráficos);
  - Suporte experimental para *rendering* de alguns nós 3D; e
  - Render áudio e mistura por software (apenas mono ou stereo).

## Pygpac

O edVO utiliza uma API que permite que as partes que foram escritas em python possam comunicar com o GPAC que, como foi explicado anteriormente, foi escrito em linguagem C. O Pygpac actua como um *interface* entre o GPAC e o edVO, permitindo utilizar scripts python para passar instruções ao GPAC. Sendo assim, foi possível adicionar funcionalidades ao GPAC recorrendo a uma linguagem de *scripting*, muito embora tenha existido partes do edVO desenvolvidas em linguagem C.

## **FOX Toolkit**

O desenvolvimento da interface gráfica (GUI) do edVO foi feito com recurso ao FOX Toolkit, um conjunto de ferramentas na sua maioria com licenciamentos baseados na GNU General Public Licence.

O Fox Toolkit encontra-se disponível para um variado número de sistemas operativos, sendo a ferramenta a usar para trabalhar sobre a interface gráfica do edVO, com destaque para o edVO.Composer onde houve necessidade de criar alguns controlos especialmente para satisfazer os requisitos deste módulo.

## **EMozilla**

O EMozilla é uma biblioteca desenvolvida especificamente para o edVO que permite integrar o Mozilla no edVO para tirar partido das suas funcionalidades como web browser, nomeadamente ao nível da interface.

## **Zope**

O Zope é um servidor de aplicações web, open-source e desenvolvido em Python. É um servidor baseado numa base de dados transaccional capaz de armazenar templates HTML dinâmicos, um motor de busca, uma base de dados relacional, para além de outros tipos de conteúdo e informações [7].

Fazem parte do Zope servidores de http, FTP, Web-DAV e XML-RPC, o que não invalida que o Zope seja utilizado em conjunto com outros servidores web como o Apache.

Ao escolher o Zope como servidor, o edVO tira partido da capacidade do Zope de gerar conteúdo HTML dinâmico a partir de *scripts* Python. Para além disso, o Zope fornece de raiz funcionalidades de FTP e XML-RPC.

## Darwin Streaming Server

O Darwin Streaming Server é uma versão open-source do Quicktime Streaming Server, permitindo a distribuição de conteúdos Quicktime, MPEG-4 e 3GPP através da Internet, recorrendo aos protocolos standard RTP e RTSP [8].

Ao ser incluído no edVO, o Darwin permite que os conteúdos MPEG-4 e 3GPP possam ser disponibilizados via *streaming* a uma aplicação cliente que envie o pedido.

## 2.2- Tecnologias e potencialidades

Em face do que podemos até agora observar existe um conjunto de factores que influem na capacidade de um utilizador de um sistema receber informação audiovisual: em primeiro lugar a natureza da própria informação. Tanto no caso de conteúdos naturais, como também no caso dos sintéticos (e em particular nestes últimos), a forma como a informação é inicialmente armazenada traça desde logo um tecto máximo de qualidade. O segundo factor que intervém neste processo é a compressão final aplicada aos conteúdos, isto é, em que forma se opta por distribuir os conteúdos audiovisuais. O terceiro elo desta cadeia é a rede na qual os conteúdos são distribuídos. A capacidade de o conteúdo desejado ser entregue em boas condições depende também por isso das características do meio. Por último, os dispositivos terminais ou dispositivos de recepção, ao serem normalmente limitados do tipo de conteúdos que conseguem tratar e na forma como os apresentam ao receptor também influenciam e limitam a capacidade dos conteúdos que podem ser percebidos por um utilizador.

O conjunto de todas estas condições constituem o ambiente sobre o qual iremos actuar, pelo que serão daqui em diante designadas por condições de utilização.

Cada etapa deste processo tem potencial para introduzir limitações ao que se pretende transmitir, sejam conteúdos vídeo, áudio ou audiovisuais, no entanto algumas dessas condições de utilização podem variar ao longo do tempo ou até mesmo serem específicas de um utilizador ou grupo de utilizadores. Actualmente o grau de adaptação que mais vulgarmente é utilizado para adequar os conteúdos às condições de utilização passa por ter diferentes versões do mesmo conteúdo, deixando ao utilizador o decidir qual a versão que irá ver, normalmente perguntando-lhe em que tipo de rede está. Este exemplo é o mais frequente quando se tratam de conteúdos *disponíveis na world wide web* e quando o equipamento utilizado para a sua recepção é um computador pessoal. No caso dos dispositivos móveis ocorre situação idêntica, mas como o emissor dos conteúdos (normalmente o operador móvel) tem conhecimento de qual o equipamento que o utilizador está a usar e das condições da rede em que este se encontra, é o próprio operador a decidir qual o conteúdo que deve disponibilizar.

Nos últimos anos temos assistido a uma crescente mobilidade dos utilizadores, acompanhada por uma oferta mais diversificada de equipamentos, conteúdos e serviços, pelo que é previsível que numa situação ideal, e no que diz respeito aos conteúdos audiovisuais, que os utilizadores pretendam aceder em qualquer dispositivo com ligação a uma rede ao mesmo conteúdo que poderiam aceder via televisão e computador pessoal. Deste modo, se já com as condições actuais, os fornecedores de conteúdos não podem ter uma versão de cada conteúdo que se adeque simultaneamente às características do equipamento de recepção e da rede utilizada, com os avanços tecnológicos ao nível dos equipamentos e das redes e com os utilizadores a pedirem que os conteúdos acompanhem o aumento da sua mobilidade, tornar-se-á virtualmente impossível existirem versões de conteúdos que cubram a generalidade das condições de utilização possíveis.

Actualmente os conteúdos são adaptados apenas em função da rede que utilizam, assumindo os fornecedores de conteúdos que os dispositivos que funcionam nessa rede têm características semelhantes. De que tenhamos conhecimento, a excepção existe apenas nas redes GPRS, onde a quantidade de equipamentos é maior e as suas características mais diferem. Neste caso, existe pelo menos um operador português a realizar adaptação de conteúdos através de uma codificação do mesmo em múltiplos níveis.

A compressão, contudo, nem sempre ocorre ao nível do ficheiro, na realidade se alargamos a definição de compressão para além do âmbito lógico podemos entender como compressão operações que ocorram ao nível da produção e da pós-produção com vista a eliminar das imagens e dos sons elementos que sejam susceptíveis de evidenciar as limitações do meio. A adaptação de conteúdos tal como é habitualmente aplicada em dispositivos móveis não é, na maioria dos casos, mais do que o corte de determinados elementos que iriam exigir uma capacidade de processamento acima daquela que normalmente se encontra na generalidade dos terminais. Raciocínio inverso mas que segue o mesmo princípio ocorre na selecção de conteúdos para plataformas móveis, onde são preferidos os conteúdos mais facilmente adaptáveis ou aqueles que não requerem quase nenhuma adaptação.

No caso português, a Optimus, pioneira na produção de conteúdos exclusivamente para telemóveis, produz desde 2004 conteúdos de natureza informativa e de entretenimento, preocupando-se desde o momento da produção dos mesmo a ter em conta factores como movimentos de câmara, enquadramentos e efeitos visuais e sonoros que sejam susceptíveis de serem imperceptíveis ou mesmo interpretados como falhas ao nível da imagem. Trata-se portanto de uma aplicação *à priori* dos princípios da codificação perceptual, isto é, apenas captar aquilo que seja susceptível de ser bem representado atendendo às limitações de utilização.

O modelo desenvolvido neste trabalho pretende aproveitar as funcionalidades do MPEG-4 e do MPEG-7 de forma a conseguir uma adaptação de conteúdos mais adequada às características de utilização no momento da sua recepção. Ao produzir conteúdos com a preocupação de que estes terão que ser visualizados em virtualmente qualquer dispositivo, estamos de acordo com a explicação dada anteriormente, a introduzir compressão desde o momento da produção, o que limita desde logo o conteúdo. Numa situação em que se pretende produzir conteúdos capazes de serem visionados em diferentes equipamentos e transmitidos utilizando redes e técnicas diversas é importante que no momento da produção as potencialidades do meio não sejam “niveladas por baixo”, de forma a não limitar também os ambientes de maior capacidade.

A utilização de um modelo que substitua a exclusão completa de elementos que exijam uma capacidade de processamento elevada para a maioria dos dispositivos móveis, por um outro modelo que (além de recorrer aos métodos mais “tradicionais” de compressão), não exclua nenhum elemento na fase da produção parece-nos dar mais garantias de que o conteúdo produzido estará sempre à altura das condições de utilização. A adaptação dos conteúdos é desta forma efectuada não só através da compressão das *streams* mas também adicionando, removendo ou substituindo elementos da cena.

Para tal, é utilizado código que utiliza indicações dadas por descritores MPEG-7 combinadas com uma análise ambiental para decidir quais os elementos que devem ser carregados para a cena MPEG-4.

### **2.2.1 - Vídeo escalável**

A chave para conseguir atingir uma melhor adequação dos conteúdos ao meio em que estes são difundidos e visionados é a aplicação ao vídeo do conceito de escalabilidade.

A IBM define no seu site[9] o vídeo escalável como a capacidade de adicionar ou remover camadas de vídeo de forma a aumentar (ou diminuir) a sua qualidade, *frame rate*, etc.

Já para a France Telecom, o aspecto mais importante envolvido na noção de vídeo escalável é a capacidade de criar *feeds* de vídeo capazes de se adaptarem às características da rede e/ou dispositivo usado pelo utilizador final[10]. Se o vídeo for codificado com os parâmetros necessários para garantir a qualidade máxima que o serviço deve prestar, de acordo com os princípios do vídeo escalável, é possível extrair partes desse mesmo vídeo de forma a reproduzir conteúdos com qualidade inferior mas que se adequem às redes e dispositivos que estão a ser utilizados.

As características desta tecnologia permitem, a nível da polivalência, utilizar um único *feed* para transmitir conteúdos adaptáveis a diferentes contextos de utilização, como por exemplo: televisão de alta-definição, televisão *standard*, vídeo para conteúdos multimédia e vídeo para telemóveis. Para além destas características o vídeo escalável apresenta uma outra importante propriedade: a sua capacidade de mobilidade dentro de uma rede. Significa isto que, com o vídeo escalável, o utilizador pode movimentar-se com um dispositivo portátil (PDA, telemóvel, etc) e no caso de ocorrerem mudanças ao nível da rede utilizada, por exemplo uma transição de uma rede wi-fi para UMTS, o conteúdo será ajustado de uma forma automática e transparente para o utilizador, sem que isso implique que o utilizador tenha que interromper a visualização do conteúdo.

Estamos portanto na presença de uma tecnologia que permite que a partir de uma única codificação, efectuada de forma a satisfazer o nível máximo de qualidade pretendido, o mesmo conteúdo possa ser visualizado em virtualmente qualquer dispositivo, sem que isso represente uma perda significativa de qualidade, face às expectativas do utilizador em função do contexto de utilização.

## Tipos de escalabilidade

Existem diversas formas de adaptar um conteúdo audiovisual para que este possa ser classificado como escalável. No que diz respeito às características de um sistema de vídeo escalável, este deve contemplar pelo menos 3 tipos de escalabilidade: temporal, espacial e qualitativa[11]. Um feed de vídeo escalável pode fazer uso de uma destas técnicas individualmente ou combina-las entre si.

### **Escalabilidade temporal**

Uma das formas de controlar o fluxo de informação numa *stream* passa por ditar qual o volume de informação que esta deverá debitar numa fracção de tempo, neste caso num segundo. Esta situação não tem que implicar necessariamente que a *bitrate* de uma *stream* tenha que ser reduzida, até porque em determinadas circunstâncias isso poderia traduzir-se num nível de degradação de imagem incompatível com os objectivos pretendidos. Deste modo, não sendo possível interferir com a estrutura interna do *frame*, resta a possibilidade de reduzir o número de *frames*. Esta técnica é particularmente eficaz quando a cena sofre poucas variações, casos em que a diminuição da qualidade percebida é praticamente nula para o receptor.

### **Escalabilidade espacial**

A escalabilidade espacial é outra das formas de adaptação de conteúdos, implicando a diminuição das dimensões do vídeo. Esta técnica de adaptação segue o princípio de que dispositivos de menor capacidade têm normalmente ecrãs com resoluções inferiores. Ao serem necessários menos *pixels* para representarem o mesmo conteúdo, a largura de banda necessária para o vídeo diminui consideravelmente.

## **Escalabilidade qualitativa**

Esta forma de escalabilidade é simultaneamente aquela que tem uma relação mais directa entre a poupança de largura de banda e a diminuição da qualidade perceptível do vídeo. O método passa simplesmente por utilizar menos bits para codificar cada amostra, reduzindo desta forma a largura de banda necessária para o sinal.

### **2.2.2 - Controlo de *streams***

Determinadas as formas através das quais se torna possível adaptar um conteúdo ao seu contexto de utilização, importa também definir de que forma e por quem é que esse controlo é feito. O que iremos seguidamente abordar são as diferentes abordagens possíveis no que toca a este aspecto.

Existem duas formas de efectuar o controlo de uma *stream*: ou no emissor ou no destinatário[12].

No caso do controlo baseado no emissor, este é o responsável por fazer as alterações necessárias no conteúdo. Para tal, o emissor utiliza as informações disponíveis sobre o estado da rede, de forma a adequar o vídeo à largura de banda disponível. Desta forma, evita-se que sejam transmitidos mais pacotes do que aqueles que a rede poderia comportar, o que resultaria em pacotes que seriam descartados, deteriorando claramente a qualidade perceptível do conteúdo. Este método pode ser aplicado tanto a situações de unicast como de multicast.

Por oposição, o controlo no receptor é mais adequado a situações de multicast, uma vez que o receptor não faz mais do que ir adicionando ou removendo canais, consoante a sua capacidade de recepção e processamento de sinal.

Perante a possibilidade de serem geradas múltiplas *streams*, a IBM propõe uma série de políticas que, a serem implementadas, irão auxiliar um sistema a decidir qual ou quais as *streams* que deverão ser usadas[9].

### ***Response time VS Best Fit tradeoff***

De acordo com este critério é dada a prioridade ao conteúdo que conseguir dar uma resposta mais rápida ao pedido. Significa isto que, se já existir um ficheiro MPEG-4 capaz de chegar mais rapidamente ao utilizador, este é seleccionado, mesmo que a geração de um novo ficheiro XMT pudesse resultar numa experiência mais adequada às circunstâncias.

### **Fidelidade**

Segundo esta política de adaptação é dada preferência a cada um dos parâmetros capazes de produzir uma melhor experiência audiovisual para o utilizador.

### **Personalização de parâmetros adaptáveis.**

Este critério dá prioridade a determinados parâmetros, apostando primeiro na melhoria daqueles a que for atribuída maior importância. Significa portanto que se, por exemplo, o utilizador considerar que a dimensão do vídeo é mais importante do que a *bitrate* o sistema tenta sempre que possível efectuar a adaptação do conteúdo seguindo essas indicações.

## **2.2.3 Aplicações**

Pelas suas características, o vídeo escalável é elegível para ser aplicado num conjunto de situações bastante heterogéneo, na verdade é nestes ambientes de características diversas que as potencialidades mais sobressaem. É contudo também nestes ambientes que mais dificuldades se apresentam quando se pretende garantir uma qualidade de serviço adequada para a visualização de conteúdos multimédia. As variações de largura de banda e a diversidade de características dos dispositivos com acesso a este tipo de redes, tornam mais difícil a criação de conteúdos capazes de responder aos requisitos de cada um dos contextos de utilização possíveis [13]. Neste contexto a possibilidade de desenvolver conteúdos multimédia adaptáveis, como o vídeo escalável, reveste-se de particular importância e utilidade.

## **Contextos de largura de banda reduzida**

Como se compreenderá, em contextos onde a largura de banda disponível para transportar um sinal audiovisual é reduzida, a eficiência que é empregue na gestão da largura de banda disponível é um factor crítico. Ao permitir que um *feed* de conteúdos audiovisuais possa ser progressivamente reduzido (ou melhorado) esta tecnologia permite garantir que o destinatário recebe sempre a mensagem, ainda que o conteúdo desta possa ser mais ou menos rico. Desta forma, os utilizadores que operam nestes contextos continuam a ter sempre acesso à informação, ainda que adaptada às possibilidades do meio.

## **Contextos com elevados requisitos de largura de banda**

Em circunstâncias em que o sinal audiovisual requer elevada largura de banda ocorre precisamente a situação inversa do cenário descrito anteriormente, muito embora os resultados e requisitos do sistema acabem por se revelar bastante semelhantes. A alta-definição, por exemplo, ao gerar *streams* com considerável peso obriga necessariamente a uma elevada eficiência na gestão das mesmas, para que não ocorra desperdício de largura de banda. Ao permitir que um *feed* de alta-definição possa ser decomposto de forma a servir também contextos de capacidade inferior, garante-se uma poupança elevada de largura de banda ao mesmo tempo que se permite uma difusão virtualmente universal.

## **Contextos de difusão virtualmente universal**

Este é contexto por excelência em que o vídeo escalável é utilizado. Ao permitir decompor um *feed* em várias *streams* adaptáveis a diferentes redes, dispositivos e até mesmo aos gostos pessoais do destinatário, o vídeo escalável permite que a informação audiovisual seja codificada apenas uma única vez, sendo posteriormente difundida e recebida adaptada ao respectivo contexto. Recorrendo novamente ao exemplo do sinal de alta-definição, este depois de ser

difundido poderá ser também visionado em dispositivos que não estejam preparados para alta definição (situação que certamente será frequente num período de transição), ou então em dispositivos móveis, como são o caso dos PDAs ou telemóveis. Deste modo, a grande variedade de redes e dispositivos com capacidades de reproduzir conteúdos audiovisuais não constituem um problema ao nível das discrepâncias entre eles, mas antes uma vantagem, uma vez que tornando o conteúdo escalável este fica de uma só vez disponível simultaneamente em qualquer contexto de utilização.

Os princípios do vídeo escalável pela sua versatilidade e pela facilidade com que os conceitos aqui expostos podem ser aplicados em *streams* MPEG-4, apresentam-se por isso como uma forma de otimizar *streams* MPEG-4. De seguida iremos ver como um sistema que combina vídeo MPEG-4 e descrições MPEG-7 pode ser usado em combinação com os conceitos de vídeo escalável.

## 3 – Adaptação e escalabilidade

Explicadas as tecnologias envolvidas, o capítulo que agora iniciamos irá abordar de uma forma mais específica o funcionamento da adaptação e escalabilidade de conteúdos ao nível do edVO. É detalhado qual é papel de cada uma das tecnologias na solução de adaptação que se pretende desenvolver. Em um segundo momento é também traçada a arquitectura geral do modelo de adaptação proposto ao nível da classificação de conteúdos. São explicadas as diferentes soluções encontradas para classificar os conteúdos com recurso ao MPEG-7. É com base na classificação proposta que todo o modelo conceptual que é traçado mais à frente foi posteriormente desenvolvido.

### 3.1 – Módulo de adaptação

O projecto desenvolvido a partir do edVO, uma *framework* desenvolvida por Luís Filipe Teixeira, que sofreu posteriormente um contributo de Moisés Medeiros ao nível do MPEG-7. O edVO, como já tivemos oportunidade de observar, inclui uma série de ferramentas para a criação, manipulação e distribuição de conteúdos MPEG-4. A ideia na base deste projecto é, tirar partido das ferramentas de descrição adicionadas ao edVO, utilizando-as para decidir qual o tipo de escalabilidade a realizar em cada conteúdo.

O conteúdo adaptável já não constitui novidade para quem está habituado a aceder a conteúdos em *streaming* publicados na internet. Contudo, após uma análise mais atenta à forma como a adaptação é feita, é facilmente observado que esta leva apenas em conta a largura de banda disponível, afectando o conteúdo por inteiro. Uma análise sintáctica do conteúdo leva-nos a concluir que diferentes conteúdos contêm elementos (ou objectos) de relevância variada para a cena em questão. A título de exemplo, pode-se pensar nos casos do entretenimento e da informação. Quando a natureza de um conteúdo é eminentemente informativa, alguma imagem adicional pode ser sacrificada de forma a reservar largura de banda para melhorar aquilo que constitui o núcleo da

mensagem: a informação. Quanto ao entretenimento, no caso de um *videoclip* poder-se-ia classificar como mais relevantes a música e a imagem do vocalista, ao passo que os restantes elementos serão de relevância inferior.

As tecnologias que são disponibilizadas pelo MPEG-4 permitem, pela própria natureza da norma, tirar de forma mais directa partido da estrutura das cenas MPEG-4 de forma a aumentar a qualidade percebida pelo utilizador de um sistema, tal como o que aqui é proposto.

As estruturas de objectos nas quais assentam as cenas MPEG-4 são ideais para este tipo de escalabilidade que propomos uma vez que permitem um controlo mais preciso sobre cada um dos elementos (ou grupos destes) presentes numa cena MPEG-4. A norma MPEG-7 surge nesta dissertação como resposta à necessidade de classificar de uma forma precisa e inequívoca cada um dos objectos, para o seu processamento possa ser efectuado de uma forma que se traduza num efectivo aumento da qualidade conteúdo que é servido.

### 3.1.1 Tecnologias

O projecto desenvolvido do âmbito desta dissertação pode ser subdividido em quatro campos ou componentes:

- MPEG-4
- MPEG-7
- Normas
- Plataforma

A componente MPEG-4 diz respeito a todo o conteúdo audiovisual, ditando a sua estrutura e representação. É na tecnologia que suporta esta norma que recai também a parte final da adaptação. Toda a operação de modificação da *stream* mpeg-4 inerente à adaptação de conteúdos só é possível graças aos mecanismos internos desta norma.

O papel do MPEG-7 está relacionado directamente com o intuito que levou à sua criação: a classificação de conteúdos. Para alcançar os objectivos propostos nesta dissertação foram utilizados especificamente dois descritores MPEG-7 para permitir o armazenamento de informação quantitativa que servisse de base às decisões que o sistema terá que tomar durante o processo de adaptação.

As normas constituem nesta dissertação o núcleo agregador destas duas últimas tecnologias. Será com base no algoritmo de decisão definido que as normas para a adaptação de conteúdo serão implementadas, permitindo utilizar informação MPEG-7 para actuar sobre *streams* MPEG-4.

Por último, os dispositivos são o elo final nesta cadeia. São os dispositivos que servem de interface entre o utilizador e o sistema e, como tal, limitam a informação que chega aos utilizadores[14].

### 3.1.2 Funcionamento

O processo envolvido na autoria de conteúdo adaptável, e posteriormente na sua subsequente adaptação, é bastante simples muito embora possua um considerável número de variáveis que por vezes o tornam algo denso no que à sua compreensão diz respeito.

A autoria do conteúdo é a primeira etapa deste processo e tudo aquilo que ficar estabelecido nesta fase influencia definitivamente toda e qualquer utilização que possa ser dada ao conteúdo.

Como já foi explicado anteriormente, uma cena MPEG-4 é composta por múltiplos objectos, podendo cada um desses objectos ter diferentes representações. No processo de autoria são definidos, entre outras coisas, o posicionamento de cada um dos objectos

na cena, a sua interacção com os restantes objectos e as diferentes representações que existem para cada objecto. É também adicionada informação MPEG-7 que permite identificar ao nível de cada uma das instâncias quais as condições que deverão ser verificadas para que seja carregada para a cena MPEG-4 a instância em questão.

Depois do processo de publicação dos conteúdos, isto é, a sua disponibilização num servidor concebido para o efeito, qualquer cliente poderá aceder aos mesmos desde que reúna as condições necessárias para a sua visualização. Neste contexto, entende-se cliente como sendo o conjunto formado pelo equipamento terminal (PC, telemóvel, PDA ou qualquer outros dispositivo com capacidade de reprodução vídeo). As condições referidas são a existência no cliente de software capaz de reproduzir vídeo MPEG-4. Para que seja possível a escalabilidade, nos moldes em que esta é proposta nesta dissertação, é ainda necessário que o *player* existente no cliente seja capaz de comunicar ao nível da camada de monitorização com o servidor. A adaptação do conteúdo depende da capacidade do *player* monitorizar as condições em que se efectua a reprodução da *stream* e de transmitir essa informação ao servidor para que este reaja em conformidade.

## **3.2 - Classificação e escalabilidade**

### **3.2.1 Descrição de cenas e objectos**

É com base em alguns dos descritores MPEG-7 que o servidor, via motor de adaptação, vai poder tomar decisões quanto às alterações a efectuar de modo a melhor adequar o conteúdo às condições em que este se encontra a ser visionado.

O MPEG-7 surge portanto, tal e qual como já estabelecemos, com o objectivo de cumprir a função primordial para que este foi criado: descrever, neste caso específico, conteúdos multimédia. No entanto, atendendo ao modelo desenvolvido, a importância deste no resultado final revela-se primordial. Para que a adaptação dos conteúdos possa ocorrer para que se traduza uma efectiva e tão eficaz quanto possível melhoria da

experiência do utilizador é necessário que por um lado o conteúdo seja classificado de uma forma inequívoca e que, por outro lado permita tanto quanto possível o seu agrupamento em tantas combinações quantos os seus contextos de utilização.

O MPEG-7 inclui um descritor que é frequentemente apontado como sendo um dos possíveis pilares para a adaptação de conteúdos: o *Variation DS*[5]. Neste descritor são enumeradas e classificadas diferentes variações do mesmo conteúdo. Estas variações de conteúdo são descritas em detalhe, sendo enumerado qual o seu grau de fidelidade relativamente ao conteúdo original, tipo de variação (por exemplo, redução de cor, compressão espacial, resumo, etc) e qual a prioridade dada a cada uma das variações. Ao enumerar e classificar as diferentes versões de um mesmo conteúdo, torna-se possível a um servidor de conteúdos ou proxy, atendendo às condições da rede e eventualmente a preferências definidas pelo utilizador, decidir qual a variação mais indicada para o dispositivo para o qual o conteúdo está a ser enviado.

No entanto as variações constituem um modelo estático de adaptação, isto é, quando ocorre o processo de autoria do conteúdo e a sua subsequente classificação de acordo com o que acabamos de descrever, o resultado é um conjunto de variações que, para ser alterado, implica a criação de novas variações. Enquanto a criação de nova variação não ocorrer o utilizador fica sempre limitado às variações existentes. Em processos de adaptação em que exista uma multiplicidade de dispositivos ou um elevado número de variáveis a considerar no processo de adaptação, a existência de um número limitado de variações do conteúdo limita sempre de alguma forma a adaptação “à medida” deste a circunstâncias que não estejam previstas directamente através da criação de uma variação conteúdo específica ou equivalente.

Em face destas limitações, propomos neste trabalho que se combinem descritores MPEG-7 já existentes para classificar o conteúdo quanto à natureza da mensagem (isto é, por exemplo, informativa, entretenimento, cinematográfica, etc) e quanto à relevância de cada objecto audiovisual para uma aproximação mais fiel ao conteúdo original.

A nível da natureza da mensagem, uma análise atenta dos conteúdos audiovisuais susceptíveis de serem visionados em redes multimédia permitiu-nos enquadrá-los nas seguintes famílias:

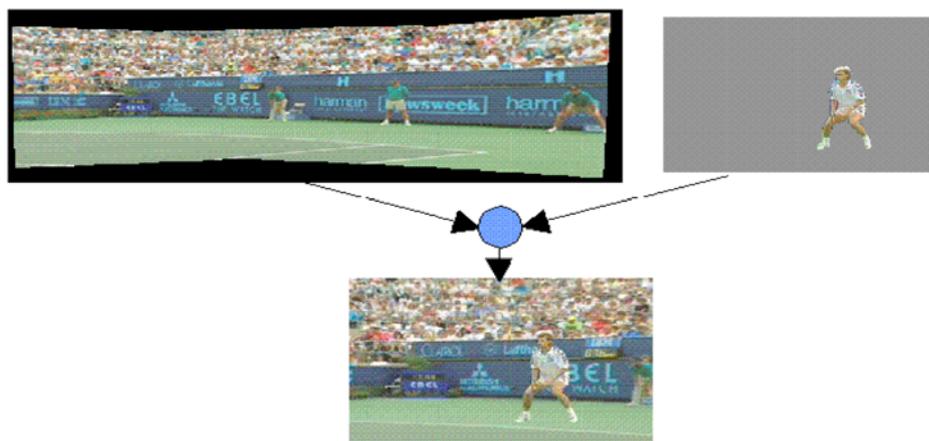
- *Talking heads*
- Informação
- Cinema/Documentários
- Desporto
- Música

**Talking heads** é o termo que é utilizado para designar as situações em que a imagem é essencialmente constituída pela imagem de um rosto de um apresentador que transmite uma mensagem de conteúdo normalmente informativo / expositivo. Para além da imagem do rosto do pivot e do som da respectiva voz, a cena pode conter também elementos auxiliares de informação, usualmente de natureza gráfica. A inclusão deste tipo de conteúdo numa categoria independente da categoria “informação” está relacionada com as especificidades de ambas as categorias. No caso dos pivots (*talking heads*) a mensagem é também ela de natureza informativa mas, desprovida de outros elementos auxiliares de informação de natureza não gráfica. Neste tipo de conteúdo não se encontram imagens que ilustrem aquilo que o emissor da mensagem transmite, além da própria imagem do pivot e dos já referidos elementos gráficos de apoio (este últimos opcionais).

A **informação** comporta todos os conteúdos cuja natureza é eminentemente informativa, mas em que existem conteúdos audiovisuais que permitem ilustrar de uma forma mais directa o conteúdo da mensagem. Um bom exemplo de um conteúdo deste tipo são os noticiários que, iniciando a leitura da notícia por parte do pivot ou do jornalista no local, vão após esta introdução ilustrar e/ou complementar a informação fornecida com imagens do acontecimento e eventualmente depoimentos de intervenientes no mesmo.

Na categoria **cinema / documentário** foram incluídos os conteúdos cujo o essencial da mensagem se encontra no próprio conteúdo audiovisual, concentrando-se por norma a atenção do espectador predominantemente na componente visual do conteúdo. Este último aspecto é extremamente importante para a adaptação, relacionando-se de forma bastante directa com o conceito de qualidade perceptível. Certamente que um utilizador que assiste a um filme no seu *PDA* não dará importância a escutar um som com qualidade inferior à de um *CD*, pois é normalmente percebido pelos utilizadores que este tipo de dispositivo não tem a mesma qualidade sonora que a sua aparelhagem de casa. No entanto o mesmo utilizador, caso sinta que a imagem não lhe proporciona uma qualidade agradável (ainda que inferior à de um ecrã de televisão) irá muito possivelmente revelar o seu desagrado.

A categoria **desporto** comporta todas as imagens relativas a um evento desportivo. Esta categoria possui, ao nível da percepção do utilizador, características que podem ser encontradas nos últimos dois grupos de conteúdos referidos. Os conteúdos desportivos têm por um lado uma forte componente informativa, uma vez que o espectador espera manter-se a par dos desenvolvimentos de um determinado acontecimento desportivo. Por outro lado estes conteúdos têm uma componente de entretenimento, uma vez que a informação que é transmitida é de natureza mais “leve”, permitindo ao espectador concentrar-se no evento em si. A inclusão desta categoria visa também encontrar uma forma de lidar com uma outra característica deste tipo de conteúdo: a inserção de elementos gráficos, que a tecnologia tem permitido que sejam cada vez mais complexos como é o caso da publicidade virtual.



**Figura 3- Exemplo de uma cena de conteúdo desportivo[2]**

A inclusão da categoria **música** deve-se à atenção particular que este tipo de conteúdo de entretenimento merece ao nível do áudio. Tal como afirmamos relativamente a outro tipo de conteúdos, a imagem desempenha um papel primordial na forma como o receptor percebe a qualidade da experiência audiovisual. No entanto, os telediscos (ou *videoclips*) ao fundirem o entretenimento com a música criam um tipo de conteúdo em que se fundem dois ingredientes a que o receptor presta especial atenção: por outro lado a música é o motor da acção e é em torno desta componente sonora que os restantes conteúdos são desenvolvidos. Contudo, tal como acabamos de ver a imagem, pelo impacto que esta tem junto do receptor, não pode ser de forma alguma descurada, justificando-se deste modo um tratamento diferenciado para este conteúdo.

No que diz respeito ao segundo aspecto a classificar (a relevância de cada instância para o conteúdo), espera-se que no processo de criação de uma cena MPEG-4 o autor efectue, num primeiro nível, uma avaliação sobre a relevância que cada objecto tem para a cena que acabou de criar, e num segundo nível uma ponderação entre os requisitos de que cada instância necessita para que seja percebida correctamente, e a importância de cada instância para o conteúdo final.

Pretende-se que o autor, efectue uma avaliação quantitativa tendo em conta os factores já referidos e utilizando uma escala (que para efeitos desta dissertação será de 0 a 10 ) que classifique cada uma das instâncias dos objectos. Espera-se também que o autor

determine quais são os objectos fundamentais para a cena, obrigando-o a que estes surjam sempre na mesma, para o efeito atribuindo a uma instância de cada um desses objectos a classificação de 0 (zero).

### 3.2.2 Descritores usados

Para classificar os conteúdos de acordo com os princípios que foram traçados no início deste capítulo foram seleccionados dois descritores MPEG-7, um para cada um dos aspectos que se pretendem descrever.

Existem dois grupos de conteúdos que se torna necessário descrever de forma independente: as cenas e os objectos.

#### Descrição de uma cena MPEG-4

Uma cena MPEG-4 é composta por diferentes objectos, cada um desses objectos possuindo propriedades distintas. Para o modelo proposto nesta dissertação iremos considerar também que as apresentações possuem uma propriedade que as coloca numa determinada categoria ou grupo de apresentações: O género. Como iremos ver mais adiante nesta dissertação, o comportamento do motor de adaptação irá variar em função da natureza do conteúdo da cena. Na descrição da cena, para efeitos da solução aqui proposta, será utilizado o descritor MPEG-7 *Classification D*, reservando-se o campo *Genre*, relativo ao género de conteúdo para a indicação da natureza deste.

O descritor *Classification D* insere-se dentro das ferramentas de criação de produção do MPEG-7. Detalhando um pouco mais a sua inserção numa descrição MPEG-7, o descritor *Classification D* é um dos componentes do esquema de descrição *CreationInformation DS*, e dentro deste esquema de descrição é um descritor opcional[5]. O campo que nos interessa observar dentro do descritor *Classification D* é o item *Genre*, que classifica o conteúdo quanto ao seu género, isto é, quanto à natureza daquilo que este representa. A adequação deste campo é perfeita para o tipo de classificação que se pretende uma vez que o que está a ser feito é classificar o conteúdo

de acordo com a sua tipologia. No entanto, é de elevada importância a existência de uma terminologia uniformizada a este nível. Apenas definindo uma classificação uniforme é possível assegurar que o software a correr do lado do servidor é capaz de efectuar a adaptação levando em conta a natureza do conteúdo.

O exemplo que se segue ilustra a classificação de um conteúdo de desporto, de acordo com o que acabamos de descrever

```
<AudioVisual xsi:type="AudioVisualSegmentType">
  <CreationInformation>
    <Creation>
      <Title>Manchester United VS Chelsea</Title>
      <Abstract><FreeTextAnnotation>Jogo entre o Manchester
        United e o Chelse na final da FA Cup 2006/2007
      </FreeTextAnnotation></Abstract>
      <Creator>Channel 4</Creator>
    </Creation>
    <Classification>
      <Genre type="main">Desporto</Genre>
      <Language type="original">Inglês</Language>
    </Classification>
  </CreationInformation>
```

**Figura 4 - Descrição parcial de uma cena em MPEG-7**

É possível introduzir ainda um maior controlo ao nível da DDL, de forma a garantir que é introduzido no descritor apenas uma das categorias que foram definidas. Deste modo assegura-se que o servidor tem sempre condições para determinar qual o tipo de adaptação que deve efectuar. A figura que se segue, ilustra o exemplo de uma DDL que restringe os parâmetros a serem introduzidos em “*Genre*”.

```
<complexType name="GenreType">
  <complexContent>
    <extension base="mpeg7:ControlledTermUseType">
      <attribute name="type" default="main">
        <simpleType>
          <restriction base="NMTOKEN">
            <enumeration value="talking head"/>
            <enumeration value="informacao"/>
          </restriction>
        </simpleType>
      </attribute>
    </extension>
  </complexContent>
</complexType>
```

```

        <enumeration value="cinema/documentario"/>
        <enumeration value="desporto"/>
        <enumeration value="musica"/>
    </restriction>
</simpleType>
</attribute>
</extension>
</complexContent>
</complexType>

```

**Figura 5 - DDL MPEG-7**

## Descrição de um objecto MPEG-4

Os objectos MPEG-4, tal como as cenas, necessitam de uma descrição para que o motor de adaptação possa saber qual o conteúdo de cada objecto, e quais as variações existentes.

A primeira necessidade ao nível da adaptação é descrever o tipo de objecto. O motor de adaptação, em função das indicações fornecidas sobre a cena, vai ter diferentes comportamentos com objectos distintos, atribuindo mais importância a uns e menos a outros em cada momento em que ocorra adaptação. Também nesta situação o descritor *Classification D* irá ser usado para descrever o objecto, recorrendo quer ao campo *Genre* quer ao campo *Abstract*. No campo *Genre* deverá ser indicado o que o objecto representa enquanto elemento da cena, enquanto que o campo *Abstract* poderá ser usado para uma descrição mais detalhada.

```

<AudioVisual xsi:type="AudioVisualSegmentType">
  <CreationInformation>
    <Creation>
      <Title>Noticias de 10 de Novembro 2006</Title>
      <Abstract><FreeTextAnnotation> Pivot lê as notícias
do bloco informativo das 20 horas
</FreeTextAnnotation></Abstract>
      <Creator>Channel 4</Creator>
    </Creation>
    <Classification>
      <Genre type="main">peessoa/Genre>
      <Language type="original">português</Language>

```

```
</Classification>  
</CreationInformation>
```

### Figura 6 - Descrição em MPEG- 7 de um objecto

O exemplo da figura anterior (figura 6 ) ilustra uma classificação de um objecto do tipo “pessoa”, neste caso o pivot de um noticiário televisivo.

A forma encontrada para descrever a fidelidade de uma instância relativamente ao conteúdo original surge através do descritor *MediaQuality D*. Este descritor insere-se dentro de esquema de descrição *MediaInformation DS* e dentro deste, como um elemento do descritor *MediaProfile D*. O descritor *MediaQuality D* destina-se a medir de forma objectiva ou subjectiva a qualidade de um conteúdo audiovisual [5]. No caso da proposta que é feita nesta dissertação esta classificação tem uma dupla função: indicar ao servidor quando é que uma instância de um objecto deve ser carregada para a cena MPEG-4, e classificar o conteúdo quanto à sua fidelidade ao original. Muito embora este descritor tenha sido pensado apenas para esta última função, atendendo ao sistema que aqui é proposto, é possível efectuar rapidamente uma extrapolação para uma função de adaptação.

Atendendo à escala já mencionada de 0 (zero) a 10 (dez), zero representa o conteúdo com o nível de fidelidade mais baixo relativamente ao conteúdo original, mas para o servidor de adaptação, representa também que este conteúdo tem o nível mínimo de qualidade para fazer parte da cena MPEG-4, independentemente do contexto de utilização. Por oposição a esta classificação, uma classificação de dez representa que o conteúdo é totalmente fiel ao conteúdo original, deduzindo-se também que será necessário um contexto de utilização muito favorável para que esta instância possa ser carregada para a cena MPEG-4. Segue-se uma descrição MPEG-7 do que poderia ser uma classificação de acordo com o que aqui é proposto.

```
<MediaQuality>  
  <QualityRating ratingType="objective">  
    <RatingValue>8</RatingValue>
```

```
</QualityRating>  
</MediaQuality>
```

### **Figura 7 - Classificação de objecto**

Atendendo ainda ao descritor *MediaProfile D*, este possui ainda dentro da sua árvore de influência o descritor *MediaInstance D*. Este descritor foi concebido para que diferentes instâncias do mesmo conteúdo possam aqui ser mencionadas através de uma descrição textual ou então de um URI. Contudo, atendendo a que a DDL MPEG-7 não nos permite descrever individualmente cada instância, associando-lhe uma descrição *MediaQuality D*, nesta dissertação iremos optar por uma outra abordagem como forma de descrever cada instância.

A solução para o que acabamos de referir está no descritor base *MediaProfile D*. A DDL MPEG-7 permite descrever vários perfis, bastando indicar qual é o perfil principal (*Master*) através do atributo *Master*. Embora menos eficiente do que a simples instanciação das diferentes versões dentro do mesmo perfil, esta solução garante todos os requisitos necessários para que possa advir correctamente a adaptação de conteúdos tal como é proposta nesta dissertação.

Estão reunidos desta forma as 3 variáveis ao nível da descrição que nos irão permitir identificar os conteúdos e cada uma das suas variações e proceder à sua classificação de forma correcta. Vejamos então:

O descritor *Classification D* permite identificar com que tipo de conteúdo o sistema está a lidar de forma a efectuar a adaptação em função desse critério.

O descritor *MediaProfile D*, possibilita que possam ser enumeradas as diferentes instâncias do mesmo objecto, através da criação de um perfil para cada uma dessas instâncias.

O descritor *MediaQuality D* , permite classificar cada instância dos objectos utilizados na cena, possibilitando desta forma que o servidor efectue a adaptação dos conteúdos seleccionando em cada momento a instância que mais se adequa a cada contexto de utilização.

## 4 – Solução proposta

Este capítulo é o resultado das análises devolvidas em capítulos anteriores. Com base nas observações que foram sendo feitas procurou-se criar um modelo conceptual no qual fossem vertidos os princípios enunciados até ao momento.

Neste capítulo é traçada a estrutura base do modelo e explicado o seu funcionamento ao nível tecnológico. São também explicadas quais as conclusões tiradas durante o processo de implementação do modelo e apontados constrangimentos, dificuldades e trabalho futuro.

### 4.1 – Adaptação de conteúdos

O processo de adaptação de conteúdo é o resultado da conjunção de dois tipos de informação: a descrição do conteúdo (que acabamos de expor) e as condições de utilização em que ocorre o visionamento do conteúdo multimédia. O modelo é desta dissertação assenta numa arquitectura cliente – servidor, isto é, há um dispositivo que efectua o pedido e um outro que fornece uma resposta a essa solicitação. Este é também o modelo em que assenta o edVO, a *framework* sobre a qual foi desenvolvido esta dissertação.

Contundo, apesar do modelo cliente – servidor se manter, a solução que é proposta implica que o software a correr no cliente tenha um papel activo. A adaptação de conteúdos, tal como é aqui sugerida, implica que durante o envio de conteúdos para o cliente, o servidor tenha conhecimento periódico das condições em que o serviço está a ser usado. A aplicação que corre no cliente deverá por isso ser capaz de realizar periodicamente um conjunto de testes às referidas condições de utilização e de comunicar os resultados ao servidor para que esta possa efectuar a adaptação.

O servidor, por seu lado, considera a informação que lhe é transmitida pelo cliente, levando também em conta eventuais preferências que o utilizador possa ter definido no sistema. As preferências do utilizador, caso existam, servem de orientação para o

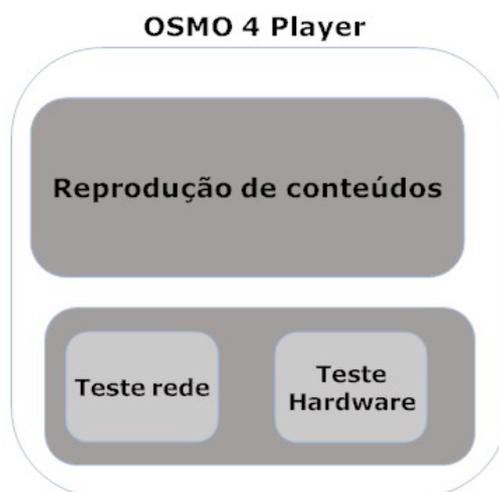
processo de geração da cena MPEG-4 sendo sempre carregadas em primeiro lugar e mantidas sempre que as condições de utilização assim o permitam.

Caso o utilizador não tenha definido qualquer preferência o servidor começa por enviar a cena no seu formato mais simplificado, melhorando progressivamente a experiência do utilizador à medida que for verificando que o contexto de utilização é favorável à realização desse mesmo incremento de qualidade.

### 4.1.1 – O cliente

O cliente referido nesta proposta é o OSMO4, um software de produção MPEG-4 integrado no GPAC. O OSMO4 foi concebido como uma versão reduzida da versão anterior que era baseada no software de referência de sistema MPEG-4 [15].

O OSMO4 possui funcionalidades que proporcionam um bom desempenho com conteúdos escaláveis. No entanto não fazem parte deste software as funcionalidades que permitem a este software aferir qual o contexto de utilização em que o conteúdo está a ser visionado e comunicar esses resultados ao servidor. Propõe-se portanto a criação de uma camada adicional no OSMO4 que contemple a realização de testes às condições de utilização e a comunicação dos resultados desses mesmos testes ao servidor.



**Figura 8 - Modelo desenvolvido para o OSMO4**

Os testes a implementar pelo software de reprodução são bastante simples e baseiam-se em dois critérios: na transferência de pacotes, (ou seja, no estado da rede) e nas características físicas do hardware utilizado. Estes dois testes pretendem traçar um retrato parcial daqueles que são os dois grandes factores de limitação à reprodução de conteúdos em redes multimédia: a capacidade da rede e a capacidade do dispositivo.

Por capacidade da rede, entende-se como sendo a característica desta que lhe permite transportar um determinado volume de informação num dado momento. Por capacidade do dispositivo, entende-se que são as características físicas deste que se relacionam directamente com a reprodução de conteúdos multimédia (resolução do ecrã, capacidade de reprodução de som, capacidade de armazenamento, etc). No que às características do dispositivo utilizado para a reprodução de conteúdos não se considera neste aspecto o software, uma vez que se pressupõe a existência prévia de um software capaz de efectuar a reprodução dos conteúdos multimédia (o próprio OSMO4).

Detalhando um pouco mais os testes a realizar pela camada construída sobre o software de reprodução, eles são como já vimos teste de rede e testes às características físicas do dispositivo. No que respeita ao teste de rede, é obtido um determinado resultado através da estimativa da transferência de pacotes ao longo do tempo. Caso o volume da transferência de pacotes aumente significativamente o software irá atribuir uma classificação superior ao estado da rede, caso se verifique a situação inversa a pontuação atribuída ao estado da rede baixará. A pontuação máxima para as condições de utilização da rede será de 10 (dez) e a mínima de 1 (um).

No que respeita à aferição das características de hardware, o software de reprodução irá apenas verificar inicialmente 2 factores: a capacidade de armazenamento do dispositivo e a resolução permitida pelo ecrã. A verificação da capacidade de armazenamento pretende aferir qual o tamanho máximo do *buffer* que o software de reprodução pode reservar. Por outro lado, a verificação da resolução do ecrã é um factor limitador a nível da dimensão dos conteúdos que possam ser reproduzidos no dispositivo, mesmo que existam instâncias com qualidade superior disponíveis para a cena MPEG-4 e que as restantes condições de utilização permitam esse esforço extra, o servidor não irá

efectuar a adaptação se verificar que o dispositivo não tem capacidade de reprodução das instâncias com resolução superior. O resultado dos testes realizados às características do hardware também resulta da obtenção de uma classificação expressa de 1 (um) a 10 (dez).

A classificação final destes dois testes é obtida através de uma ponderação entre os testes realizados às condições da rede e os testes realizados às características físicas do dispositivo. No decorrer do desenvolvimento desta dissertação foi entendido que a média obtida deveria ser uma média aritmética, uma vez que, quer um, quer outro factor podem influenciar significativamente o contexto de utilização. O resultado dos testes é enviado ao servidor, sendo que quando da realização do primeiro teste será também enviada para o servidor a informação da resolução máxima permitida pelo dispositivo onde o conteúdo está a ser visionado.

#### **4.1.2 – O servidor**

O módulo de adaptação do edVO proposto nesta dissertação apoia-se no edVO.Server, o módulo de servidor da *framework*. O servidor inclui um repositório de *media objects* que podem ser utilizados na criação de cenas MPEG-4, estando associada a cada *media object* uma descrição MPEG-7, conforme foi explicado no capítulo anterior. O mesmo acontece com as cenas MPEG-4, que depois de criadas no edVO.Composer, são armazenadas no servidor, acompanhadas da respectiva descrição em formato MPEG-7. Para a elaboração desta deste trabalho foi projectada uma camada de adaptação, que efectue a interface entre o cliente e o servidor, realizando adaptações ao conteúdo sempre que tal se verifique ser a melhor opção.



**Figura 9 - integração do módulo de adaptação do edVO.Server**

Tal como já foi explicado aquando da definição das regras para a descrição de conteúdos, a adaptação assenta em aspectos relacionados com as propriedades de cada objecto, assumindo comportamentos diferentes consoante a classificação que for dada à cena MPEG-4. Seguidamente descreve-se o processo de adaptação para cada uma das famílias de conteúdos definidas para a descrição das cenas.

Os conteúdos com a classificação *talking heads* possuem características específicas, quer pela sua baixa complexidade ao nível da descrição da cena, quer pela natureza do próprio conteúdo. Em relação a este último aspecto, a natureza do conteúdo é exclusivamente informativa com nenhum ou muito pouco apoio ao nível de imagem ou gráficos que ilustrem a informação que está a ser transmitida pelo pivot. A baixa complexidade da cena introduz uma grande preponderância no pivot, enquanto centro das atenções e no conteúdo da informação que este transmite. A adaptação leva por isso em conta esta preponderância e tanto no primeiro momento de adaptação como num segundo momento de adaptação faz incidir a adaptação exclusivamente sobre o pivot. Apenas após estes dois níveis de adaptação, se ainda existirem condições que permitam

enriquecer mais o conteúdo, é melhorado o fundo da imagem e seguidamente, num outro momento de adaptação, acrescentados gráficos (caso estes existam.)

Concretizando aquilo que acabou de aqui ser exposto, numa cena MPEG-4: existiria uma pivot a ler as notícias, um fundo que seria colocado recorrendo a uma técnica de keying e um elemento gráfico. Na imagem a surge-nos a cena MPEG-4 apenas com a pivot e o fundo (ambos na sua resolução máxima).



**Figura 10 - Talking head sem grafismo**

A cena MPEG-4 em formato XMT-A que se segue descreve parcialmente a imagem.

```
<OD>
  <ObjectDescriptor objectDescriptorID="VideoStream"
  binaryID="3">
    <Descr>
      <esDescr>
        <ES_Descriptor ES_ID="VideoStream" binaryID="3">
          <decConfigDescr>
            <DecoderConfigDescriptor
            objectTypeIndication="MPEG4Visual" streamType="Visual"/>
          </decConfigDescr>
          <slConfigDescr>
            <SLConfigDescriptor><predefined
            value="2"/></SLConfigDescriptor>
          </slConfigDescr>
          <StreamSource url="pivot.mp4"/>
        </ES_Descriptor>
      </esDescr>
    </Descr>
  </ObjectDescriptor>
</OD>
```

**Figura 11 – Descrição parcial em XMT de uma cena do tipo "talking head"**

Após a adaptação realizada à cena, neste caso com a adição de um elemento gráfico, a descrição XMT-A da mesma alterou-se, uma vez que houve necessidade de incluir mais um objecto. Esta situação encontra-se ilustrada nas figuras 12 e 13, correspondendo à imagem obtida e à descrição da cena correspondente ao objecto adicionado, em formato XMT-A, respectivamente.



**Figura 12 - Cena após inserção de grafismo**

Após a inserção de grafismo, a descrição da cena é alterada passando a incluir uma descrição adicional, correspondente ao objecto que se acrescentou à cena MPEG-4. Esta transformação encontra-se ilustrada na figura que se segue.

```

<ObjectDescriptorUpdate>
  <OD>
    <ObjectDescriptor objectDescriptorID="JPEGImage"
binaryID="4">
      <Descr>
        <esDescr>
          <ES_Descriptor ES_ID="JPEGImage" binaryID="4">
            <decConfigDescr>
              <DecoderConfigDescriptor
objectTypeIndication="MPEG4Visual" streamType="Visual"/>
            </decConfigDescr>
            <slConfigDescr>
              <SLConfigDescriptor><predefined
value="2"/></SLConfigDescriptor>
            </slConfigDescr>
            <StreamSource url="Logo.jpg"/>
          </ES_Descriptor>
        </esDescr>
      </Descr>
    </ObjectDescriptor>
  </OD>
</ObjectDescriptorUpdate>

```

**Figura 13 - Descrição corresponde ao objecto "grafismo"**

Para os conteúdos da categoria **informação**, existe um comportamento que, embora privilegie os objectos que têm como elemento primordial pessoas (jornalistas, entrevistados, etc), atribui importância semelhante às restantes imagens que servem para ilustrar a notícia. Esta opção é tomada porque, ao contrário dos *talking heads*, a informação que é transmitida é significativamente complementada por um conjunto de imagens de apoio que ilustram a notícia que está a ser transmitida.

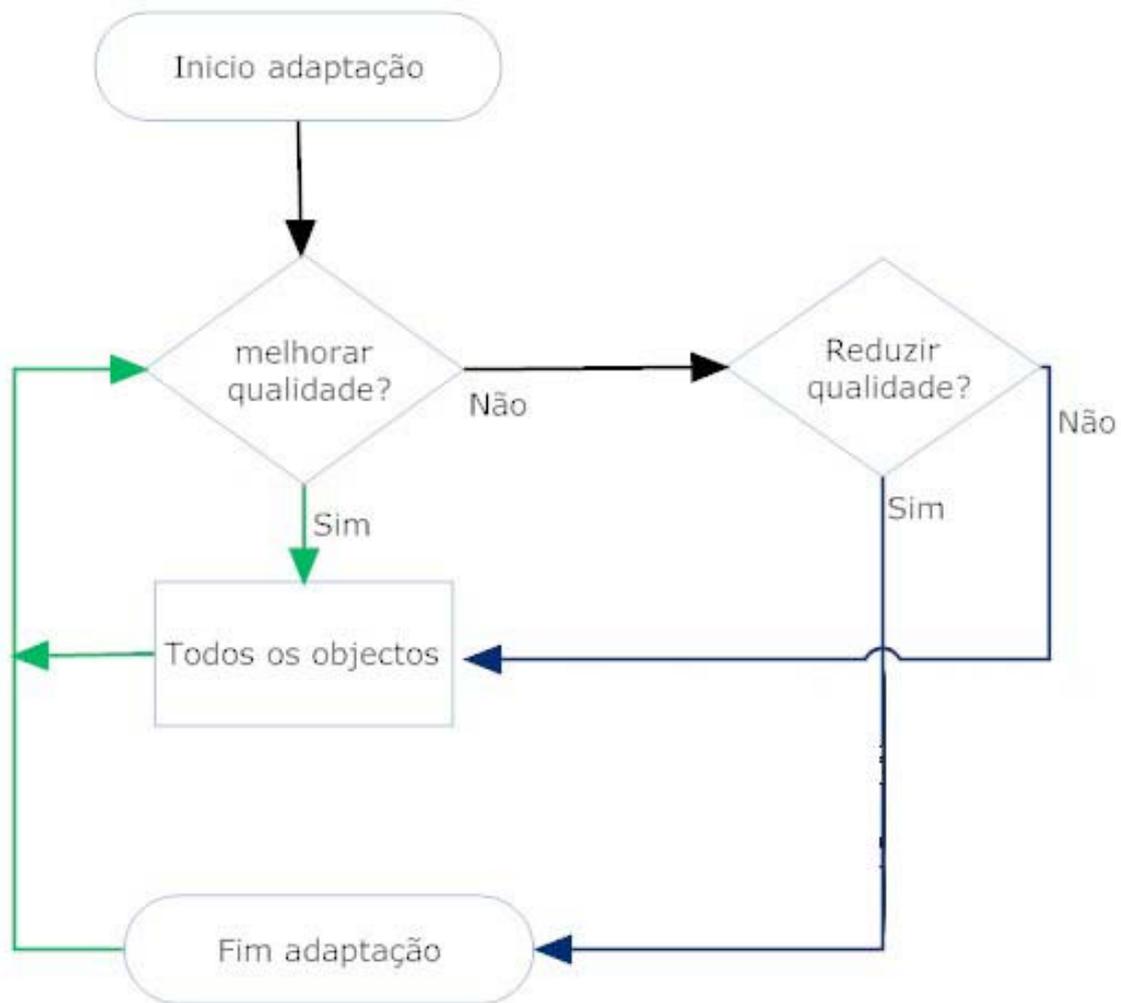
Deste modo, assume tanta importância o *media object* correspondente à pessoa que dá a notícia, como aquele que ilustra a informação que está a ser transmitida. Em termos de processo de adaptação de conteúdos, este raciocínio traduz-se num primeiro momento por uma adaptação dirigida ao objecto que contém as pessoas que transmitem a informação. Num segundo momento de adaptação, é escalado o conteúdo

correspondente às restantes imagens que acompanham a notícia. Caso ainda exista espaço para incrementar a qualidade da cena, todo o processo até agora descrito é repetido uma segunda vez. E só numa etapa posterior de adaptação são introduzidas outras *streams* com conteúdo de apoio, como por exemplo, grafismo.

A categoria **cinema / documentário** é também um caso específico de adaptação uma vez que, atendendo à natureza destes conteúdos, colocar uma maior ênfase num ou outro elemento do conteúdo implicaria o risco de desvirtuar a intenção do autor da obra.

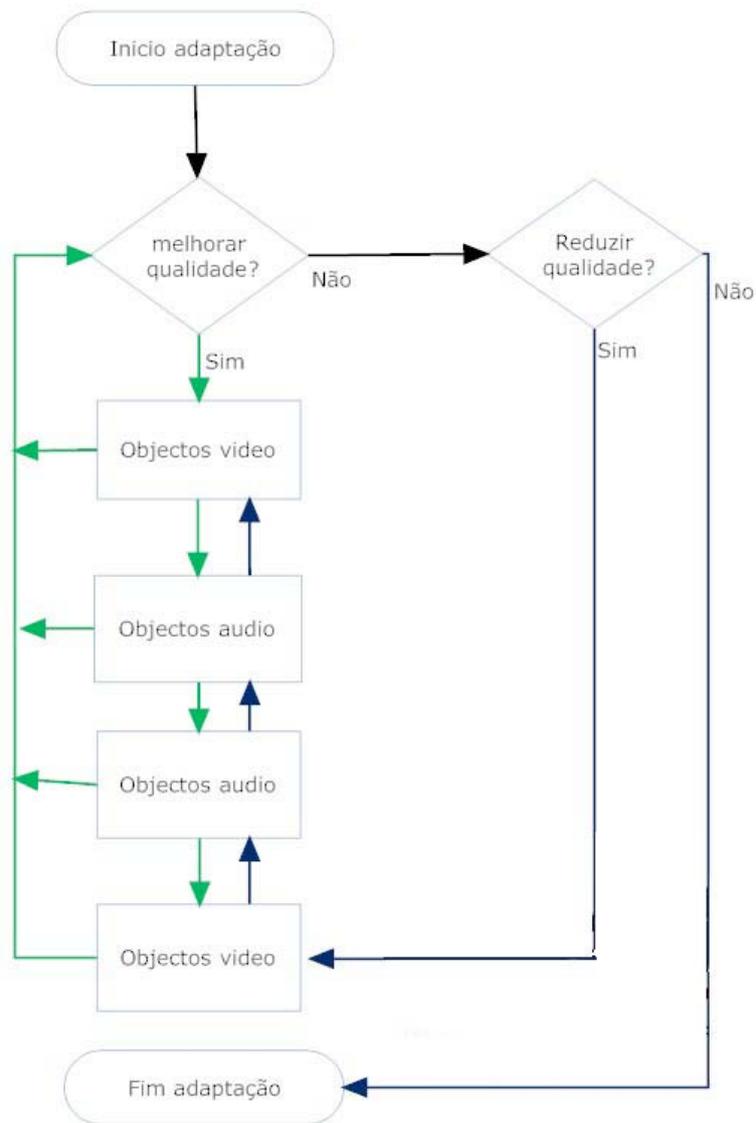
No caso do cinema, existem aspectos com a iluminação e o enquadramento que poderiam ser afectados de uma forma significativa, afectando negativamente a intenção do autor da obra e podendo mesmo prejudicar o seu sentido. No documentário, sendo também um tipo de conteúdo eminentemente expositivo e de fortíssima componente visual, o favorecimento de um determinado elemento em detrimento de outro também dos pareceu desaconselhável. Considerando estas especificidades, o modelo de adaptação proposto para este tipo de conteúdos implica que, a existirem condições para se verificar adaptação, esta deve afectar de igual forma todos os objectos que compõem a cena.

Nos conteúdos que se integram na categoria “desporto”, a adaptação considera apenas dois grupos de objectos: os atletas e todos os restantes objectos. Este critério é feito considerando que o desporto combina de uma forma muito específica informação com entretenimento. Este aspecto confere aos atletas um estatuto tanto de “núcleo da notícia” como de estrelas. Assim sendo, entende-se que tudo aquilo que ocorre para além dos atletas e da actividade desportiva em si é um complemento ao evento desportivo em si.



**Figura 14 - Processo de adaptação do conteúdo do tipo cinema /documentário**

Traduzindo esta interpretação em termos de adaptação de conteúdo, a adaptação no primeiro, segundo e terceiro momentos centra-se exclusivamente no(s) objecto(s) que representa(m) o(s) atleta(s), apenas introduzindo adaptação dos restantes objectos já num quarto momento de adaptação. Os restantes objectos poderão ser de natureza diversa como infografismo ou publicidade.

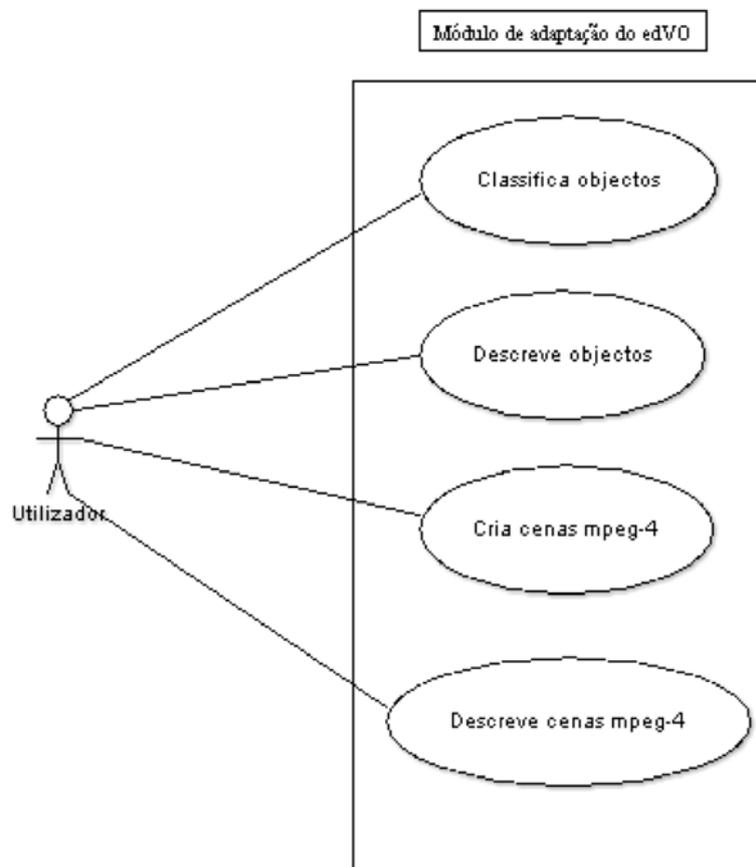


**Figura 15 - Processo de adaptação de conteúdo do tipo “música”**

Por último, a categoria música contempla de uma forma especial o aspecto sonoro do conteúdo, não descurando também a vertente da imagem muito característica da indústria do espectáculo. Por este motivo é proposta uma adaptação que ocorra num primeiro nível sobre todos os objectos de natureza visual (vídeo), num segundo e terceiro nível uma adaptação centrada nos objectos de natureza sonora (áudio) e num quarto nível, novamente uma adaptação centrada nos objectos vídeo da imagem.

Importa ainda referir que em todos estes casos foi descrito um processo de incremento de qualidade, no entanto o algoritmo torna também possível a situação inversa, conforme pode ser verificado pelas figuras 14 e 15. No caso de o processo em causa implicar uma adaptação no sentido da diminuição da qualidade, o algoritmo irá actuar pela ordem inversa, tentando garantir a qualidade de serviço através da diminuição de qualidade ou eliminação dos objectos pela ordem inversa de relevância.

A ferramenta descrita desta dissertação de mestrado acrescenta à *framework* do edVO a capacidade de os utilizadores criarem cenas MPEG-4, e associando descrições MPEG-7 a essas mesmas cenas e aos objectos que as compõem, criar conteúdos escaláveis.

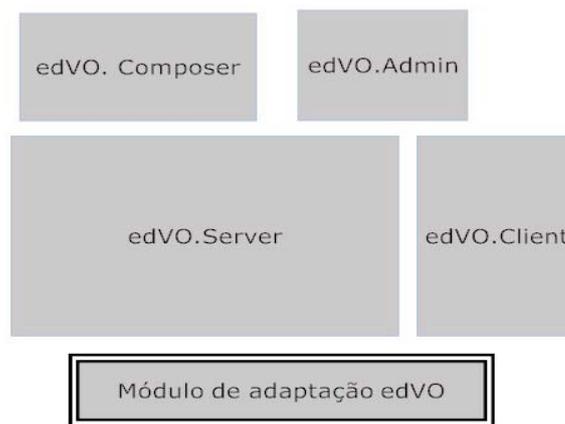


**Figura 16 - Diagrama de casos de uso**

As funcionalidades acrescentadas ao edVO pelo módulo de adaptação de conteúdos encontram-se explicitas na figura 16 sob a forma de um diagrama de casos de uso.

O módulo de adaptação de conteúdos que aqui é proposto para o edVO, actua quer ao nível do cliente quer ao nível do servidor, acrescentando uma camada adicional em ambos os módulos (e actuando também sobre o edVO.Composer) que permite que cliente e servidor troquem informações que irão contribuir para a tomada de decisão no processo de adaptação.

A figura 17 ilustra os diversos módulos do edVO , acrescentando-lhe o módulo de adaptação desenvolvido no âmbito desta dissertação.



**Figura 17 - Estrutura do edVO já contemplando o módulo de adaptação**

## 4.2 - Modelo conceptual

Os conceitos expostos anteriormente foram vertidos num modelo de teste, criado com o objectivo de validar as premissas e soluções propostas nesta dissertação. Este modelo foi construído em python de forma a facilitar a interacção com a *framework*, e tirando deste modo partido das ferramentas postas ao dispor para trabalhar sobre as aplicações que a compõem.

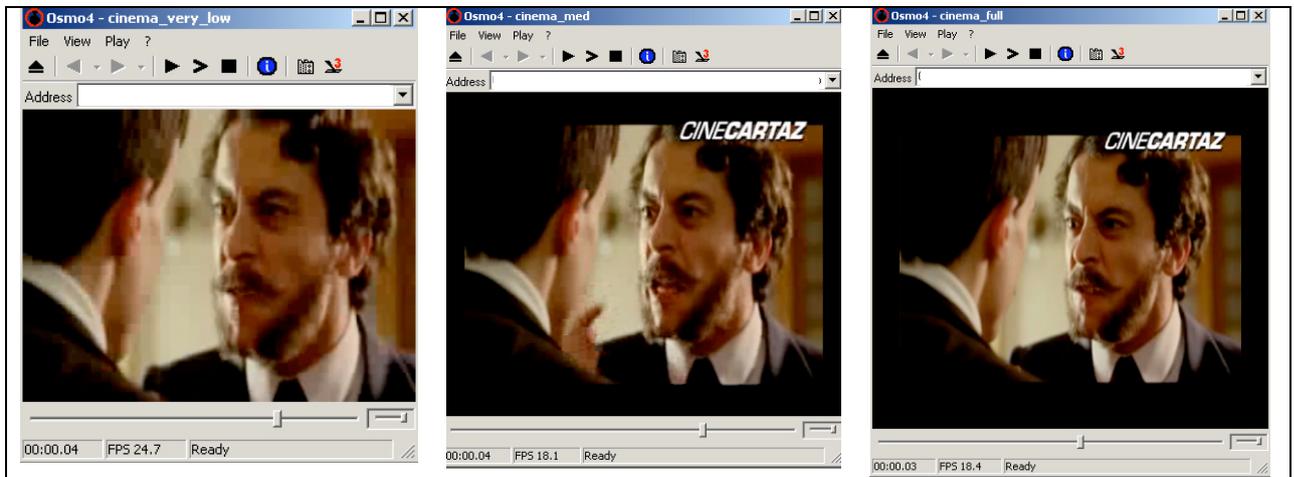
O modelo conceptual é composto por diferentes componentes que visam emular contextos de utilização nos quais possam ser aplicados os conceitos expostos nesta dissertação. O módulo de simulação e controlo recria as condições de utilização ao nível da rede (largura de banda e latência) e dispositivo (capacidade de armazenamento e resolução).

O módulo de *parsing* efectua o processamento das descrições das cenas mpeg-4 de forma a estar apto a fornecer ao módulo de adaptação uma descrição da cena que está a ser servida e quais as possíveis adaptações que poderão ser feitas sobre essa mesma cena.

O módulo de adaptação é o agregador das informações fornecidas pelos módulos de simulação e de *parsing*, aplicando o algoritmo de adaptação de conteúdos em função da avaliação que é feita a cada instante sobre as condições de utilização e capacidade de adaptação da *stream* exibida.

Para efeitos de adaptação foram não só considerados os factores relacionados com a rede através da qual é passada a *stream*, mas também as próprias capacidades do dispositivo. As duas variáveis que se relacionam com este último aspecto têm uma dupla função ao nível da adaptação do conteúdo. Considerar os factores inerentes ao dispositivo permite não só uma melhor adequação ao mesmo num primeiro momento de adaptação do conteúdo, mas também poderá permitir num momento futuro a transferência de uma *stream* entre dispositivos, dando a possibilidade de o utilizador

trocar a cada momento de equipamento terminal, tirando sempre total partido das capacidades do mesmo.



**Figura 18- A mesma cena em diferentes streams MPEG-4, com conteúdos e codificações diferentes**

Na figura 18 são apresentadas diferentes exemplos do mesmo conteúdo. Enquanto que na imagem mais à esquerda é apresentada a *stream* codificada a apenas 60 Kbps, na imagem do meio a mesma *stream* é codificada a 128 Kbps e foi-lhe adiciona informação gráfica. A diferença entre a imagem do meio e a imagem mais à direita é que esta última está codificada a uma *bitrate* de 512 Kbps.

Com este exemplo pretende-se ilustrar sobretudo a flexibilidade do MPEG-4 quando aplicado a contextos de utilização diversos. A mesma *stream* pode adaptar-se ao dispositivo não apenas em função do número de bits, mas também do próprio conteúdo de cada *stream*. A qualidade de uma *stream* deriva não só da qualidade perceptível da imagem, mas também de um adequado equilíbrio entre esta e a experiência que o dispositivo de visualização é capaz de proporcionar.

O modelo desenvolvido actua da seguinte forma para efectuar a adaptação de conteúdos:

O módulo de *parsing* utiliza as bibliotecas do GPAC (e em particular as da aplicação MP4BOX) para criar uma descrição da cena em XMT. Essa descrição é depois passada ao módulo de adaptação que obtém as variáveis de ambiente e as utiliza para formar uma descrição em XMT. A descrição gerada pelo módulo de adaptação é sempre a

melhor representação possível, considerando as variáveis ambientais e o algoritmo de decisão. O módulo adaptação devolve a nova descrição ao módulo de *parsing* que realiza o processo inverso para gerar a nova cena MP4.

Enquadrando o módulo de adaptação na arquitectura do edVO, este módulo irá actuar directamente sobre o edVO.Composer para efectuar a adaptação de conteúdos. Existe também no entanto comunicação com o edVO.Server para a publicação dos conteúdos adaptados e com o edVO.Client para recolher variáveis de ambiente.

#### **4.2.1 – Análise de desempenho**

Durante as experiências realizadas com o modelo conceptual foi possível aferir uma conclusão sobre os principais pontos fortes e fracos da solução proposta nesta dissertação. A capacidade de adaptação dinâmica dos conteúdos (isto é, a sua adaptação em função da categoria em que estes se inserem), permite criar adaptações muito mais fieis e em que os elementos essenciais de cada conteúdo são privilegiados relativamente aos restantes.

No entanto, o trabalho desenvolvido apenas permitiu chegar a um modelo de adaptação off-line, isto é, implicando sempre a geração de um novo ficheiro mpeg-4, ao invés de actuar sobre o ficheiro já existente.

Esta limitação encontra a sua raíz em dois aspectos do trabalho desenvolvido: por um lado, a necessidade de gerar uma nova descrição leva sempre à necessidade de gerar um novo ficheiro, enquanto que por outro lado, não foi possível encontrar até ao momento da escrita desta dissertação uma forma de garantir uma transição perfeita entre as duas versões da descrição. Em relação a este último aspectos, os factos que estão na base destas limitações estão para além da própria necessidade de gerar uma nova descrição, abrangendo também as dificuldades ao nível da implementação de um contador capaz de manter uma informação fiável sobre o ponto de visualização de cada uma das streams que compõem a cena no momento da adaptação.

## 5 – Conclusão

Existem muitos factores que intervêm no processo de adaptação de um conteúdo de forma a o tornar acessível a virtualmente qualquer dispositivo com capacidades de reprodução de uma *stream* de vídeo.

O primeiro desses factores, senão talvez o mais importante, é o entendimento que é feito sobre a importância que é dada à natureza do próprio conteúdo a ser adaptado. Como podemos constatar pelo que foi exposto nesta dissertação, existem diversas abordagens a esta problemática. Centrar as atenções na adaptação do conteúdo, independentemente da natureza deste é talvez a forma de adaptação mais comum, muito embora como também tivemos oportunidade de verificar, existam ferramentas que permitem uma solução intermédia, isto é, que não permitindo uma adaptação totalmente “à medida”, possibilitam que o mesmo conteúdo possa ter um número limitado de adaptações em função dos variados contextos de utilização.

O módulo de adaptação que é proposto nesta dissertação, sendo fruto de uma reflexão aprofundada sobre as problemáticas da adaptação de conteúdos, reflecte as preocupações resultantes dessa mesma reflexão. O modelo elaborado permite uma adaptação de conteúdos à medida de cada dispositivo, independentemente de as características desse dispositivo se encontrarem ou não previstas no servidor de adaptação.

A combinação de características objectivas, como o estado da rede e as características de um determinado dispositivo, com uma análise dos conteúdos (ainda que subjectiva) permite que a adaptação seja feita de uma forma mais ponderada e controlada. A flexibilidade desta adaptação é, talvez também o seu ponto mais fraco. Ao assentar grandemente na informação fornecida pelos descritores MPEG-7, obriga à correcta

utilização destes por parte do autor do conteúdo, sob pena de induzir o sistema num comportamento de adaptação não adequado ao conteúdo produzido.

Existem outras tecnologias que se propõem atingir fins semelhantes aos pretendidos com esta dissertação. O MPEG-2 assenta num modelo baseado em XML e que incorpora do design da norma características que permitem aos utilizadores efectuarem adaptação de conteúdos com o objectivo destes serem transversais ao nível das várias plataformas multimédia existentes. O MPEG-J ou MPEG4-Java é outra tecnologia vocacionada para a adaptação de conteúdos MPEG-4, que optamos por não utilizar nesta dissertação uma vez que durante o trabalho realizado não foram encontradas ferramentas que pudessem satisfazer todos os requisitos traçados inicialmente. Ambas estas tecnologias, por serem *standards*, podem servir para a implementação das ideias e objectivos que estiveram na génese desta dissertação.

Uma opção MPEG-J poderá fazer até mais sentido num cenário que não seja baseado em MPEG-21, se for tomado em consideração que tecnologias baseadas em java poderão chegar mais facilmente a um maior número de plataformas. Ainda assim, no trabalho de investigação realizado à partida para esta tese, as bibliotecas java (J2ME) capazes de extrair informações sobre os dispositivos de telecomunicações móveis revelaram-se insuficientes quando comparadas com as ferramentas já desenvolvidas para plataformas de maior capacidade.

Futuramente, importará por isso desenvolver sistemas que por um lado, assegurem que os conteúdos são correctamente classificados, e que por outro introduzam um factor de precisão ainda maior na descrição, e conseqüentemente na adaptação. Poderá também ser desenvolvido um trabalho na portabilidade do software de reprodução para um maior número dispositivos móveis, de forma a assegurar um maior alcance deste modelo de adaptação ao mesmo.



**Figura 19 – Diferentes streams de notícias.**

A prioridade contudo deverá incidir sobre um modelo de adaptação que permita efectuar esta adaptação sem que isso implique a geração de uma nova descrição MPEG-4, e conseqüentemente a geração de uma stream. Em casos como por exemplo o da figura 19, sucessivos processos de geração de novas descrições ao longo do mesmo serviço noticioso implicariam que o espectador visionasse as notícias sempre desde o início. Em alternativa, continuando no modelo actual, a exploração de um caminho que utilize um código de tempo (*timecode*) para sincronizar as diferentes versões da cena e minimizar o efeito de “salto” resultante da nova descrição poderá ser uma alternativa viável e que convém explorar.

## Lista de referencias

- [1] TEIXEIRA, Luís Filipe Pinto de Almeida – **Edition and Description Framework for Vídeo Objects**, Porto, FEUP, 2004
- [2] PEREIRA, Fernando e EBRAHIMI, Touradj – **The Mpeg 4 Book**, New Jersey, IMSC Press Multimédia, 2002
- [3] KOENEN, Rob – MPEG-4 Overview [Em linha], Março 2002 [consultado em 25 Agosto 2005]. Disponível em: [www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm](http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm)
- [4] MEDEIROS, Moisés Emanuel Adrega – **Descrição e Manipulação de Objectos para Composição de Multimédia**, Porto, FEUP, 2005
- [5] MARTÍNEZ, José M., MPEG-7 Overview (Version 10) [em linha], (2004) [consultado a 12 Julho 2006]. Disponível em : [www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm](http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm)
- [6] Sitio web do GPAC, [em linha], [consultado em 27 Agosto 2005], Disponível em: <http://gpac.sourceforge.net>
- [7] Sitio web do Zope,[em linha], [consultado em 27 Agosto 2005], Disponível em: [www.zope.org](http://www.zope.org)
- [8] Sitio web do Darwin Streaming Server,[em linha] [consultado em 27 Agosto 2005], Disponível em: [http:// developer.apple.com/darwin/projects/streaming](http://developer.apple.com/darwin/projects/streaming)
- [9] Sitio web da IBM- projecto enterprise MPEG-4,[em linha],[consultado a 15 de Julho de 200]. Disponível em : <http://researchweb.watson.ibm.com/mpeg4/Projects/emb.htm>
- [10] France Telecom, Dossier du mois de mai 2006 [em linha],(2006), 6 páginas, Maio de 2006 [Consultado a 2 se Setembro de 2006]. Disponível em : [www.francetelecom.com/en/group/rd/news/thematique/dossier\\_mois/ddm200605/att00038637/ddm200605va.pdf](http://www.francetelecom.com/en/group/rd/news/thematique/dossier_mois/ddm200605/att00038637/ddm200605va.pdf)
- [11] Queen Mary University of London, Department of Electronic Engineering, Scalable Vídeo Coding, [em linha], [consultado a 17 de Agosto 2006]. Disponível em: [www.elec.qmul.ac.uk/mmv/svc.html](http://www.elec.qmul.ac.uk/mmv/svc.html)

- [12] MRAK, Marta et alli, Scalable Vídeo Coding in Network Applications, [em linha] (2002), 7 páginas, Junho de 2002 [consultado a 16 de Agosto de 2006]. Disponível em: [ieeexplore.ieee.org/iel5/7973/22057/01026656.pdf](http://ieeexplore.ieee.org/iel5/7973/22057/01026656.pdf)
- [13] MAGALHÃES, João Miguel da Costa, Universal Access to Multimedia Content Based on the MPEG-7 Standard, Lisboa, Instituto Superior Técnico, 2002
- [14] JONATAN, Tierno, Celeste Campo, "Smart Camera Phones: Limits and Applications," *IEEE Pervasive Computing*, vol. 04, no. 2, pp. 84-87, April-June, 2005.
- [15] Sitio web do OSMO4, [em linha], [consultado em 10 Setembro 200], Disponível em: <http://www.comelec.enst.fr/osmose/>

## Bibliografia

- [1] TEIXEIRA, Luís Filipe Pinto de Almeida – **Edition and Description Framework for Vídeo Objects**, Porto, FEUP, 2004
- [2] PEREIRA, Fernando e EBRAHIMI, Touradj – **The Mpeg 4 Book**, New Jersey, IMSC Press Multimédia, 2002
- [3] KOENEN, Rob – MPEG-4 Overview [Em linha], (Março 2002) [consultado em 25 Agosto 2005]. Disponível em: [www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm](http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm)
- [4] MEDEIROS, Moisés Emanuel Adrega – **Descrição e Manipulação de Objectos para Composição de Multimédia**, Porto, FEUP, 2005
- [5] Sitio web do GPAC,[em linha] [consultado em 27 Agosto 2005]), Disponível em: <http://gpac.sourceforge.net>
- [6] Sitio web do Zope, [em linha],[consultado em 27 Agosto 2005], Disponível em: [www.zope.org](http://www.zope.org)
- [7] Sitio web do Darwin Streaming Server,[em linha] [consultado em 27 Agosto 2005], Disponível em: [http:// developer.apple.com/darwin/projects/streaming](http://developer.apple.com/darwin/projects/streaming)
- [8] Sitio web da IBM- projecto enterprise MPEG-4,[em linha], [consultado a 15 de Julho de 2006]. Disponível em : <http://researchweb.watson.ibm.com/mpeg4/Projects/emb.htm>
- [9] France Telecom, Dossier du mois de mai 2006 [em linha],(2006), 6 páginas, Maio de 2006 [Consultado a 2 se Setembro de 2006]. Disponível em : [www.francetelecom.com/en/group/rd/news/thematique/dossier\\_mois/ddm200605/att00038637/ddm200605va.pdf](http://www.francetelecom.com/en/group/rd/news/thematique/dossier_mois/ddm200605/att00038637/ddm200605va.pdf)
- [10] Queen Mary University of London, Department of Electronic Engeneering, Scalable Vídeo Coding, [em linha] [consultado a 17 de Agosto 2006]. Disponível em: [www.elec.qmul.ac.uk/mmv/svc.html](http://www.elec.qmul.ac.uk/mmv/svc.html)
- [11] MRAK, Marta et alli, Scalable Vídeo Coding in Network Applications, [em linha] (2002), 7 páginas, Junho de 2002 [consultado a 16 de Agosto de 2006]. Disponível em: [ieeexplore.ieee.org/iel5/7973/22057/01026656.pdf](http://ieeexplore.ieee.org/iel5/7973/22057/01026656.pdf)

- [12] Sitio web da IBM- projecto Autonomic MPEG-4,[em linha] [consultado a 15 de Julho de 2006]. Disponível em : <http://researchweb.watson.ibm.com/mpeg4/Projects/embAutonomic.htm>
- [13] Sitio web do OSMO4, [em linha] [consultado em 10 Setembro 2005], Disponível em: <http://www.comelec.enst.fr/osmose/>
- [14] Apple Computer, MPEG-4 The new standard for multimédia on the internet, powered by Quicktime, [em linha] (2003), 6 páginas [Consultado em Agosto de 2005]. Disponível em: [http://images.apple.com/au/quicktime/MPEG4\\_V3.pdf](http://images.apple.com/au/quicktime/MPEG4_V3.pdf)
- [15] MAGALHÃES, João Miguel da Costa, Universal Access to Multimedia Content Based on the MPEG-7 Standard, Lisboa, Instituto Superior Técnico, 2002
- [16] JONATAN, Tierno, Celeste Campo, "Smart Camera Phones: Limits and Applications," *IEEE Pervasive Computing* ,vol. 04, no. 2, pp. 84-87, April-June, 2005.