

Linguagens XML de Descrição de Cenas com Conteúdos Multimédia

Pedro Pinto¹, Isidro Vila Verde²

¹ Departamento de Ciências Básicas de Computação
Escola Superior de Tecnologia e Gestão,
Instituto Politécnico de Viana do Castelo
Apartado 574, s/n 4901-908 Viana do Castelo
pedropinto@estg.ipvc.pt

² Departamento de Engenharia Electrotécnica e de Computadores
Faculdade de Engenharia da Universidade do Porto,
Rua Dr. Roberto Frias, s/n 4200-465 Porto
jvv@fe.up.pt

Abstract. Neste artigo fazemos uma análise comparativa de formatos de linguagens de descrição de cenas dinâmicas baseadas em XML. Apresenta-se uma descrição resumida de vários formatos, focando as suas principais características. A necessidade de análise comparativa surge num contexto de desenvolvimento de um ambiente de e-monitorização de provas electrónicas de avaliação. Nesta perspectiva, são definidos critérios de avaliação adequados a este contexto e com base nestes critérios é realizado um estudo comparativo, identificado o formato mais adequado para o fim especificado.

1 Introdução

No contexto aplicacional existe já um número significativo de produtos multimédia, quer 2D ou 3D. Por forma a realizar seus objectivos, estes sistemas possuem várias componentes que trabalham para a produção, manipulação e apresentação dos objectos nas cenas multimédia. Na composição destas cenas são utilizadas linguagens específicas para o efeito chamadas linguagens de descrição de cena, que representam uma entidade essencial no processo e tem a intenção de descrever de uma forma eficaz e adequada uma cena multimédia, descrevendo todos os objectos que a compõem e definindo todos os seus estados nesse ambiente. Neste momento podemos encontrar várias linguagens possuindo características e objectivos muito próprios.

É objectivo deste artigo fazer uma análise comparativa de algumas dessas linguagens numa perspectiva particular do desenvolvimento de uma aplicação de monitorização de provas electrónicas. Na secção seguinte é apresentado o contexto no qual surgiu esta necessidade. Na secção 3 são descritas de forma resumida as linguagens analisadas. Na secção 4 formulam-se os critérios de análise orientados para o contexto da aplicação. Na secção 5 apresentam-se os resultados da análise considerando os critérios definidos e as características particulares das linguagens.

Por último, são tiradas conclusões desta análise e é feito o enquadramento da opção escolhida na aplicação em desenvolvimento.

2 Contexto

A necessidade desta análise surge no contexto do desenvolvimento de uma aplicação específica, desenvolvida no âmbito de uma tese de Mestrado. Esta tese tem por base de trabalho, o desenvolvimento de um modelo de monitorização remota, vocacionado especialmente para ambientes de formação. A arquitectura de monitorização para este modelo está apresentada na Fig.1.

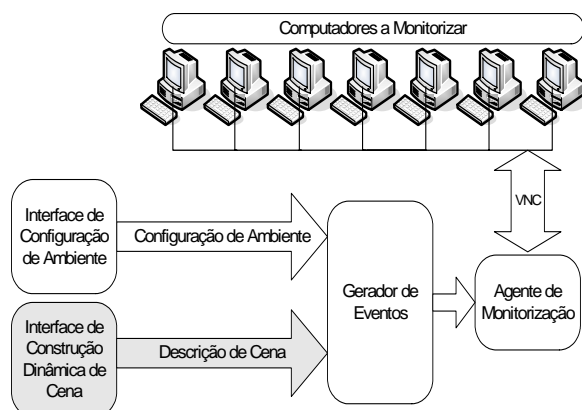


Fig. 1. Diagrama de blocos da arquitectura de monitorização de provas electrónicas de exame

Nesta arquitectura, a responsabilidade de monitorização pertence ao “Agente de Monitorização”. A monitorização consiste em duas vertentes, a monitorização dos ecrãs remotos e a monitorização dos processos.

A monitorização dos ecrãs é realizada recorrendo ao protocolo Remote Frame Buffer (RFB)[1] disponibilizado pelo software Real Virtual Network Computing (RealVNC)[2] que trata de estabelecer fluxos deste protocolo para cada computador a monitorizar, fluxos estes que transportam todos os pixels do ambiente de trabalho visto por cada utilizador desses mesmos computadores. Na Fig.2 é apresentado um exemplo de um cena do processo de monitorização.

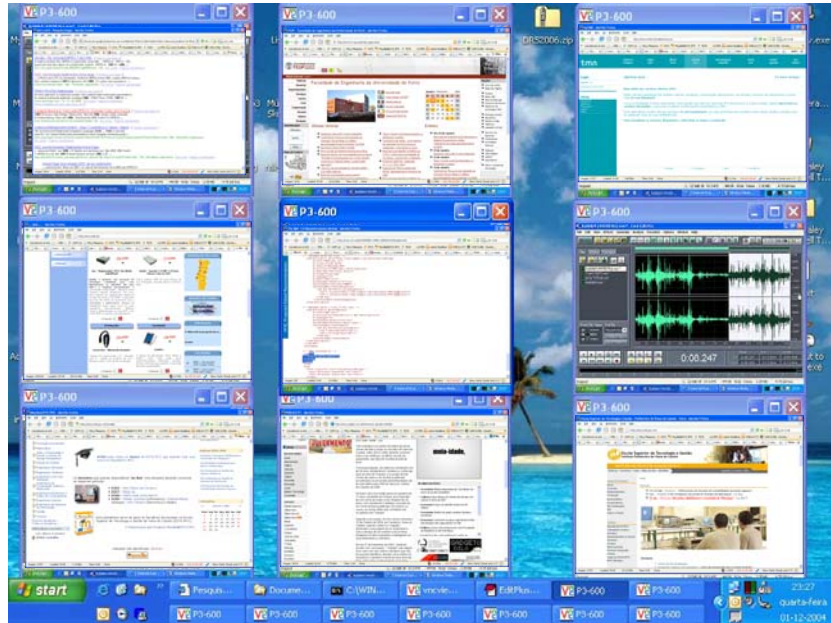


Fig. 2. Exemplo dos objectos presentes na cena multimédia da aplicação a desenvolver

A monitorização dos processos recorre à interface Windows Management Interface (WMI)[3] para configurar a geração de alertas a enviar ao “Agente de Monitorização”. Estes alertas são disparados quando são realizadas acções relevantes, por exemplo, quando são iniciadas aplicações ou quando existem acessos não autorizados a determinadas páginas Internet.

Desta forma, implementa-se uma monitorização activa e passiva das provas electrónicas que estão a ser realizadas pelos formandos num determinado ambiente.

Em ambientes de larga escala esta monitorização não pode ser permanente e terá de haver programas de monitorização que automatizem as “rondas” periódicas pelos diversos postos.

O programa de monitorização é gerado automaticamente na entidade “Gerador de Eventos” à custa de um ficheiro de descrição de ambiente e um fluxo de descrição dinâmica da cena. O ficheiro de descrição de ambiente é criado previamente à realização do exame e descreve o ambiente da realização da prova. Esta descrição é feita num formato XML simples, validado segundo um XSD[4] definido, de acordo com a seguinte semântica:

- ambiente a monitorizar: “Sala”
 - informação de identificação do aluno: “Pedro Pinto”
 - número identificador do computador (IP): “192.168.0.20”
 - posição relativa do computador no espaço: “15”

Quanto ao fluxo de descrição de cena este apresenta-se como um fluxo constante que traduz em cada instante os objectos que devem ser monitorizados e como eles se apresentam na cena de monitorização. Esta descrição apresenta-se como parte fundamental deste sistema e deve estar definida numa linguagem de descrição de cenas com suporte para cenas dinâmicas. Isto é, cenas onde os objectos visualizados mudam em função da decisão humana, em cada instante.

Esta linguagem de descrição de cena terá, por um lado, de traduzir do modo mais fiel construções de cenas possíveis realizadas na “Interface de Construção de Cena” e por outro terá de ser uma linguagem adequada e consistente para a leitura no “Gerador de Eventos”. Existem duas informações principais enviadas sob a linguagem de descrição de cena. A primeira informação define o espaço de monitorização, com a definição da divisão do ecrã e a posição de cada janela de monitorização de acordo com a seguinte semântica:

- Espaço de divisão do monitor: "4 Colunas 3 Linhas"
- Janela número "1" com posição na "Linha 1 Coluna 3"
- Janela número "2" com posição na "Linha 3 Coluna 2"

A segunda informação transporta dados sobre a dinâmica espacial e temporal da cena, apresentados com a seguinte semântica:

- imagem do computador da posição "1" visualizado na janela "2" com início aos "4" segundos e duração de "2" segundos
- imagem do computador da posição "15" visualizado na janela "4" com início aos "2" segundos e duração de "4" segundos

O conjunto destes elementos constitui um fluxo contínuo que traduz em cada instante a posição de todos os elementos na cena de monitorização. Por conseguinte, existe a necessidade de identificar a melhor escolha no formato desta linguagem num conjunto de linguagens possíveis de linguagens de descrição de cena e, por isso, é realizado neste documento um estudo comparativo neste contexto.

3 Linguagens

Das várias linguagens que podem ser encontradas para a componente essencial da aplicação referida, existe um subconjunto específico que é relevante para este estudo: as linguagens de descrição de cena baseadas em XML. O fundamento para o foco do estudo nas linguagens de descrição de cena baseadas em XML é induzido pelas características que herdaram desta linguagem base, nomeadamente residente no facto do XML:

- ser uma linguagem de norma não proprietária
- ser independente dos vários tipos de plataformas e sistemas, permitindo um alto grau de interoperabilidade;
- oferecer capacidade descritiva adequada para ferramentas de geração de descrições automáticas ou para uma edição simples em modo texto com linguagem perceptível para o ser humano;

- permitir uma descrição estruturada de meta-informação para selecção, avaliação, e integração de objectos;
- ser apropriada para hierarquias de elementos que representam vários tipos de objectos e conceitos adequados a sistemas gráficos e cenas;

Nesta esfera podem ser encontradas várias linguagens de cena, mas apenas vão ser referenciadas neste estudo aquelas que, por um lado, possuem um uso mais corrente e são aplicadas num número considerável de aplicações com um grau de desenvolvimento e maturidade bem definido e, por outro lado, que se enquadram sob os aspectos gerais nesta aplicação de monitorização em desenvolvimento. Cada uma das linguagens analisadas incorpora um conjunto de mecanismos próprios, que resultam na obtenção de diferentes resultados, no âmbito da aplicação específica descrita na Fig. 1. Nesta fase torna-se imperativo descrever de uma forma minimalista algumas linguagens de descrição de cena que, pelo estudo dos pré-requisitos efectuado, se apresentam como as soluções mais adequadas ao sistema, realçando as suas características individuais relevantes ao contexto. Segundo este propósito, seleccionaram-se as seguintes linguagens:

- XMT-A[5]
- XMT-O[6]
- SMIL[7]

Segue-se uma descrição breve destas linguagens e suas características relevantes para o estudo. Nos tópicos seguintes a designação SMIL depreende a referência à versão 2.0, ou seja, SMIL 2.0.

3.1 Linguagens desenvolvidas no contexto da norma MPEG

A norma Moving Picture Experts Group (MPEG)[8] define uma estrutura onde é possível transmitir vários tipos de objectos numa plataforma comum. Esta norma divide-se em duas partes fundamentais: a codificação dos objectos intervenientes e a possibilidade de definição da apresentação dos objectos pertencentes à cena multimédia. A apresentação, dos objectos que compõem a cena audiovisual, é especificada através de formatos próprios para descrição das cenas. O fluxo que passa este tipo de interacções e que contém as informações das relações entre os objectos multimédia, é chamado fluxo descritor de cenas e é especificado através da linguagem BInary Format for Scene (BIFS)[9]. O BIFS foi inspirado na linguagem VRML[10], e possui características próprias, adequadas para o sistema de transporte MPEG-4. A linguagem BIFS permite a eficiente representação de apresentações dinâmicas e interactivas, incluindo gráficos 2D & 3D, imagens, texto e material audiovisual. Esta representação inclui a descrição de organização no espaço e no tempo dos diferentes elementos de cena, animações e interacções com o utilizador. Ao contrário do VRML, onde os seus objectos e eventos são declarados em modo texto, em BIFS esta estrutura é especificada em código binário. A adopção do formato binário é fundamental para a compressão e posterior distribuição das cenas no MPEG-4, no entanto esse formato dificulta o processo de criação e alteração do conteúdo, embora este facto possa ser contornado com a utilização de ferramentas de edição gráfica, onde o formato binário já se torna transparente para os editores destes conteúdos.

Com intuito de ultrapassar a limitação do formato binário na criação e alteração de conteúdos na linguagem BIFS, o formato eXtensible MPEG-4 Textual (XMT)[11] foi proposto como uma estrutura para especificar descrições de cenas MPEG-4 através de uma sintaxe textual. O XMT é uma extensão textual usada na norma MPEG-4 e tenta combinar vantagens da representação binária BIFS e da representação de alto nível de abstracção de cenas multimédia, SMIL. Sendo assim, o XMT apresenta características novas no contexto MPEG-4 como:

- um intercâmbio transparente entre editores e ferramentas de produção
- interoperabilidade com formatos. O formato XMT pode ser convertido em SMIL, VRML ou em BIFS (para os leitores MPEG-4)
- possibilidade de edição manual do conteúdo

Esta estrutura é instanciada por duas linguagens, com diferentes sintaxes e semânticas XMT-A e XMT-O apresentadas de seguida.

Linguagem XMT-A. A linguagem XMT-A permite a descrição de baixo nível de conteúdos multimédia baseada em Extensible 3D (X3D)[12] e é uma versão em formato texto da linguagem BIFS. A sua estrutura baseia-se em construções XML com representação directa das expressões existentes no formato binário do MPEG-4, e por representar construções existente no BIFS, o XMT-A também se aproxima do tipo de estruturas definidas em VRML. Também no aspecto semântico, esta linguagem apresenta uma equivalência estrita com o BIFS, não acrescentando nenhuma outra funcionalidade adicional. Assim sendo, apresenta o mesmo conjunto de características centrais já oferecidas pelo formato binário, que se apresentam enumeradas desta forma:

- Permite a integração e controlo de conteúdos áudio/vídeo diferentes na mesma cena.
- Tem um conjunto rico de construtores gráficos 2D/3D.
- Possui um conjunto de elementos que permitem a interactividade local e remota.
- Permite animações locais e remotas como por exemplo a alteração de posições de objectos, de cores, e formas.
- Permite a reutilização de conteúdos através de referências para fluxos.

De seguida é apresentado um exemplo de uma descrição de cena no formato XMT-A.

```
<Shape>
  <geometry>
    <Rectangle size="40 30"/>
  </geometry>
  <appearance>
    <Appearance>
      <texture>
        <MovieTexture url="http://site.pt/obj.mpg" loop="false"
          startTime="0.0" stopTime="-1.0"/>
      </texture>
    </Appearance>
  </appearance>
</Shape>
```

Como se pode observar, existe um alto grau de verbosidade para o tipo de exemplo simples que se quer descrever.

Linguagem XMT-O. Com o intuito de diminuir a complexidade existente na linguagem XMT-A na especificação de cenas MPEG-4, foi desenvolvida a linguagem XMT-O. Nesta linguagem, os objectos audiovisuais e as suas transformações são descritas num alto nível de abstracção, facilitando o processo de edição, ao contrário do que ocorre em XMT-A. A linguagem XMT-O tende a uma aproximação forte à linguagem SMIL e por isso apresenta algumas das suas características próprias, conseguindo assim uma interoperabilidade entre estes dois sistemas. Isto é conseguido pelo XMT-O recorrendo a alguns módulos presentes na linguagem SMIL. A estrutura do documento XMT-O também é dividida em duas partes, um cabeçalho e um corpo. No cabeçalho encontram-se as informações gerais do documento, como a forma da apresentação, a meta-informação para definição semântica do documento e as referências para conjuntos de elementos. De seguida é apresentado um exemplo de um cabeçalho no formato XMT-O.

```
<head>
  <layout metrics="pixel" type="xmt/xmt-basic-layout">
    <topLayout width="300" height="300" backgroundColor="branco">
      <region id="regiao_video" size="91 27"/>
    </topLayout>
  </layout>
</head>
```

Em XMT-O, de forma análoga a SMIL, as composições possuem semântica de sincronização e é no corpo do documento XMT-O que se descreve a semântica temporal e sequencial da cena, onde várias composições podem ser definidas, através dos elementos *par*, *seq* e *excl* (tipo de semântica temporal implementada - paralela, sequencial ou exclusiva). Ainda no corpo do documento, as características para apresentação dos objectos multimédia são especificadas nos próprios elementos, através dos seus atributos. Além das características para apresentação, as relações entre esses objectos também podem estar incluídas nos seus atributos. Dessa forma, além de composições com semântica de sincronização, estas relações podem ser estabelecidas através de eventos, especificados nesses atributos. De seguida é apresentado um corpo associado ao cabeçalho acima descrito. Neste exemplo, estes dois elementos formam um documento que descreve uma cena em formato XMT-O.

```
<body><par>
  <video src="video.mp4" region="regiao_video" begin="0s"
    dur="indefinite"/>
  <audio src="audio.mp4" begin="0s" dur="indefinite"/>
</par></body>
```

3.2 SMIL

A linguagem SMIL apresenta-se como uma boa candidata neste estudo por ser uma linguagem normalizada pelo W3C[13] e por ser desenvolvida desde início como uma linguagem para apresentações multimédia. De modo geral, o SMIL permite a gestão de apresentações de audiovisuais interactivas que integram fluxos de áudio, imagens e texto, ou qualquer outro tipo de objecto multimédia. Para isso, possui funções para a definição de sequências, duração, posição e visibilidade dos vários elementos

(objectos) constituintes da cena multimédia, oferecendo a possibilidade de descrever o comportamento temporal da apresentação e descrever a localização espacial dos objectos no ecrã.

A segunda versão desta linguagem (SMIL 2.0) veio acrescentar dados novos importantes neste contexto, nomeadamente a introdução de conceitos como modularização e perfis, que permitem uma expansão da linguagem de um modo relevante. Assim, existem módulos que podem ser utilizados com elementos, atributos e valores de atributos adequados à descrição de uma aplicação particular. Existem vários módulos disponíveis neste contexto, como por exemplo, o módulo de ligação, o módulo de controlo de conteúdo e o módulo de apresentação. Além disso, é possível nesta linguagem seleccionar um perfil que descreve o conjunto de módulos que estão a ser utilizados num determinado documento, com a função de compatibilizar esta descrição com a versão anterior (SMIL 1.0) e assegurar a compatibilidade com a semântica e conteúdos SMIL entre os vários módulos.

No código seguinte, é apresentado um exemplo de descrição de cena em SMIL, com a possibilidade de interacção dos conteúdos .

```
<smil xmlns="http://www.w3.org/2005/SMIL20/Language">
  <head>
    <layout>
      <region id="geral" width="30" height="50"/>
      <region id="especifica" width="300" height="500"/>
    </layout>
  </head>
  <body><par>
    <a href="videol.avi" target="especifica" accesskey="a">
      <video region="geral" src="videol.avi" begin="0s"
        dur="indefinite"/></a>
    <a href="video5.avi" target="especifica" accesskey="a">
      <video region="geral" src="video5.avi" begin="0s"
        dur="indefinite"/></a>
  </par></body></smil>
```

De seguida são apresentados os critérios de análise.

4 Critérios de Análise

Antes da análise comparativa destas tecnologias, torna-se necessário distinguir o conjunto de critérios que se revelaram preponderantes no desenvolvimento da aplicação específica de monitorização. A recolha destes critérios de comparação e análise é de grande importância para a decisão final na escolha da linguagem, uma vez que já se mostrou que a linguagem de descrição de cena é uma componente essencial na aplicação em foco. Para a aplicação de visualização e monitorização em desenvolvimento, são enumerados os seguintes critérios:

1. Conjunto de elementos adequados: Dando lugar à monitorização existirão elementos a monitorizar num determinado ambiente. Na linguagem de descrição terá de existir um conjunto de elementos, atributos e valores associados adequados à descrição e manuseamento dos objectos na cena. Se estes dados não estiverem definidos, terão de ser definidos fora da norma das linguagens.

2. Semântica própria para representar alterações temporais: É necessário que a linguagem possua formas de descrição de alterações temporais na monitorização, assim sendo, é valorizado o modo como estas alterações são representadas para, da melhor forma, se traduzir todas as sequências temporais de monitorização possíveis neste tipo de aplicações. Seria conveniente que esta sequência fosse lógica e simples para melhor compreensão de autores e editores da aplicação referida.
3. Semântica adequada para representar alterações espaciais: O modo como são expressas as transformações na cena também será um critério relevante neste contexto. Torna-se importante a existência de formas básicas para a transformação espacial dos elementos, nomeadamente na sua posição e tamanho, e uma semântica própria também na representação fácil e intuitiva destas alterações.
4. Interpretação na edição de conteúdos: É importante para o autor/editor dos conteúdos ter uma noção da representação da cena apenas por inspecção do ficheiro que a compõem. À partida as descrições terão de ser exprimidas em modo texto para possibilitarem a fácil edição e deverão ter um grau de complexidade baixo para, dessa forma, serem intuitivas e facilmente entendidas na interacção humana.
5. Baixo volume de informação: Para existir facilidade no envio de informação e baixo tempo de atraso entre alterações à monitorização, um dos critérios que podem ser definidos passa pela baixa complexidade e tamanho do ficheiro que contém a linguagem da descrição de cena. Na aplicação em causa exige-se uma transferência semi-contínua de conteúdos que pode ter frequências bastante altas, o que coloca este factor como relevante neste ponto.
6. Escalabilidade e Modularização: Como último critério e não menos importante, convém referir que é intenção no desenvolvimento desta aplicação, utilizar linguagens que possam ser aplicadas no contexto referido e que possam ser extensíveis, permitindo algumas modificações, adaptações ou a utilização de incrementos modulares adequados.

Definidos os critérios que serão avaliados, podemos definir também uma escala de valores que quantifique, de forma objectiva, as linguagens em análise. Sendo assim teremos valores de 1 a 3 que devem corresponder à seguinte interpretação:

1. Linguagem que não satisfaz o critério estabelecido.
2. Linguagem que satisfaz o critério estabelecido com necessidade de algumas alterações adicionais.
3. Linguagem que satisfaz o critério estabelecido sem necessidade de alterações adicionais.

Seguidamente mostra-se os resultados da análise efectuada.

5 Resultado da Análise

Ao realizar este estudo, em primeira instância pôde-se verificar que o XML está a influenciar de várias formas o formato MPEG, formato importante na esfera das aplicações multimédia. Neste momento, tanto o formato BIFS (norma MPEG-4) como o formato BiM[14] (norma MPEG7) estão a ser complementados com esquemas de representação XML levando o MPEG a utilizar conceitos de algumas especificações como SMIL e X3D.

Numa análise mais pormenorizada, pode-se referir que a linguagem de descrição de cena BIFS numa fase anterior à conversão e compressão apresenta um conjunto vasto de elementos de desenho facial e de figuras geométricas, que não são necessários numa aplicação de monitorização no contexto apresentado. Esta linguagem possui também um conjunto variado de funções construtoras de gráficos 2D/3D, que são, no âmbito da análise, igualmente dispensáveis. Mas convém referir nesta altura que o principal factor que leva a rejeitar esta opção como válida no desenvolvimento da aplicação referida, está relacionada com o formato nativo da linguagem. Uma vez que a linguagem BIFS é convertida em formato binário, podemos verificar que não permite a edição ou reedição directa dos conteúdos.

O MPEG redireccionou o formato desta linguagem para outro formato próprio de XML - o XMT que, na norma XMT-A apresenta a mesma forma da versão BIFS mas em formato texto, já permitindo ao autor/editor uma criação e alteração de conteúdos mais directa. Apesar desta vantagem note-se que continua a existir um conjunto de elementos não adequados, herdados das construções geométricas do BIFS. Na realidade, por representar directamente todos os elementos presentes na arquitectura MPEG-4, o XMT-A continua a apresentar-se como uma linguagem algo complexa e extensa para a edição fácil na criação e alteração de conteúdos.

Sendo uma linguagem de menor complexidade, o XMT-O apresenta-se como uma boa solução em relação ao XMT-A. Esta norma é baseada em SMIL e por isso apresenta um conjunto mais adequado de elementos. De qualquer forma, este formato apresenta alguma compatibilidade com o formato de representação da cena em árvore como o BIFS e o XMT-A e, com isto, algumas das características do SMIL são implementadas de forma parcial, como por exemplo na sincronização, nas transições e controlo de conteúdos ou até não implementadas como no caso do tratamento de eventos de entrada (por ex. teclado).

O SMIL apresenta-se como uma linguagem para descrição de cenas de princípio simples e bastante eficiente. É de referir que esta é a linguagem que apresenta os elementos mais adequados ao sistema em desenvolvimento, principalmente porque implementa uma modularização abrangente e, baseado neste conceito, podem ser utilizados alguns módulos orientados para este tipo específico de aplicação em desenvolvimento. Para este efeito os seguintes módulos demonstraram boa adequabilidade:

- Módulo de Ligação
- Módulo de Controlo de Conteúdo
- Módulo de Apresentação

Após o estabelecimento dos critérios de análise e após esta análise particular da adequabilidade destas linguagens, irá ser definido o grau de adequabilidade e a medida da integração destas linguagens nestes sistemas. Ao providenciar esse conjunto possível de critérios e ao cruzar esta informação com as características inerentes às linguagens em análise, é apontada uma solução que será a mais adequada no contexto explicado. Na Tabela 1 é apresentado o conjunto de linguagens em análise e a sua correspondente adequabilidade aos critérios formulados utilizando a escala de valores estabelecida na secção 4.

Tabela 1. Tabela com análise comparativa da adequabilidade das linguagens submetidas

Linguagens	Critério 1	Critério 2	Critério 3	Critério 4	Critério 5	Critério 6
XMT-A	1	2	3	1	3	1
XMT-O	2	3	3	3	3	3
SMIL 2.0	3	3	3	3	3	3

Desta tabela sobressai que as linguagens XMT-O e o SMIL se apresentam como boas candidatas à aplicação em desenvolvimento. Note-se a diferença existente entre o formato XMT-A e XMT-O baseada em primeira instância no critério de adequação de elementos (critério 1), onde se verifica que o XMT-O, oferece um conjunto de elementos mais adequados do que o XMT-A. Em segunda instância também se verificam diferenças notórias na facilidade de interpretação na edição de conteúdos e consequente complexidade (critério 4), onde o XMT-A revelou ser mais complexo do que o formato XMT-O. Quanto ao critério 6 podemos verificar que existe uma discrepância entre o XMT-A e os outros dois formatos. Isto deve-se à estrita ligação do XMT-A com o formato binário, não permitindo qualquer modularização ou capacidade de adaptação por parte deste formato à aplicação em desenvolvimento. Sendo assim, o XMT-O e o SMIL são apresentados como formatos válidos, embora com algumas pequenas diferenças entre eles. Ambos diferem no conjunto de elementos que disponibilizam (critério 1), sendo o SMIL o formato mais adequado, uma vez que, para além dos elementos principais (objectos multimédia) e seus atributos, este permite um conjunto vasto de possibilidades de interacção com o conteúdo e transições adequadas ao contexto em causa. Sob outra perspectiva, no critério 6 também podemos observar algumas diferenças. Neste critério, o SMIL apresenta-se com um conjunto vasto de módulos que permitem adaptações a vários tipos de sistemas. Alguns destes revelaram ser bastante úteis para a aplicação, nomeadamente os módulos de ligação e apresentação. Apesar de alguns destes conceitos também estarem presentes no XMT-O, o SMIL apresenta uma maior adequabilidade à aplicação a desenvolver. Em suma, o SMIL possui vantagens adicionais em relação ao XMT-O, que passam pela existência de elementos primários adequados (com os seus atributos respectivos) e pela capacidade de modularização disponível, oferecendo um conjunto adicional de elementos que permitem descrever melhor a cena da aplicação em causa e permitindo uma futura escalabilidade do sistema.

6 Conclusões

A decisão na escolha da linguagem da descrição de cena é determinante para o projecto que está a ser desenvolvido e teve como base as conclusões do estudo evidenciado neste documento. Tal como é indicado na secção do resultado da análise, a solução que se demonstrou mais adequada aos critérios exigidos foi a linguagem SMIL. Tendo por base este resultado, no âmbito referido, já foram formulados e testados alguns exemplos de possíveis descrições de cena que poderão existir. O seguinte código apresenta um exemplo simples de uma descrição de cena, onde os objectos centrais da cena são dois objectos vídeo que representam o fluxo das imagens do ambiente de trabalho de cada computador a monitorizar.

```
<smil>
  <head>
    <layout>
      <root-layout width="320" height="240"/>
      <region id="janela1" left="170" top="110"/>
      <region id="janela2" left="50" top="50"/>
    </layout>
  </head>
  <body>
    <par>
      <video src="comp1.avi" begin="5s" region="janela1"/>
      <video src="comp5.avi" begin="2s" region="janela2"/>
    </par>
  </body>
</smil>
```

Neste formato linguagem SMIL, depois de definidos os objectos e seus atributos, poderemos ter ainda interacções entre eles e transformações espaciais e temporais adequadas. Estas transformações, poderão ser definidas no futuro para traduzir níveis de importância de monitorização, de acordo com critérios definidos antes do início da monitorização, ou ao longo do processo de monitorização, de acordo com valores de variáveis que se pretendam monitorar.

Referências

1. "Remote Frame Buffer Protocol", acedida pela última vez em 17 de Novembro de 2005, <http://www.realvnc.com/docs/rfbproto.pdf>
2. "Software RealVNC", acedida pela última vez em 17 de Novembro de 2005, <http://www.realvnc.com/>
3. "Windows Management Interface Reference", acedida pela última vez em 17 de Novembro de 2005, http://msdn.microsoft.com/library/en-us/wmisdk/wmi/wmi_reference.asp
4. "XSD", acedida pela última vez em 17 de Novembro de 2005, <http://www.w3.org/TR/xmlschema-0/>

5. “MPEG-4, eXtensible MPEG-4 Textual Format - XMT-A”, acedida pela última vez em 17 de Novembro de 2005,
<http://www.digitalpreservation.gov/formats/fdd/fdd000156.shtml>
6. “MPEG-4, eXtensible MPEG-4 Textual Format - XMT-O” , acedida pela última vez em 17 de Novembro de 2005,
<http://www.digitalpreservation.gov/formats/fdd/fdd000156.shtml>
7. “SMIL - *Synchronized Multimedia Integration Language*”, acedida pela última vez em 17 de Novembro de 2005,
<http://doc.async.com.br/w3-recs/REC-smil/introduction.html>
8. “Moving Picture Experts Group”, acedida pela última vez em 17 de Novembro de 2005, <http://www.mpeg.org/MPEG/index.html>
9. “BIFS - BInary Format for Scenes”, acedida pela última vez em 17 de Novembro de 2005, http://www.chiariglione.org/mpeg/tutorials/papers/icj-mpeg4-si/05-BIFS_paper/5-BIFS_paper.htm
10. “VRML - Virtual Reality Modeling Language”, acedida pela última vez em 17 de Novembro de 2005, <http://www.w3.org/MarkUp/VRML/>
11. “MPEG-4, eXtensible MPEG-4 Textual Format (XMT)”, acedida pela última vez em 17 de Novembro de 2005,
<http://www.digitalpreservation.gov/formats/fdd/fdd000156.shtml>
12. “X3D”, Extensible 3D: XML meets VRML, Len Bullard,
<http://www.xml.com/pub/a/2003/08/06/x3d.html>
13. “W3C - The World Wide Web Consortium”, acedida pela última vez em 17 de Novembro de 2005, <http://www.w3.org/>
14. “Les formats du multimédia et des hypermédias”, acedida pela última vez em 17 de Novembro de 2005, <http://www.enseiht.fr/~dayre/hypermedias/Formats.html>
15. ISO/IEC International Organisation for Standardisation. 14496-11:2004. Coding of Audio Visual Objects – Part 11: Scene description and application engine/Amd 4. XMT & MPEG-J extensions. 2004.