7. information system architecture

learning objectives

After reading this chapter you should be able to dicuss the considerations that play a role in developing a multimedia information system, characterize an abstract multimedia data format, give examples of multimedia content queries, define the notion of virtual resources, and discuss the requirements for networked virtual environments.

From a system development perspective, a multimedia information system may be considered as a multimedia database, providing storage and retrieval facilities for media objects. Yet, rather than a solution this presents us with a problem, since there are many options to provide such storage facilities and equally many to support retrieval. In this chapter, we will study the architectural issues involved in developing multimedia information systems, and we will introduce the notion of media abstraction to provide for a uniform approach to arbitrary media objects. Finally, we will discuss the additional problems that networked multimedia confront us with.

1

7.1 architectural issues

The notion of *multimedia information system* is sufficiently generic to allow for a variety of realizations. Let's have a look at the issues involved.

As concerns the database (that is the storage and rerieval facilities), we may have to deal with homegrown solution, commercial third party databases or

(even) legacy sources. To make things worse, we will usually want to deploy a combination of these.

With respect to the information architecture, we may wish for a common format (which unifies the various media types), but in practice we will often have to work with the native formats or be satisfied with a hybrid information architecture that uses both media abstractions and native media types such as images and video.

The notion of media abstraction, introduced in Subrahmanian (1998), allows for uniform indexes over the multimedia information stored, and (as we will discuss in the next section) for query relaxation by employing hierarchical and equivalence relations.

Summarizing, for content organisation (which basically is the information architecture) we have the following options:

content organisation

- *autonomy* – index per media type
- *uniformity* – unified index
- *hybrid* – media indexes + unified index

In Subrahmanian (1998), a clear preference is stated for a uniform approach, as expressed in the *Principle of Uniformity*:

Principle of Uniformity

> *... from a semantical point of view the* content *of a multimedia source is independent of the source itself, so we may use statements as meta data to provide a description of media objects.*

Naturally, there are some tradeoffs. In summary, Subrahmanian (1998) claims that: metadata can be stored using standard relational and OO structures, and that manipulating metadata is easy, and moreover that feature extraction is straightforward. </lp> Now consider, is feature extraction really so straightforward as suggested here? I would believe not. Certainly, media types can be processed and analysis algorithms can be executed. But will this result in meaningful annotations? Given the current state of the art, hardly so!

research directions– *the information retrieval cycle*

When considering an information system, we may proceed from a simple generic software architecture, consisting of:

software architecture

- a database of media object, supporting
- operations on media objects, and offering
- logical views on media objects

However, such a database-centered notion of information system seems not to do justice to the actual support and information system must provide when considering the full information retrieval cycle:

information retrieval cycle

1. specification of the user's information need
2. translation into query operations
3. search and retrieval of media objects
4. ranking according to likelihood or relevance
5. presentation of results and user feedback
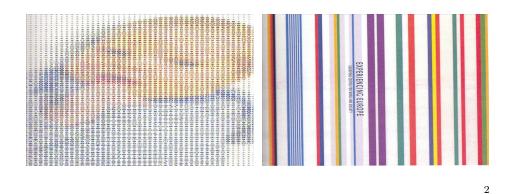6. resulting in a possibly modified query

When we look at older day information retrieval applications in libraries, we see more or less the automation of card catalogs, with search functionality for keywords and headings. Modern day versions of these systems, however, offer graphical userinterfaces, electronic forms and hypertext features.

When we look at the web and how it may support digital libraries, we see some dramatic changes with respect to the card catalogue type of applications. We can now have access to a variety of sources of information, at low cost, including geographically distributed resources, due to improved networking. And, everybody is free to make information available, and what is worse, everybody seems to be doing so. Hence, the web is a continuously growing repository of information of a (very) heterogeneous kind.

Considering the web as an information retrieval system we may observe, following Baeza-Yates and Ribeiro-Neto (1999), that:

- despite high interactivity, access is difficult;
- quick response is and will remain important!

So, we need better (user-centered) retrieval strategies to support the full information retrieval cycle. Let me (again) mention someof the relevant (research) topics:*user interfaces, information visualisation, user-profiling and navigation.*

2

7.2 media abstractions

Let's have a closer look at media abstractions. How can we capture the characterization of a variety of media types in one common media abstraction. A definition

of such a media abstraction is proposedin Subrahmanian (1998). Leaving the formal details aside, a media abstraction has the following components:

<div align="right">*media abstraction*</div>

- *state* – smallest chunk of media data
- *feature* – any object in a state
- *attributes* – characteristics of objects
- *feature extraction map* – to identify content
- relations – to capture state-dependent information
- (inter)relations between 'states' or chunks

Now, that characterization is sufficiently abstract, and you may wonder how on earth to apply this to an actual media database.

However, before giving some examples, we must note that the *feature extraction map* does not need to provide information about the content of a chunk of media data automatically. It may well be a hand-coded annotation.

Our first example is an image database.

<div align="right">*example – image database*</div>

states: { pic1.gif,...,picn.gif }
features: names of people
extraction: find people in pictures
relations: left-of, ...

In an image database it does not make much sense to speak about relations between 'states' or chunks of media data, that is the images.

For our next example though, video databases, it does make sense to speak about such relations, since it allows us to talk about scenes as sequences of frames.

<div align="right">*example – video database*</div>

states: set of frames
features: persons and objects
extraction: gives features per frame
relations: frame-dependent and frame-independent information
inter-state relation: specifies sequences of frames

Now, with this definition of media abstractions, we can define a simple multimedia database, simply as

<div align="right">*simple multimedia database*</div>

- a finite set M of media abstractions

But, following Subrahmanian (1998), we can do better than that. In order to deal with the problems of *synonymy* and *inheritance*, we can define a structured multimedia database that supports:

<div align="right">*structured multimedia database*</div>

- *equivalence relations* –to deal with synonymy
- *partial ordering* – to deal with inheritance

- *query relaxation* – to please the user

Recall that we have discussed the relation between a 'house of prayer' and 'church' as an example of synonymy in section 4.3. As an example of inheritance we may think of the relation between 'church' and 'cathedral'. Naturally, every cathedral is a church. But the reverse does not necessarily hold. Having this information about possible equivalence and inheritance relationships, we can relax queries in order to obtain better results. For example, when a user asks for cathedral in a particular region, we could even notify the user of the fact that although there are no cathedrals there, there are a number of churches that may be of interest. (For a mathemathical characterization of structured multimedia databases, study Subrahmanian (1998).)

3

query languages Having media abstractions, what would a query language for such a database look like? Again, following Subrahmanian (1998), wemay extend SQL with special functions as indicated below:

SMDS – functions

Type: object $\mapsto$ type
ObjectWithFeatures: $f \mapsto \{o \mid \text{object } o \text{ contains } f\}$
ObjectWithFeaturesAndAttributes: $(f, a, v) \mapsto \{o \mid o \text{ contains } f \text{ with } a = v\}$
FeaturesInObject: $o \mapsto \{f \mid o \text{ contains } f\}$
FeaturesAndAttributesInObject: $o \mapsto \{(f, a, v) \mid o \text{ contains } f \text{ with } a = v\}$

Having such functionswe can characterize an extension of SQL, which has been dubbed SMDS-SQL in Subrahmanian (1998), as follows.

SMDS-SQL

SELECT – media entities

- m – if m is not a continuous media object
- $m : [i, j]$ – m is continuous, i, j integers (segments)
- $m.a$ – m is media entity, a is attribute

FROM

- \<media\>\<source\>\<M\>

WHERE

- term IN funcall

As an example, look at the following SMDS-SQL snippet.

```
SELECT M
FROM   smds source1 M
WHERE  Type(M) = Image AND
     M IN ObjectWithFeature("Dennis") AND
     M IN ObjectWithFeature("Jane") AND
     left("Jane","Dennis",M)
```

Note that *M* is a relation in the image database media abstraction, which contains one or more images that depict Jane to the left of Dennis. Now, did they exchange the briefcase, or did they not?

When we do not have a uniform representation, but a hybrid representation for our multimedia data instead, we need to be able to: express queries in specialized language, and to perform operations (joins) between SMDS and non-SMDS data.

Our variant of SQL, dubbed HM-SQL, differs from SMDS-SQL in two respects: function calls are annotated with media source, and queries to non-SMDS data may be embedded.

As a final example, look at the following snippet:

```
SELECT M
FROM smds video1, videodb video2
WHERE M IN smds:ObjectWithFeature("Dennis") AND
     M IN videodb:VideoWithObject("Dennis")
```

In this example, we are collecting all video fragments with Dennis in it, irrespective of where that fragment comes from, an (smds) database or another (video) database.

research directions– *digital libraries*

Where *media abstractions*, as discussed above, aremeant to be technical abstractions needed for uniform access to media items, we need quite a different set of abstraction to cope with one of the major applications of multimedia information storage and retrieval: digital libraries.

According to Baeza-Yates and Ribeiro-Neto (1999), digital libraries will need a long time to evolve, not only because there are many technical hurdles to be overcome, but also because effective digital libraries are dependent on an active community of users:

> *Digital libraries are constructed – collected and organized – by a community of users. Their functional capabilities support the information needs and users of this community. Digital libraries are an extension, enhancement and integration of a variety of information institutions as physicalplaces where resources are selected, collected, organized, preserved and accessed in support of a user community.*

The occurrence of digital libraries on the web is partly a response to advances in technology, and partly due to an increased appreciation of the facilities the internet can provide. From a development perspective, digital libraries may be regarded as:

> *... federated structures that provide humans both intellectual and physical access to the huge and growing worldwide networks of information encoded in multimedia digital formats.*

Early research in digital libraries has focussed on the digitization of existing material, for the preservation of our cultural heritage, as well as on architectural issues for the 'electronic preservation', so to speak, of digital libraries themselves, to make them "immune to degradation and technological obsolescence", Baeza-Yates and Ribeiro-Neto (1999).

To bring order in the variety of research issues related to digital libraries, Baeza-Yates and Ribeiro-Neto (1999) introduces a set of abstractions that is know as the 5S model:

digital libraries (5S)

- *streams*: (content) – from text to multimedia content
- *structures*: (data) – from database to hypertext networks
- *spaces*: (information) – from vector space to virtual reality
- *scenarios*: (procedures) – from service to stories
- *societies*: (stakeholders) – from authors to libraries

These abstractions act as "a framework for providing theoretical and practical unification of digital libraries". More concretely, observe that the framework encompasses three technical notions (streams, structures and spaces; which correspond more or less with data, content and information) and two notions related to the social context of digital libraries (scenarios and societies; which range over possible uses and users, respectively).

For further research you may look at the following resources:

> D-Lib Forum – http://*www.dlib.org*
> Informedia – http://*www.informedia.cs.cmu.edu*

The D-Lib Forum site gives access to a variety of resources, including a magazine with background articles as well as a test-suite that may help you in developing digital library technology. The Informedia site provides an example of a digital library project, with research on, among others, video content analysis, summarization and in-context result presentation.

4

7.3 networked multimedia

For the end user there should not be much difference between a stand-alone media presentation and a networked media presentation. But what goes on *behind the scenes* will be totally different. In this section, we will study, or rather have a glance at, the issues that play a role in realizing effective multimedia presentations. These issues concern the management of resources by the underlying network infrastructure, but may also concern authoring to the extent that the choice of which media objects to present may affect the demands on resources.

To begin, let's try to establish, following Fluckiger (1995), in what sense networked multimedia applications might differ from other network applications:

networked multimedia

- real-time transmission of continuous media information (audio, video)
- substantial volumes of data (despite compression)
- distribution-oriented – e.g. audio/video broadcast

Naturally, the extent to which network resource demands are made depends heavily on the application at hand. But as an example, you might think of the retransmission of television news items on demand, as nowadays provided via both cable and DSL.

For any network to satisfy such demands, a number of criteria must be met, that may be summarized as: throughput, in terms of bitrates and burstiness; transmission delay, including signal propagation time; delay variation, also known as jitter; and error rate, that is data alteration and loss.

For a detailed discussion of criteria, consult Fluckiger (1995), or any other book on networks and distributed systems. With respect to distribution-oriented multimedia, that is audio and video broadcasts, two additional criteria play a role, in particular: multicasting and broadcasting capabilities and document caching. Especially caching strategies are of utmost importance if large volumes of data need to be (re)transmitted.

Now, how do we guarantee that our (networked) multimedia presentations will come across with the right quality, that is free of annoying jitter, without loss or distortion, without long periods of waiting. For this, the somewhat magical notion of *Quality of Service* has been invented. Quoting Fluckiger (1995):

> Quality of Service *is a concept based on the statement that not all appli-
> cations need the same performance from the network over which they run.
> Thus, applications may indicate their specific requirements to the network,
> before they actually start transmitting information data.*

Quality of Service (*QoS*) is one of these notions that gets delegated to the other
parties, all the time. For example, in the MPEG-4 standard proposal interfaces
are provided to determine *QoS* parameters, but the actual realization of it is left
to the network providers. According to Fluckiger (1995) it is not entirely clear
how *QoS* requirements should be interpreted. We have the following options:
we might consider them as hard requirements, or alternatively as guidance for
optimizing internal resources, or even more simply as criteria for the acceptance
of a request.

At present, one thing is certain. The current web does not offer *Quality of
Service*. And what is worse, presentation formats (such as for example *flash*) do
not cope well with the variability of resources. More specifically, you may get
quite different results when you switch to another display platform

5

virtual objects

Ideally, it should not make any difference to the author at what display platform a
presentation is viewed, nor should the author have to worry about low-quality or
ill-functioning networks. In practice, however, it seems not to be realistic to hide
all this variability from the author and delegate it entirely to the 'lower layers' as
in the MPEG-4 proposal.

Both in the SMIL and RM3D standards, provisions are made for the author
to provide a range of options from which one will be chosen, dependent on for
example availability, platform characteristics, and network capabilities.

A formal characterization of such an approach is given in Subrahmanian
(1998), by defining *virtual objects*.

virtual objects

- $VO = \{(O_i, Q_i, C_i) \mid 1 \leqslant i \leqslant k\}$

where

- $C_1, \ldots, C_k$ – mutually exclusive conditions
- $Q_1, \ldots, Q_k$ – queries
- $O_1, \ldots, O_k$ – objects

In general, a virtual object is a media object that consists of multiple objects, that may be obtained by executing a query, having mutually exclusive conditions to determine which object will be selected. Actually, the requirement that the conditions are mutually exclusive is overly strict. A more pragmatic approach would be to regard the objects as an ordered sequence, from which the first eligible one will be chosen, that is provided that its associated conditions are satisfied.

As an example, you may look at the Universal Media proposal from the Web3D Consortium, that allows for providing multiple URNs or URLs, of which the first one that is available is chosen. In this way, for instance, a texture may be loaded from the local hard disk, or if it is not available there from some site that replicates the Universal Media textures.

6

networked virtual environments

It does seem to be an exageration to declare *networked virtual environments* to be the ultimate challenge for networked multimedia, considering that such environments may contain all types of (streaming) media, including video and 3D graphics, in addition to rich interaction facilities. (if you have no idea what I am talking about, just think of, for example, Quake or DOOM, and read on.) To be somewhat more precise, we may list a number of essential characteristics of networked virtual environments, taken from Singhal and Zyda (1999):

networked virtual environments

- *shared sense of space* – room, building, terrain
- *shared sense of presence* – avatar (body and motion)
- *shared sense of time* – real-time interaction and behavior

In addition, networked virtual environments offer

- *a way to communicate* – by gesture, voice or text
- *a way to share ...* – interaction through objects

Dependent on the visual realism, resolution and interaction modes such an environment may be more or less 'immersive'. In a truly immersive environment, for example one with a haptic interface and force feedback, interaction through

objects may become even threathening. In desktop VEs, sharing may be limited to the shoot-em-up type of interaction, that is in effect the exchange of bullets.

Networked virtual environments have a relatively long history. An early example is SIMNET (dating from 1984), a distributed command and control simulation developed for the US Department of Defense, Singhal and Zyda (1999). Although commercial multi-user virtual communities, such as the *blaxxun* Community server, may also be ranked under networked virtual environments, the volume of data exchange needed for maintaining an up-to-date state is far less for those environments than for game-like simulation environments from the military tradition. Consider, as an example, a command and control strategy game which contains a variety of vehicles, each of which must send out a so-called *Protocol Data Unit* (PDU), to update the other participants as to their actual location and speed. When the delivery of PDUs is delayed (due to for example geographic dispersion, the number of participants, and the size of the PDU), other strategies, such as *dead reckoning*, must be used to perform collision detection and determine possible hits.

To conclude, let's establish what challenges networked virtual environments offers with respect to software design and network performance.

challenges

- *network bandwidth* – limited resource
- *heterogeneity* – multiple platforms
- *distributed interaction* – network delays
- *resource management* – real-time interaction and shared objects
- *failure management* – stop, ..., degradation
- *scalability* – wrt. number of participants

Now it would be too easy to delegate this all back to the network provider. Simply requiring more bandwidth would not solve the scalability problem and even though adding bandwidth might allow for adding another hundred of entities, smart updates and caching is probably needed to cope with large numbers of participants.

The distinguishing feature of networked virtual environments, in this respect, is the need to

manage dynamic shared state

to allow for real-time interaction between the participants. Failing to do so would result in poor performance which would cause immersion, if present at all, to be lost immediately.

example(s) – *unreal*

Unreal Tournament[1] is a highly popular multiplayer game. The storyline is simple, but effective: *It's the year 2362. The most anticipated Tournament ever is about to take place, dwarfing the spectacle and drama of previous events. The finest competitors ever assembled prepare to lay waste to their opponents and claim the Tournament Trophy for themselves.*

There a a number of roles you can associate with:

> the corrupt, thunder crash, iron guard, juggernauts, iron skull, sun blade, super nova, black legion, fire storm, hellions, bloof fist, goliath

An interesting feature of the Unreal Tournament games is that they can be andapted and even be re-programmed[2] by the users themselves, has has been done for example for the *Mission Rehearsal Exercise* discussed in section 9.2.

scripting: www.gamedev.net/reference/list.asp?categoryid=76

research directions– *architectural patterns*

Facing the task of developing a multimedia information system, there are many options. Currently, the web seems to be the dominant infrastructure upon which to build a multimedia system. Now, assuming that we chose the web as our vehicle, how should we approach building such a system or, in other words, what architectural patterns can we deploy to build an actual multimedia information system? As you undoubtly know, the web is a document system that makes a clear distinction between *servers* that deliver documents and *clients* that display documents. See Eliens (2000), section 12.1. At the server-side you are free to do almost anything, as long as the document is delivered in the proper format. At the client-side, we have a generic document viewer that is suitable for HTML with images and sound. Dependent on the actual browser, a number of other formats may be allowed. However, in general, extensions with additional formats are

---

[1] www.unrealtournament.com
[2] www.unrealtournament.com/ut2004/screenshots.php

realized by so-called *plugins* that are loaded by the browser to enable a particular format, such as *shockwave*, *flash* or *VRML*. Nowadays, there is an overwhelming number of formats including, apart from the formats mentioned, audio and video formats as well as a number of XML-based formats as for example SMIL and SVG. For each of these formats the user (client) has to download a plugin. An alternative to plugins (at the client-side) is provided by Java *applets*. For Java applets the user does not need to download any code, since the Java platform takes care of downloading the necessary classes. However, since applets may be of arbitrary complexity, downloading the classes needed by an application may take prohibitively long.

The actual situation at the client-side may be even more complex. In many cases a media format does not only require a plugin, but also an applet. The plugin and applet can communicate with eachother through a mechanism (introduced by Netscape under the name LiveConnect) which allows for exchanging messages using the built-in DOM (Document Object Model) of the browser. In addition, the plugin and applet may be controlled through Javascript (or VBscript). A little dazzling at first perhaps, but usually not to difficult to deal with in practice.

Despite the fact that the web provides a general infrasructure for both (multimedia) servers and clients, it might be worthwhile to explore other options, at the client-side as well as the server-side. In the following, we will look briefly at:

- the Java Media Framework, and
- the DLP+X3D platform

as examples of, respectively, a framework for creating dedicated multimedia applications at the client-side and a framework for developing intelligent multimedia systems, with client-side (rich media 3D) components as well as additional server-side (agent) components.

Java Media Framework The Java platform offers rich means to create (distributed) systems. Also included are powerful GUI libraries (in particular, Swing), 3D libraries (Java3D) and libraries that allow the use and manipulation of images, audio and video (the Java Media Framework). Or, in the words of the SUN web site:

<div align="right">java Media Framework[3]</div>

> *The JavaTM Media APIs meet the increasing demand for multimedia in the enterprise by providing a unified, non-proprietary, platform-neutral solution. This set of APIs supports the integration of audio and video clips, animated presentations, 2D fonts, graphics, and images, as well as speech input/output and 3D models. By providing standard players and integrating these supporting technologies, the Java Media APIs enable developers to produce and distribute compelling, media-rich content.*

However, although Java was once introduced as the *dial tone of the Internet* (see Eliens (2000), section 6.3), due to security restrictions on applets it is not

---

[3]java.sun.com/products/java-media/jmf/reference/api

always possible to deploy media-rich applets, without taking recourse to the Java plugin to circumvent these restrictions.

DLP+X3D In our DLP+X3D platform, that is introduced in section **??** and described in more detail in appendix **??**, we adopted a different approach by assuming the availability of a generic X3D/VRML plugin with a Java-based External Authoring Interface (EAI). In addition, we deploy a high-level ditributed logic programming language (DLP) to control the content and behavior of the plugin. Moreover, DLP may also be used for creating dedicated (intelligent) servers to allow for multi-user applications.

The DLP language is Java-based and is loaded using an applet. (The DLP jar file is of medium size, about 800 K, and does not require the download of any additional code.) Dua, again, to the security restrictions on applets, additional DLP servers must reside on the site from where the applet was downloaded.

Our plugin, which is currently the *blaxxun* VRML plugin, allows for incorporating a fairly large number of rich media formats, including (real) audio and (real) video., thus allowing for an integrated presentation environment where rich media can be displayed in 3D space in a unified manner. A disadvantage of such a unified presentation format, however, is that additional authoring effort is required to realize the integration of the various formats.

7.4 development(s) – living in a virtual economy

Mashups on the Web are interesting representatives of what one may call a *virtual economy*, with a business-model that is not grounded in traditional *production* and *trade* values, but rather consists of value-added services with an indirect, albeit substantial, financial spin-off, due to recommendations and referrals. The basic mechanisms in a recommender economy are, according to Kassel et al. (2007):

- cross sale – users who bought A also bought B
- up sale – if you buy A and B together ...

Where the principles underlying this virtual economy have definitely proven their value in first (ordinary) life economy, what are the chances that these principles are also valid, for example, in Second Life?

According to the media companies selling their services to assist the creation of presence in Second Life, there are plenty *New Media Opportunities In The Online World Second Life*[4], to a possibly even greater extent, as they boldly claim, as in what they call *the predecessor of Second Life, the World Wide Web.*

To assess the role web services, including semantic web services, may play in Second Life, it seems worthwhile to investigate to what extent web services can be deployed to deliver more traditional media, such as *digital TV*. To support the business model of digital TV, which in outline may be summarized as *providing additional information*, *game playing* and *video on demand*, with an appropriate payment scheme, Daskalova & Atanasova (2007) argue in favor of the use of a

---

[4]www.youtube.com/watch?v=8NOHRJB9uyI

SOA (Service Oriented Architecture), to allow for a unified, well-maintainable approach in managing collections of audio-visual objects. Such services would include meta-data annotation, water-marking for intellectual property protection, and search facilities for the end-user. Atanasova et al. (2007) even propose to wrap each individual audio-visual object in a (semantic) web service and provide compound services based on semantic web technologies such as OWL-S[5] (OWL-based Web Service Ontology) and WSMO[6] (Web Service Modelling Ontology) using semi-automatic methods together with appropriate semantic web tools[7], for the description and composition of such services. Obviously, there is a great technical challenge in creating such self adjusting service environments.

With respect to the application of web services in Second Life, however, a far more modest aim, it seems that nevertheless the business model associated with the delivery of media items through digital TV channels may profitably be used in Second Life, and also the idea of wrapping media items in web services has in some way an immediate appeal.

In Eliens & Wang (2007), we introduced the notion of *serial recommender*, which generates guided tours in 3D digital dossier(s) based on (expert) user tracking. See section 6.4. To incrementally refine such tours for individual users, we used a behavioral model originally developed in Oard et al. (2006). This model distinguishes between:

$$U = user$$
$$I = item$$
$$B = behavior$$
$$R = recommendation$$
$$F = feature$$

and allows for characterizing observations (from which implicit ratings can be derived) and recommendations, as follows:

- observations – $U \times I \times B$

- recommendations – $U \times I$

In a centralized approach the mapping $U \times I \times B \to U \times I$ provides recommendations from observations, either directly by applying the $U \times I \to I \times I$ mapping, or indirectly by the mapping $U \times I \to U \times U \to I \times I$, which uses an intermediate matrix (or product space) $U \times U$ indicating the (preference) relation between users or user-groups. Taken as a matrix, we may fill the entries with distance or weight values. Otherwise, when we use product spaces, we need to provide an additional mapping to the range of $[0, 1]$, where distance can be taken as the dual of weight, that is $d = 1 - w$.

In a decentralized approach, Oard et al. (2006) argue that it is better to use the actual features of the items, and proceed from a mapping $I \times F \to U \times I \times R$.

---

[5]www.daml.org/services/owl-s/

[6]www.wsmo.org/

[7]composing-the-semantic-web.blogspot.com/

Updating preferences is then a matter of applying a $I \times B \rightarrow I \times F$ mapping, by analyzing which features are considered important.

For example, observing that a user spends a particular amount of time and gives a rating r, we may apply this rating to all features of the item, which will indirectly influence the rating of items with similar features.

B = [time = 20sec, rating = r]
F = [artist = rembrandt, topic = portrait]
R = [artist(rembrandt) = r, topic(portrait) = r]

Oard et al. (2006) observe that B and R need not to be standardized, however F must be a common or shared feature space to allow for the generalization of the rating of particular items to similar items.

With reference to the CHIP project, mentioned in the previous section, we may model a collection of artworks by (partially) enumerating their properties, as indicated below:
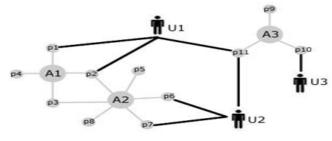
A = [$p_1, p_2, ...$]
where p_k = [$f_1 = v_1, f_2 = v_2, \dots$]

with as an example

$A_{nightwatch}$ = [artist=rembrandt, topic=group]
$A_{guernica}$ = [artist=picasso, topic=group]

Then we can see how preferences may be shared among users, by taking into account the (preference) value adhered to artworks or individual properties, as illustrated in the figure below.

users, artworks and properties

9

As a note, to avoid misunderstanding, Picasso's Guernica is not part of the collection of the Rijksmuseum, and does as such not figure in the CHIP studies. The example is taken, however, to clarify some properties of metrics on art collections, to be discussed in the next section.

To measure similarity, in information retrieval commonly a distance measure is used. In mathematical terms a distance function $d : X \rightarrow [0,1]$ is distance measure if:

$$d(x, y) = d(y, x)$$
$$d(x, y) \leqslant d(x, z) + d(z, y)$$
$$d(x, x) = 0$$

From an abstract perspective, measuring the distance between artworks, grouped according to some preference criterium, may give insight in along which dimesnion the grouping is done, or in other words what attributes have preference over others. When we consider the artworks

$a_1 = [$ artist $=$ rembrandt, topic $=$ self-portrait $]$
$a_2 = [$ artist $=$ rembrandt, name $=$ nightwatch $]$
$a_3 = [$ artist $=$ picasso, topic $=$ self-portrait $]$
$a_4 = [$ artist $=$ picasso, name $=$ guernica $]$

we may, in an abstract fashion, deduce that if $d(a_1, a_2) < d(a_1, a_3)$ then $r(topic) < r(artist)$, however if $d(a_1, a_3) < d(a_1, a_2)$ the reverse is true, that is then $r(artist) < r(topic)$. Somehow, it seems unlikely that a_2 and a_4 will be grouped together, since even though their topic may considered to be related, the aesthetic impact of these works is quite different, where *self portrets* as a genre practiced over the centuries indeed seem to form a 'logical' category. Note that we may also express this as $w(artist) < w(topic)$ if we choose to apply weights to existing ratings, and then use the observation that if $d(a_1, a_3) < d(a_1, a_2)$ then $w(artist) < w(topic)$ to generate a guided tour in which a_3 precedes a_2.

For serial recommenders, that provide the user with a sequence of items $\ldots, s_{n-1}, s_n, \ldots$, and for s_n possibly alternatives $a_1, a_2, \ldots$, we may adapt the (implied) preference of the user, when the user chooses to select alternative a_k instead of accepting s_n as provided by the recommender, to adjust the weight of the items involved, or features thereof, by taking into account an additional constraint on the distance measure. Differently put, when we denote by $s_{n-1} \mapsto s_n/[a_1, a_2, \ldots]$ the presentation of item s_n with as possible alternatives $a_1, a_2, \ldots$, we know that $d(s_{n-1}, a_k) < d(s_{n-1}, s_n)$ for some k, if the user chooses for a_k In other words, from observation B_n we can deduce R_n:

$B_n = [$ time $=$ 20sec, forward $= a_k]$
$F_n = [$ artist $=$ rembrandt, topic $=$ portrait $]$
$R_n = [$ $d(s_n, a_k) < d(s_n, s_{n+1})$ $]$

leaving, at this moment, the feature vector F_n unaffected. Together, the collection of recommendations, or more properly revisions R_i over a sequence S, can be solved as a system of linear equations to adapt or revise the (original) ratings. Hence, we might be tempted to speak of the $R4$ framework, *rate, recommend, regret, revise*. However, we prefer to take into account the cyclic/incremental nature of recommending, which allows us to identify revision with rating.

measures for feedback discrepancey So far, we have not indicated how to process user feedback, given during the presentation of a guided tour, which in

the simple case merely consists of selecting a possible alternative. Before looking in more detail at how to process user feedback, let us consider the dimensions involved in the rating of items, determining the eventual recommendation of these or similar items. In outline, the dimensions involved in rating are:

dimension(s)

- positive vs negative
- individual vs community/collaborative
- feature-based vs item-based

Surprisingly, in Shneiderman (1997) we found that negative ratings of artworks had no predictive value for an explicit rating of (preferences for) the categories and properties of artworks. Leaving the dimension *individual vs community/collaborative* aside, since this falls outside of the scope of this paper, we face the question of how to revise feature ratings on the basis of preferences stated for items, which occurs (implicitly) when the user selects an alternative for an item presented in a guided tour, from a finite collection of alternatives.

A very straightforward way is to ask explicitly what properties influence the decision. More precisely, we may ask the user why a particular alternative is selected, and let the user indicate what s/he likes about the selected alternative and dislikes about the item presented by the recommender. It is our expectation, which must however yet be verified, that negative preferences do have an impact on the explicit characterization of the (positive and negative) preferences for general artwork categories and properties, since presenting a guided tour, as an organized collection of items, is in some sense more directly related to user goals (or educational targets) than the presentation of an unorganized collection of individual items. Cf. van Setten (2005).

So let's look at $s_{n-1} \mapsto s_n/[a_1, a_2, \ldots]$ expressing alternative selection options $a_1, a_2, \ldots$ at s_n in sequence $S = \ldots, s_{n-1}, s_n$. We may distinguish between the following interpretations, or revisions:

interpretation(s)

- neutral interpretation – use $d(s_n, a_k) < d(s_n, s_{n+1})$
- positive interpretation – increase $w(feature(a_k))$
- negative interpretation – decrease $w(feature(s_{n+1}))$

How to actually deal with the revision of weights for individual features is, again, beyond the scope of this paper. We refer however to Eliens (2000), where we used feature vectors to find (dis)similarity between musical fragments, and to Schmidt et al. (1999), on which our previous work was based, where a feature grammar is introduced that characterizes an object or item as a hierarchical structure, that may be used to access and manipulate the component-attributes of an item.

questions

1. (*) What are the issues in designing a *(multimedia) information system architecture.* Discuss the tradeoffs involved.

concepts

2. What considerations would you have when designing an architecture for a multimedia information system.

3. Characterize the notion of *media abstraction.*

4. What are the issues in *networked multimedia.*

technology

5. Describe (the structure of) a video database, using *media abstractions.*

6. Give a definition of the notion of a *structured multimedia database.*

7. Give an example (with explanation) of querying a *hybrid multimedia database.*

8. Define (and explain) the notion of *virtual objects* in *networked multimedia.*

projects & further reading As a project, you may implement a multi-player game in which you may exchange pictures and videos, for example pictures and videos of celebrities.

Further you may explore the development of a data format for text, images and video with appropriate presentation parameters, including postioning on the screen and intermediate transitions.

For further reading you may study information system architecture patterns[8], nd explore the technical issues of constructing server based advanced multimedia applications in Li and Drew (2004).

the artwork

1. examples of dutch design, from Betsky (2004).

2. idem.

3. screenshots – from *splinter cell: chaos theory*, taken from Veronica/Gammo[9], a television program about games.

4. screenshots – respectively *Sekken 5*, *Sims 2*, and *Super Monkey Ball*, taken from insidegamer.nl[10].

5. screenshots – from Unreal Tournament[11], see section 7.3.

6. idem.

7. idem.

8. *resonance* – exhibition and performances, Montevideo[12], april 2005.

---

[8]www.opengroup.org/architecture/togaf8-doc/arch/p4/patterns/patterns.htm
[9]www.gammo.nl
[10]http://www.insidegamer.nl
[11]www.unrealtournament.com/ut2004/screenshots.php
[12]www.montevideo.nl

9. CHIP – property diagram connecting users.

10. signs – sports, van Rooijen (2003), p. 274, 275.

Opening this chapter are examples of *dutch design*, taken from the book *False Flat*, with the somewhat arrogant subtitle *why is dutch design so good?*. It is often noted that dutch design is original, functional and free from false traditionalism. Well, judge for yourself.

The screenshots from the various games are included as a preparation for chapter 9, where we discuss realism and immersion in games, and also because multiplayer games like *Unreal Tournament* have all the functionality a serious application would ever need.