Future User Centric Media: Research Challenges and the Road Ahead

Oscar Mayora*

NextMEDIA Coordination Action
oscar.mayora@create-net.org
http://www.gatv.ssr.upm.es/nextmedia/

Abstract. This work presents relevant excerpts from two white papers recently published with the support of the Networked Media Unit of the European Commission, outlining a number of research challenges and the way ahead towards providing user centric media, as well as rich multimedia and multimodal content, via the Internet and Ambient Systems over the next decade. It will focus on key advances in Future Media technologies that will enable innovative applications and services, engaging new experiences, where multimedia digital content will be more immersive and closely related to the physical world. According to the consolidated opinion of international experts in the field, high quality research on these challenges will be able to enhance our communication experiences, as well as the way we work and live in the next future.

Keywords: Future Media, User Centric Media, Ambient Media Challenges.

1 Introduction

This work reports relevant excerpts from recent forums and publications supported by the Networked Media Unit of the European Commission, aimed to understand which future challenges need to be addressed by the research and professional community in order to provide user centric media and innovative applications and services through the Internet and Ambient Systems over the next decade.

The first part of the paper will introduce and outline main research challenges in the design of user centric media services in the extended home, where the focus must not be only on the devices and the technologies within the physical home space, but to applications where the user is placed in the centre and allows the home-based services and content to follow the user, regardless of the physical location or the device used for content consumption[1]. It will also address main challenges in the design of personalized access to Media Systems.

The second part of the paper will reflect the consolidated opinion of 25 experts from the EU, USA and Korea consulted by the Future Media Internet Task Force (FMI-TF) on challenges regarding Future Media Technologies, along with the potential impact these challenges might have for the development of Ambient Systems and Media[2].

* Supported by the Networked Media Unit of the DG Information Society and Media of the European Commission.

[1] ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/netmedia/user-centric-media_en.pdf

[2] http://www.future-internet.eu/uploads/media/FMI-TF-White_paper_042010_01.pdf

S. Gabrielli, D. Elias, and K. Kahol (Eds.): AMBI-SYS 2011, LNICST 70, pp. 78–106, 2011.

1.1 User Centric Media Services in the Extended Home

The 'connected home' concept has been long researched and solutions that combine communication, audio-visual content and home automation are beginning to appear on the market, promising to simplify the everyday life at home. However, until recently the focus has been on the devices and the technologies within the physical home space. A major breakthrough for user centric media services is taking place nowadays through the delivery and sharing of media content.

Up to now, content was typically broadcasted, to the home, and the users were able to consume it on different home devices, in quite a static manner. The 'extended home' concept places the user in the centre and allows the home-based services and content to follow the user, regardless of the physical location or the device used for content consumption. The user is able to remotely connect to the home and, in a seamless and smooth manner, access the services according to context and profiling. One can consider extending the home network by [1]:

- *Spatial extensions*
  - Users 'taking home with them' when not at home (e.g., access to home information, remote home control)
  - Extending the (closed) home network to multiple homes or home-like sites (office, car, etc.)
- *Functional extensions*
  - Extending the home interface to external services

User centric media services are services created and delivered taking into account the user's preferences and context but also services built by the users themselves.

Such a system is able to assure strong user involvement to co-create their applications by uploading and sharing their user-generated content. The user is not only a content consumer, but also a content creator and distributor.

Sharing of self-created media items can occur among a group of family and friends, or a much wider community/social network.

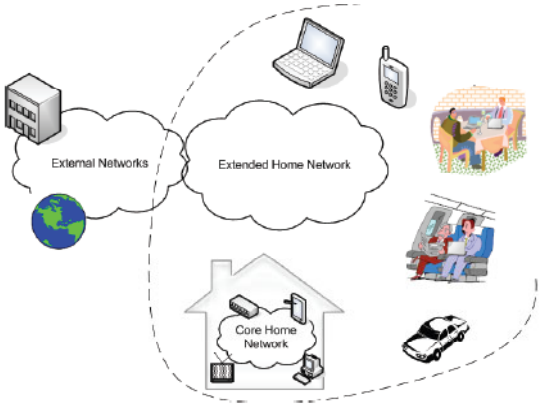

Fig. 1. The Extended Home Concept

1.2 State of the Art

Content distribution channels are changing and the typical broadcast 'one-to-many' model is being redefined.

Large media companies are using alternative methods of distributing their content, including methods that were considered as a threat, in the past. Most music labels sell their content secured with Digital Rights Management (DRM) systems, but some started distributing it also DRM-free, in order to allow users enjoy their legally purchased content on any of their home/mobile devices, without being 'locked' on a specific DRM platform. Traditional TV channels are starting to distribute their programs via Internet peer-to-peer systems.

The Joost service allows users to get a 'TV-like' experience on their personal computers, by leveraging the peer-to-peer distribution, for enjoying content on demand, regardless of their geographical location, or access network. But even the typical television broadcasting experience has become more user-centric by allowing consumers to have control of their preferred time for watching the content. Time-shifting devices and services digitally record programs, from the broadcast television, and have it readily available for users to enjoy at some later point in time, whenever they choose to do so.

Moreover, on-line web-based services, like You-Tube, allow users to easily become active content distributors and reach global audience, by uploading their creations and sharing them with friends or the whole Internet community. Similar initiatives for community oriented content-creation and distribution have been taking place worldwide considering different media channels such as the internet, and most recently through iDTV. Example of this is SAMBA EU project that focuses on generation and distribution of community- oriented content for enhancing inclusion of individuals in society.

As the size of home digital storage devices constantly increases, while their prices keep decreasing, users have the ability to store huge amounts of digital content, both commercial and user-generated, in their homes. Terabyte storage in the home will soon be reality, with such a low cost that no on-line storage service will be able to provide.

In the near future, users will host most of their content in their home network and will access it when away from home, using a mobile device or a remote computer.
The extended home concept, which enables the home content to be accessed from any location depending on the user's preference, is already a reality. Solutions such as the Orb [2] allow the remote access of the home hosted content, from any device that has a web browser. As a PC based solution, it assumes that the home computer is running constantly all day, so that it can serve the remote clients.

At the same time, dedicated stand-alone consumer electronics devices enable the concept of content place-shifting. They are getting the content from the live broadcasted channels, that the user has subscribed to and which are delivered to the home, and encode them to be streamed over the Internet to other personal devices, such as mobile clients. This allows the users to enjoy their favourite television channels, which are available in their homes, as well as in remote locations, in real-time.

Research on personalization, profiling and context awareness has been going on for many years for delivering services and content adapted to the user preferences. [3,4,5]. The mobile world through the 3GPP activities deals with personalization, built on the concept of a Personal Service Environment (PSE), describing how a user can manage and interact with his or her services, and a Virtual Home Environment (VHE), specifying how personalized information can cross network boundaries, thus creating consistency, across multiple devices, with respect to personalized services and features.

In the home environment, there are solutions that allow adaptation of the content to the specific characteristics of the device that will be rendering, mainly focusing on screen size/resolution, supported codecs, etc.

1.3 Gaps and Constraints

Currently, users host their self-generated content in many different locations and devices. Usually, content to be shared with a wider group of friends is uploaded on a 3rd party service, and people are invited to access it from there. More personal content, to be shared only within the family members, is hosted on home devices and there is always content stored on some of the creation devices, like camera phones. For the end users, this fragmentation of content repositories is confusing and difficult to manage.

Hosting content on 3rd party services is efficient, from bandwidth point of view, when the content needs to be shared with a large group of people, but there are always privacy issues.

On the other hand, content that is stored on a home device, which can be accessed via remote access to the home network, is stored on a safe, user-managed, place but, it does not scale for large amounts of visiting users, since remote accessing visitors would cause consumption of the available bandwidth that the home network has. In this way it is clear that there is need for a system that can efficiently and automatically handle the sharing of the user generated content.

Due to the fact that the home network would need to be accessible from the public Internet, for enabling the remote access use-cases, there is a need for a standardized solution. All home devices, regardless of vendor, should be accessible from any kind of personal remote and mobile device.

The access has been given regardless of the Internet Service Provider, that provides the home connectivity to the Internet, or the access network that the remote device uses. Current solutions are proprietary and include many restrictions. Most of them relay either on the fact that homes have publicly accessible Internet Protocol addresses, or that the traffic (between the two end points) is tunnelled via a 3rd party service.

The next generation of the Internet Protocol (IPv6), which has been promising a solution for those two problems for many years, has not been widely deployed yet, and it would bring a new level of administrative complexity in the home, since the direct connectivity to any home device would require much more strict firewall and privacy rules.

The research work that has been done on the personalization area has been focusing on the delivery of the content, which is in the best interest of the user,

according to his/her profile. However, the current work on adaptation of the content seems to be focused only on the required transformations to render the media of the preferred device. There is no mechanism to allow, for example, adaptation of the content to fit the time constraints of the user. The reason is that multimedia content is usually treated as a block entity, with just some metadata enhancements. Even worse, if the content is user-generated it is usually lacking metadata, which makes it very difficult to find and adapt after publication.

Recently found European projects are addressing the problems mentioned about in various ways. An example of this can be found in the 6th FP, CONTENT NoE (Content Network and Services for Home Users) whose research goal is to provide seamless, context-aware, end-to-end content delivery in a heterogeneous environment in a community context. This should be possible by addressing the key areas identified as community networks that cover the residential users. A detailed CONTENT research roadmap describes the research challenges in content services, overlay networks and community networks from the point of view of the residential users: CONTENT Deliverable D2.1. 'CONTENT Research Vision and Roadmap' [6].

1.4 Research Challenges

A) Seam less access to content, regardless of its location /repository

User-created content is distributed in many places (online services, home, mobile devices, etc.), and it will always be like this, since every solution has its own advantages and disadvantages, depending on the required usage. One of the key roles that the home intelligence systems could play in the future is the coordination of those resources and repositories. Enabling seamless sharing of all media repositories, between the owner and potential visitors, no matter where the content is really hosted, would be the target. The home, Internet and mobile domain need to merge towards the ultimate user experience, hiding from the users the protocols, interfaces and bearers used for transferring and sharing content. Especially as user generated content size is getting larger in size, due to high resolution cameras, high-definition, etc., the traditional sharing methods (such as MMS, or email) are becoming very limited and unusable.

b) Protecting the digital user experiences, for the future

Clever systems are needed for helping the users automatically backup and archive their content. Nowadays people happily keep a lot of their life-time experiences in digital format, but do not really know how to take care of them, for ensuring that they will not lose them, in case of an accident or a hardware failure.

Unless there is a system in place that can protect the users, and act on their behalf for securing their digital experiences, one will face the sad incidents that people will lose digital experiences (images, videos, etc.) that may have been accumulated over time and the new technology will prove to be much less time resistant than the old traditional methods of storing experiences (e.g. photo paper). Smart homes, mobile devices, storage device and on-line services must form a safety grid for all the user-generated content.

c) How to protect content distribution in the extended home

Digital technology has turned the vision of on-demand content into a reality. With high quality content being so easily transferable, digital content can be enjoyed anytime and anywhere - on any devices of the extended home (TV, DVR, DVD or portable devices).

To ensure that simple and flexible content usage models can be effectively implemented, the industry needs a content protection standard that reflects the needs of all its players:

- *Content providers* want to protect their intellectual property rights against unauthorized redistribution while providing business models that encourage consumption in a secure chain of trust.
- *CA/DRM vendors* want to ensure that the content rights and usage terms defined by their systems are applied to all connected devices in the home network.
- *Device manufacturers and chip vendors* seek an open, low cost solution that promotes device interoperability and a level playing field for competition.
- *Network operators* want to maintain their networks' integrity while ensuring that their consumers can enjoy content on as many connected and unconnected devices as possible.

At the same time, as users start generating more personal content, and share it with their friends, they would like to make sure that it will not leak to others outside their circle of trust. Thus, there is potential need for a DRM system that focuses on user generated content, allowing the creators to set the access and consumption rights for a specific group of people.

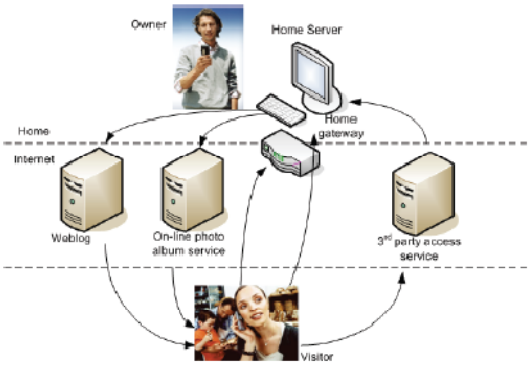

Fig. 2. Current Plethora of Content Offering Methods

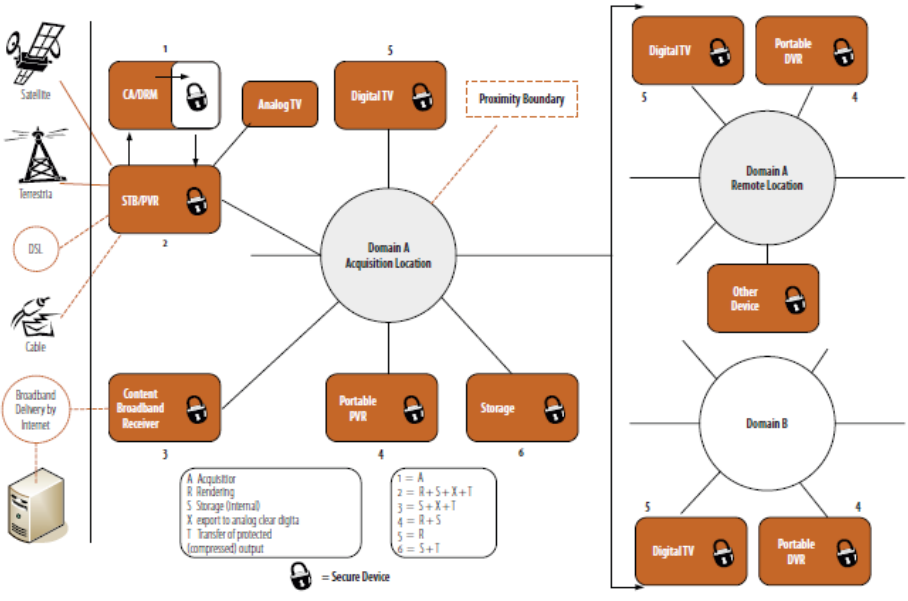

Fig. 3. Secure Domains

d) Content summarization

A new challenge for content adaptation is the method of content summarisation according to the user's profile and time constraints. As users consume content on their mobile devices, away from home, they might have a limited amount of time for viewing some media items. Thus, there is need for intelligent systems that would be able to summarise the content that the user wants to consume, to fit the available time slot that the user has for this. Different media items (e.g. video) can be summarized in order to fit in a smaller time period, without losing essential parts. In such a system, metadata should be inserted through the whole duration of media items, commercial and user created, so that intelligent summarization and optimisation could be carried out based on user mood, available time, and environment context.

e) New user interaction methods in the home

As the activities that the users perform in the digital home are getting complex, there is need for new interaction methods that would allow them to access services and enjoy their content more easily and intuitively. A far more natural interaction between the users and the home devices is required, in order to deliver a personalized experience within the home and the social networks, taking into consideration multiple devices and environments. Research should be focussed on how different kinds of sensors and actuators, such as acceleration, gyroscopes, motion detectors and cognitive cameras, can be used in the home environment for advanced interactive content and context based services. This information, combined with the profiling and

context data would allow user interaction via simple, more natural interfaces based on gestures, touch, speech, etc.

1.5 Business Cases / Use Cases

As the home environment becomes more aware about its users and their habits, it will begin making clever decisions on their behalf. Identifying the people that are in the home, their mood and, for example, their next day schedule, the home entertainment system could suggest a list of the most appropriate tv programmes to watch. In the future, users will be able to interact with their media systems, not using a traditional remote control, but maybe just a natural gesture.

These new sets of use cases will require the cooperation and interoperability of multiple in-home and external services. The model of selling stand-alone products and services will be (and is already) transformed to the model of offering experiences, instead.

Users consume content on the mobile devices when they are away from home, for example when commuting to work by public transportation. Depending on the length of their daily trip they might want to watch something on their portable devices, exactly for their duration of their trip, without losing essential information. A clever adaptation system in the home could convert the content in shorter version, to fit exactly the time duration that the user has available.

For example, watching a 45-minute TV series could be cleverly converted into a 35-minute session. Or, watching an originally long personal video clip could become shorter, on demand, by removing parts that are not interesting or are in a way repeated. There is a brand new market of users who are prepared to purchase content if it fits their needs and environment constraints.

New products and services will be offered to the users for automatic management of their content. People will not need to decide anymore what should be stored on the personal computer, what should be backed up on a home storage server, and what needs to be put on on-line services, in order to be shared also with others. The home will be able to support the user by making suggestions and taking some decisions on his behalf, always protecting the content and the privacy of the creators.

Automated backup and archiving will ensure that data is always replicated on a secure location, while clever synchronization mechanisms will make sure that the latest content will be available to the users when using their mobile/portable devices when outside the home. Also, remotely connecting to the home network will become an intuitive and transparent task for users that will not notice explicit differences when accessing data regardless where the content that they consume is located. New market space is created for service providers that will be handling the management and security complexity of the home networks, on behalf of the users.

1.6 Potential Research Themes

The areas which are of particular interest, in the extended home domain, could be some of the following:

- Seamless content access and sharing, regardless of the location of the content (home or externally hosted).

- Mash-up of home and external service. The home providing web-service interfaces to other Internet services, for ultimately personalized experiences.
- The home as the last place of privacy. How to keep it this way?
- Systems for automated backup and content loss prevention.
- New user interaction methods, in the home environment.
- Ambient telepresence and hybrid physical/digital presence.

1.7 Personalised Access to Media Systems

As the available amount of data is constantly increasing, the need for finding accurate and relevant information quickly becomes a necessity. Given that knowledge worker spend a large amount of time in searching for the right information, any mechanism to reduce this time spent will have a large impact. This aim has lead to a growing interest in personalization techniques.

Personalization is the use of technology and user information to tailor interactions between a service or task and each individual user according to the user's perceived or stated preferences. Using information either previously obtained (e.g. interests, occupation, hobbies, etc) or provided in real time about the user (relevance feedback), the offerings of the system are altered to fit the user's stated or inferred needs.

Personalized access is generally described as the system's ability to customize the user interface, the information channels and the services provided according to the individual user's needs, personal interests and preferences.

Personalization is one of the key elements towards achieving success of user centric media services, both in terms of producing and sharing (and/or searching and retrieving) content. The vision is to create a user centric media landscape, in which user creativity and user experience is placed on the forefront.

User creativity is encouraged, and user created content reaches its envisaged audience, by means of developing new ways to find, access and consume ready-made or individually tailored media. New forms of entrepreneurship might arise and existing enterprises will be able to refine their offering enhancing their business or making them more competent by focusing on specific niches.

1.8 State of the Art

The average European user is already in possession of several intelligent devices; this can be seen as an essential shift in modern European lifestyle. It is important to remember that the success of these technologies is strongly related to sociological factors. For each device, a user usually receives and uses content tailored to the specific device.

We come to a point where one can assume that personalization can be categorized into aspects of personalized service and personalized interface. Personalization till now has mostly been perceived as a way to change the look or behaviour of the device (personalized interface), for example ring tones or themes for mobile phones or customisation of popular video and audio players.

However, a personalized service environment would present the key for mass adoption of networked media services, as it would offer users very strong

'socializing' tools. User-produced content would reach its audience easier and User Centric Media approaches would bridge the gap between available technology and 'social' needs.

Several Technologies and approaches, which are already available today, have to be considered when it comes to developing new aspects and trends of User Centric Personalization.

These include aspects of the user device, aspects of the user preferences, requirements (content and context), privacy and aspects of media management (device and content).

Semantics and context also play a critical role in personalization. In today's increasingly dynamic media environments, adequate information about media files and their domain relations and services is essential to allow for user centric media to be accessible and to maintain a certain degree of navigation between the huge amounts of available material.

Ongoing research in technologies related to user, context and device profiling have already resulted in some degree of commercially mature products, for example user profiling in several sites, is using data collected from a number of different sites, which can result in the creation of a personalized web page. Service, content and device providers can leverage this user-specific knowledge to offer products that allow individualized and navigable content query and retrieval. Technologies used include content-based and collaborative filtering, in which a filter is applied to information from different data sources to select relevant data that may apply to the specific (usually e-commerce) experience of a customer or specific group of customers.

The difference between the two lies in the process of identifying the content that is suitable for the user. When using content-based filtering, the suitability of a particular content for a user is estimated through direct comparison of content meta-data and the model of the active user. When using collaborative filtering, the similarity between users is estimated first, and content that is liked by users similar to the active user, is recommended. Both approaches can be used as complementary methods, as both have advantages and drawbacks. Finally, data analysis tools are also used to predict likely future interactions.

User Profiling has also become the prime task of specific consulting companies that collect information about user habits to allow content providers to better target their audiences.

Since personalization depends on gathering and using personal information about the user, adequate protection of users' privacy is a major concern.

1.9 Gaps and Constraints

Currently, personalization is exploited mostly by businesses in order to allow for making consumer-provider relationships more closely tailored to the individual. The perception is solely driven by means of marketing and sales. The user is perceived simply as a consumer, and a sound personalization strategy exists only to improve customer service.

Although profitable businesses have the means to enrich their customer base by using existing technologies, user centric media approaches have different needs and

different goals. It's not sufficient to personalize the end-user and remember their name, but to place context in the equation and to ensure both creative and interface are appropriate for the individual. Personalization then would allow tailoring the level of richness and productivity to create a well-rounded approach for users to interact with each other and also for businesses to interact with the users.

Another important issue is the search and retrieval of the requested material. Users get annoyed by situations that limit their control on the content they receive. The user will, for example, not be satisfied with a slow service response due to networks with low transmission rate, or a service presentation layout that does not fit the display of the user terminal, or get thousands of document titles as a result of an information search. From the service provider's side it is not possible and economical to make different versions of a service for every different terminal type and network.

Unlike conventional text-based search, media (audiovisual) search tasks usually aim at retrieving something similar to a query. In addition, the query specification and resources to be searched are often of different media types, e.g., the query is formulated in terms of keywords (text) and the retrieved resources are images, sound, video, etc. On the other hand, the query itself may be complex consisting of multi-components of different media types. As a consequence, it is usually difficult the express the query precisely. There are often ambiguous mappings between query description and its underlying interpretation.

To capture the user's real intention of audiovisual search, the query usually needs to be properly interpreted in an interactive way. It is more often a result of several iterations of interpretation and refinement.

Recommender Systems were originally defined as systems in which 'people provide recommendations as input, which the system then aggregates and directs to appropriate recipients'. Current generation of recommender systems still requires further improvements to make recommendation methods more effective and applicable to an even broader range of real-life applications, including recommending multimedia content or products to purchase in a store made by a 'smart' shopping cart.

These improvements include better methods for representing user behaviour and the information about the items to be recommended, more advanced recommendation modelling methods, incorporation of various contextual information into the recommendation process, utilization of multi-criteria ratings, development of less intrusive and more flexible recommendation methods that also rely on the measures that more effectively determine performance of recommender systems. Furthermore, most algorithms focus on single item recommendations and do not consider any more complex recommendation structures.

Finally, since personalization depends on personal information about the user, which must be collected and stored by the system, there is a major issue regarding adequate protection of the user's privacy. Up to now, the real identity of the user is mostly concealed behind a pseudonym, which is associated with all user-related information stored by the system.

This implies that the user has somehow registered with the system and must log on to it in order to access his personalized interface. Even though no real identity can be connected with the pseudonym, this doesn't prevent misuse of user information in all cases. An example is targeted advertising, which does not need to know who the user

really is, but is only interested in knowing that the user that is now logged on is known and has a specific profile.

1.10 Research Challenges

a) Adaptive Multimedia Content and Dynamic Personalized Composition of Multimedia Content

Content Adaptation up to now refers to the action of transforming content to adapt to device capabilities.

Content adaptation is usually related to mobile devices that require special handling because of their limited computational power, small screen size and constrained keyboard functionality.

However, effective personalization schemes have to be researched in order to lead to the creation of adaptive content not only in terms of content format but also in terms of content context. Research Items include:

- Semantic-based, personalized delivery concept.
- Adaptation of content based on user experience and perceptional factors. Adaptation of either content or service to achieve the perceptional quality of experience required by each individual user.
- Shifting of multimedia adaptation functionality to Web Services accessible through standard interfaces that allow for multimedia format conversion and composition leading to flexible, application independent adaptation.
- Agent technologies for automated Web service discovery, execution, composition, and interoperation. They present a logic-based agent technology for service composition, predicated on the use of reusable, task-specific, high-level generic procedures and user-specific customizing constraints.

Future research should also aim at the development of technologies and tools to continuously monitor and improve the overall quality of media services.

The outcome of this research (technology and tools) is intended to be implemented in both popular, commercial (e.g. triple-play market, mobile services) and specialized (e.g. medical digital video libraries) services.

The most innovative and prominent functionality is expected to be the introduction feedback mechanisms allowing not only for Quality of Experience (QoE) assessment and monitoring but also for QoE improvement (long-term QoE improvement with network infrastructure and QoS parameters modifications and short-term QoE improvement with flexible content compression, transcoding and adaptation).

In terms of content adaptation, expected diversity of end user content reception devices should also facilitate introduction of content adaptation services, such as compression schemes with new generation time-varying Compression Ratios (CR). Varying in time compression ratio can be applied to a recording in which some parts are of higher importance than others.

b) Efficient Content Search and Retrieval

User centric personalized access to media systems should allow for a very efficient way of locating the desired information based on user preferences and user profiles.

Given the option to fine tune search and retrieval preferences, user centric media approaches will ease the way of locating and retrieving information in environments characterized by immensely large amounts of data and content, either professional or user created and available in single or distributed environments.

Today, even the most advanced methods of content retrieval cannot prevent avalanches of query results coming down to the user, whereas the user would like only the most 'relevant' content displayed. In order to overcome these issues, the following technologies should be further researched:

Methods for modeling users
User interaction mechanisms should provide observation mechanisms, based on which conclusions can be obtained about user preferences. Related to these mechanisms is the type of user relevance feedback that is obtained. Users should be able to evaluate the suitability of a particular content or the system may have to make implicit assumptions based on user actions on the content.

Semantic-based content search with support for content distribution networks
P2P computing is a potentially powerful model for information sharing between ad-hoc groups of users because of its low cost of entry and natural model for resource scaling. As P2P communities grow, however, locating information distributed across the large number of peers becomes problematic. Further research is required in adapting state-of-the-art algorithms, to content sharing environments and to the formulation of multiple communities with cross-interests.

Content-based search is expected to become the next big trend in search engines. These technologies allow the indexing (annotation) and querying of real visual (image, 3D) or acoustic information inside multimedia content and are based on very advanced recognition and segmentation algorithms.

User Centric Media systems will especially benefit from this technology, since it would allow classification of content that is lagging completely metadata annotations or is incorrectly tagged (sometimes on purpose).

Combinations of semantic and content-based search
The introduction of automatic tagging procedures based on low level optic/acoustic characteristics of media files. The latter assumes specific research on the combination of low-level characteristics, which constitute the basis of content-based retrieval, with high-level semantic interpretations of the content in a structured manner (using for example specific ontologies).

Most of the aforementioned research topics concerning content/context based search and retrieval of 3D content are covered by the project VICTORY (www.victory-eu.org), while project DIVAS (www.ist-divas.eu) is effectively researching indexing and direct search techniques for compressed audio and video files.

c) User Privacy

Future systems should address *User Privacy* issues carefully than currently considered. User information should be treated according to easily understandable

policies that users have easy access to. Trust mechanisms should exist, that allow users to verify that systems use their information according to the agreement.

Users should be able to look into the data that has been collected about them, modify it, remove it, and even personalize the data collection policies themselves. Development of standards in this direction would help greatly, as they intend to ensure interoperability between devices and trust by the users.

Convenience and security are always competing aspects of systems of any kind. There is a great challenge in designing systems that are both secure and easy to use. Novel user centric media systems should be both.

d) Content Mobility, Accessibility and Convergence of Networks

Today, mobility is of significant importance. With the introduction of wireless networks the user has become increasingly mobile, freeing himself from the bounds of a fixed connection to the internet, and thus, to the system. Therefore, content produced by the user becomes mobile as well, a fact that raises issues when needing to keep it always accessible to other users of the system.

Content Mobility and Accessibility is therefore a research challenge on its own, in three main areas:

- Provision of tools, services and systems for supporting mobile user creativity and content production (produce and share on the move).
- Efficient content adaptation, transmission, storage and sharing for addressing the largest possible audience.
- Value added services of a converged media and networks. This is a challenge for content distribution networks, especially for personalized access to content, because the ways into which content can be acquired and used are multiplied by this convergence.

e) Content protection

Copyright protection (or Ownership appreciation) of the content and insurance of the content ownership are of significant importance in creating user-centric repositories. So far two main complementary research areas seem to provide the necessary protection both for the content and for the owner: watermarking and Digital Rights Management (DRM). Watermarking techniques have long been used for the provision of robust copyright protection of multimedia material as well as for multimedia annotation, with indexing and labeling information.

The main challenge in content watermarking is the robustness of the watermarking techniques against several types of attacks that do not substantially degrade the model's visual quality such as rotation, translation and uniform scaling, points reordering, remeshing, polygon simplification, mesh smoothing operations, cropping, local deformations.

DRM allows intellectual property owners to express policies for content usage with confidence that these policies will be respected once the content is distributed in the network. Integration of P2P technology with DRM is the latest and possibly the most important technological and business frontier and it is partly faced in the project VICTORY.

1.11 Business Cases/Use Cases

As the users become the centre of interest and gain more power, both user communities and business communities gain from this experience. Combining the best of both worlds', users will finally be able to personalize their 'experience'. From interface and device to content a new media production and consumption environment arises, that serves technological and social advancement alike.

Tailored multimedia content is delivered seamlessly to multiple devices. Users will be able to indicate a preference for a particular content type or types, and content will be chosen or even created, based on user preferences.

Users may further refine content preferences as content is experienced, thus having a more enjoyable experience.

Search results, based on content and semantic characteristics, will be more efficient and more productive. In commercial scenarios users will be able to find and indicate a desire to purchase content or learn more about specific content. Searching through a vast amount of information will become easier and huge niche markets will evolve in which all content can reach its intended audience.

User feedback on the 'Experience' will be considered to enhance future services, from content production to sharing and consumption.

2 3D Content Generation Leveraging Emerging Acquisition Channels

Content creators always look for new forms and ways for improving their content and adding new sensations to the viewer experience. High Definition video has been the latest innovation in the area of content enrichment. 3D is the next single greatest innovation in film-making. Recent film releases such as 'Avatar' have revolutionalised cinema by the extensive use of 3D technology and 3D content production along with real actors creating a new genre at the outset of the 2010s. The box office tickets show that audiences have very quickly embraced this by making Avatar the most successful in cinema ticket sales film in the history of digital cinema (beating even the Titanic).

However, today's young society is becoming increasingly content art and design "literate" as a result of technological advances and lower costs in photography, cinematography, 2D/3D graphics design and animation technologies and as a result of increased emphasis on media design in education. As a result, novel forms of 3D content, should also find its way into small and medium size content creation companies, moving the experience from cinema halls and cinema projectors to the everyday household environments and computers, providing increased number of audiences with a taste of the versatility and power of 3D as both consumers and producers. 3D Internet is a concept that has recently come into the spotlight in the R&D arena, catching the attention of many people, and leading to a lot of discussions. Several research challenges such as: visualisation and representation of information, and creation and transportation of information, among others, will need to be investigated and solution found for 3D internet to become a reality.

The success of 3D cinema has led to several major consumer electronics manufacturers and broadcasters to announce plans to launch 3D-capable TVs and offer 3D content in 2010. 3DTV will require the integration of a diversity of key technologies from computing to graphics, imaging to display, and signal processing to communications.

There are a number of competing 3D technologies available, and the decision to support 3D will require an understanding of the relative merits of each in the context of the home. The provision of 3D content into the home will require significant cooperation between content providers, service providers and consumer electronics manufacturers to ensure consumer confidence in the technology and avoid a repeat of the confusion surrounding the introduction of HD technologies.

Today's 3DTV technology is based on stereo vision where left and right eye images are presented to the viewer through temporal or spatial multiplexing by wearing a pair of glasses. Usually the content is captured using two cameras mounted on a rig. Recently there are few consumer electronics manufacturers that provide a single camera set up for the capture of the left and right eye images. The next step in the 3D TV development could be the multiview autostereoscopic imaging system, where a large number of pairs of video signals are recorded and presented on a display that does not require glasses for viewing. Although, several autostereoscopic displays have been reported, there are still limitations on resolution and viewing position. Furthermore, stereo and multiview technologies rely upon the brain to fuse the two disparate images to create the 3D sensation. As a result such systems tend to cause eye strain, fatigue and headaches after prolonged viewing as users are required to focus to the screen plane but converge their eyes to a point in space, producing unnatural viewing.

With recent advances in digital technology, some human factors which result in eye fatigue have been eliminated. However, some intrinsic eye fatigue factors will always exist in stereoscopic 3D technology.

Creating a truly realistic 3D real-time viewing experience in an ergonomic and cost effective manner is a fundamental engineering challenge. Future 3D technology should seek to advance the current existing technologies not only in capturing and manipulating 3D content but also in creating a new 3D content format which offers fatigue free viewing with more than one person independently of the viewer's position.

3D holoscopic and holography are two technologies that overcome the shortcomings of stereoscopic imaging, but their adoptions for 3D TV and 3D cinema are still in their infancy. Holographic recording requires coherent light but offer the ultimate 3D viewing experience. Holoscopic video uses microlens arrays to recording a 3D scene and can operate under incoherent illumination, which is in contrast with holography, and hence it allows more conventional live capture and display procedures to be adopted.

Future 3D video could use different technologies for 3D content creation and display.

2.1 Future Research Challenges

Communicating information using images plays a major role in today's society. There are a significant number of applications where the ability to display and visualize a

3D image comfortably confers real benefits. Examples from the professional domain include medical imaging, scientific visualization, security/defense, education, computer-aided design, and remote inspection. While in consumer markets 3D video games and 3D multimedia offer a rich experience to the consumer.

In very recent years 3D technology has become an extremely hot topic of research and there is a real feeling of excitement surrounding the technology. With the success of 3D movies such as Avatar burned in their memories, content creators, distributors, consumer electronics manufacturers and Hollywood studios have all expressed serious interest in wowing their own audiences. As a result, research in the 3D technology has intensified to progress its introduction to the home consumer be it 3D TV and/or 3D internet. However, for the 3D technology to be fully adopted by the home consumer, solution to several research challenges need to be investigated and solutions are needed to simplify the generation of 3D content and provide the users/ producers similar hardware and software facilities as those enjoyed today by 2D video makers and users. Among other research challenges, visualization and representation of information, and transportation of information, remain key despite the numerous advances made in the field of stereo vision.

3D content generation

Today, the capture of 3D images and video relies on stereo vision where a number of cameras located in different positions are used. However, this kind of setup for 3D content production is cumbersome and requires correct calibration of all cameras. To that effect an important research challenge is to develop novel technologies which allow the 3D content to be captured using a single 3D camera and hence simplifies 3D video production and allows adoption of the well known techniques used today in 2D video production.

Due the increasing in computing power, computer generated graphics are becoming more and more an essential part of today visual content. Hence another research challenge is provide repositories of reusable and adaptable 3D assets (animated characters, background/environments, props, etc) that can communicate/ interact with each other and systems that share/distribute rendering and processing requirements over intelligent networks. Computer graphics plug-ins have been developed for the generation of stereo images. However, if a new 3D content genre is defined to avoid the shortcomings of stereo imaging, then similarly true 3D graphic tools need to be made available.

3D Content Editing, Authoring and Sharing

The dramatic explosion of the user-generated content culture on the web has illustrated that any new form of 3D content should be able to take off on non professionals online resources such as YouTube. This would allow both professional and non-professional authors to jointly develop 3D content with a more realistic sensation. Tools are required for users, whether professionals or not, to be able, via open source authoring environments, to create intelligent content but also to share it across open networks and remix it by using various elements of different distributed content items to produce in turn new content.

There exist a wealth of software tools available today to home consumer for editing and authoring of 2D audio visual content. Hence, it is a requirement that

similar tools are made available to the home consumer as well as professionals to allow editing and authoring for seamless compositing of any new form of 3D content. The interconnecting backbone of today's World Wide Web is still the hypertext. Although a plethora of multimedia content is published on the internet, non-hypertext interactivity is sandboxed on the site. In order to create a real 3D Internet experience, hyperlinks must not be restricted to the hypertext, but extended to any kind of media available on the internet (hypermedia). In 3D video, extraction of objects is potentially easier, as parallax effects clearly separate foreground objects from the background. These objects can be used to link to further 3D video clips, allowing for nonlinear video experiences.

3D Visualisation

Binocular vision feed the human brain at all time with slightly offset views of a scene: the greater the disparity, the closer the object. At the same time, converging nearby or far away gives our brain hints of distances. This principle has been exploited by stereo vision technique where two discrete views are presented to the left and right eye of the viewer via colour, polarisation or time separation techniques requiring special glasses. Over the last few years a number of autostereoscopic multiview 3D displays have been demonstrated where lenticular optical elements or parallax barriers technologies are used to separated the left and right view.

A particularly contentious aspect of stereoscopic displays for entertainment applications is the human factors issue. Furthermore, due to the lack of perspective continuity in 2D view systems, objects in the scene often lack solidity (cardboarding) and give rise to an 'unreal' experience. Hence for 3D TV and 3D internet to become a reality, another research challenge that needs investigating and resolving is the development of 3D content format and related technologies which offers fatigue free viewing with more than one person independently of the viewer's position.

3 Immersive Multimedia Experiences

Traditionally, research on multimedia has provided information to the user primarily through just two sensory modalities - sound and vision. Hence up to now, research on different aspects involving multimedia from coding, to transmission to evaluation of the quality of experience and interaction has mainly been focused on these topics. Although there is a growing interest in other types of sensory interaction such as haptics, sound and in general more immersive experiences of multimedia, the topic is still under researched. This statement is even more valid for research on aspects related to the integration, display and transmission of multisensory information enriching the multimedia experience.

Research on immersive multi-sensory environments has proven that taking into account multisensory data such as vibro-sensory (e.g. floor vibration) and low-frequency subsonic effects could improve distance communication in applications like remote music performance and telemedicine. However, still no advanced strategies for data compression and transmission of these alternatives modalities have been adopted.

Other research is focusing on alternative and novel ways for interacting with media, e.g. through tangible interactions and use of natural elements.

Regarding the development of new user interfaces, recent research on immersive displays has made relevant progress related to instruments for surround-view and high definition cinema for designing more engaging immersive experiences. Other approaches of immersion consider a shift from highly structured settings of interfaces (like virtual reality and large screens) to more portable ones such as mobile devices augmenting the real world. An important research aspect in all these new interactive systems is pushing interaction designers to exploit the intrinsic characteristics of immersiveness to improve intuitiveness in interaction to enhance user experiences.

Nowadays, the research on immersive interactions has provided a series of applications at the level of the state-of-the-art with high potential of providing social benefit to users and better life quality. Examples of these are in the field of therapeutic systems and rehabilitation, exertion games, connectedness, locative media for education and digital TV among others. To not underestimate is the important issue that is rising among social scientist regarding immersive environments is the perception and relevance of aesthetics in such environments or more critical, the social implications of these new applications amongst which potential addictions to technology-based experiences.

Other relevant work relies on the concept of Social Immersive Media. The notion of this concept has been formalized recently5 in the context of multimedia immersive museum exhibits, by extrapolating a trend enabled by advances in the technology of interacting games, direct manipulation interfaces, and interactive arts.

3.1 Future Research Challenges

We live in a multimedia world. Users are increasingly looking for new multimedia experiences: new ways to capture, share and consume their multimedia content. In recent times, both research and industry has focused on designing technology that could enable immersive experiences with multimedia content, especially suited for home environments. These environments might enable the users to interact with multimedia content even when they are not co-located.

Despite many years of research on Media Spaces, we are still far from developing technologies that would allow people to virtually interact at distance with the same efficiency and ease than when face-to-face. Ethnographical observations from real work settings show that many solutions developed to support collaborative interactions at a distance are flawed as they “fracture the relation between action and the relevant environment”. For example, using many video cameras to capture and share different points of view between two remote locations might seem to be an improvement over the use of a single camera. However, users might feel lost in the attempt to understand which view is the partner currently looking at or how to adapt common communication strategies to this multitude of perspectives. We need to find more subtle technological solutions to translate communication mechanisms which are effective in presence but not available when conversational partners are not co-located. These solutions should allow recreating the same functions using different but equivalent strategies. We highlight a number of these mechanisms that might potentially enable unexplored interaction capabilities in media spaces.

Gesture recognition

It is well known that gestures complement verbal interactions and help humans to disambiguate references used during the interaction (e.g., discussing blood test reports from different patients) or to better support comparisons between various information media (e.g., combining a broken leg x-ray result with a plastic leg miniature so that the physician can point specific articulations in the former and manipulate the latter while explaining the injury cause). Supporting and developing new approaches to gesture recognition is of primary importance for immersive multimedia experiences.

Advanced Gesture recognition also helps in enhancing the user QoE.

Modeling the focus of attention

One of the challenges that designers of interactive immersive environments are constantly facing is the detection of the users' focus of attention. This roughly corresponds to the point in the shared workspace that is currently looked at by the user and/or the objects s/he is interacting with.

Detecting this element is extremely complicated. However, this information is extremely valuable as it can help designers design better support systems for the users' interactions in the system. Future immersive multimedia spaces will benefit from sensors and models able to recognize the users' focus of attention. In this context, the eye contact could be advantageous, where most of the existing means for remote interaction (e.g., video conferencing, web cams, etc.) do not allow the eye contact, where each user should focus on the camera, which does resemble the nature communication.

Context detection and modeling

Future Internet Media devices should react in a dynamic fashion to changes in the user's situation, for instance switch automatically from visually displaying text to reading the text out loud when the user starts another activity that keeps him/her from looking, yet not from listening. Time-saving applications would be very helpful as well. Any modality of content (including 3D, haptic content, games, etc.) and its delivery should be intelligently substitutable at any time, for reasons of convenience, time-saving, filtering for relevance, or improved understanding - not to mention bandwidth bottlenecks or the benefits of people with disabilities.

3D body reconstruction

Advancements in the field of video analysis might allow in the near future to design immersive environments where the 3D model of the body of the user is fully reconstructed from the video feed of several cameras available in the environment. From this information more elaborated task models can be derived and used to allow forms of interactions with computational devices and multimedia content. With the development of 3D human body scanners, it becomes possible to generate hundreds of accurate body measurements as well as an accurate 3D shape (without skeletal knowledge) in the form of point clouds from a specific subject. From this information a complete pipeline will allow capturing the shape of real people with parameterization techniques for modeling to animation. Furthermore, comprehensive biomechanical models can offer accurate mechanical body deformability through an accurate, however

complex representation of the biologic materials below the skin (muscles, fat tissues, bones) and how these interact between each other.

Integration of digital and physical space

In media spaces, humans tend to interact with both digital and physical artifacts. However, connections between these two realms have been limited so far. One of the challenges for the future of these environments will be to build better support and easy translation between "bits and atoms".

Furthermore, the goal would be to not only recreate and emulate face-to-face setting but to design interaction capabilities that can be augmented by technology enabling forms of interactions that are not possible in standard co-located interactions.

Persuasive multimedia experiences

Immersive multimedia environments might not only provide new ways of creating and consuming media for entertainment or work purposes. Combined with persuasive technologies, these environments might stimulate users towards a positive change in their behavior (e.g., embracing a more active lifestyle, etc.).

Multisensory immersive experiences

Today immersive systems are mostly focused on audio-visual applications while other sensorial modalities are still largely uncovered. While for audio and video there exist extensive studies employing models of human perception (e.g. for advanced compression, interaction, etc.), multisensory data is only marginally considered.

Social Immersive Multimedia (SIM)

SIM is a form of computer generated/mediated augmented reality that focuses on the aspects of social (interpersonal) interactions and aims at overcoming the still dominant GUI (Graphic User Interface) metaphor in multimedia. SIM aims at providing immersive experiences (including for example: communication, education, entertaining, training, and socialization) to users by means of body controlled interactivity, using lightweight (even invisible) sensing infrastructures, rather than cumbersome wearable technologies.

Ultimately the goal is to design user interactive behavior so to attain highly effective and engaging experiences at a social level. SIM is an emerging research area in the Human Computer Interaction (HCI) research community whose value and impact (economic and social) will be boosted by its incorporation in the framework of the Future Internet Multimedia. In particular the open, scalable, ubiquitous Future Media Internet infrastructure coupled with Social Immersive Multimedia will open the way to the formation of a new generation of Internet based services for consumers and of user-centric social networks.

4 Multimedia, Multimodal and Deformable Object Search

Multimedia content, which is available over the Internet, is increasing at a rate faster than the respective increase of computational power and storage capabilities. Internet capacity will approach the amount of yota (1024) bytes in 2010. Such a tremendous

amount of content cannot be processed and indexed by the current computational power unless personalised and user-centric mechanisms are implemented so that only the content of interest is delivered to the end-users.

This growth of popularity of media is not accompanied by the rapid development of media search technologies and the existing solutions still lack several important features, which could guarantee high-quality search services and improved end-user experience. These features are listed below:

Multimodal content search and retrieval in a unified manner

Currently, information is perceived, stored and processed in various forms leading to vast amounts of heterogeneous multimodal data (ranging from pure audiovisual data, to fully enriched media information associating also data originating from real world sensors monitoring the environment, etc.). User perception and interpretation of the information is in most cases on a conceptual level, independently of the form this content is available. Assuming the availability of an optimal, user-centric, search and retrieval engine, when users search for content they should be able to:

- express their query in any form most suitable for them;
- retrieve content in various forms providing the user with a complete view of the retrieved information;
- interact with the content using the most suitable modality for the particular user and under the specific context each time.

Sophisticated mechanisms for interaction with content

Secondly, social and collaborative behaviour of users interacting with the content should be exploited at best, which will enable them to better express what they want to retrieve. We need to pay a great research attention on increasing the content utilization efficiency, measured as the fraction of the relevant delivered content (i.e., content which satisfies their information needs and preferences) over the total amount of delivered content.

Towards the direction of delivering relevant multimedia content to users, another barrier that needs to be overcome is that the vast amount of information is not actually annotated.

Efficient presentation of the retrieved results

Search results suffer in most cases from the sequential presentation and the information overload, i.e. the presentation of huge amounts of information which is in most cases irrelevant to the query or not optimally presented to the user. This becomes significantly more important when information search and retrieval is performed from mobile or notebook devices with limited presentation capabilities. Appropriate filtering of the retrieved results is needed combined also with advanced visualisation techniques to compress information utilising the visual space and novel interfaces for fast and easy information access, based on context aware information, such as location and device specific performance indicators.

Efficient methods for deformable objects search and retrieval

In most of the multimedia object retrieval approaches presented so far, search is performed by using as query a similar object. Low-level feature extraction methods are applied to the object, providing a description of the object's global shape. These methods are not sufficient for Future Media Internet applications for three main reasons:

- An input multimedia object is not always available and it cannot be created from scratch by a non-expert user. On the other hand, using as query an image/ video or a hand-drawn sketch is more convenient (e.g. a 2D photo can be taken from the mobile device's digital camera; a sketch can be drawn by using a PDA touch screen interface, a low-quality audio excerpt can be recorded from the mobile phone, etc.).
- Neither partial matching nor articulation invariance is supported by methods applied to the (3D) objects' global shape. Articulation invariance is essential for applications where deformable objects exist (usually fashion applications).
- Presently introduced multimedia search engines commonly do not allow the user to form complex and multi-modal queries. While the multimedia object retrieval is becoming popular is of utmost importance to allow the user to formulate query which consist not only of a multimedia object, but at the same time –a textual or voice description.

Methods for measuring Quality of Experience

The introduction of more complex search methods than the users are used to create a new problem – the user perceived satisfaction and the method of its measurement.
When issuing a multimedia query it is hard to judge whether the result is relevant, and, furthermore whether the user will be satisfied with such result. The situation encountered here is much more complex than with traditional textual queries, as the opinion regarding the relevance of the query result will vary from user to user depending on the query context.

The information on the level of satisfaction of the user with a search service is critical, both during the development and deployment stages. A common methodology and toolbox for measurement of the Quality of Experience of search services would be also an aid for the academic community allowing for comparison of results achieved by different search engines.

4.1 Future Research Challenges

Defining the next generation of media search technologies requires major research efforts in multimedia, multimodal and deformable objects search. Indeed, the continuous and rapid growth of multimedia content available on the Internet has not been accompanied by a similar development of efficient, multimodal and intuitive cross media search capabilities.

Despite the significant achievements, the current technology suffers from a number of strong limitations which prevent the user to efficiently access the desired (and theoretically available) information. This leads to several research challenges which are grouped below in three interlaced categories:

Towards truly multimodal search

A first Research Challenge in this domain consists of the expansion of current schemes to truly multimodal capabilities (e.g. exploiting all information of all available modalities when searching for media content). When current search engines are mostly limited to text queries, future search engines will need to use richer and more diverse sources of information including combined data from speech, audio, video, images, 3D, social tags (either automatically or user generated), physiological signals providing information of the emotional state or activity of the user (heart beat, brain waves, etc) or geolocation information. An important aspect of this research challenge is the possibility of performing cross-modality search such as for example audio-based video retrieval (a typical example of the latter would be to retrieve a video –e.g. sequence of images - that semantically matches a given audio content such as music).

Other key aspects in multimodal search include the management of diversity and uncertainty in search to achieve richer and more personalized results, and the use of contextual information as a way to create a common ground to model the relationships among different content sources.

This challenge calls for new paradigms for content-based high-level (or semantic) signal representations that would permit cross modal navigation. It also appears obvious that major challenges remain in the field of machine learning and in particular with respect to multiple heterogeneous signals fusion, content and context adaptation using limited training data, and audio source separation. Finally, it is necessary to increase the quality of search by eliminating or grouping all multimedia content that does not add extra information to the user (because it is duplicated or of less quality than existing content) and implementing personalized search systems.

Towards search of multimodal and deformable content

Another important research challenge will target the access of complex, possibly multimodal and deformable information from a rudimentary query.

Indeed, current search experiences may be very tedious and especially in the case where it is not straightforward to describe the searched information with simple text queries. It is, here, highly desirable to be able to define a query by associating rudimentary and imperfect – possibly cross modal- descriptions of the searched information (e.g. search from hand drawn sketches, query by humming or low quality audio recordings, search for multiple videos of the same scene - but taken from different angles - search of 3D objects from 2D views or simple drawings or search of 2D projections of an object from the relevant 3D description, possibly considering 3D deformation models). This represents major theoretic research challenges in multi-level hierarchical content representation, in complex multimodal scene analysis and in decomposition models on known or unknown dictionaries or object bases.

Towards efficient user interaction

The third research challenge directly targets the development of new paradigms for *user interaction* (and satisfaction) in a multimodal context. It is, indeed, essential that future generation search tools can propose intuitive and rich query interfaces (multimodal, multi-level – that range from a full picture or only sketch – and intuitive

– e.g. by means of, for example, natural language), efficient presentation of the retrieved results and finally efficient means for user interaction with the retrieved information for successive searches. This includes the possibility of multimodal navigation across different media, for instance, jumping from a piece of soundtrack to the relevant movie, and then to other movies involving the same actor or to the e-book version of the book from which the movie originated.

It is critical to develop methods and tools for assessment of the quality of experience of the user. These methods were developed and standardized for voice and video services – now need to be moved forward to 3D content and multimedia services. This will put the user in the center of attention and will forge a true user-centric media environment.

This represents major challenges for the characterization of mono-modal and cross modal media similarity and its use in the selection of the retrieved information. Other research challenges of the same kind include the presentation of the retrieved information on heterogeneous devices and better user-feedback exploitation for multi query search in a fully multimodal context.

5 Content with Memory and Behaviour

In the area of 3D, virtual worlds and gaming, advances are needed to increase the level of realism and interactivity. By adding to virtual characters and virtual objects memory and behavior will lead to the transition from the “smart content” to the “intelligent content”. The positioning and the advantages of the intelligent content may be seen in the following figure

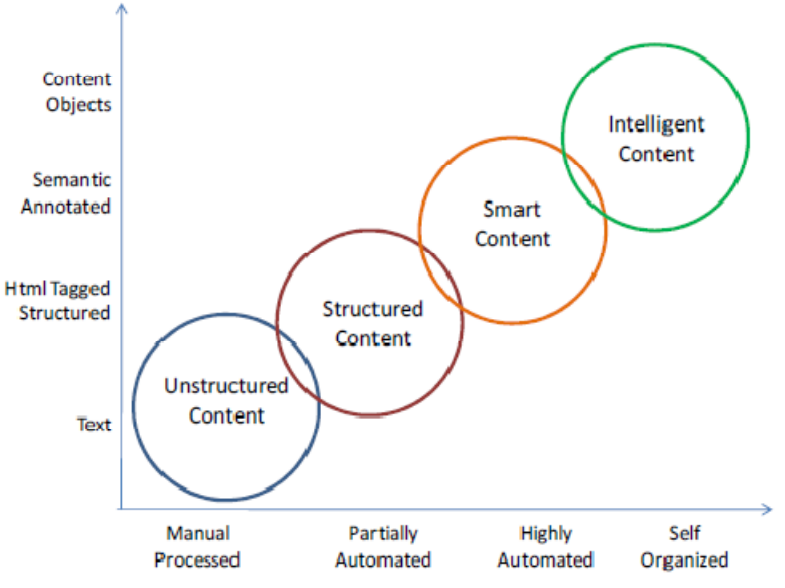

Fig. 4. Intelligent Content Positioning

We define as *Content Objects* polymorphic/holistic containers, which may consist of media, rules, behaviour, relations and characteristics or any combination of the above. *Media* can be anything that a human can perceive/experience with his/her senses, c*haracteristics* are meaningfully descriptions of the object, *Rules* refer to the way an object is treated and manipulated by other objects or the environment (discovered, retrieved, casted, adapted, delivered, transformed, presented), *behaviour* can refer to the way the object affects other objects or the environment, *relations* between an object with other objects can refer to time, space and synchronisation issues. As can be seen the integration of memory and behavior to content objects can in some sense be said to be the equivalent of personality for virtual characters. The behavior rules their interactions, while memory provides them with the capacity to remember past interactions and to learn from their experiences. As a result these virtual characters can interact socially and serve as companions for people and even potentially amongst each other. While the level of physical realism of these characters has vastly improved (improved movement, more realistic facial gestures, etc.), the biggest challenge not only making them look like humans but make them behave as natural as humans for which they must have social and cognitive intelligence, emotions and memory. To date this level of realism has not been achieved in any virtual worlds or games. The only characters which, today, have such a high level of social and cognitive skills are those which have a human being behind them. But providing emotions and personality for virtual characters isn't just a challenge for virtual worlds and gaming – it can have real world benefits in areas such as education and health care.

5.1 Future Research Challenges

Future Media must build on the new capabilities offered by new web technologies to provide an improved user experience. Content will adapt to user context and purpose. Such content exposes 'a behaviour'. It remembers, reacts, interacts and thereby actually becomes bi-directionally immersive, i.e. immerses the user as well as immerses itself into the user's environment.

Recent years brought us neologisms such as blobjects, blogjects, tweetjects and the Spime. All these terms denote a kind of object which is capable to converse with its environment – the real one as well as the virtual one. This is not just the Internet of things. Today, the Simple Object Access Protocol (SOAP) talks to systems, but in the future we will access objects – no matter what systems manage those objects or where they are managed.

Future Media will be composed of such autonomous content objects or content object mash-ups. The autonomous objects will *travel over the network, split and combine* to generate the new service or a 'virtual world object'.

The above ideas include a number of research challenges that have already been described. For example media encoding, media search etc. Yet, there is a number of research areas related to the content objects themselves:

Ontologies and Semantic Description

This set of challenge includes research on the modeling and semantic description of the content objects as media containers, along with descriptions of their structure, characteristics, behavior, rules and relations.

Decomposition and Reconstruction

This set of challenges includes research on the decomposition of scenes and scenarios, along with real-time component-preserving capturing technologies: advanced capturing systems using stereoscopic cameras, camera arrays, time-of-flight cameras or 3D scanners preserve three-dimensional information and thereby the components and their spatio-temporal relationships. Moreover, assembly and reconstruction of complex scenes and scenarios as content objects mash-ups.

Network Support for Content Objects

This set of challenges includes research on the network architecture (including distributed network intelligence, network topology and traffic awareness, content distribution including caches), and issues like routing and streaming of the content objects.

6 New Application Areas

The influence of the Internet in today's interpersonal communication and interaction with information has reached a level never imagined by past generations. The penetration of Internet services and applications has reached a stage that is starting to have a profound impact on different dimensions of people's lifestyles, including their everyday habits. The advances in Future Media technologies will enable innovative applications and services, engaging new experiences, where multimedia digital content will be more immersive and closely related to the physical world. In fact, the Future Internet will allow a new generation of online ubiquitous applications sustained by new enablers of the Internet amongst which the Internet of Services, the Internet of Things, the Internet of Media & Content and the Social Internet (see figure below).

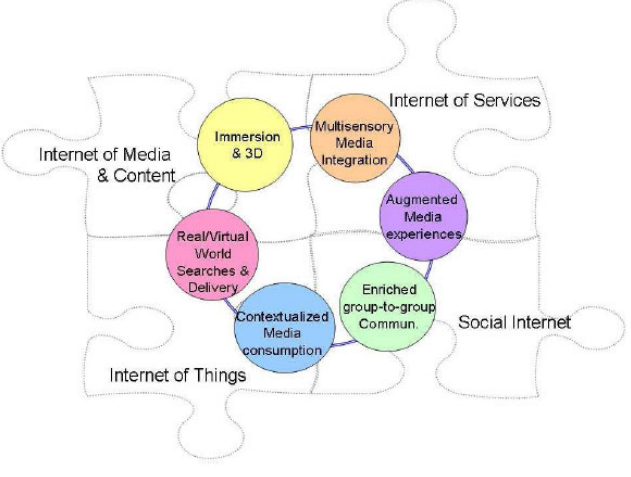

Fig. 5. New Dimensions of Applications Enablers of Future Internet

On top of these new dimensions of the Internet, the applications that will be created will allow users to enhance human-media and human-human communications as follows:

Immersive and 3D Applications

The envisioned new applications in this field will include revolutionary ways of interacting with media through sophisticated representations of real and virtual worlds. Among the kind of applications enabled by immersive media it is expected the appearance of new physically strong experiences such as exertion games, simulation and training of real life situations in realistic way (e.g. chirurgical interventions), new entertainment experiences (e.g. tele-reality concerts), and personal enriched communications (e.g. sophisticated tele-presence).

Multisensory Media Integration

The Future Internet will enable applications dealing not only with data and information but also will enhance the perception of digital contents by exploring new ways of impacting Human perception. In fact the Future Internet will include a new set of revolutionary applications focusing on enriching Human sensations. This will be achieved by stimulating simultaneously the different senses beyond audiovisual content by including haptics or smell.

Augmented Media Experiences

A set of new applications of the Future Internet is envisioned to impact the way people interact with the physical world. In fact, there are expected applications dealing with augmented media that will enrich elements in the physical world with digital content in a way that will modify the conception of reality. In such context, a new kind of reality will be conceived as the sum of physical and augmented information that will go beyond early augmented reality prototypic developments into a well established discipline rich of applications and services.

Enriched Group to Group Communications

Traditional one-to-one and group-to-group communications will be transformed in the context of the Future Internet. The kind of applications envisioned in this concern, will include synchronous and asynchronous interactions including multi-channel sources to enhance the way people communicate.

Typical applications and services in this domain will include engaging of distant family members into end-to-end games and collaborative work with remotely located teams, involving verbal and nonverbal communications and interaction with tangible devices and objects.

Contextualized Media Consumptions

Future Media will enable applications where information not only will be available at any time but moreover its consumption will be adaptive in order to be selective in terms of when, where and how to present it.

This will imply a level of intelligence embedded in media and in its composition in order that users will be able to access content in its best possible way and in the right

moment when it is needed. This will allow for applications contextualized according to personal profiles, location, types of media, available devices, resources and QoS needs.

Real-Virtual Worlds Searches and Delivery

The Future Internet will include a rich set of applications dealing with real, virtual and mixed information and consequently its identification and retrieval. The envisioned applications in this respect will include searching facilities allowing for cross modal search in both, input and output. This means that the input mechanisms enabled by search applications will allow not only textual queries but also multimodal ones while the output will be given also in combined modalities. Moreover, the searches will not be limited to the virtual world but will retrieve information regarding physical objects, their location, their state and their relationship and enrichment with digital contents.

7 Conclusion

This paper has presented a series of new research trends for the design and study of Future User-Centric Media technology. As a result of previous discuss sessions among experts coordinated by the FMI-TF, research challenges were identified and highlighted with the goal of determining the current trend and future perspectives of the research community on the Future Media technology, with emphasis on their impact and potential application in the area of Ambient Systems and Media.

References

1. Nokia Research Center, Extended Home Team, http://research.nokia.com/extendedhome/
2. Orb, http://www.orb.com/
3. IST Amigo Project, Deliverable D2.2: State of the art analysis including assessment of system architectures for ambient intelligence (April 2005)
4. IST ePerSpace Project, Report of State of the Art in Personalisation. Common Framework (February 2004)
5. IST ePerSpace Project, D4.1: Content Management Architecture design (November 2004)
6. CONTENT D2.1 Deliverable CONTENT Research Vision and Roadmap (March 2007), http://www.istcontent.eu