VISNET: NOE on NETWORKED AUDIOVISUAL MEDIA TECHNOLOGIES

A. H. Sadka (VISNET Coordinator)

Centre for Communication Systems Research, University of Surrey, Guildford GU2 7XH, UK
Tel: +44 (01483) 689490; Fax: +44 (01483) 686011; Email: a.sadka@eim.surrey.ac.uk

ABSTRACT

VISNET is a network of excellence (NoE) retained for funding from call 1 of the 6th framework of the IST programme. The network [1] brings together 15 leading European institutions in the field of audiovisual (AV) media processing and networking technologies and integrates over 100 researchers covering a wide spectrum of AV media technologies. Visnet has a clear plan for integration as well as a vision for durability and sustainability following the 5-year duration of EC funding. Visnet was ranked the best NoE submission in its strategic objective 2.3.1.8 on networked audiovisual systems. Following the successful negotiations and contract preparation stages, Visnet was launched at the beginning of December 2003 to be the first FP6 call 1 project to inaugurate in the whole of the DG directorate. This paper addresses the major technical issues that Visnet will tackle in the first 2 years of its lifetime by describing major elements of the joint programme of activities (JPA).

1. INTRODUCTION

The main objective [2] of action line 2.3.1.8 on networked AV systems and home platforms as defined in the 2003/4 workprogramme was to develop end-to-end networked AV systems and applications, and open trusted and interoperable multimedia user platforms and devices, notably for broadcasting and in-home platforms with full interactivity capacity. In response to this preset strategic objective, Visnet has put forward a corresponding technical objective around which its proposed JPA revolved. The technical objective of Visnet is defined as the creation and emergence of new technologies aiming to improve the status quo of research and development in the following disciplines that are related to networked AV systems:

- AV processing: Covering creation, coding, storage and transport of AV content.
- Networking technologies for immersive home platforms and mobile systems.
- QoS control for AV transmission over heterogeneous networks.
- Content and network security.

Visnet partners have an internationally proven track record and a mass expertise in all areas related to networked audiovisual systems. In particular, many Visnet partners have been instrumental in the development of international standards for audio/video compression, networking protocols and multimedia communication systems. Many partners have also made significant contributions to the advance of this field of technology through prestigious publications, world patents and successful ventures aiming at technology transfer. The synergy of partners working directly for the release of international standards and production of valuable technological output puts the VISNET consortium in a key position to influence the future of audio-visual media technologies.

2. SCOPE OF RESEARCH

The research activities of Visnet aim at consolidating the excellence of European institutions in the field of networked AV media technologies. Therefore, the focus of the research activities will be on developing intelligent media processing techniques that are capable of creating fully networked virtual and augmented reality environments e.g. home platforms, with high level of interactivity and immersion. The scope of research in Visnet has been set in order to enable and facilitate the following technical achievements:

- Shaping the future of research in networked AV systems and home platforms and the creation of a strong research platform for the execution of world-class research in this rather fertile field of technology.

- Creation of new media-rich and interactive AV applications and services to the European citizen in home platforms and to the mobile user.

- Ensuring a sustainable excellence in scientific research in the areas defined by the scope of action line 2.3.1.8 on networked AV systems on both the short and long terms. The focus of research activities will be on the following technical areas:

    - Creation and coding of AV content for immersive platforms: This includes the process of capturing and acquisition of AV material suitable for the creation of media-rich 3D stereographic AV scenes and the creation of immersive 3D home platform scenes. Coding includes more importantly researching new and efficient methods for compression and scalability of created content for storage and transmission purposes. Adaptive transcoding will also be investigated for inter-network communications. Coding also involves joint source and channel coding of complex AV scenes in addition to compression of 3D audio and video material.

    - Storage and transport of compressed AV information over heterogeneous networks: This technical area is concerned with all the technologies aimed at the provision of end-to-end QoS control for the transmission of AV information over heterogeneous networks. This includes media transcoding, rate control and error resilience algorithms in addition to traffic modelling for optimal resource management and allocation. This area also includes the development of fast and efficient metadata generation techniques for easy access and search of AV databases.

    - AV processing techniques for immersive communications: This area is concerned with the development of automatic scene analysis techniques such as segmentation, AV recognition, feature detection and object tracking in addition to interactivity (inter-user, user-object) and interfacing techniques for mixed/augmented reality systems.

    - Security: The activities focus on the security of both the content and the network. This includes securing the integrity of data, confidentiality and digital rights management in addition to adaptive security functions for optimal QoS control in networked AV services.

3. RESEARCH PROGRAMME

The joint research activities are organized in a number of Workpackages that deal with the different stages of the content value chain. The JPA for the first 2 years of EC-funded period consists of 8 main workpackages WP0-WP7 where WP0, WP6 and WP7 are for integration, dissemination and management activities respectively. WP1-WP5 are the research WP's and consist of sub-WP's that span across the whole spectrum of research covered by the network. A description of the research activities planned for each workpackage is given below:

WP1 deals with the issue of content creation for AV home platforms. This WP consists of 2 sub-WP's, namely WP1.1 and WP1.2. In WP1.1, topics investigated include the automatic modelling of natural 3D objects using mono, stereo, and multiple camera views. Moreover, the work will look into methods for the creation of high-quality, photo-realistic and accurate 3D models of manufactured objects. Voice models and audio recordings will be used in the modelling of natural audio objects. This WP will also look into audiovisual synchronization techniques aiming at synchronizing the audio objects with their corresponding visual objects. These synchronization tools will also be used for the synthesis of realistic AV scenes. WP1.2 will be concerned with the implementation of a hybrid AV scene rendering engine capable of merging natural AV objects with synthetic AV objects in a realistic and immersive way. The engine will make use of hybrid (natural/synthetic) camera and microphone models, image based rendering techniques and will be able to provide 3D sound with efficient spatialization of sound sources. The rendering engine will also be complemented by an animation and interaction control engine that will provide synthetic object control and facilitate authoring of hybrid immersive AV scenes.

Content coding and adaptation are two important aspects of every AV network. Both aspects are covered in WP2 which is divided into 2 sub-WP's, namely WP2.1 and WP2.2. In WP2.1, The coding of AV material will include the development of object-based scalable and error resilient methods aiming at optimizing the perceptual quality through the efficient utilization of available resources. Joint source-channel coding will also be used with fine granular scalability (FGS). Adaptive error

control will be investigated with view to improve the performance of H.264/AVC under error-prone conditions. WP2.1 will also focus on the scalable and error-resilient compression of 3D audiovisual content. Research will also explore the use of metadata information in the encoding of still pictures and video sequences in order to improve the coding efficiency of standard compression algorithms. WP2.2 addresses content adaptation for inter-network AV communications. In this WP, compressed-domain downscaling techniques for object-based video data will be developed aiming at scalability of shape, motion and texture data. New transcoding algorithms will be developed with view to convert object-based AV data from DVB quality (MPEG-2) to 3G quality (e.g. MPEG-4, H.264). This activity will also look into efficient techniques for 2D/3D conversion so that the AV media will be transported and displayed in line with the user's requirements, preferences and capabilities.

The problems arising during the transmission of AV data will be considered in WP3.1. The main focus will be on optimizing end-to-end QoS of AV transmissions over heterogeneous networks with the use of adaptation techniques and gateways. Network adaptation layers for specific codecs will be implemented with view to adapt the encoding and system parameters to the monitored network conditions. The content adaptation techniques developed in WP2.2 will be a vital input to the work of this WP. WP3.2 will develop efficient storage, indexing and search techniques that facilitate user-content interaction as well as interoperability between data and metadata formats. These techniques will aim at presenting content to the end-user in an intelligent way that allows easy search and accessibility.

In WP4, new tools for AV content analysis will be developed. This WP consists of 4 sub-WP's each having a slightly different focus. Research on speech and audio analysis tools will be carried out within WP4.1. WP4.2.1 will deal with topics such as probabilistic frameworks for face detection and tracking, facial expression recognition and derivation of facial animation parameters from video data, facial feature (eyes) tracking. Techniques for detection of video shots that depict people (which can be a useful preprocessing stage for face and body analysis tools) will be also considered. WP4.2.2 will deal with segmentation and tracking of generic objects and human body from natural video sequences. A robust video analysis system for supervised and unsupervised segmentation and tracking of generic objects that will integrate a number of advanced techniques will be developed. Model based techniques for articulated human body tracking in two and three dimensions from monocular video sequences will be also derived. WP4.3 will develop a set of multimodal (joint audio and video) analysis tools, with special emphasis on the detection and recognition of people in audiovisual sequences. Tasks in this workpackage include detection of shots where a person is speaking, using both audio and video data and joint audiovisual detection and recognition of persons.

WP5 is concerned with the challenging issue of managing and protecting intellectual property rights in digital AV data by combining traditional methods (e.g. encryption, scrambling, user authentication) with techniques used to protect the content "in the clear" i.e. watermarking and fingerprinting. New methods for watermarking of AV data (including joint watermarking of audio and video data and fusion of detection results) that are robust to geometric and collusion attacks will be proposed. Implementation of new methods for robust fingerprint extraction on AV data along with efficient fingerprint matching strategies and appropriately designed fingerprint databases will be pursued. The design of new asymmetric watermarking techniques, and protocols for anonymous purchase based on asymmetric watermarks will also be among the aims of this WP which will additionally deal with the specification of benchmarking procedures for audio & video watermarking and fingerprinting techniques.

4. INTEGRATION AND DISSEMINATION

Obviously, efficient integration that will lead to the creation of a world-recognized workforce in the area of networked AV systems would not be possible only by conducting joint research activities. Additional actions that seek to further enhance the quality of integration and strengthen the ties among Visnet partners form an important part of the JPA. Recognizing the importance of efficient integration mechanisms, especially at the beginning of the NoE's lifecycle, Visnet has included in its JPA a special workpackage, namely WP0, that is devoted only to integration activities. These include holding internal workshops and training programmes, exchange of research personnel at various levels, sharing of research material and resources in addition to the creation of research integration tools such as common software platforms.

Furthermore, since the aim of Visnet is to create a knowledge base that will be maximally diffused in order to help enahnce the status of science and technology in Europe, WP6 has been set in the JPA as a vehicle towards achieving this aim. WP6 consists of a number of activities including external short courses and training programmes, joint publications, a number of exhibitions and

demonstrations etc that were specifically designed for spreading the accruing results and ensuring a maximally large dissemination zone.

5. MANAGEMENT AND RESPONSIBILITIES

In order to ensure a smooth progress of the network and guarantee successful integration, an efficient management structure that oversees the activities of the network, takes the appropriate strategic and day-to-day decisions required and proceeds to appropriate corrective actions when needed has been set up in Visnet. A multi-layer management hierarchy was established with view to monitor the progress of the network activities at various levels of coordination and administration. Visnet highest management layer consists of the Supervisory Committee which is the governing body of the network. This Committee sets the visionary policy of the network and makes decisions on matters of strategic importance. It also liaises with the NoE executive board whose main responsibility is the operational direction of the NoE. An integration executive is appointed with the key responsibility of ensuring that Visnet produces the required products, to the required quality, and within the specified constraints of time and costs. An integration executive looks after and coordinates the integration activities of the network from a financial as well as process point of view. He also assesses the quality of integration within Visnet and reports his findings to the SC. Efficient management, coordination and monitoring of the network will benefit from the introduction and use of a number of quantitative and qualitative indicators for scientific performance and integration. In case these indicators pinpoint to a certain dysfunction, latency or weakness, appropriate corrective actions will then be taken. Quality assurance teams will be formed to conduct regular assessment of the progress of Visnet activities and ensure that the network is being compliant with EC objectives by adhering to formal quality assurance standards.

6. CONCLUSIONS

Visnet is an FP6 network of excellence consisting of leading research institutions in Europe whose main strategic objective is to establish a world force in the area of networked audiovisual media technologies. For the first 2 years of the funded NoE lifetime, Visnet has a detailed JPA covering a wide spectrum of networked audiovisual media technologies aiming at making significant contributions to a sustainable advance of research in this field. This paper has shed the light on the scope of research in Visnet and some of the major research directions envisaged in the first 2 years of the network. In this paper, it was shown that in addition to its comprehensive research programme, Visnet has a feasible plan for the close integration of partners as well as a clear vision for dissemination and spreading of excellence to the largest possible audience worldwide.

7. REFERENCES

[1] ftp.cordis.lu/pub/ist/docs/b_pgnoe_200201_en.doc
[2] ftp.cordis.lu/pub/ist/docs/wp2003-04_final_en.pdf