The roll-out of DTT in France

— not just SD ... but HD and mobile TV services as well

Aurélien Louis and Matthieu Roger
Conseil supérieur de l'audiovisuel (CSA)

Despite the rather late launch of DTT services in France compared to some neighbouring countries, they are already being adopted widely by the viewers, thanks to a long tradition of terrestrial television reception. By increasing the number of services and the transmission quality, the digital television landscape is maintaining its rapid evolutionary pace.

This article gives a brief description of the current DTT situation in France and of its soon-to-come evolution: HDTV and mobile TV launches. In the longer term, terrestrial broadcasting networks will develop further, as a result of a national scheme to re-allocate the frequencies freed up by the digital switchover process.

Digital Terrestrial Television in France

Digital terrestrial television in France will soon be three years old, as the first services started broadcasting in 2005. Its roll-out brought immediate value for the French consumer: more channels with better quality pictures, at an affordable price (only the cost of an MPEG-2 set-top box) and trouble-free everyday use.

This technological breakthrough had however been in preparation for many years. The first call for tender organized by the *Conseil supérieur de l'audiovisuel* (CSA), the French independent broadcasting regulator, indeed took place in 2001. This followed many years of international discussions at the technical level, and intense regulatory and legal thinking.

An unusual regulatory framework for the main television broadcasting platform

The French regulatory framework for DTT is quite specific among the European member states and pertains to the French "*exception culturelle*".

On the one hand, the selection procedure applies to each TV channel editor, to whom the resource is allocated, and not to the platform, multiplex or technical operators. The goal behind this direct selection of each DTT broadcaster is to preserve the political, social and cultural pluralism of the French audiovisual landscape.

On the other hand, the selection is achieved through a "beauty contest" and leads to no financial charge for spectrum usage. This aims at maintaining high standards of content quality for the selected channels. Indeed, the selection of the broadcasters and the free allocation of spectrum enable the regulator to be more demanding regarding the applicants' broadcasting obligation and contribution to the production of European or French programmes.

On the technical side, the broadcasting and compression standards are defined through government decisions. Spectrum planning is under the responsibility of the CSA, in partnership with the *Agence Nationale des Fréquences* (ANFr), the French spectrum management agency that is responsible for coordination with neighbouring countries.

In 2007, terrestrial TV accounted for 57% of French television viewing and 18% of that was digital viewing. These figures show that terrestrial is the main television platform, which explains why such attention is being provided by the regulator to the diversity and quality of the terrestrial programming.

Planning and organization of a DTT multiplex

In France, as in the majority of the European countries, six digital multiplexes have been planned and coordinated, according to the Chester 97 agreement.

Selecting a broadcasting modulation scheme implies, on the one hand, a trade-off between the robustness of the transmissions and the cost of the infrastructure and, on the other hand, the data rate to be used. The decision taken in France was to maximize the number of channels while keeping a fair image quality. Those considerations led to the choice of 64-QAM modulation [1].

Taking into account this modulation, it was foreseen that the 113 main TV transmitters would provide DTT coverage to 85% of the population.

As far as the channel line-up is concerned, *Fig. 1* shows the present distribution of services across the multiplexes.

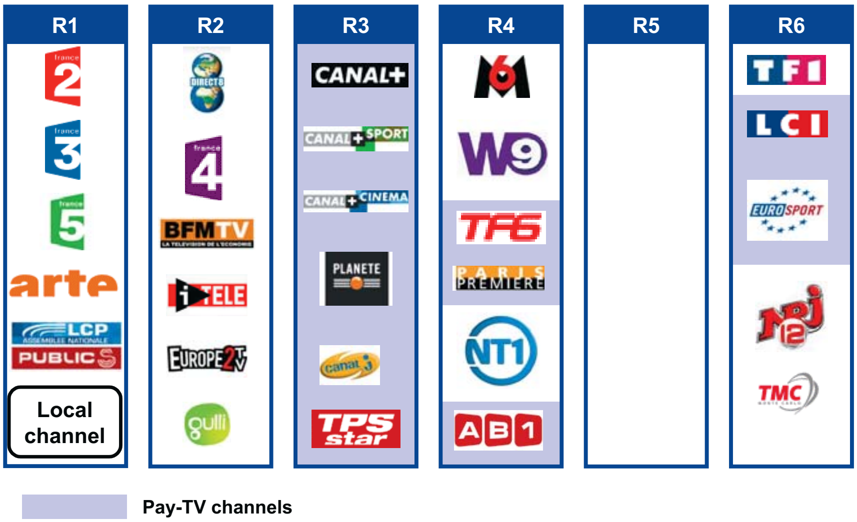


Figure 1
The six current DTT multiplexes in France

Multiplex R1 was specifically designed to enable the public broadcaster France 3, which operates local and regional programmes, to broadcast regional variations.

Although this situation implies additional constraints and, hence, a more extensive use of spectrum, those regional variations are quite popular and are important in regard to the role of digital television as a community link.

From the beginning, multiplex R5 has not carried any services, being dedicated to future or innovative uses. But as detailed below, it is now being allocated to the first free-to-air HD services on the terrestrial platform.

The debate on compression standards: MPEG-2 or MPEG-4

The French situation is also specific in the way that multiplexes use both MPEG-2 and MPEG-4 compression standards. Once again, the reasons for this technological choice is to be found in social and political orientations.

1. To be precise and complete, usually with FEC 2/3, guard interval 1/32 and carrier type 8k.

In 2004, a public debate was opened regarding the opportunity to launch DTT in MPEG-4 rather than in MPEG-2, taking into account the recent developments in MPEG-4.

MPEG-4 was an opportunity to derive benefits from compression gains in the future and was considered as a first step towards HD. However, it was still a very expensive technology at the time, with no significant efficiency gains over MPEG-2 in the next two years to come. Above all, the availability of multi-standard terminals (MPEG-4 SD and HD) was not secured at that time, which made the possibility of a seamless switch to HD unrealistic.

On the other hand, MPEG-2 was largely acknowledged as an affordable and available technology. It had been selected for the majority of DTT platforms around the world, leading to large economies of scale, particularly thanks to an earlier launch of the technology in neighbouring countries, such as the United Kingdom, seven years before. The downside of MPEG-2 was its maturity: no additional compression gains were to be expected, while MPEG-4 technology claimed to be able to produce a 50% improvement. Moreover, it implied a replacement of the set-top boxes deployed if switching to HD.

After a fierce debate, a hybrid choice was made by the government: MPEG-2 to be used for free-to-air services and MPEG-4 for pay-TV and HD services.

This unprecedented decision had a strong rationale behind it. By choosing MPEG-2 for the FTA services, a wide public accessibility to those services, at limited costs, was achievable and the launch of the free DTT services was expected to be on time. And indeed it was: on 31 March 2005, when set-top boxes retailed at around 100€, the 18 free DTT services were ready to broadcast, as scheduled.

The choice of MPEG-4 for pay-TV services was meant to facilitate the introduction of new services at a later date, thanks to future compression gains which would free up some spectrum used by the first pay-TV services. In that regard, this decision has also been a success. Since the 2005 MPEG-4 pay-TV launch, compression gains of 30% have been achieved, thus enabling the development of an HDTV terrestrial offer in France.

The second reason for this choice was the question of the MPEG-4 HD terminals: there was some hope that the pay-TV market could create a customer base ready to receive HD services, and also that this would help to decrease the costs of the MPEG-4 technology. However, the slow take-off of the DTT pay-TV market in France (a few hundreds of thousands of subscribers) has limited the size of the MPEG-4 user base. Nevertheless, the simultaneous deployment of HD-capable, DTT-compatible ADSL set-top boxes (combined DTT and IPTV) has significantly increased this installed base.

The success of DTT: coverage and take-up

At the end of 2007, DTT was available to 85% of the population of France, from the 113 main transmitting sites.

The roll-out of DTT has been progressive and organized in six phases. For the regulator, dealing with the huge spectrum re-engineering and refarming that such a development implied – and protecting the analogue terrestrial broadcasting networks from interference caused by the digital transmissions – has been a huge challenge. Indeed, the goal has been to deploy the six new national digital networks in addition to the existing national [2] analogue ones. To cope with the workload, the technical and planning capacity of the CSA has been strongly increased.

2. Three analogue terrestrial networks (TF1, France 2 and France 3) cover around 99% of the population in the UHF bands, and two other networks (France 5/Arte and M6) serve around 85% of the population, also in the UHF bands.

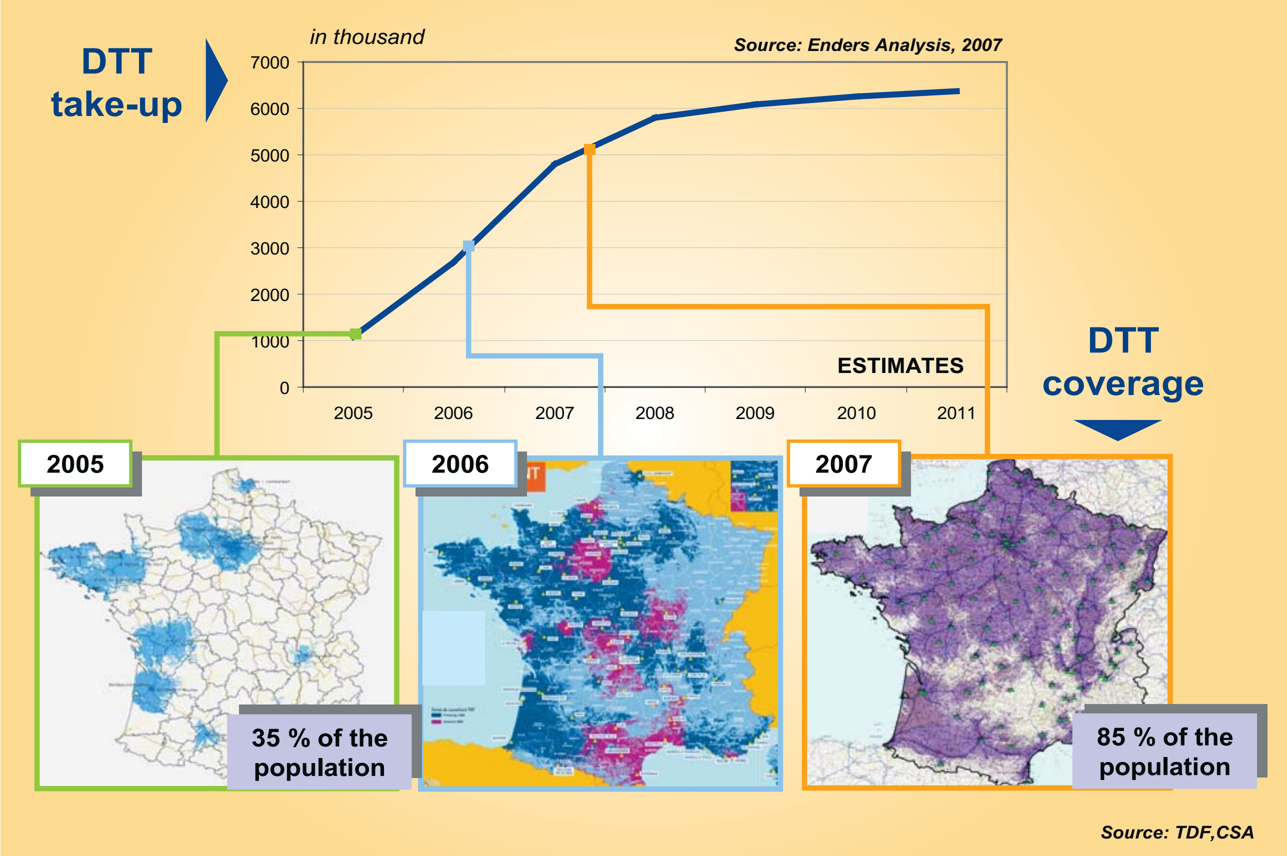


Figure 2
DTT take-up in France: 2005 to 2007

However, the latest audiovisual bill, voted on the 5th March 2007, now implies that the broadcasters should further extend the DTT coverage to at least 95% of the population for every "analogue incumbent channel".

In order to comply with this obligation, the CSA has paved the way for this extension by detailed obligations for the broadcasters, in order to prevent any digital divide. These complementary obligations aim at levelling the coverage of each *département* (French administrative sub-region), so that DTT coverage in geographically difficult regions with a low density of population keep in line with those benefiting from an easy coverage over large urban areas.

Those two obligations, national and "departmental" coverage, will represent an important workload in terms of the transmitters to switch on. It is generally agreed that more than 1500 new transmitters will be needed to match that coverage.

Finally, for the last 5% share of the population, satellite coverage is recommended. French law requires that a satellite bouquet, accessible without paying subscription fees or having to rent a terminal, should be put into operation. Today, such an offer has been provided by CanalSatellite with the service called TNTSat, which delivers the 18 free DTT services.

These three coverage extensions, although financially demanding for the broadcasters and implying an important planning workload, are conditional to the success of the DTT platform in the future. It will surely drive the DTT take-up and also increase the popular demand for new services, such as HDTV.

A strong expectation for HD content

Aside from the receiver costs and the rapidly expanding geographical coverage, the two main key drivers for the adoption of DTT by the public were certainly the improved image quality and the greater number of services. Eighteen free-to-air digital services against five analogue ones, as well as eleven pay services against one, contributed to a momentum that was supported by the crisp and sharp images provided by the digital transmissions. No more "snow" on the screen, and no more crackling sounds.

In parallel, the success of the DVD and CD against VHS and audio magnetic tapes has led to a growing expectation of quality. Home theatres, multichannel sound, audio effects, improving video games, and higher resolutions on personal computers have all contributed in the meantime to create a need for more quality.

HDTV therefore comes as the next step and is often considered as the future of television, in the same way that colour replaced black and white, several decades ago.

Existing HDTV offer over other bearers

French-speaking HDTV services are already proposed over satellite, cable and ADSL TV (IPTV). Some are free-to-air, but most are pay services, often billed as "options" in the bouquets, even when the corresponding SD services are free-to-air. In particular, the first free-to-air private French channels, such as M6 or TF1 which get most of their audiences from terrestrial broadcasting (whether analogue or digital), are never available for free when they are simulcast in HD.

Some new services, such as NRJ12 HD, are provided as part of the basic bouquet of some triple play offers. Looking a bit more like a free-to-air service, since no additional fee is needed, it nevertheless requires a subscription to an internet access provider.

Moreover, ADSL – unlike traditional broadcast networks – quickly bumps into bandwidth limitations and therefore HD accessibility over 5 Mbit/s, if proposed, is limited to just a few customers.

HDTV needs to be terrestrial

Taking into account the broad take-up of terrestrial reception for TV services in France, the introduction of HDTV over DTT is a must. Otherwise, this would limit the ability of most people to gain access to such services. Moreover, since the terrestrial platform is still leading the audience-driven market for advertisements, this is probably the only place where free-to-air HDTV can be first implemented and then developed: if HDTV is to be provided to every French citizen, then it must be terrestrial.

Abbreviations

64-QAM	64-state Quadrature Amplitude Modulation
AAC	Advanced Audio Coding
ADSL	Asymmetric Digital Subscriber Line
AVC	(MPEG-4) Advanced Video Coding
CAS	Conditional Access System
CSA	*Conseil supérieur de l'audiovisuel* (French broadcasting regulator)
CGTI	*Conseil général des technologies de l'information* (French Council for Information Technology)
DTT	Digital Terrestrial Television
DVB	Digital Video Broadcasting http://www.dvb.org/
DVB-H	DVB - Handheld
ESG	Electronic Service Guide
ETSI	European Telecommunication Standards Institute http://pda.etsi.org/pda/queryform.asp
FTA	Free-To-Air
HD	High-Definition
HE-AAC	High Efficiency AAC
IPDC	IP DataCasting
IPTV	Internet Protocol Televison
MPEG	Moving Picture Experts Group http://www.chiariglione.org/mpeg/
OMA	Open Mobile Alliance http://www.openmobilealliance.org/
SD	Standard-Definition
TNT	*Télévision Numérique Terrestre* (Digital Terrestrial Television, DTT)

But terrestrial broadcasting implies having to deal with some constraints. The spectrum scarcity is non-negotiable, particularly if the bands allocated to television services today are partly reduced in the future.

There is therefore a strong need to get more into less space. The MPEG-4 part 10 codec (also named H.264) matches this target well. Further work in the DVB Forum, such as on scalable video codec support or the DVB-T2 modulation system, could also help to squeeze in more. Unfortunately, this would probably delay the wide deployment of HDTV in France by some years.

Once the codec and the modulation scheme have been agreed, the next problem to be solved concerns the number of services per multiplex or the service bandwidths to be used.

Thanks to a report from the *Conseil général des technologies de l'information* (CGTI) – a department of the Ministry of Finance in charge of information technology – and also as a result of a consultation launched by the CSA in December 2006, it is considered feasible to broadcast three HDTV services in a multiplex, with a combined bitrate of 24 Mbit/s, with the help of statistical multiplexing (statmux).

Some HDTV experiments in Paris, Lyon and Marseille, which allowed the testing of different bitrates, were extended to more cities. These new trials were subject to a very stringent obligation – only native HD content could be broadcast (i.e. no upconverted HD programmes). As expected, the feedback from the people able to watch these events in HD (Rugby World Cup, *Tour de France*, Concerts, Series, etc.) was excellent. At the time, three services could be broadcast in a single multiplex.

HDTV but where and when?

One national multiplex, called R5, was identified to carry the HD services. In June 2007, a beauty contest was organized by the CSA to allocate this multiplex to two HD services, the third one being already reserved by the government for a public television service. Among the four candidates, M6 HD and TF1 HD – which are simply HD simulcasts of their SD channels – were selected. There are now ongoing negotiations to finalise their obligations before the actual launch can occur between now and the end of the year. The roll-out rate for extending coverage of this network, aiming finally to become national, will be set soon. The CSA is very keen to extend coverage quickly to 85% of the population as a minimum target.

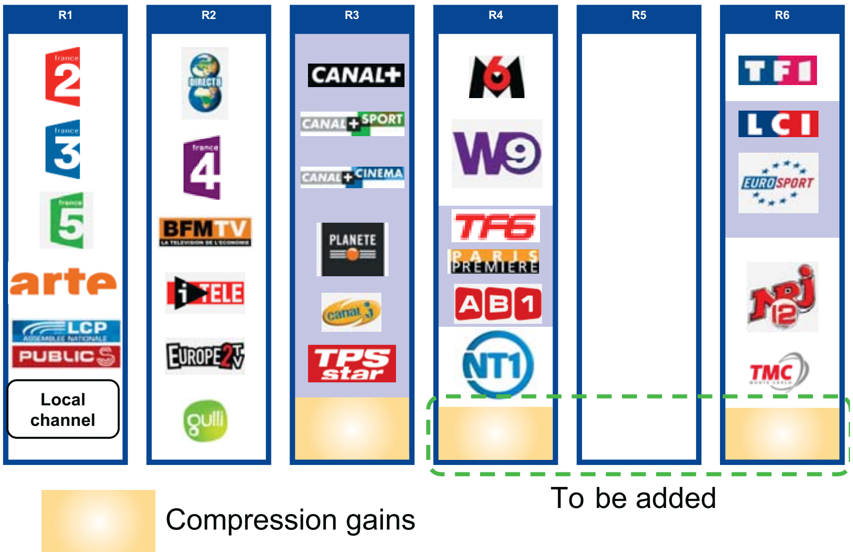


Figure 3
Making space in the multiplexes for new HD services

In parallel with this process, with thanks to the aforementioned report from CGTI, it was recently concluded that a further compression gain on SD MPEG-4 pay services was achievable. This led to a recent decision of the *Conseil* to provide sufficient resources for two new HD services: a pay-service converted from SD to HD and a brand new service. The latter will become available when the new freed-up space has been consolidated onto a single multiplex, i.e. by transferring some existing services to other multiplexes. A new beauty contest is being issued for the conversion of the pay service.

The other side of HD broadcasting: reception

Last but not least, some work was undertaken by various French groups (HD Forum, Simavelec and other partners) to widen the reception means for HDTV by updating the IEC EN-62216-1 standard and by preparing a test bed. Among other topics, this involved the audio components and the signalling for simulcast SD/HD services.

Closely working with them, as the owner of the signalling profile, the CSA will take care of signals compliance so that future, as well as the already-installed, HDTV-compatible equipment performs well.

Indeed, more than a million receivers of different kinds are already able to handle HDTV services. Most of the latest triple-play boxes embed a digital tuner and are MPEG-4 HD compatible. Various types of set-top boxes have been sold over the last two years, and the first integrated television sets were proposed to the customers last summer. In addition, the majority of DTT pay-TV set-top boxes distributed across the French territory are also HDTV compatible.

If some difficulties can be foreseen, they should easily be overcome thanks to the high level of involvement of the industry, broadcasters and network operators. Despite there being no French DTG or Freeview organization, as in the UK, to check the conformity of receivers and richness of signalling (EPG, etc.), this looks like a very important and shared step, targeting a high level of services for the delivery of HDTV programmes to everyone.

And even the law, as voted on 5th March 2007, will contribute to a skyrocketing deployment of HDTV receivers, since it is mandated that all the “HD Ready” television sets will actually embed the parts needed to natively support HDTV reception from 1st December 2008.

But if HDTV is a major project for DTT, other projects are also being implemented in order to continue improving the public offer.

The new frontiers of digital broadcasting

If digitalization has brought more quality to the public, especially with the introduction of HDTV, digital broadcasting frequencies are able to deliver much more for the benefit of the whole population: among others, an increase in the proximity of television with the development of local television on a much larger scale than is possible in analogue and, of course, consumption of audiovisual content in mobile situations on handheld terminals.

The bill of 5th March 2007 tackles these new challenges and urges the CSA to move forward in these fields.

The development of local digital television

In comparison with its closest neighbours, France does not have a developed local broadcasting sector over its terrestrial networks. For instance, Spain has more than one thousand local television stations, where France currently has only twenty-six of them. Adding such new services is therefore at the heart of the CSA policy, in order to keep building a high level of media pluralism and cultural diversity by allowing easier access to these services.

Digitalization is therefore a real opportunity to improve the situation. Indeed, by dividing transmission costs, it opens new economic spaces for local television to develop within.

As far as spectrum is concerned, a reorganization of the multiplexes took place on 13th September 2007 and freed up a slot on the public DTT multiplex (identified as “R1” on the illustrations). Thanks to its unique engineering and frequency structure which enables local programme variations, this multiplex is particularly suited to local television. Where appropriate, this resource will be complemented – in some specific areas where priority has been given to an increased availability of local variations of France 3 or other public services – by frequencies which will constitute an 8th DTT multiplex.

In 2008, the CSA will launch an unprecedented wave of calls for tender over the whole territory for the digitalization of local broadcasting, which will hopefully lead to a stronger local audiovisual sector.

The launch of DVB-H services

The CSA has also launched, on 6th November 2007, a call for tender for mobile television services using DVB-H technology. With this initiative, France will join the group of European countries which are leaders in this area: Finland and Italy but also, very soon, Germany, Switzerland, Austria and Spain.

As detailed below, this call is the result of more than one year of intense work for the main supporters of mobile television in France and the Administration. These efforts have lead to a quick implementation of the regulatory process which progressively removed the main uncertainties for the roll-out of mobile TV: the availability of spectrum, the framework for the authorization procedure, the technical standards.

The first phase consisted of spectrum planning, overseen by the CSA. This work resulted in the identification of a set of frequencies covering around 80 of the main urban areas in France, allowing a coverage of more than 30% of the population. This new multiplex, nicknamed M7, will extend its coverage when the analogue switch-off has been completed.

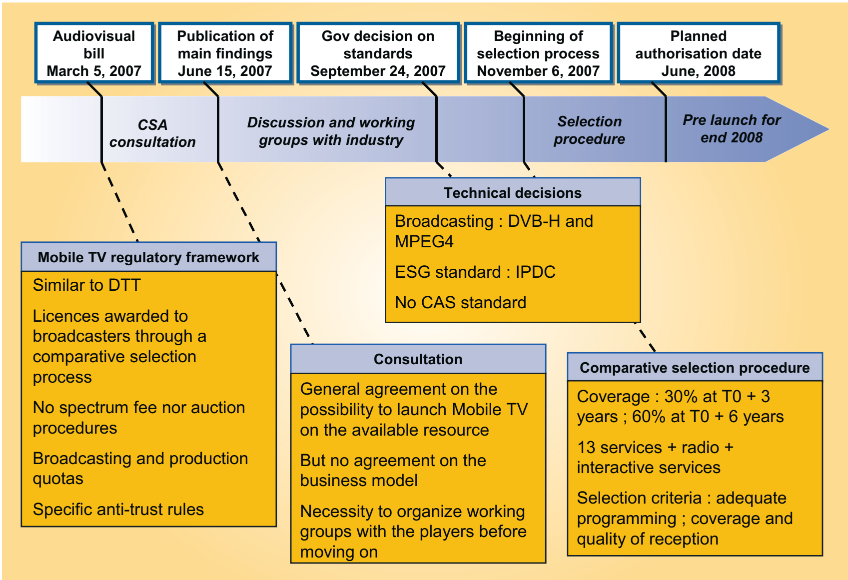


Figure 4
Planning for mobile TV broadcast services in France

With a guaranteed availability of spectrum for mobile television, the definition of the regulatory framework for mobile television was the next big challenge. The bill of 5th March 2007 was an opportunity to adapt the existing regulations on digital television in order to tackle the specific convergence issues that this new medium was creating.

Concerning the standards, the choice of MPEG-4 AVC, HE-AAC and DVB-H by the Ministry of Industry and the Ministry of Culture, after a wide consultation, was well received. Their "open approach" to conditional access, consisting of imposing one scrambling algorithm (Ismacryp) and pushing for a simulcrypt of the services was also well supported.

However, the early choice of DVB-IPDC as the ESG standard is now becoming problematic, considering the recent evolutions in the OMA-BCAST smartcard profile which is to become available at the end of 2008, and the strong support it is receiving from the mobile operators. An evolution of the technical framework might therefore be needed in order to launch mobile services in the best conditions.

The CSA took the initiative to organize technical working groups with the industry, to solve some issues over what was needed to launch a call for tender. The most important of these were:

- **Estimating the relevant bitrates for the different services.** It was a necessary step for the CSA to decide on the number of services to be broadcast in the multiplex. The CSA decided to open the DVB-H multiplex to 16 television services, 4 to 9 radio services and to reserve 120 kbit/s for interactive services;

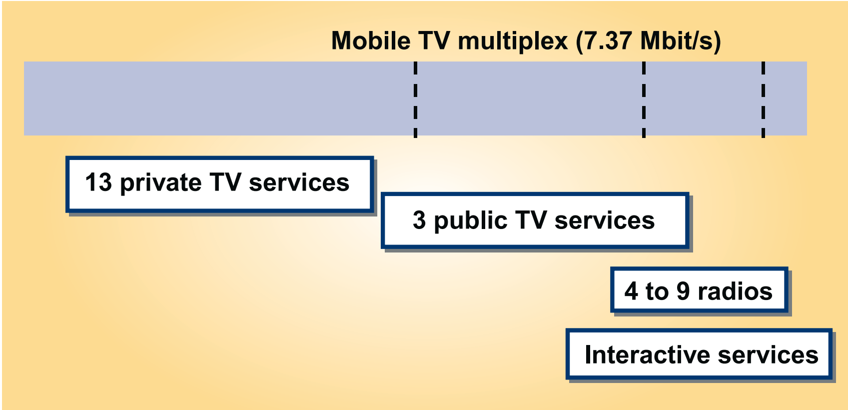


Figure 5
Allocation of mobile services in a single multiplex

- **Defining common terms of references for the quality of coverage and the minimal field strengths needed.** The estimation of the scale of the related network infrastructure has guided the CSA in the definition of minimal obligations in term of coverage (30% of the population to be offered outdoor coverage over the next three years and 60% after six years);

- **Assessing the risk of interference to existing services due to the introduction of DVB-H networks.** This point is still under study, with various field trials taking place in France.

36 projects have been registered with the CSA, less than half of which are pure simulcasts. The results reveal a strong involvement of newcomers in the audiovisual business, such as the mobile phone operators. Indeed, Orange for instance filed two projects.

As in 2007, 2008 will be full of challenges for the regulator in the field of mobile TV. Besides the necessary technical work (planning the protocols, signalling profile, etc.. .), new issues will have to be addressed. Some examples are: interactive services, which will require specific regulations; the extension of the existing regulations concerning the protection of minors from harmful content ...

Before joining the technology division of the CSA, **Aurélien Louis** worked at Siemens medical solutions in the field of nuclear imaging, as a reseach scientist in applied mathematics and statistics.

At the CSA, he is project manager for mobile television and European spectrum policy in the “development & forecast” team.

He is a graduate of *Ecole Polytechnique* in Paris and holds a Masters degree in information and communication technology from *Telecom ParisTech*. He is a civil servant of the *corps des telecommunications*.

Matthieu Roger is head of the department “development & forecast” in the technologies directorate of the *Conseil supérieur de l'audiovisuel* (CSA). He holds a Masters degree in telecommunication from the *Ecole Nationale Supérieure des Télécommunications* in Paris (*Telecom ParisTech*).

He has been active in the fields of broadcasting and telecommunications over the last ten years and has been involved in various different projects : DTT promotion in its early stages; standardization follow-up; project management for private networks deployment; conditional access systems; DTT signalling; spectrum policy …

Conclusions

After a late launch of DTT in 2005, France is now quickly developing a host of new digital television services. With up to eight HD DTT services using MPEG-4 compression planned for the end of 2008, to DVB-H mobile TV, plus digital radio in T-DMB ... the evolution of the whole French audio-visual landscape is strongly moving forwards, also supported by cable, satellite and a fast developing TV-over-ADSL (IPTV) platform.

Bringing this digital revolution to everyone will also be one of the main challenges in the future for the regulator and the government. The extension of DTT coverage to 95% of population is set to be completed before November 2011 and a CSA report is soon to delivered to the government on how to develop DTT in the French overseas territories.

At some point, all these developments will require more spectrum to be further deployed: switching off the analogue TV transmissions will therefore be part of the main priorities for the CSA in the coming years.

ANTS

— a complete system for automatic news programme annotation based on audiovisual content and text analysis

Giorgio Dimino, Alberto Messina and Roberto Borgotallo
RAI Centre for Research and Technology Innovation

This article describes an integrated system for the automatic annotation of television news programmes named ANTS (Automatic Newscast Transcription System). It consists of several analysis components, integrated within a unified architecture. Users have the possibility of accessing a large daily-growing database of news stories from the main national channels – all identified, categorised and published in a fully automatic way. The system identifies story boundaries, extracts texts from spoken content, classifies stories by subject and links external relevant information coming from the web.

The system's performance has been evaluated in a real-life scenario by a panel of professional users inside RAI. The strength of the approach behind ANTS is its ability to integrate several heterogeneous tools in a performant and ready-for-production environment. ANTS is capable of elaborating many hours of material per day, without significant service drops and with sufficiently good accuracy for industrial deployment in large broadcasting facilities.

Introduction and related work

Automatic programme segmentation is one of the most challenging and complex subjects of research.

Although being able to produce a correct segmentation is a key factor for improving the accessibility and precision of search and retrieval, we cannot count on an established approach at solving the problem in general. The common base of the approaches for news is constituted by the use of a combination of visual, audio and speech features.

The TRECVID initiative [1] had news segmentation among its tasks in 2003 and 2004. The works described in [2] and [3] illustrate several different approaches, identified and developed by the TRECVID participants in those two series. The best performing approaches presented at TRECVID 2004 included video and audio analysis, alone or supplemented with automatic speech-to-text transcripts, and showed an F-measure of between 0.6 and 0.7. The baseline features employed in several of the cases are (i) visual similarity between shots within a time window and (ii) the temporal distance between shots [4]. Other heuristics – such as the similarity of faces appearing in the shots and the detection of the repeated appearance of anchor person shots [5][6][7] – can add a supplemental layer of information to improve the accuracy.

The audio channel contribution can be employed to detect pauses, potential boundaries for topic changes [4][6][8], or to detect changes in audio classification patterns (e.g. from music to speech [6]), or to detect speaker changes [8].

As a third information source, text from transcripts or automated speech recognition is very often used, either by searching similar word appearances in different shots or by detecting text similarities between the shots [5][6].

The use of automatic speech-to-text transcripts introduces several issues in the news segmentation task, due to some typical errors such as missing words, word deletions and insertions, and wrongly transcribed words. The current system does not use Newsroom Computer System (NRCS) data, although recently we started investigating how to integrate this information as well, both for improving transcripts and for improving the performance of automatic segmentation.

Architecture

The system, the main components of which are a centralised workflow engine and a collection of generic *AntsClients*, has been designed to be highly distributed and scalable.

Each *AntsClient* is configured to carry out a specific task of the overall process, such as the speech-to-text activity and the news story segmentation. Running as daemons, the *AntsClients* communicate with the workflow engine via HTTP protocol to get jobs and to notify success/failure. The requested process is then executed by a specific command, launched locally by the *AntsClients*. Such an approach, described in *Fig. 1*, allows for deployment on multiple hosts within the network and/or multiple instances supplying the same service so that it is possible to scale the system as much as needed to reach the required throughput and to provide failover capabilities.

All the metadata produced are collected within a centralised repository until final delivery to the publication platform. The search & retrieve subsystem supports full text search, filtering by categories and/or by named entities, besides identification and publication information. Finally, as shown in *Fig. 2*, the page that is presented displays all the resulting components and allows them to be

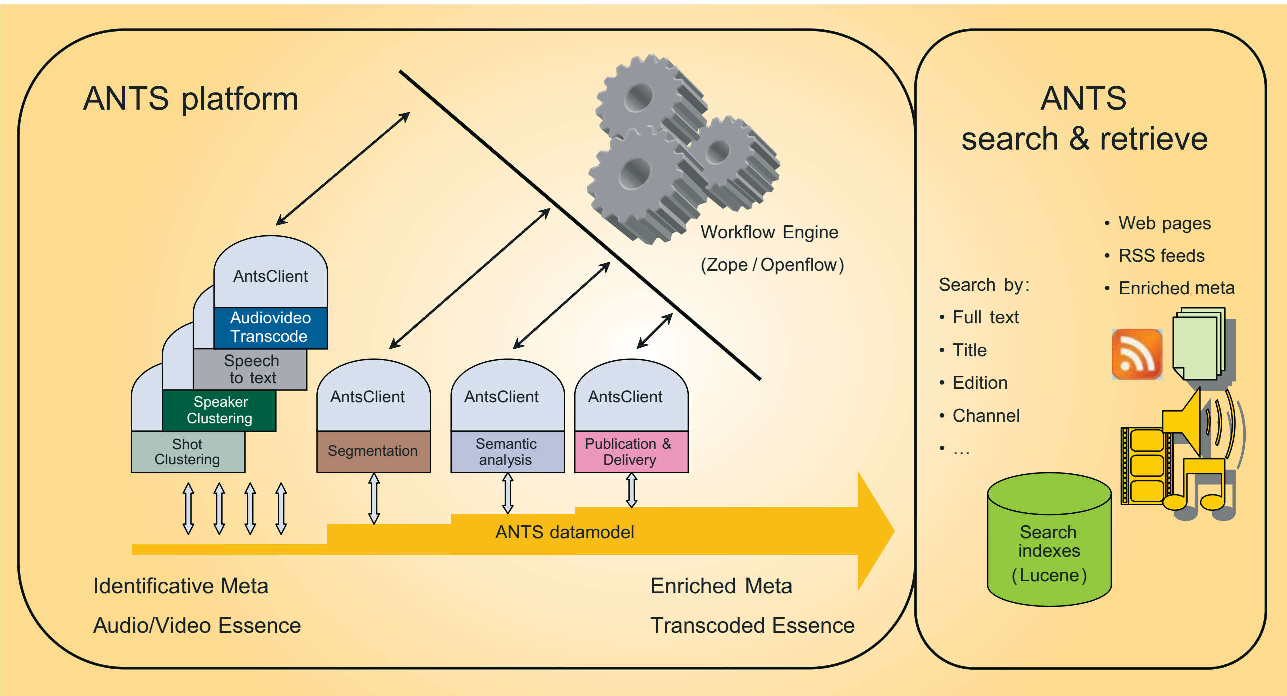


Figure 1
Ants architecture

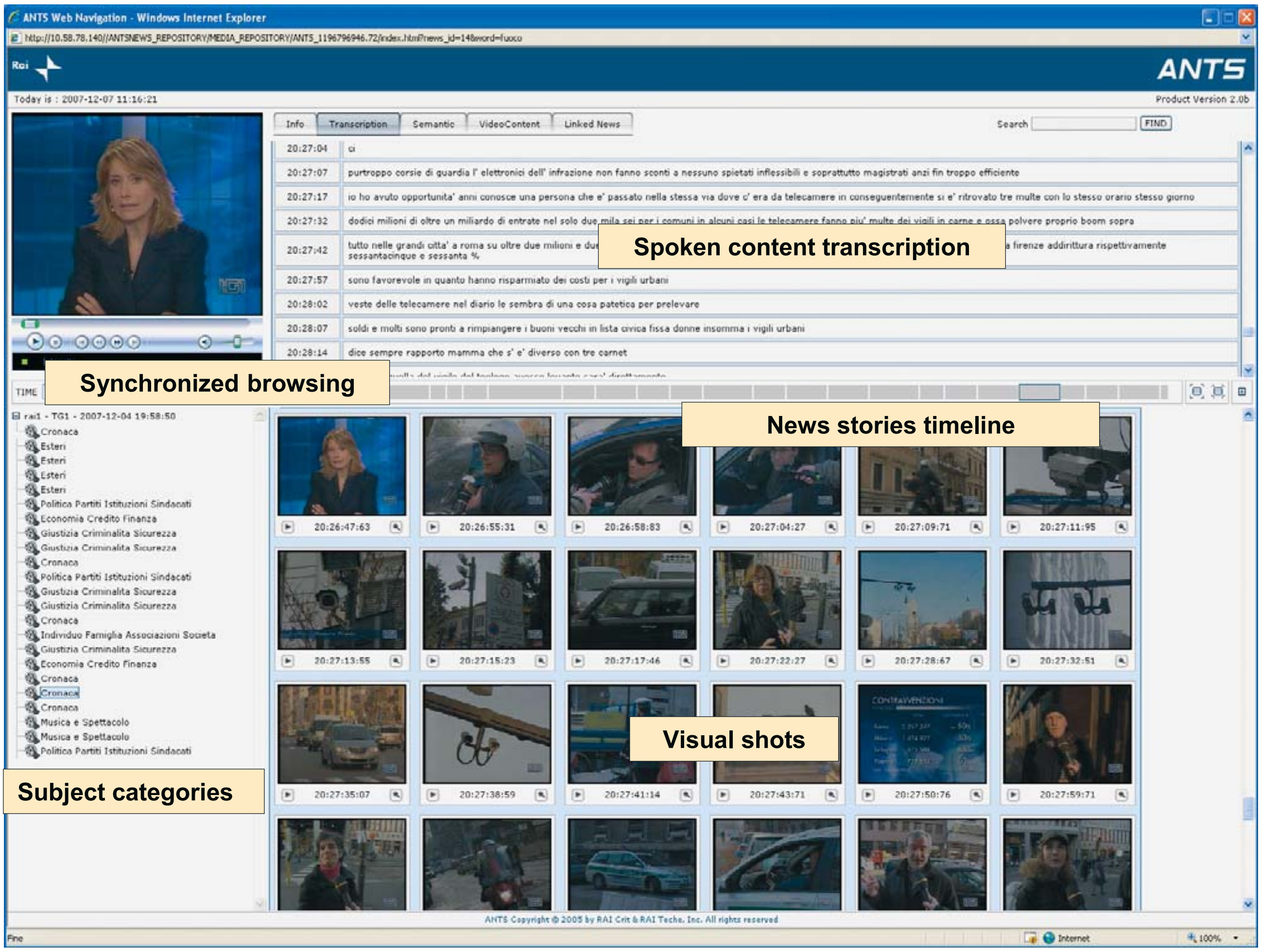


Figure 2
ANTS browsing interface

included synchronously along the common timeline. All the retrieval – as well as monitoring and administration – functionalities can be accessed over an IP network using a web browser.

Moreover, ANTS delivers the finished materials and the metadata, collected in XML format, to the RAI longterm archive catalogue system.

Analysis tools and automatic annotation services

Fig. 3 illustrates the functional block diagram of ANTS, which includes semantic analysis of spoken text and automatic editorial segmentation.

Videoclip matching

To achieve automatic segmentation of live streams into pro-

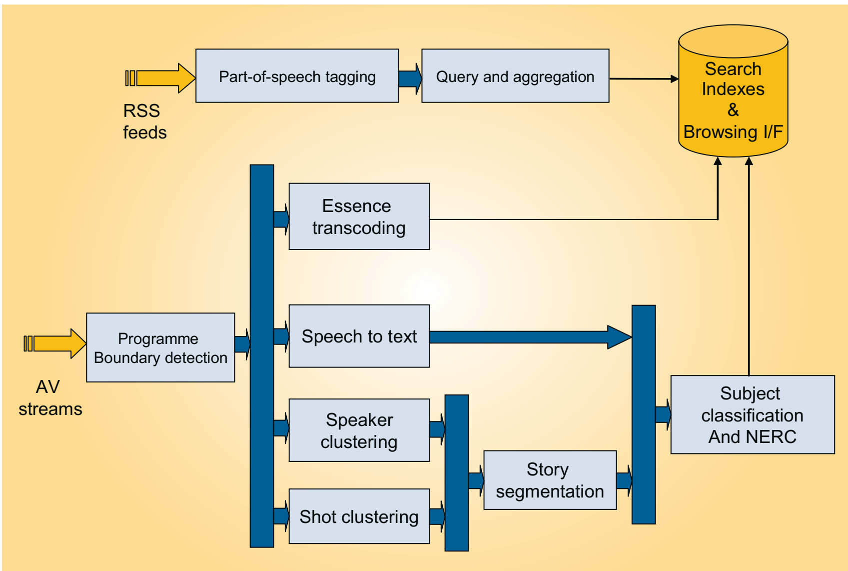


Figure 3
Ants functional block diagramm

grammes which can be further analysed in the subsequent chain, ANTS makes use of a videoclip-matching technique.

Starting and ending programme jingles are used as references to be found in acquired streams, through the use of a shot-clustering algorithm.

As an offline process, feature signatures are extracted from each shot of the reference clips, which includes HSV colour space[1] and Luminance histograms, texture-signature histograms (contrast and directionality), and temporal-activity histograms.

All histograms have 65 bins, to form a 65 x 7 x N feature vector for each shot, where N is the number of pictures in the shot. During the online phase, clipshot signatures are used as multidimensional fixed centroids in a clustering algorithm.

At the end of the clustering, incoming shots aggregated with sufficient strength (i.e. closeness to the centroid) are classified as instances of the clipshot associated with the centroid. The average detection accuracy of the process is 0.86, while recall is about 0.87. (Precision is defined as the ratio between the correctly-identified items and the total number of detected items, while recall is defined as the ratio between the correctly-identified items and the effective number of items.)

Shot clustering

Shot clustering is done following an optimized bottom-up clustering method which makes use of the intersection of feature histograms as the core distance measurement.

The whole process is divided in the following steps:

- feature histograms extraction;
- shot detection;
- partial clustering on fixed-length segments;
- selection of relevant clusters found in the partial clustering phase, and;
- re-clustering on selected clusters.

Thus, the process has a sequence of images as the input and a set of cluster labels associated with each of the input images as the output. This means that the process performs shot detection as a side effect of the clustering process.

Audio clustering

Audio clustering is performed using the mClust tool [9], a free software under the GNU General Public Licence, developed at the LIUM laboratories of the University of Maine – Le Mans, France.

Elaboration results consist of a set of labelled clusters, each of which points to a different person talking in the analysed audioclip. The single cluster is a set of time intervals identified with relative boundaries from the beginning of the clip.

Segmentation of news stories

Segmentation of news programmes into news stories is done by exploiting both aural and visual cues with the help of a three-layered heuristic framework, deduced by the observation of editorial styles of a statistically significant set of programmes, spanning approximately 40 hours (~80 programmes).

1. For an explanation of HSV, see Wikipedia article: **http://en.wikipedia.org/wiki/HSV_color_space**

Abbreviations

ANTS	(RAI) Automatic Newscast Transcription System	**NRCS**	NewsRoom Computer System
HTTP	HyperText Transfer Protocol	**RSS**	Really Simple Syndication
IP	Internet Protocol	**XML**	eXtensible Markup Language

The basic heuristics, widely adopted in literature – e.g. by [10] – is that being able to detect boundaries of shots containing the anchorman is equivalent to detecting news story boundaries.

To detect anchorman shots we use another heuristics, namely that the most frequent speaker is the anchorman and that he/she speaks for periods of time spreading right along the programme timeline. This allows us to select the most probable candidate speaker among the ones identified by a speaker-clustering process.

This approach doesn't permit the system to discern situations in which the anchorman introduces several brief stories in sequence without external contributions (e.g. reportages). To overcome this limitation we use the third heuristic, consisting of the knowledge that, in the great majority of observed cases, the introduction of a new brief story is accompanied by a camera shot change (e.g. from a close-up shot to a wider one). To optimize the accuracy in selecting the camera shot changes, we perform a videoshot-clustering process based on the same features explained in *Fig. 3*. This allows us to detect and classify shot clusters as pertaining to studio shots containing the anchorman, following the same frequency/extension heuristic used for detecting the candidate speaker. This double clustering process (both on audio and on video) enables a very simple and effective recursive algorithm which selects alternatively video and audio clusters on the basis of their mutual coverage percentage.

Finally, story boundaries are identified as those in which either an audio or a video cluster boundary occurs among the clusters selected by the recursive algorithm, with an adaptive threshold to avoid oversegmentation. An outline of the experimental evaluation of the segmentation algorithm is presented on the next page.

Subject classification and link to external sources

In ANTS, spoken-content extraction is performed using a speech-to-text engine based on [11], which is capable of transcribing both Italian and English. Subject classification of segmented stories is done using a naive Bayesian classification model trained on a corpus made up of items of extracted text and annotated with a standard subject taxonomy of 28 classes. The corpus counts 25'000 items, 4/5 of which are used for training and the remaining 1/5 for testing. The overall observed subject classification accuracy on the running system is 0.82, while programme-level accuracy, i.e. the average classification accuracy calculated for the set of items belonging to the same programme, is 0.88.

Links to external information sources are implemented through linguistic analysis of the RSS feeds of a pool of six major newspapers, which are polled periodically. On each individual RSS item title, a part-of-speech tagging is performed in order to extract the most significant word cues, which are in turn used to perform a full text query on the text automatically extracted from the television news items. Query results are arranged in a browse list associated with the item title, and each individual news story is linked to the RSS item for which search scores are above a certain threshold. Aggregated news stories under a certain title can themselves be seen as RSS services provided by ANTS.

The result is that users can have a multimedia integration of RSS feeds coming from the major newspapers, made with relevant news stories collected from the major television newscasts. On the other hand, television news items can be automatically annotated using the RSS titles.

The mean average precision of the news-item aggregation process is 0.97, calculated as the ratio between the relevant news items and the total ones, with respect to the RSS feed title in a certain aggregation.

Experimental evaluation of the news stories segmentation algorithm

We tested our news story segmentation algorithm against a set of test programmes running to about 40 hours of material. The test set had been manually tagged, i.e. all true story boundaries had been identified. To assess the system performance we used an alignment measurement taking into account starting boundaries and ending boundaries with different weights, as well as considering missing material as having more impact than excessive material on the measurement.

In a first phase, we randomly selected a subset of the test material and then optimized empirically the parameters of evaluation in order to achieve a good match between the users' assessment of the segmentation quality and the objective measurement. We thus obtained a user-validated quality measurement. In the second phase, once the measurement has been verified according to the described procedure, we adjusted the segmentation model parameters in order to optimize the user-validated measurement output.

Table 1 shows the precision obtained for four subclasses of programmes.

Table 1 – Precision figures of automatic segmentation

Class	Tg1	Tg2	Tg3	TgR
Precision	0.81	0.69	0.80	0.73

The role of open source components

The system described in this article would not be viable without the availability of open source tools. Firstly, because almost all computers in service within this architecture run on the Linux operating system, which was also the development and testing platform. From components written in C programming language, compiled with the GNU Compiler Collection (GCC), to other components written in much higher-level programming languages, such as python, perl or ruby, or simple scripts in any variant of the Unix Shell ... the various ingredients were prepared and integrated in the open source domain.

To manipulate audio and video material, the *MJPEG Tools* [12], were successfully adopted; the workflow management relies on the couple *Openflow* over *Zope* [13][14]; the publication service is built as a web application which makes use of *Postgresql* [15] together with the *Apache* web servers http and *Tomcat* and the search engine *Lucene* [16].

The tool used for speaker segmentation, within the editorial segmentation process, is *mClust* [9], while for the subject categories, *Categorizer* [17] was adopted.

Beyond the concept of an open source toolkit, it is the open source environment, including the experience of the people involved, which made possible the achievement of such a complex system with continuous adjustments and requests for new features.

Conclusions

In this article we have given an overview of an automatic news programme annotation system named ANTS, developed at the RAI Research Centre in Turin. The strength of the underlying approach of ANTS consists in offering to the users a good global performance by integrating several analysis tools into a single fully-automated product.

Documentation of TV news items must provide usable results in a considerably shorter delay than for other television programmes. While with a manual annotation process – although assisted by automatic acquisition and shot detection – one item took a couple of working days before being

Giorgio Dimino received a degree in electrical engineering from the Polytechnic of Turin in 1987. In 1988 he joined the Research Centre of RAI - Radiotelevisione Italiana in Turin, working in the fields of digital audio and video processing, and archiving. His interests include the design of automated digital archives and the application of information technology in television production.

Mr Dimino has been leader of the work area on Metadata, Access and Delivery of the IST 6th Framework project PrestoSpace. He is also an active member of the EBU PMC and FIAT/IFTA.

Alberto Messina works for the Research Centre of RAI - Radiotelevisione Italiana in Turin: *Centro Ricerche e Innovazione Tecnologica* (CRIT). He is involved in several internal and international research projects in the field of digital archiving, with particular emphasis on automated documentation, and automated production. His current interests range from file formats and metadata standards to the domain of content analysis and information extraction algorithms, which is now his main focus. He is also author of various technical and scientific publications in this subject area. He collaborates extensively with the local University of Turin - Computer Science Department, which involves common research projects and students' tutorship.

Mr Messina is an active member of several EBU projects including P/TVFILE, P/MAG and P/CP, and is chairman of the P/SCAIE project dealing with automatic metadata extraction techniques.

Roberto Borgotallo graduated in Telecommunication Engineering at *Politecnico di Torino* in 1999. Since 2001, he has been working for RAI - Radiotelevisione Italiana at the R&D department (*Centro Ricerche e Innovazione Tecnologica*) in Turin. Initially, he was involved in several projects gravitating around the RAI multimedia catalogue, called CMM, regarding metadata ingestion and transformation.

More recently, Mr Borgotallo has been working in a team that is developing an automatic metadata extraction platform which is actually used extensively in RAI for experimental purposes, and even for real in the production evironment. His major professional interests are metadata and essence transformation, system integration and workflow management.

available to users, after the introduction of ANTS a newscast becomes searchable at single-story level within two hours after publication.

The ANTS system is currently in use in RAI to index the main newscast editions of the three national generalist channels RAI1, RAI2 and RAI3. Users can directly query the system or subscribe to personalized RSS feeds on specific topics and be notified when new relevant content is available. The service will be extended progressively to cover also the regional news.

Another interesting application of ANTS – which is being provided to some Italian regional administrations – is the monitoring of local stations' transmissions. In this case, all the relevant channels' emissions are recorded and indexed. Administration officers can then browse the content for statistical analysis purposes or for verification of compliance to the Authority's regulations on TV emissions.

Future works will be directed towards extending editorial segmentation to other kinds of programmes and to the development and integration of emerging techniques in metadata extraction.

Acknowledgements

The authors wish to thank Laurent Boch and Daniele Airola, from RAI CRIT, for their outstanding contributions to the design and development of the ANTS system.

References

[1] Trec video retrieval evaluation. Internet site: http://www-nlpir.nist.gov/projects/t01v/

[2] T. Chua, S. Chang, L. Chaisorn and W. Hsu: **Story boundary detection in large broadcast news video archives techniques, experience and trends**
In *Proc. of ACM Multimedia 2004*.

[3] W. Kraaj, A. Smeaton and P. Over: **Trecvid 2004: An overview**
In *Proc. of TRECVID Workshop 2004*.

[4] D. Eichmann and D.-J. Park: **Boundary and feature extraction at the university of Iowa**
In *Proc. of TRECVID Workshop 2004*.

[5] M.J. Pickering, L. Wong, and S.M. Rueger: **Anses: Summarization of news video**
In *Proc. of International Conference on Image and Video Retrieval (CIVR)*, 2003.

[6] T. Volkmer, S.M.M. Tahahoghi and H.E. Williams: **Rmit university at trecvid 2004**
In *Proc. of TRECVID Workshop 2004*.

[7] Y. Zhai, X. Chao, Y. Zhang, O. Javed, A. Yilmaz, F. Rafi et al: **University of central Florida at trecvid 2004**
In *Proc. of TRECVID Workshop 2004*.

[8] G.M. Qu'enot, D. Mararu, S. Ayache, M. Charhad and L. Besacier: **Clips-lis-lsr-labri experiments at trecvid 2004**
In *Proc. of TRECVID Workshop 2004*.

[9] Internet site: http://www-lium.univ-lemans.fr/tools

[10] M. De Santo, G. Percannella, C. Sansone and M. Vento: **Unsupervised news video segmentation by combined audio-video analysis**
In *MRCS*, pages 273–281, 2006.

[11] F. Brugnara, M. Cettolo, M. Federico and D. Giuliani: **A system for the segmentation and transcription of Italian radio news**
In *Proc. of RIAO, Content-Based Multimedia Information Access*, 2000.

[12] Internet site: http://mjpeg.sourceforge.net

[13] Internet site: http://openflow.sourceforge.net

[14] Internet site: http://www.zope.org

[15] Internet site: http://www.postgresql.org

[16] Internet site: http://www.apache.org

[17] Internet site: http://search.cpan.org

Microphone systems used for Surround Sound pickup
— and their use at Wimbledon tennis and The Proms

Bill Whiston
BBC Outside Broadcasts

This article briefly describes some of the microphones developed specifically for Surround Sound pickup, along with several of the main Surround acquisition systems on which the majority of the dedicated Surround mics are based. It offers some personal advice on whether a particular system is suitable for use in this recording environment or that. Some microphone systems are obviously more intrusive "in shot" than others, depending on the location.

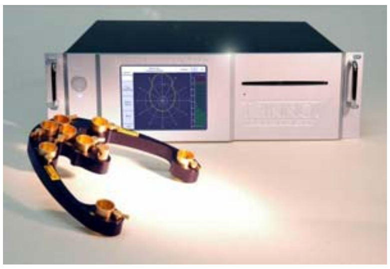
The author also describes two major outside broadcasts that have involved Surround Sound mixes – the Wimbledon Tennis Championships and the BBC Proms Concerts from the Royal Albert Hall in London.

The types of microphones and the style of their use for Surround acquisition are as varied as there are sound mixers! Much thought has been given to this subject over many years by many different people, some with an engineering base, some of them practitioners, and others who are theoreticians. Most methodologies are developments of stereo recording techniques with added components to derive the Surround information.

I'll start with three systems designed specifically for Surround. One requires the use of a computer to generate and adjust the output to obtain the required Surround soundfield. Another requires a matrixing box to generate the Surround. The third produces separate discrete outputs but, because it is built as a complete system, I won't include it in this section; instead, it will be discussed later in the section on "discrete" systems.

Surround microphones

Trinnov SRP



This "High Spatial Resolution" microphone array consists of eight near-coincident omni-directional (omni) capsules in a horseshoe-shaped mount. The outputs of these microphones are passed through a computer-controlled matrixing system, which converts timing differences produced by the array into amplitude differences. It also allows spatial (directional) control of the microphone to control the reverberant field. The manufacturer claims, among other things, optimized channel separation, an enhanced "sweet spot", production of good phantom images over the full 360° soundfield and good quality downmixing to stereo and mono.

This is a complex and expensive microphone, difficult to use in the outside broadcast external TV environment (as opposed to inside buildings) due to the necessarily large windproofing arrangements required.

For further information – **http://www.trinnov-audio.com/products.php**.

SoundField

The SoundField microphone is based on the concept that any Surround audio event can be represented by four basic bits of information:

- the front-to-back or depth information (X);
- the left/right information (Y);
- the up/down information (Z);
- the central point (W) to which the other three elements are referenced.

Collectively, these four elements are referred to as B-format and are combined via a matrix to produce any Surround audio format from 4.0 all the way to 22.2, with height information if required. It just depends on the design of the matrix.

The microphone head consists of 4 x figure-of-eight capsules and 1 x omni capsule, electronically combined to form an absolute coincident array. Because the array is always referenced to the same central point, any time- or phase-related anomalies don't exist, giving extremely good fold-down to stereo and mono signals.

It is easy to use and quick to set up, is remotely controlled from the accompanying matrixing box and is comparatively cheap.

For further information – **http://www.soundfield.com/soundfield/soundfield.php**.

Holophone H2-PRO

The Holophone H2-PRO claims to emulate the characteristics of a human head. Sound waves "bend around" the head and are picked up by eight omni microphone "elements" to provide spatial audio imaging. The individual elements combine with the spherical body to act as an "acoustic lens". The information relating to the complete soundfield can be replicated without the use of additional microphones.

The Holophone H2-PRO is capable of recording up to 7.1 channels of discrete Surround Sound (Left, Right, Centre, LFE, Left Surround, Right Surround, Top (Height) and Centre Rear). These co-relate to the standard 5.1 channels and add a top channel for formats such as IMAX and a centre rear channel for extended Surround formats such as Dolby EX and DTS ES. You can have total flexibility over the discrete Surround audio signals and choose to use whichever channels any particular Surround project requires, as the channels remain discrete all the way through the process.

It is simple and relatively cheap to use as there is no processing involved, providing just the eight discrete outputs from the "elements". However, the device relies on the quality of those microphone "elements" for the acceptability of its audio performance.

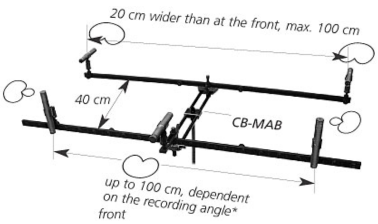
For further information – **http://www.Holophone.com/products.html**.

Discrete microphone systems

This is the most contentious and widely discussed area of Surround recording. Most microphone setups in this category are variations on three basic formats ... and those formats themselves being based on the Decca Tree stereo array. They are based on combinations of spaced or near-coincident omni and directional microphones.

OCT Surround – supercardioids / cardioid / omnis

This is probably the simplest of the setups that can be used for spaced or near-coincident Surround acquisition and the basis for many different techniques. It's a derivative of the "Decca Tree" system with the long forward arm replaced by a very short one and the omni mics used in the Decca Tree replaced with microphones of different polarities.



The basic OCT Surround array consists of three microphones arranged in a line on a bar, the two on the extremes being supercardioid microphones pointing +/- 90° to the side. The central, forward-facing microphone is a cardioid, usually mounted about 8 cms forward of the bar. Variations on this setup include mounting omni microphones (low-pass filtered) at the hypercardioid microphone positions, or at the central cardioid mic position, to extend the lower frequency response of the whole setup. This OCT "front" system is combined with one of the many "rear pickup" systems, such as the Hamasaki Square or the IRT Cross, or even a simple pair of spaced cardioids pointing backwards, to pick up the required rear image.

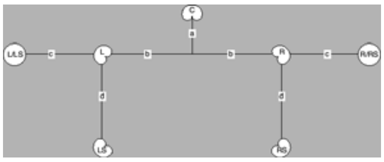
For further information –
http://www.hauptmikrofon.de/theile/Multich_Recording_30.Oct.2001_. PDF

"Williams" Array – cardioids

This method has also generated a swathe of variations but is based on a single principle. That principle was put forward by audio consultant Michael Williams and says that, when putting an array together, a "critical link" is made between the angles of coverage of adjacent pairs of microphones in that array, so that their coverage angles don't overlap but just touch – in effect, a modified one microphone-per-speaker format. In theory this gives a very even representation of the soundfield.

The array consists of five cardioid microphones, the L/R mics are one metre apart and the centre mic is mounted 25cms in front of the bar. The rear facing cardioids are about 60cms apart and some 250cms behind the bar. The front L/R capsules are +/- 70° and the rear mics are +/- 150° with reference to the centre.

Fukada Tree – cardioids and omnis



Akira Fukada, a senior recording engineer from NHK in Japan, proposed this array. It is also based on the Decca Tree approach in that it retains the long front arm but changes the omni capsules in the original to cardioids angled between +/- 130° and +/- 150° relative to the centre. Each microphone is evenly spaced from the centre point at about 1 to 1.5 metres. Two omni outrigger mics are added to the array, in line with the L/R mics and a similar distance from their corresponding cardioid partners.

Two cardioid microphones are added to cover the rear, in line with the forward-facing L/R microphones, coincidentally if required but no more than 2 metres behind them, and at an angle of between +/- 60° and +/- 90°.

Abbreviations

L/R	Left/Right	**M+S**	Mono and Stereo
LFE	Low-Frequency Extension	**PA**	Public Addressing
LS	Left Surround	**RS**	Right Surround

When mixing the five main front and rear elements, there appears to be a separation of the front and back soundfields. The omni outriggers are used to blend the two together by being panned front to back and mixed into the soundfield appropriately.

For further information – **http://www.tonmeister.ca/main/textbook/node840.html**

Hamasaki Square – figure-of-eights

Kimio Hamasaki of NHK introduced this technique. All of the above discrete techniques have included frontal pickup in their design. The Hamasaki Square was invented as a means of picking up the diffuse soundfield in a reverberant environment. It complements any of the standard stereo recording techniques such as OCT or Decca Tree for reproducing Surround ambience.

It consists of four figure-of-eight microphones fixed in a square (dimensions usually between 2 and 4 metres), with the axis of each of the four microphones at right angles to the direct sound. Because microphones are aimed sideways, they effectively prevent unwanted direct sound from appearing in the Surround channels. In a situation with an audience below the array, the off-axis presentation of the microphones also helps to control excessive audience pickup.

The four microphone signals are routed discretely to the Left, Right, Left Surround and Right Surround channels.

For further information – **http://www.tonmeister.ca/main/textbook/node843.html**.

Schoeps "Head" – omni + figure-of-eight

This technique was devised by an American audio engineer, Jerry Bruck. The microphone consists of the Schoeps KFM 360 sphere microphone with two forward-facing figure-of-eight capsules placed near-coincidentally to the omni capsules in the sphere. The outputs of these capsules are fed into a matrix box, from which are derived four virtual forward and rear facing mics. A four-channel Surround output, together with a matrixed Centre channel and an LFE channel, are thus produced from the matrix box.

The polar patterns of the virtual microphones can also be changed to give a variety of pickup characteristics if required.

If the four unprocessed mics are recorded directly, the Surround effect can be manipulated at a later date by being played back through the matrix box. This makes the system very similar to the SoundField B-Format system.

For further information – **http://www.schoeps.de/PDFs/SCHOEPS_surround-brochure.pdf**.

MSM - cardioids / hypercardioids + figure-of-eight

This is a near-coincident technique, deriving a Surround effect with three capsules (usually two cardioids and a figure-of-eight), the outputs of which may be simply matrixed in a sound-mixing desk to provide a four-channel soundfield.

Of all the techniques, this must be the simplest and cheapest to set up and use. In my view, it is certainly the easiest system to use and is very portable – usually being small and light – and thus is eminently suitable for boom and/or rod work. It requires only three channels of recording media to obtain 4.0 Surround Sound!

Summary

Nearly all of the systems discussed have varying out-of-phase components and therefore different phase performance when folded down to stereo and mono. Compatibility with stereo and mono is something that broadcasters are very much aware of, as most of their audience still listen that way, be it via radio or television transmissions. The Surround audio can sound as wonderful as you could possibly wish in the recording control room but, if the listener at home receives a poor quality stereo

The BLITS 5.1 tone generator

The BLITS 5.1 tone generator was designed in conjunction with BSkyB to provide line identification of Surround Sound signal sources. Provision is made for identifying stereo, channel phase and the six 5.1 audio channels: L, R, C, LFE, LS and RS.

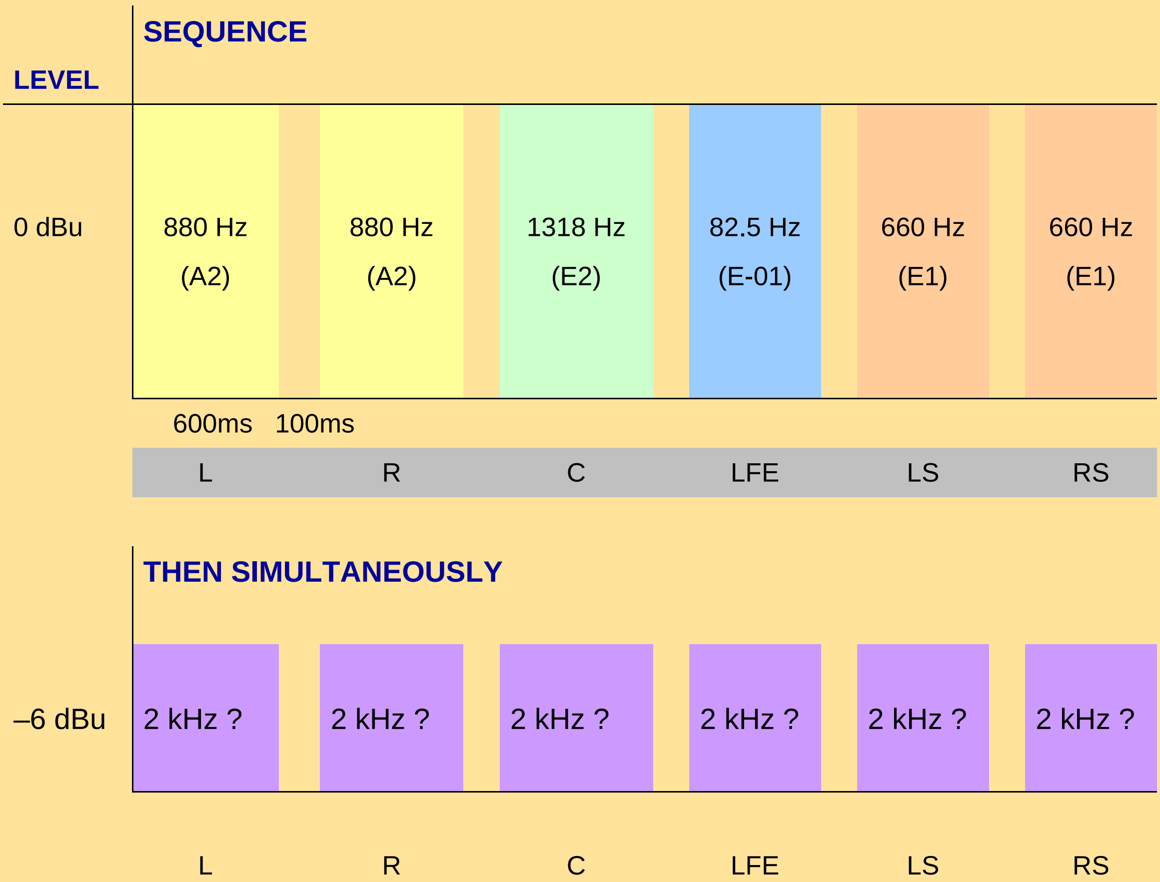
The tone generator may be switched between the 5.1 and the stereo oscillator sequences, and individual 5.1 sequences may be “looped”. The unit also includes a discreet EBU tone source.

With “Tone output Select” selected to 5.1 and “5.1 cycle” selected to Normal, the BLITS tone generator will output the following tone sequences:

- Left/Right identification: 1 kHz 0 dBu tone on both Left and Right channels, with four interruptions on the Left channel. The four interruptions will differentiate the ident from the standard EBU or GLITS tones.
- Phase check: 2 kHz tone at –6 dBu on all audio channels.
- Channel identification: The sequence is:
  (Front) Left – 880 Hz
  (Front) Right – 880 Hz
  Centre – 1318.51 Hz
  LFE – 82.407 Hz
  Left Surround – 659.255 Hz
  Right Surround – 659.255 Hz.

These frequencies correspond to notes A4, A4, E4, E0, E3 and E3 respectively.

The Blits generator may be locked into any one of these cycles using the 5.1 cycle push button.



transmission because your Surround mix produces a poor stereo fold-down, then the complaints will come flooding in.

The best fold-down is produced by those systems that are coincident or near-coincident. The more widely spaced the array, the harder it is to get the two different mixes to correlate.

4. How do you choose?

This very much depends on circumstances. What am I covering? What is the programme content? How and where can I mount the microphones? Is it a fixed situation or will it be moveable? Is it for television or radio, or for producing a CD recording? If it's for television, will a particular array be too obvious "in shot" and what compromises will I have to make, to get a decent result?

What surround format am I going to use? 4.0/5.0/5.1 ... all the way up to 22.2 and beyond? What am I recording onto – multitrack, videotape with limited tracks, broadcasting live? Will there be any post-production involved? Is it a mixture of all of these?

The decision has to be based on training, experience, experimentation and, in the end, trial and error. And production colleagues have to be involved in the decision-making at the earliest possible moment. If they don't understand what is involved in what you are trying to attempt, you will find the whole process becoming an uphill struggle. Most production people know very little about Surround Sound and they need to be educated and encouraged to find out more – otherwise interest in Surround Sound from that quarter will wither away.

There are a host of techniques out there to learn and choose from, plus some of your own – given the time for thought, experimentation, practice and, last but not least, Trial and Error!

I was very fortunate in that I had used Dolby Surround at Wimbledon tennis for years before I was ever asked to do 5.1, and I had plenty of time to experiment without any pressure from anyone, other than myself!

Two programme examples

Wimbledon Centre Court

Audio coverage at Wimbledon for television was developed from early BBC Radio broadcasts, and started in the 1930s. The original style of tennis coverage was to place an omni mic on the chair for the effects, an omni over the umpire's head to hear the scores and, as a later addition, an omni mic above the crowd in the stands.

"Modern" mono coverage, starting from the late 1960s, consisted of a pair of shotgun mics looking into each side of the court from the umpire's chair, a switched cardioid mic for the umpire (also sent to the PA system) and a crowd mic (shotgun or omni depending on the venue) away from the action to keep prying ears away from anything the players at the chair might say to the Umpire. This is still the basic coverage for most tennis matches, with a little more attention being paid to what the players are saying to the Umpire!

In the 1980s, stereo coverage evolved to provide a fixed, stable, "one-end-of-the-court" image for the viewer. Ping-pong stereo, much beloved of hi-fi buffs at the time, was not considered the correct way to go!

Stereo started by using a well wind-proofed AKG C24 (set to cardioid) as a crossed pair but this was soon changed to an M+S system using a Neumann RSM 191. This was preferred because it was easier to mount in the venue, it was smaller in-shot and also gave better compatibility for viewers still listening in mono. This was in effect a hypercardioid and figure-of-eight combination.

Additionally, a stereo crowd mic was placed on a bar above the spectators, together with a camera giving a fixed-view, wide-angle shot of the court, which the mic matched perfectly. A net mic was added for "let calls" and for better coverage of the centre of the court during doubles matches. "Let calls" don't happen now but the mic remains for its effects contribution. A pair of courtside mics were added at either end of the base line for a little more crowd "intimacy" to bring home Wimbledon's unique atmosphere, the court itself being much smaller than it looks on television.

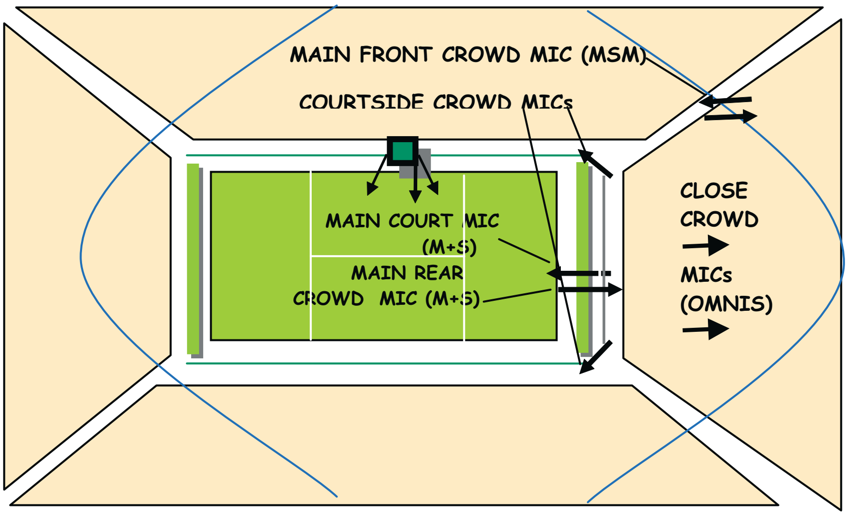


Figure 1
Position of the main Dolby 5.1 FX mics at Wimbledon

With the advent of Dolby Surround coverage, the rear channel offered a chance to involve the listener-viewer even more in "being in the crowd", even though the rear channel was only mono!

To this end, an additional single omni rear-crowd mic was placed up above with the stereo crowd mic, and some "close crowd" mics were placed in among the spectators for an even more realistic "in the middle of the crowd" experience, these three new mics being sent to the rear channel. The courtside pair were split and fed to both the front and rear of the mix to attempt to "join up" the front and back to give a smoother Surround effect. The front stereo coverage had progressed to an M+S setup using a Schoeps CCM cardioid and figure-of-eight system for better width coverage of the court and an even more discrete presence on court.

In spite of the limitations of the Dolby Surround format, the Surround effect proved to be very successful, and the time and effort taken to experiment and produce it were a very beneficial step towards producing further, more adventurous Surround mixes.

The current full 5.0 coverage adds a rear-facing M+S stereo pair (Cardioid) down at the court level (see *Fig. 1*) to provide the main rear crowd effects. Due to mounting constraints, this rear-facing microphone has to be mounted some 40 cms behind the forward-facing M+S pair facing the court and, in spite of initial worries about timing and phase issues, it has proved not to be a problem when folded down to stereo. The courtside mics are still fed to both front and rear channels to fill in the side "blank" that can be a feature of some Surround mixes. The close crowd mics are still used as a "spaced-pair". An MSM pair replaces the stereo and mono crowd mics in the Dolby Surround setup to give a high Surround crowd mic. There is no LFE channel produced, as there is little content at that low frequency worth broadcasting.

In effect my Surround array for Wimbledon acts like a very modified Fukada Tree.

My 5.0 Wimbledon coverage actually started as 4.0 as I didn't like using the Centre channel, especially for the court effects, as I felt it collapsed the stereo. However, I currently feed some of the commentary to the Centre channel (at about 8 dB lower than the L/R mix) just to give it some anchorage for viewers who might be outside the "sweet-spot" at home.

The stereo and Surround mixes are done on the same board. Previously, during the HD/5.1 trials in 2006, the stereo transmission to SD viewers travelled a completely separate path to that of the 5.1 mix for the HD viewers. The SD stereo mix was derived from the front of the 5.1 Surround mix and the metadata of the Dolby system adjusted so that no Surround or Centre channel contributed to the stereo fold-down on the HD service. This is because I wanted the stereo mix to be identical for both. This has now changed and, in future, the SD/stereo transmission will be derived from the HD/5.1 transmission automatically in the transmission suite, which means that, this year, I will have to decide whether or not to include some of the Rear and Centre channels in the fold-down, using the Dolby metadata.

BBC Proms 2007/8

The basic microphone setup in the Royal Albert Hall for the Proms is designed and put in by BBC Radio and we, in Television Outside Broadcasts, add any additional microphones that we may want. The microphones used vary from day to day, depending on the orchestras and ensembles playing, hence the large number of microphones rigged for the season.

The system we used in 2007 was designed by my colleague Andy Payne and developed by trial and error over two years of recording the Last Night of the Proms for 5.1 post production, as well as his experiences recording opera in Surround at the Royal Opera House in London.

The sound field is split into three zones – the Stage, the Arena and the Space. The Stage zone uses a Decca Tree combined with two cardioid mics on the main bar pointing backwards and upwards. The Arena zone consists of an array of four Schoeps MK 4s, slung in the centre of the arena in a modified IRT Cross formation. The Space zone is made up of a pair of widely-spaced omni microphones set just below the dome and towards the rear of the Hall, combined with a pair of widely-spaced wide-cardioid mics set above the stage canopy.

Our production colleagues on the Proms are concerned about the quality of the stereo derived from the fold-down of the HD Surround mix, as this affects the majority of their viewers who watch and listen to the SD transmission. This has caused us to separate the stereo and Surround mixing areas into two different areas to enable us to create separate mixes for SD and HD. This has resulted in the main music-mixing truck providing Surround "stems" of the three zones. The Centre and LFE feeds, and separate soloists feeds if necessary, are sent to a second truck where the Surround mix is put together and any in-vision links and interviews added. The stereo mix is then made in the main music truck during the actual performance and passed to the second truck where the same links and interviews are added, the two mixes taking separate paths out to air.

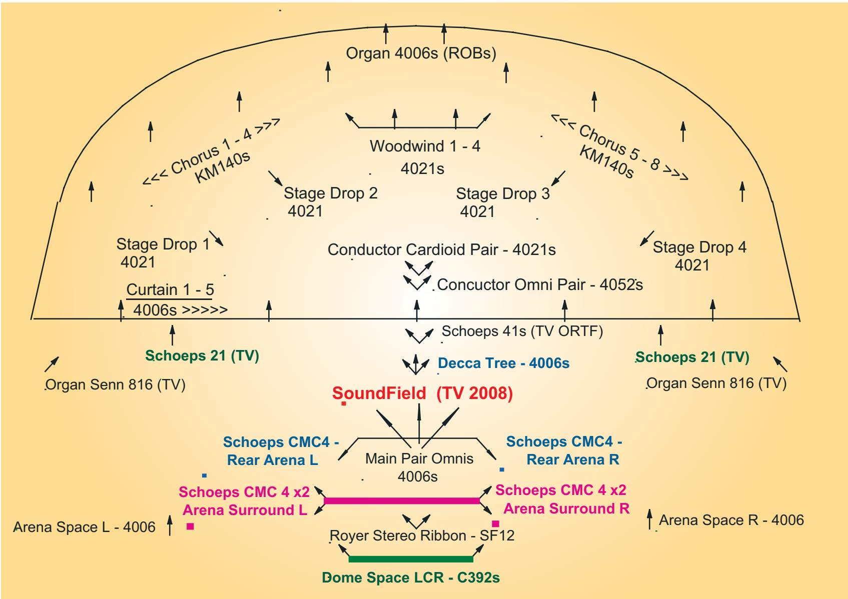


Figure 2
The Proms 2008 mic layout at the Royal Albert Hall, London

This is an expensive and ultimately redundant way of operating. It is redundant because, from this year on, the SD pictures and stereo sound will be derived from the HD pictures and Surround mix, so there will be no opportunity to transmit a separate stereo mix for SD viewers.

This year (2008) we are hoping to rig a SoundField microphone in the Hall for the duration of the Proms. As mentioned above, this microphone's Surround output is referenced to a single point giving extremely good stereo and mono compatibility. Hopefully this will fully address our production colleagues' fears, as well as simplifying the whole process.

Bill Whiston joined BBC Television in September 1970 and BBC Outside Broadcasts in 1978. He has been involved in a whole range of different types of programmes, covering all aspects of Outside Broadcast's output, from major events such as the Queen's Golden Jubilee celebrations for which he won a Royal Television Society Award, to Sport, including fifteen years mixing Centre Court at Wimbledon, where he first became involved in Surround mixing, and two Olympics Tennis events, in Atlanta in 1996 and Athens in 2004.

Mr Whiston have been involved in major BBC Drama productions over many years, including the first live drama transmitted on the BBC for twenty years when they produced "Quatermass" from a location just outside London in April 2005 which was also recorded for Surround post production. In recent years he has regularly balanced a number of BBC Promenade Concerts from the Royal Albert Hall for both BBC1 and BBC2. He has also mixed a wide variety of programmes, both in the UK and abroad, including the Handover of Hong Kong to China, live interactive natural history programmes from the Rift Valley to schools across the United States and the UK, religious programmes from Jordan and a documentary series on aircraft from France and the USA.

Summary

Each of the programmes discussed above has its own particular problems. Because I have been mixing Wimbledon in Surround for so long, it has pretty much settled as a concept, although that's not to say that changes and challenges don't come along every year. I was lucky enough to have had the freedom and time to think, invent and play with it before anyone became aware of what I was doing.

The Proms is a very different proposition, relatively recent in design and with little or no time allocated to think about or modify the system, to try different arrays. In this situation one needs to multitrack the microphones and spend time after the event practising. Productions are reluctant to spend money this way but, in a large organization, this is exactly what the training budget should be used for. Problems and techniques will only be resolved and formulated by practice and reflection.

Practitioners need to talk to producers. They need to understand what's involved, what are the problems and, most of all, what are the benefits that Surround Sound brings to their programmes.

There should be time for discussion, experimentation, practice and trial and error!

... It's the only way.

And so, where to start?

- Start from stereo and add to that.
- You don't have to reinvent the wheel. Use any of the techniques available as a start. Experience and practicality will create change.
- There doesn't have to be something in ALL the channels ALL the time.
- 4.0 can work just as well as 5.0 or 5.1.
- Do whatever suits the subject. Make it as natural as possible unless you are after a specific effect, and above all ...
- **Use your ears!**

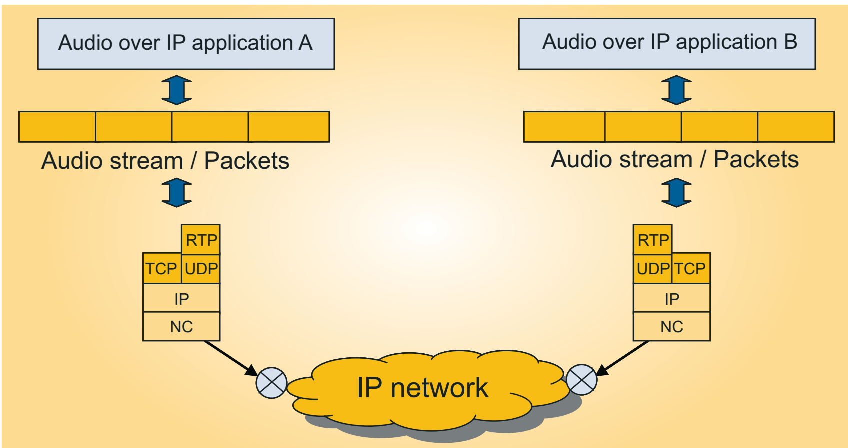
Streaming audio contributions over IP — a new EBU standard

Lars Jonsson
Swedish Radio

Mathias Coinchon
EBU Technical Department

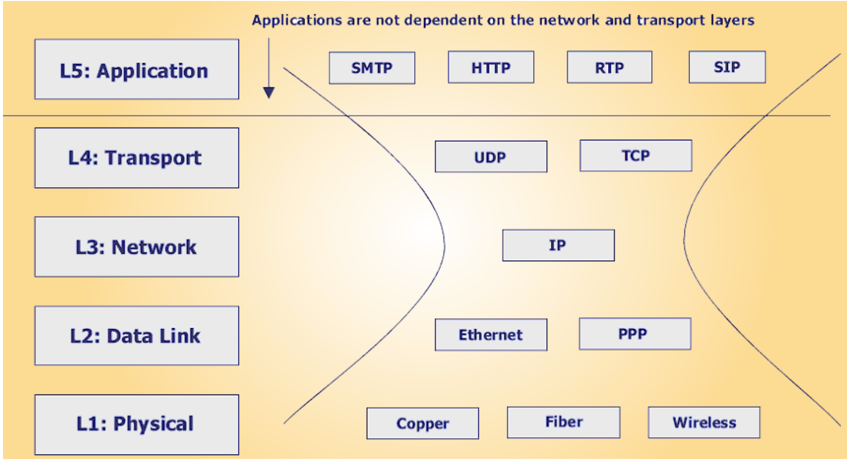
Audio-over-IP end units are increasingly being used in radio operations for the streaming of radio programmes over IP networks, from remote sites or local offices into main studio centres. The IP networks used can be well-managed private networks with controlled Quality of Service. However, the open Internet is increasingly being used also for various types of radio contribution, especially over longer distances. Radio correspondents will have the choice in their equipment to use either ISDN, the Internet via ADSL or other available IP networks to deliver their reports. ISDN services used in broadcasting will be closed down in some countries.

The EBU has created a standard for interoperability in a project group, N/ACIP (Audio Contribution over IP). This standard, which has been jointly developed by members of the EBU group and manufacturers, is published as EBU Tech 3326-2007. The standard has quickly been implemented by the manufacturers. A "plug test" between nine manufacturers, held in February 2008, proved that earlier incompatible units can now connect according to the new standard.



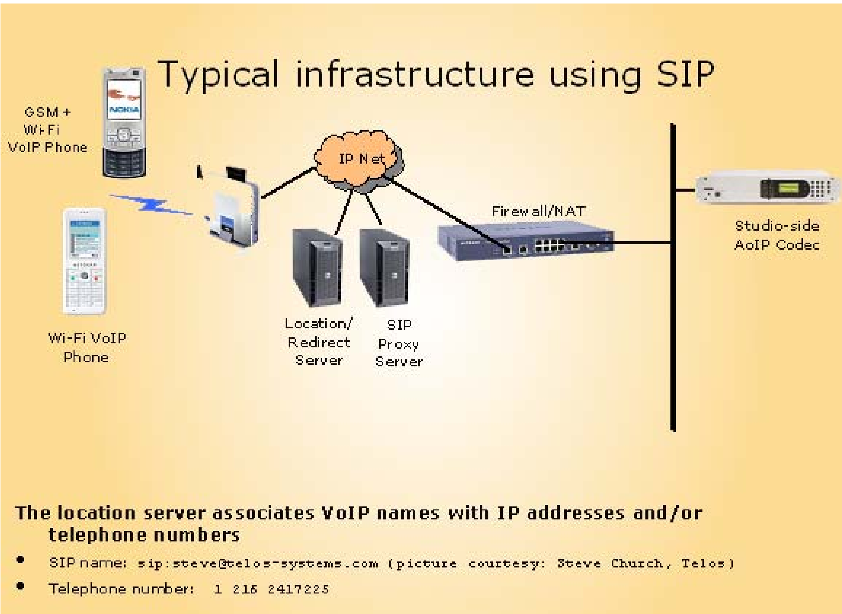
Internet Protocol (IP) is used worldwide on the Internet and also on the IP-based corporate or private networks used by broadcasters. It is independent of the underlying data transmission technology and many IP adaptations exist for the physical layers, e.g. Ethernet, ATM and SDH over copper, fibre or radio links.

Applications can communicate in a standardized way over interconnected IP networks. Host computers can be reached almost instantly wherever they are located. Fixed or temporary connections for audio contributions can share similar types of applications. The connection is established by dialling a number or an e-mail-like name. The audio stream is then sent using standardized protocols (SIP, RTP, UDP). Many types of audio coding formats can be used at various



bitrates. Higher bitrates will allow stereo audio or multi-channel with linear PCM coding. The permitted maximum audio bitrate will depend on the bandwidth and if the network has a good QoS.

The open Internet, as well as closed IP networks, are constantly being improved and are moving towards higher bandwidths, which will allow for the transfer of high-definition TV pictures and high quality audio to all consumers. It will be possible to use both corporate IP networks and the Internet for audio-only contribution and distribution. The cost of using networks will decrease and the Quality of Service will be improved. NGN (Next Generation Networks) is the common name for improved public networks that will offer considerably higher bandwidth.



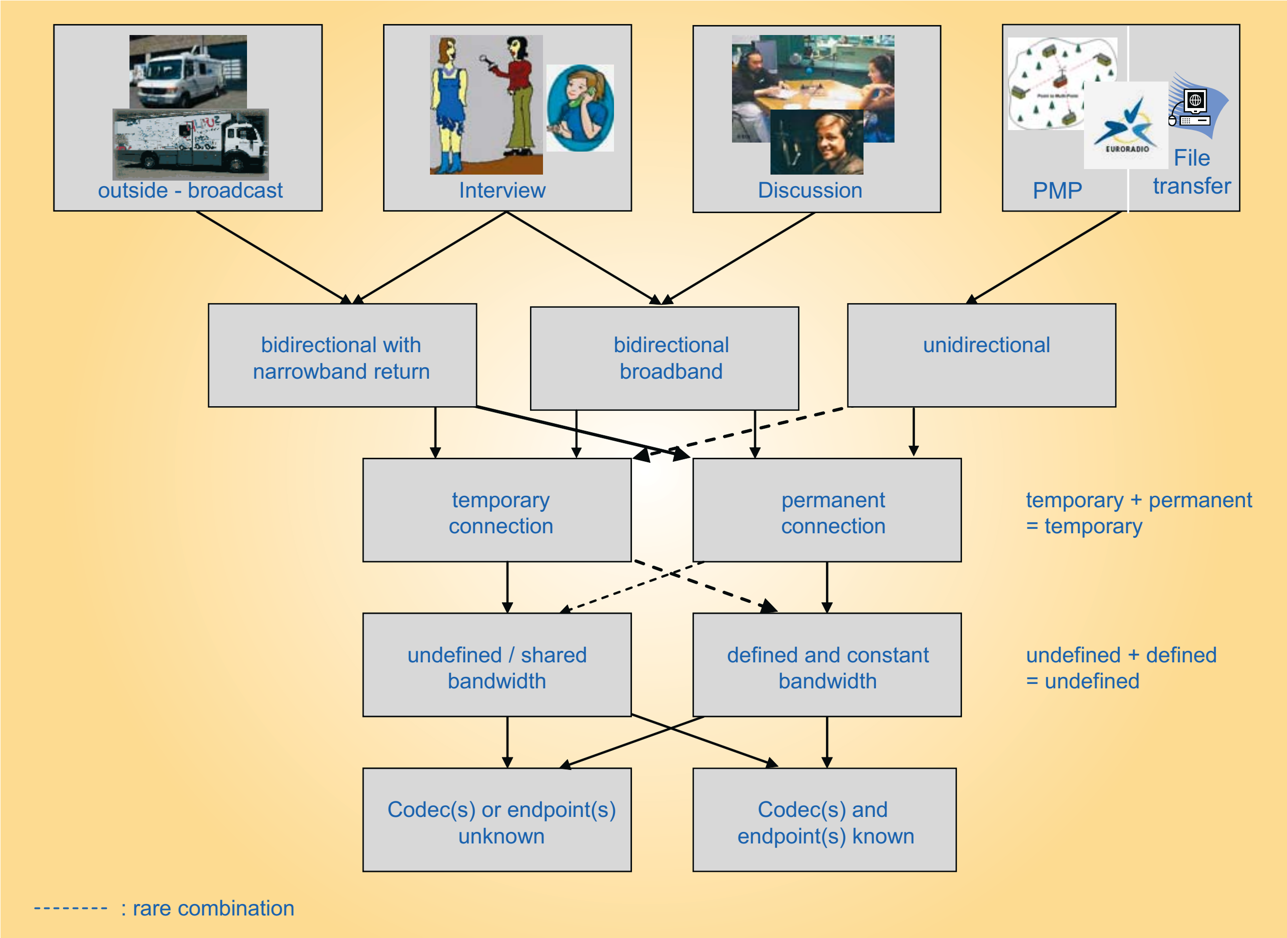
Mobile IP networks will be available with good population coverage in many countries. 3G/ UMTS and LTE systems (4G) will soon offer higher bitrates upstream. WiMAX and WLAN hotspots are other possible evolutions which may offer solutions for radio contribution. However, the robustness required for live contribution may not always be guaranteed in these systems because of interference and shared access. The added value of these new networks will be near-instant access for news reporters in all urban areas. The improvements gained by using the presence functionality in SIP (Session Initiation Protocol), will be that the reporter can easily be reached almost anywhere in the world with one identity, irrespective of the platform being used, such as a mobile phone or Audio over IP codec.

Types of connections

Two types of connections can be used:

- **Permanent connections** – which are generally based on managed private networks with constant and well-known bandwidth and Quality of Service. For permanent connections, the audio codec types are usually known in advance.

- **Temporary connections** – which may be based on previously-unknown networks with shared and unknown bandwidth over the Internet or over temporary leased private networks. The codecs and endpoint may be unknown. The audio codec type can be found through negotiation using the SIP and SDP protocols.



Some of the different types of operations for Audio over IP

EBU standard

In the past, Audio over IP end units from different manufacturers have not been compatible. Based on an initiative coming from German vendors and broadcasters, the EBU started a project group called N/ACIP (Audio Contribution over IP). One of its tasks is to suggest a method for interoperability. After a joint meeting between broadcasters and manufacturers in September 2007, an interoperability recommendation was published, which manufacturers have already implemented [1].

In a plug-test held at the IRT in Munich in February 2008, nine manufacturers demonstrated interoperability. The EBU is working on a reference implementation software, which can be used to verify compatibility.

The requirements for interoperability are based on the use of RTP over UDP for the audio session and SIP for signalling. The packet payload audio structure is defined by IETF RFC documents for commonly-used audio formats in radio contribution, such as G.722, MPEG Layer II and linear PCM. By using SIP, a negotiation can be made to automatically find a common audio coding system on unknown units at each end.

The EBU standard suggests using RTP over UDP rather than TCP. A one-way RTP stream with a small header (low overhead) is more suitable for audio transfers. Moreover, RTP over UDP sometimes has higher priority than TCP in routers.

Further information on the EBU N/ACIP project can be obtained at : **http://www.ebu-acip.org**

Types of equipment

Different types of equipment can be identified:

- General contribution equipment – meant for all types of contribution (fixed or remote)

Example of an ISDN and Audio over IP fixed end unit

- Portable contribution equipment – used mainly for monophonic speech contribution at low bitrates

Example of a portable ISDN and Audio over IP unit

Networks

Telecom operators are offering an increasing number of IP-based services. Traditional services based on ATM, PSTN, ISDN, SDH and PDH will gradually be phased out or will become expensive niche products. They will tend to be replaced by all-over-IP services carried over copper, fibre or wireless links.

Abbreviations

3G	3rd Generation mobile communications	**QoS**	Quality of Service
ADSL	Asymmetric Digital Subscriber Line	**RFC**	Request For Comments (IETF standard)
AoIP	Audio over IP: broadband audio	**RSTP**	Real-Time Streaming Protocol
ATM	Asynchronous Transfer Mode	**RTCP**	Real-Time Control Protocol
HSDPA	High-Speed Downlink Packet Access	**RTP**	Real-time Transport Protocol
HSUPA	High-Speed Uplink Packet Access	**SAP**	Session Announcement Protocol
IETF	Internet Engineering Task Force http://www.ietf.org/	**SDH**	Synchronous Digital Hierarchy
		SDP	Session Description Protocol
IP	Internet Protocol	**SDSL**	Symmetric Digital Subscriber Line
ISDN	Integrated Services Digital Network	**SIP**	Session Initiation Protocol
LTE	Long Term Evolution (4th generation mobile networks)	**SMTP**	Simple Mail Transfer Protocol
		TCP	Transmission Control Protocol
PCM	Pulse Code Modulation	**UDP**	User Datagram Protocol
PDH	Plesiochronous Digital Hierarchy	**UMTS**	Universal Mobile Telecommunication System
PPP	Point-to-Point Protocol	**VoIP**	Voice-over-IP: narrowband audio
PSTN	Public Switched Telephone Network	**WiMAX**	Worldwide interoperability for Mobile Access

Some EBU members have experience of real-world testing of Audio over IP codecs over various types of IP networks. Determining the most suitable type of network from the solutions offered by providers requires careful evaluation.

In particular, measurements and real-world tests with audio must be made by the broadcaster in order to analyse the service level agreement and to verify the performance offered by the network provider. The network must be tested with the applications that are to be used. Long-term testing (months) of uninterrupted audio throughput is recommended.

The distinction between well-managed IP networks and the open Internet is important. On the open Internet, no mechanisms yet exist to achieve a good QoS. The Internet is a “best effort” network with no guaranteed Quality of Service at all. Over a ten-year period the packet loss, delay and jitter over the Internet have slowly improved, but the network performance still poses a major problem to the developers of Audio over IP units.

Last mile access

There are many access solutions to connect the end user to the Internet or private IP networks. Here’s an overview:

Fibre optics.

- This is the highest quality access solution, offering low error rates and low delays. It is ideal for contribution purposes but is still expensive and not widespread.
- Some cities have started to deploy FTTH (Fibre to the Home) access.

Copper with xDSL (ADSL, ADSL2+, VDSL, SDSL).

- This type of access is now in widespread use to make Internet connections. Business providers also have solutions for connecting to private IP networks using xDSL with a guaranteed Quality of Service. This type of solution is preferred to straightforward Internet access for contribution purposes.
- ADSL: Asymmetrical uplink/downlink bitrate.
- SDSL: Symmetrical uplink/downlink bitrate. This type of access is preferred for contribution because of the higher uplink bitrate for sending.

Mobile communication: 3G/UMTS, HSDPA, WiMAX

- Mobile high bitrate accesses to the Internet are starting to emerge. However the problem is the lack of guaranteed Quality of Service. In many cases the access is shared among users of the same radio cell. So, at the moment, this is not ideal for contribution despite the great advantage of mobility. Many operators also filter the traffic and block access for Voice/Audio over IP.
- Some operators plans to offer access to private IP networks in the future. We may expect that future solutions will have better Quality of Service.

Satellite

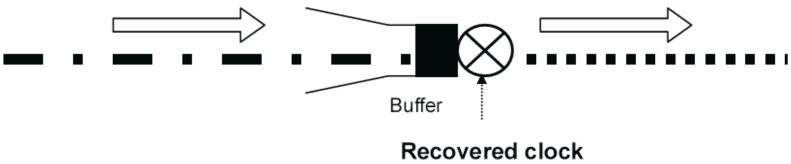
- It is possible to get Internet access through a satellite by using a transmitter system for the return channel. It is used in remote location where all other access technologies are not available. DVB-RCS (Return Channel over Satellite) is used in most cases. The access is generally shared by users and so it is difficult to have a guaranteed Quality of Service. Inmarsat BGAN is also another option.
- It may be possible in the future to have access to private satellite networks with enhanced Quality of Service.
- The delay for the transmission over satellite is generally long (about 500ms roundtrip delay). It is also necessary to have a direct view to the satellite.

Wi-Fi

- Wi-Fi is not really a last mile access but more a home network solution. It is however available as a last mile access in some cities (sometimes for free).
- The frequency band is shared without coordination by many users and also other systems (microwave ovens, DECT telephones). So it is impossible to have guaranteed Quality of Service on such accesses but they may give good results for a local access in not-too-crowded places.

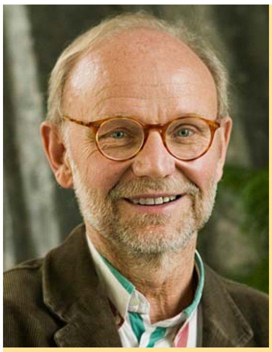
Synchronization

Variations in the delivery time of packets occur, mainly due to varying delays in routers and the sharing of available capacity with other data traffic. Buffering is required to compensate for this variation *(see the diagram)*. There is generally no clock transported so it must be reconstructed at the receiving end. Many different clock recovery algorithms exist. The difficulty is to estimate the slow clock drift correctly and separate it from the short-time network jitter. Streaming audio with high quality is dependant on a guaranteed and stable clock rate at both the sending and receiving ends. Another possibility is to use an external clock source, such as GPS or from a common non-IP network clock.



Delay

Buffers at the receiving end can introduce a considerable amount of delay. The delay buffer size is a trade-off between an acceptable delay and a reliable transmission. In addition, the IP network itself has a delay, from a few tenths of milliseconds in well-managed networks up to 500 ms or more on very long distances over the open Internet. The audio encoding itself may give delays of just a few milliseconds for PCM up to more than hundreds of milliseconds for some bitrate-reduced coding



Lars Jonsson was born in 1949 and received an M.Sc in Electronic Engineering at the Royal Institute of Technology in Stockholm in 1972. He joined the Research Team of Swedish Radio and, in the early days, worked on the development of video and RF systems. Later he worked in Operation and Development at the Local Radio Company. He rejoined the SR development team in 1992.

During the last decade, Mr Jonsson has worked on digital audio quality issues, archiving and the audio computer infrastructure within Swedish Radio. He is a member of several working parties within the Audio Engineering Society and the EBU.

Lars Jonsson is currently the chairman of the EBU Audio Contribution over IP working group, N/ACIP.

Mathias Coinchon was born in 1975 and graduated in 2000 in communication systems engineering from the Swiss Institute of Technology in Lausanne (EPFL), Switzerland, and the Eurecom Institute in Sophia-Antipolis (France). He developed his diploma thesis at BBC R&D in Kingswood Warren, studying and developing propagation analysis solutions for Digital Radio Mondiale (DRM) field trials. He then joined as technical project manager, a startup company called Wavecall, which is active in the development of a physical propagation prediction tool for the mobile telecommunication industry.

After Wavecall, Mr Coinchon spent four years at RSR public swiss radio, first with responsibility for contribution and then in charge of a group dealing with distribution, contribution and IT networks. During this period, he was also actively involved in the technical study group of the Swiss broadcasting corporation (SRG-SSR idée suisse) for the re-launch of DAB digital radio in Switzerland.

Since 2006 Mathias Coinchon has been a senior engineer in the EBU Technical Department. He is currently secretary of the N/ACIP and N/VCIP groups dealing with audio and video contribution over IP. He is also involved in digital radio matters and is vice-chairman of the WorldDMB technical committee. His other areas of work include audio, distribution, open-source software, IP-based TV studios and traffic information. In his spare time, he is involved in helping a community radio station.

formats. The time to fill in packets must also be considered: longer packets will mean increased delay, especially at low bitrates. In the case of a two-way conversation, a total round-trip delay which is less than 50 ms is generally preferred, otherwise a conversation becomes difficult, especially when persons from the general public are interviewed. Experienced reporters may be less sensitive to higher values of delay. When using Audio over IP in combination with video contribution, lip sync will be an issue.

Conclusions

The continuous development of IP networks, combined with more sophisticated Audio over IP end units will lead to more use of this technology in the future. The EBU group N/ACIP has made a proposal for interoperability. Connections over the Internet with different types of telephony and professional units for broadcasting will improve telephone audio quality and worldwide access for reporters. Small handheld units and also software codecs in laptops or mobile phones will provide very efficient tools for reporters. SIP will provide a very powerful way of finding the other end, and negotiate a suitable audio coding format. Fixed Audio over IP units will begin to replace older synchronous point-to-point equipment for contribution of stereo or multichannel audio.

The transition process in relation to Wi95revCo07

Darko Ratkaj
EBU Techncial Department

The first T-DAB planning meeting – held in Wiesbaden, Germany, in 1995 – produced an allotment Plan for T-DAB and a new agreement called "Wiesbaden Special Arrangement, 1995" (Wi95). Most of the countries in Europe obtained two coverages across their whole national territories.

Subsequently, the need for additional T-DAB services intended to cover smaller areas was identified, which led to the second T-DAB planning meeting in Maastricht (The Netherlands) in 2002. In additional to the two coverages already available in the Wi95 Plan, each CEPT country obtained one additional coverage. The original Wi95 had to be revised and it was now called "Wiesbaden 1995 Special Arrangement, as revised in Maastricht 2002" (Wi95revMa02).

Following the GE06 Agreement that resulted from RRC-04 and RRC-06, the T-DAB Special Arrangement was further revised at a meeting in Constanta, Romania, in July 2007. The revised document is now called "Wiesbaden 1995 Special Arrangement, as revised in Constanta 2007" (Wi95revCo07).

The first T-DAB planning meeting took place at Wiesbaden (Germany) in 1995, under the umbrella of CEPT. The meeting produced an allotment Plan for T-DAB and a new agreement called "Wiesbaden Special Arrangement, 1995" (**Wi95**). Most of the countries in Europe obtained two coverages across their whole national territories.

The Wi95 Special Arrangement covered several frequency bands that were at that time considered suitable for the introduction of T-DAB. However, the majority of the agreed allotments were accommodated in the frequency bands 174 - 230 MHz (also called VHF or Band III), 230 - 240 MHz (VHF channel 13) or 1452 - 1467.5 MHz (also called the 1.5 GHz band or simply "L-band"). It is Band III that we are concerned with here.

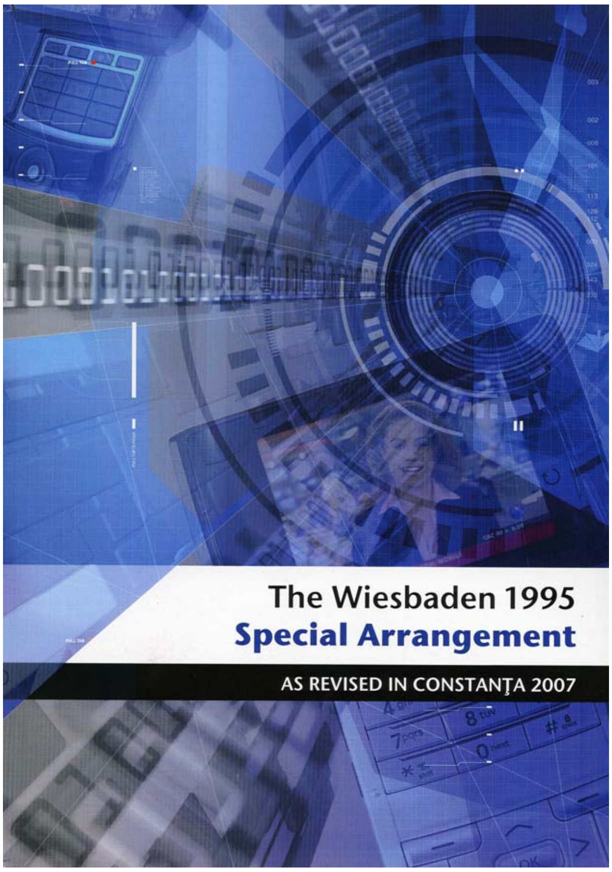
Subsequently, the need for additional T-DAB services intended to cover smaller areas was identified, which led to the second T-DAB planning meeting in Maastricht (The Netherlands) in 2002. In additional to the two coverages already available in the Wi95 Plan, each CEPT country obtained one additional coverage. New allotments were to be found only in the 1.5 GHz band where T-DAB partition was extended to the 1452 - 1479.5 MHz band (a total of 16 T-DAB blocks). Moreover, a new Maastricht Special Arrangement 2002 (**Ma02**) was established to cover all T-DAB allotments in the 1.5 GHz band, including those agreed at Wiesbaden in 1995 which were to be transferred into the new Ma02 Plan.

As a result, the original Wi95 had to be revised and it was now called "Wiesbaden 1995 Special Arrangement, as revised in Maastricht 2002" (**Wi95revMa02**).

That revision had no impact on the allotments in Band III which continued to be implemented, eventually resulting in more than 2000 T-DAB transmitters being put into operation.

Putting the latest revision in context

After the Maastricht meeting broadcasters and administrations turned their attention to the fast-approaching ITU Regional Radiocommunications Conference which was to take place in two sessions; the first in 2004 (RRC-04) and the second in 2006 (RRC-06).



RRC-06 produced a new GE06 Agreement [1] and a set of associated frequency plans. The most important was the new plan for DVB-T and T-DAB in the frequency band 174 - 230 MHz and for DVB-T in the band 470 - 862 MHz. The Conference was declared to have been a great success, in particular for European countries.

However, the Wi95revMa02 Special Arrangement was still in force, signed by more than 30 countries, all of which have also signed the new GE06 Agreement. There was an overlap between the two agreements, i.e. both covered T-DAB in the frequency band 174 - 230 MHz. Something clearly had to be done to avoid a “regulatory conflict” between the two agreements and so the Electronic Communications Committee (ECC) of the CEPT decided to abrogate the parts of Wi95revMa02 relevant to the frequency band 174 -230 MHz.

The actual meeting to revise the Special Arrangement took place on 04 July 2007 in Constanta (Romania). The revised document is now called “Wiesbaden 1995 Special Arrangement, as revised in Constanta 2007” (**Wi95revCo07**).

The Wi95revCo07 document covers the frequency bands 47 - 68 MHz, 87.5 - 108 MHz and 230 - 240 MHz, whereas T-DAB in the frequency band 174 - 230 MHz is governed only by the GE06 Agreement.

What has been revised and how

Abrogation of the frequency band 174 - 230 MHz

The main objective of the 2007 revision was to align the European Special Arrangement for T-DAB with the outcome of the RRC-06, i.e. the GE06 Agreement. This was achieved by removing all references to the frequency band 174 - 230 MHz from the text of the Special Arrangement. The corresponding technical information has also been removed from the relevant technical annexes.

Furthermore, T-DAB allotments in the frequency band 174 - 230 MHz have been deleted from the Wi95revCo07 Plan (Annex 1 to the Special Arrangement).

1. The ITU-R is the guardian of the GE06 Agreement (see **www.itu.int/ITU-R/terrestrial/index.html**).

Abbreviations

CEPT	*Conférence Européenne des Postes et Télécommunications* (European Conference of Postal and Telecommunications Administrations)
DAB	Digital Audio Broadcasting (Eureka-147) **http://www.worlddab.org/**
DVB	Digital Video Broadcasting **http://www.dvb.org/**
DVB-T	DVB - Terrestrial
ECC	(CEPT) Electronic Communications Committee
ERO	European Radio Office of the CEPT
RRC	(ITU) Regional Radiocommunication Conference
T-DAB	Terrestrial - DAB

Finally, bilateral agreements relevant to the frequency band 174 - 230 MHz were deleted. These agreements were concluded between individual administrations at the Wiesbaden 1995 meeting and were concerned with specific implementation conditions. No presumption has been made as to whether or not administrations will want to retain some of these agreements for the transition period.

Protection of the existing T-DAB transmissions

More than 2000 T-DAB transmitters were brought into operation over the years in accordance with Wi95 and, subsequently, Wi95revMa02. It was important to protect these transmitters, at least for a period of time, so that the transition to the new regime (e.g. GE06) can take place without disruption of services. The transition arrangements consist of several elements:

- It was decided that only those T-DAB transmitters will be protected that were formally notified to ERO [2] by 02 July 2007 (shortly before the revision took place). These transmitters were assumed to have been operational at the time of the revision. That cut-off date was announced several months in advance to allow sufficient time for the administrations to update their records.

- The transmitters which met the first condition were recorded in a new Annex II to the Final Acts [3] which was created for that purpose.

- A new Article 2 was included in the Final Acts which reads:

  "The T-DAB assignments in the frequency band 174-230 MHz recorded by 02 July 2007 in the Assignment List in accordance with Article 6 of the Wiesbaden, 1995 Special Arrangement, as revised in Maastricht 2002, as provided in Annex II, shall be protected, taking into account the relevant bilateral agreements reached at the RRC-06, until the date to be agreed by the administrations concerned but not later than 01 January 2012."

- The new Annex II also contains an introductory text as follows:

  "According to Article 2 of the Final Acts, these T-DAB assignments shall be protected, taking into account the relevant bilateral agreements reached at the RRC-06, until the date to be agreed by the administrations concerned but not later than 01 January 2012."

  "Administrations may agree bi-laterally to protect actual service areas of individual assignments or SFN. Protection of the fully implemented allotments can be based on the original allotment parameters."

  "The original allotments cannot be developed any further as the provisions of the Special Arrangement are withdrawn for these frequencies."

---

2. The European Radiocommunications Office (ERO) is the Plan Management Body for Wi95revCo07 (see also **www.ero.dk/tdab/wi95revco07**).

3. The Wi95revCo07 Final Acts are an overlay document that reflects the history and the context of the Special Arrangement. Annex I to the Final Acts is the Special Arrangement itself (including five technical annexes). Annex II to the Final Acts contains a list of T-DAB assignments in the frequency band 174 - 230 MHz recorded on 02 July 2007 in the Assignment List in accordance with Article 6 of the Wi95revMa02.

Darko Ratkaj holds a degree in electrical engineering from the University of Zagreb, Croatia. For more than 15 years he has specialised in radio spectrum management. He worked for several years in the Croatian Ministry of Maritime Affairs, Transport and Communications, and then for eight years in the European Radiocommunications Office (ERO) in Copenhagen.

In the ERO, Mr Ratkaj's responsibilities included support for the CEPT/ECC in the area of frequency management and spectrum engineering, and in the T-DAB Plan management. He was actively involved in CEPT activities concerning the introduction of T-DAB and DVB-T. He took part in the Maastricht conference in 2002 as well as in RRC-04 and RRC-06.

In January 2008, Darko Ratkaj joined the EBU Technical Department where he now works as a member of the spectrum group.

Portable indoor reception

The original Wi95 technical planning parameters are suitable for mobile (outdoor) reception which is not always sufficient – as more and more T-DAB networks nowadays aim at providing portable *indoor* reception. Therefore, an additional provision was included in Annex 2 of Wi95revCo07 which allows for portable indoor reception.

However, it was decided not to include any additional technical parameters for portable indoor reception. Instead, the administrations concerned shall bilaterally agree the bases to be used for co-ordination.

Transition process

Transition provisions are necessary only for those T-DAB networks which operate in the frequency band 174 - 230 MHz and which are not in accordance with the GE06 Plan. Protection of such networks is based on the parameters of the original allotments (i.e. those allotments which existed in the Wi95revMa02 plan before the latest revision in 2007).

The revised Wi95revCo07 Special Arrangement does not contain any specific provisions for modification of the existing T-DAB transmitters as contained in the new Annex II. However, the original technical information has been retained in the relevant technical annexes (which continue to apply in the remaining frequency bands covered by the Wi95revCo07) and can be used in bi- and multi-lateral negotiations.

For the existing T-DAB assignments that are, or will be, included in the GE06 Plan, the relevant provisions of the GE06 Agreement can be used.

For the T-DAB assignments that are not in accordance with the GE06 Plan, there are several possible approaches:

- The networks can be modified / aligned with the GE06 Plan before the end of the transition period (i.e. the end of 2011); or
- The networks continue to operate without change after the end of the transition period, provided that they are either:
  - co-ordinated in accordance with the GE06 rules and included in the GE06 Plan; or
  - protected by means of bi- or multi-lateral agreements, as necessary;
- The transmission ceases not later than 01 January 2012 (the end of the transition period).