IMPROVED SOUND FOR DOMESTIC TV

Ted Fletcher President, Orbitsound Ltd. London.

INTRODUCTION

This paper acknowledges over 50 years of observation, experiment and implementation of sound reproduction with particular reference to broadcast television from the point of view of a professional in the field, leading to the commercial introduction of patented technologies for improving the viewer enjoyment of television in the home.

HISTORY

Even though radio broadcasting is a lot older than television, TV sound was always an 'ad-on'….
By the time regular TV broadcasts were being made, the sound of 'radio' had become sophisticated, with spacious loudspeaker cabinets offering a warm friendly sound…. In 1936 the BBC started the very first so called 'high definition' television broadcasting from Alexandra Palace and the technology used for the sound was the same as for the regular radio broadcasting, it was Amplitude Modulation; notorious for high levels of distortion and subject to atmospheric magnetic interference, but because the vision part of the spectrum needed good signal strength, the sound quality was, in general, acceptable to the viewer; certainly not noticeably inferior to the vision content!

The first TV receivers in the US did not even have loudspeakers fitted to them….. Instead, they had an output to connect to a radio to get the sound!
But quickly, TV design engineers incorporated latest developments in how good quality could be achieved, Broadcasters started to use FM (Frequency Modulation) for the sound so that the limitations to quality were mostly in the receivers.

MONO OR STEREO

The origins of stereo sound go back as far as the mid 1930's and early research into how to make cinema sound more realistic.
Systems were developed both in the US and at EMI laboratories in the UK aimed at making the sound of the actor come from the correct place on the screen.
With a wide viewing angle of a cinema screen it was not difficult to achieve a useable effect of direction and depth, particularly as sound could be projected physically through the screen.
However, with television the problem was more difficult; it was clear from the start that in the interests of realism, TV sound should come from a source close to the screen, and in the early days screen sizes were limited by the practical limits of Cathode Ray Tubes…… a 24 inch screen was a monster, very expensive, fragile and heavy!
One of the interesting problems with early TVs was that the magnetic flux from loudspeaker drivers, by definition, had a seriously bad effect on Cathode Ray Tubes, particularly colour CRTs, so while loudspeakers needed to be close to the screen, proximity was a problem.
But as with so many technical problems, the answers came swiftly, first with magnetically screened loudspeaker drivers, and then, when flat screen TVs started to become economically viable, magnetic flux ceased to be a problem.

From the 1970s onwards, all TVs (apart from a few portable types) have been fitted with so-called 'stereo' loudspeakers, sometimes in a small loudspeaker area below the screen and in a few cases, with those magnetically screened loudspeakers on each side of the screen.

One can only assume that the marketing demand must have been there for ‘stereo’ sound, even though 99% of broadcast output was mono.
But over the years there has been a subtle change in the attitude of broadcasters….. while the main components of the sound from transmissions continues to be mono, sound engineers have been using both stereo and binaural mixes of sounds in order to achieve atmosphere and space; this is very much in evidence in wild-life programmes and often in drama.
But the broadcasters have been fighting a losing battle with the commercial needs of the consumer….. Styling and cabinet/enclosure size have controlled domestic TVs for decades; the loudspeaker drivers have consistently and steadily got smaller and closer together and in spite of manufacturers’ strident claims of ‘perfect sound’ reproduction and resounding bass, the truth is that quality has steadily deteriorated.

QUALITY INNOVATION

In 1986, after more than 15 years of research, the BBC introduced a form of digital transmission called NICAM. This enabled broadcasting of a 2-channel audio at improved quality over the conventional FM system. (In modern terms the system equated to 750bps, 14 bit 32KHz sample rate).
All transmissions were in ‘stereo’ although the BBC did not publicise this until 5 years later in 1991!
Although the technology was superior, it largely ‘fell on deaf ears’ as by then, ‘flat screen’ TVs were being introduced with their severe limitation of room for sound reproduction.

Digital television switchover, with its further improvement in sound transmission, started in the UK in October 2007 and was completed on 24th October 2012 and throughout that period the problems of TV sound reproduction went from a minor irritation for many viewers, to a serious and widely recognised problem. All the TV channels were able to transmit sound of excellent quality while only a handful of television receiver manufacturers made any attempt to take advantage of the available quality except perhaps to allow for the connection of external audio equipment by fitting both analogue and digital outputs.

TECHNICAL INNOVATION IS NOT ALWAYS POSITIVE

Over the years since the early 1970s many manufacturers made what seemed like an innovative decision to try to produce ‘immersive’ sound systems by introducing surround sound. The basic idea of reproducing dialogue from a central loudspeaker and music and effects via strategically placed satellite loudspeakers seemed logical and was compelling enough to make the introduction of ‘5.1’ and similar systems popular for a while…. But there are serious drawbacks.
In any 5.1 system, the only place where any sort of reasonable directional sound can be appreciated is in a single ‘sweet spot’. Any deviation creates different path lengths for the sound from the different loudspeakers and produces interference and cancellations.
The problem is made worse by acoustic reflections in the listening room to an extent that even the so called finest and most expensive systems are disappointing and most are so poor that they can only be called toys.
The performance problems coupled with the required complexity of the installation (wires all over the place) have meant that surround sound continues to be a diminishing market.
(This should not be confused with binaural recording and reproduction, which can be highly successful and applicable to spatial sound for television and cinema.)

COMMERCIAL INNOVATION

It was in 2003 that I first had the idea of producing spatial sound from a single point loudspeaker, and it had absolutely nothing to do with television sound. I was recording some backing vocals; a small choir, using ‘sum and difference’ technique where you can record a wonderfully realistic ‘sound stage’ with a combination of microphones all in the same place.

(For the sake of completeness, for those who may not be familiar with this technology; when recording say a choir, place a cardioid (unidirectional) microphone in front of the choir positioned so that it 'hears' the whole choir. Then place a figure-8 (bi-directional) microphone close to it at 90 degrees to the direction of the choir. The cardioid microphone picks up the whole choir; that is both the left and the right. The figure-8 microphone also picks up the L and R but the two sides are out of phase, or more correctly, reversed polarity, so it is picking up L minus R. With some simple manipulation on a mixer, by adding the outputs of the two microphones you achieve (L+R) + (L-R) which sums to Left only. Similarly by subtracting the sound of the figure-8 microphone from the cardioid microphone, you achieve (L+R) – (L-R) which resolves to Right only! So you have achieved separate left and right signals from a single microphone position.)

It seemed quite obvious to me that the same simple manipulation could be used to create 2-channel so-called 'stereo' sound from an arrangement of loudspeakers placed very close together.
In the space of a single day I cut up some cardboard tubes and using lots of gaffer tape created the reverse of the recording microphones, using small full-range loudspeaker drivers.
Then results were encouraging in that suddenly I could discern a sense of space around the loudspeaker.

It actually took a further five years to turn the original concept into a commercial reality. The development, financing and organisation of a new technology company is quite another story and not relevant to this paper.

Meanwhile the quality of sound on increasingly slender television receivers had deteriorated to a point where it could not be ignored; there were mutterings in both the technical and the mainstream press about how no-one seemed to care about sound any more, how digital sound is so much worse than we were all used to in the latter years of 625 line UHF colour TV…… missing the point entirely that the problems were at the receiving end, not the transmissions.

A further and even more serious problem was the rapid take-up of digital media via the domestic TV receiver; the huge increase in sales of DVD and similar good quality media as well as changes in the way TV was being used in the household. The demand for a very much higher sound quality had become a clamour!

THE SOUNDBAR

The introduction of the 'soundbar' was inevitable; at Orbitsound we possibly could claim to have invented it in around 2005, but it is likely that a number of manufacturers were led to the same basic idea at the same time due to the obvious demand for a solution to the problem.
Early 'soundbars' were little more than re-proportioned hi-fi loudspeaker enclosures, elongated so that there was a loudspeaker at each end in an attempt at 'stereo' separation.
Later 'bars have mainly stuck to the principle of 2-speaker stereo convinced that better sounding television is just a matter of using slightly bigger loudspeakers.
A few manufacturers (specifically Yamaha) applied much more thought and research to the problems of TV sound and decided on the novel approach of Wave Field Synthesis (WFS).
This technology can successfully produce a realistic sound stage, but unfortunately to work at all it requires something like 30 or more individual audio channels each with its own processor and loudspeaker, and once it has been set up, it further requires an acoustically 'dead' environment to be effective.

AN OBJECTIVE VIEW OF REQUIREMENTS FOR TELEVISION SOUND

The primary requirement is obviously that the viewer should be able to hear as well as see the broadcast content; these are the requirements in order of importance:

1) The viewer must hear dialogue clearly and understandably.
2) The sound should be loud enough for (1) without being overpowering.
3) The 'timbre' of the sound must be suitable for the listening environment.
4) The type of sound must suit the content; normal TV or feature film.
5) The sound must be as immersive as possible.
6) The sound must be controllable.

The first and most important requirement seems obvious, but dialogue is degraded by having the same signal source reproduced from loudspeakers that are separated by even a small distance; the 30 to 50 cms distance that is normal for TV loudspeaker drivers offers the worst possible listening experience for all except a viewer exactly equidistant from the loudspeakers. For an 'off-centre' listener, the time difference from the two drivers introduces acoustic interference and cancellations that, while not sounding obvious, have a tiring effect on the listener and a general dissatisfaction with the sound.
Technically, the velocity of sound is 340 metres/sec. The half-wavelength of a sound of 4KHz (a frequency that is most important for intelligibility) is 42.5 cms…. About the distance between the loudspeaker drivers on a modern TV, so to the listener, there will be increasing cancellation of mid and high frequencies 'off axis' of the loudspeakers.

The second requirement again seems obvious, but many smaller televisions have seriously limited audio capability.

The 'timbre' of the sound is important because it would not be acceptable in a domestic setting to offer realistic reproduction over the whole audible range; too much of the extremes of low frequency or high frequency would immediately cause distress as well as testing the acoustic insulation properties of the room!
This might not seem obvious; one might assume that the highest possible 'quality' should be the goal of a designer of TV sound systems, but our own experiments and experience clearly show that clarity of the mid audio ranges is of greater importance for viewer/listener comfort than extended frequency response.
The BBC carried out similar work in the 1960s and showed that radio listener comfort could be retained even when both top end and bottom end frequencies were controlled together; that is, if the response at LF is restricted to say 150Hz, then as long as the HF response was restricted to say 6KHz, listening remained 'comfortable' and balanced.
At Orbitsound we have taken this experimentation much further and have shown that this is a factor affecting intelligibility.

Factor No 4 could be taken as a part of factor 3; it is clear that while the severely restricted range of little in-built loudspeakers might be adequate for listening to the news or the score in a game show, a well-produced drama or a feature film demands good volume, good dynamic range and reasonable frequency response.

Factor 5 follows on…. For a feature film to be properly enjoyable it must be immersive; the viewer must become involved with both the vision and the sound, and here is where there is a need for what is loosely called 'stereo'; there has to be a believable effect of 'depth' to the sound.
Pin-point accuracy of image definition is not necessary and actually not desirable as the television screen occupies such a narrow viewing angle, but if a feeling of space is provided then the experience and reality is considerably enhanced.

Factor 6 may seem trivial, but control of the main aspects of the sound has to be made by the viewer from his viewing position; the remote control must be an efficient part of the system.

A UNIQUE SOLUTION TO THE PROBLEMS

In 2008 as a newly formed company, Orbitsound Ltd., we applied for patents for a number of implementations of my work with single-point spatial sound, we christened the system 'airSOUND' and were able to demonstrate a practical new 'soundbar' based on this technology to the commercial world; we demonstrated it to the John Lewis Partnership buyers and to a few other independent retail outlets. It was universally acclaimed and the Orbitsound 'T12' product appeared on the shelves.

While the soundbar was by no means iconic looking, it addressed the requirements from first principles and incorporated the fundamentals of 'airSOUND'; The main intelligible sound (the sum of left and right) is reproduced from a central front loudspeaker driver, while the spatial information (basically the left minus right) is reproduced from a dipole driver (pair of drivers) across the axis of the 'main' signal.

The theory of co-incident microphones is fine as a first approximation, but there are flaws in the logic in that the 'lefts' and 'rights' referred to, have differences between the outputs of the cardioid and the figure-8 microphones. Similarly, in an airSOUND loudspeaker, subtle filtering is needed to retain a perfect realistic sound from such a variety of source material that comes via our TV receivers.

Since the introduction of the original T12 soundbar we have continued active research and development so that present day airSOUND systems perform to satisfy the requirements:

DIALOGUE…. Dialogue is by definition a 'mono' source and unlike conventional systems, it is reproduced from a single central loudspeaker driver; it is a 'real' signal rather than a 'virtual' image being reproduced from somewhere between a pair of loudspeakers. This means that the sound is clear and undistorted by audio cancellations. If the dialogue has spatial information with it (depth and position) this information is played out via the 'spatial' dipole drivers. The effect of this is that dialogue is subjectively clear, intelligible and in its place in the audio sound-field.

POWER…. The power required for normal television viewing is remarkably small; actual power to the loudspeaker drivers is normally lower than a single watt, however, to appreciate the dynamic range of film soundtracks, a power reserve of something like 100 watts may be needed. With such an overload margin provided, the sound can remain clean and undistorted at all volume levels and handle the huge dynamic range necessary for that immersive sound.

TIMBRE… Many potential buyers of sound systems mourn the days of the 'radiogram', a piece of furniture characterised by polished wood. The reality is that such equipment did possess a unique 'timbre' of sound, a soft warm effect that is usually missing from equipment moulded from plastics. Our early development showed without any doubt that manufacturing from wood composites could retain that pleasing character.

TYPE AND IMMERSION… Conventional systems still stick to producing a left hand and a right hand signal from a pair of loudspeakers close together even though the production of any stereo effect is negligible. AirSOUND technology is able to reproduce the intended spatial effects, not just to a single sweet spot, but to the entire listening/viewing area. With present day high definition pictures and the undistorted spatial sound this vastly improves the ability of the viewer to become immersed in the programme.

While a buyer of a sound system may demand a sub-woofer with thumping bass response, our practical research shows that extreme frequency range can be detrimental.

In practice, for television and film viewing, extending the frequency ranges below say 60Hz and above say 8KHz offers no subjective advantage, however, the degree and type of filtering does have effect.

AIRSOUND VERSUS STEREO, THE DIFFERENCES.

2-speaker stereo reproduction tries to achieve realistic spatial sound by providing left signals to a left ear and right signals to the right. In a carefully controlled environment the depth and positioning effects are achieved, although any secondary reflections in the room degrade the imaging. If the loudspeakers are positioned 1.6 metres apart and the listening position is

about 2 metres back from the axis of the loudspeakers, then a good realistic sound picture is possible, but any deviation of the parameters degrades the sound picture.
AirSOUND produces the 'main' information from a single loudspeaker driver in the centre, and the 'spatial' information from a dipole loudspeaker. The spatial information mixes with the 'main' information and the listener hears an acoustic mix that is subtly different across the audio sound-field. The key factor is that as all the reproduced sound comes from a single point then there is no timing error from different loudspeaker drivers; the sound is effectively the same from ANY listening position.
Conventional stereo attempts to 'override' the acoustics of the listening room and impose the acoustic environment of the recording. If the listening acoustic environment is at all 'lively' and less than perfect, acoustic reflections will constantly 'muddy' the sound.

AirSOUND is able to involve the acoustic deficiencies of the room; the recorded sounds actually stay sounding natural because all the sounds originate from a single point, and so a true spatial image is reproduced in any environment.

Ted Fletcher
August 2016