REPORT 1081-1*

THE RELATIVE TIMING OF SOUND AND PICTURE SIGNALS

(Question 35/11)

(1986 - 1989 - 1990)

1. Introduction

A television programme may be impaired by a lack of synchronism, that is a relative timing error, between the sound and picture signals.

Loss of time relationship between sound and picture can occur in two principal areas:

- in equipment used for studio effects or interconnect and pre-emission processing; such as video digital production effects units, synchronizers, standards converters and audio digital encoding equipment. This equipment will of necessity be different for the sound and picture signals and therefore may introduce appreciable differential delay between these components. Although the amount of differential delay introduced by individual items of equipment may be relatively small, the cumulative effect occurring when such equipment is cascaded in the programme path may be significant.
- where sound and picture are sent over separate circuits, particularly if one circuit is terrestrial and the other satellite.

2. Effects of variations in the relative timing of sound and picture

In real life it is unnatural for a sound stimulus to be perceived before the corresponding visual stimulus. Sound is normally delayed with respect to what is seen but the degree of acceptable delay is based on subjective experience. In television broadcasting, this allows a certain tolerance of the onset of the point of subjective annoyance with increasing delay. There will be inherent delays due to fundamental factors such as source to microphone and reproducing loudspeaker to listener distances. The subjective stimuli will be affected by the choice of camera shot e.g., long, medium or close-up, the latter often provided by long focal length telephoto lenses.

Studies have been carried out by the EBU [CCIR, 1986-90a, b] and in Australia [CCIR, 1982-86] which suggest that delay of sound with respect to picture when in the order of 40 ms is often detectable. This corresponds to a path length in air from source to listener in the order of 13 metres.

---

* This Report should be brought to the attention of Study Group 10 and the CMTT.

Table I summarizes the conclusions of these studies:

TABLE I

CCIR input documents	Criteria	Sound advanced with reference to vision	Sound delayed with reference vision
Australia [CCIR, 1982-86]	"detectable" "subjectively annoying"	20 ms 40 ms	40 ms 160 ms
UK [CCIR, 1986-90a]	recommended limit (EBU) for outputs intended to feed broadcasting transmitters	40 ms	60 ms

These results compare with those reported in Report 412.

3. Methods of adjustment

In practice, adjustment is performed by introducing additional delay into the sound or picture signal. It is generally preferable to delay the sound because this is less expensive.

Analogue sound signals, unfortunately, contain no timing information suitable for controlling the operation of equipment which can automatically correct timing differences that may accumulate in a programme's path to final transmission. Manual correction can be performed with difficulty by a skilled operator by observation of the audible and visual cues. Ideally, alterations of the relative delay of audio signals should be made only during silence as manual changes can create disturbing effects. It is therefore more desirable to carry out such correction before commencement of transmission, possibly during a rehearsal, e.g., by means of a manual "clapper board" such as used in the film industry or an equivalent test signal. Where changes of the relative video timing varies during the course of a transmission it would be desirable to produce a control signal to automatically insert a corresponding audio delay. An example of one type of such equipment, which is now available, in at least one television standard, is a frame synchronizer incorporating a matched audio delay unit.

An Australian contribution proposes that one method could use the channel status data which is specified in Recommendation 647 for the digital audio interface. The Document (CCIR, 1986-90c) describes a method for deriving an 11-bit code which could be added to the video signal for the use in downstream restoration of the video and audio time relativity.

In the case of digital sound signals, timing information could be incorporated into the data to control automatic correction equipment.

Digital processing equipment treating both sound and picture should be capable of matching process delays to ensure there is no relative timing error at the output.

4. Accumulation of delays

In longer circuits involving several organizations, it is not uncommon for small delays to be introduced by each organization. While each individual delay may be acceptable, the cumulative error can become excessive. Moreover, when a programme is compiled from different segments received from different origination points each having different relative timing errors, subsequent retiming is impractical.

5. Conclusion

It is important that the relative timing of sound and picture be kept to a minimum to avoid unacceptable impairment to the programme. It is desirable that each organization processing the signal ensures that the timing is correct at the point of exchange.

Even with relatively uncritical material, experience suggests that a sound advance in the order of 40 ms, or a delay in the order of 160 ms will be disturbing to a significant number of viewers. Where the programme contains frequent audible and visual cues, e.g., lip synchronization, the relative timing error exceeding 20 ms advance or 40 ms delay has found to be detectable.

Further studies are required to minimize these timing errors and to standardize the methods of correction.

REFERENCES

CCIR Documents

[CCIR, 1982-86]: 11/364 (Australia).

[CCIR, 1986-90]: a. 11/33 (United Kingdom); b. 11/54 (EBU); c. 11/564 (Australia).

REPORT 1237

SATELLITE NEWS GATHERING

(1990)

The text of this Report as jointly approved by Study Groups 10, 11 and CMTT can be found in the Annex to Volume XII.
