

RECOMMANDATION UIT-R BS.1548

Spécifications utilisateur en matière de systèmes de codage audio pour la radiodiffusion numérique

(Question UIT-R 208/10)

(2001)

L'Assemblée des radiocommunications de l'UIT,

considérant

- a) que le système de son stéréophonique multicanal avec ou sans image associée est l'objet de la Recommandation UIT-R BS.775;
- b) que le codage audio pour la radiodiffusion de télévision numérique par voie de Terre est l'objet de la Recommandation UIT-R BS.1196;
- c) que le codage audio à faible débit binaire est l'objet de la Recommandation UIT-R BS.1115;
- d) que le système de codage préconisé dans la Recommandation UIT-R BS.1115 offre des modes de codage monophonique et stéréophonique à deux canaux;
- e) que la qualité sonore de base et la qualité de l'image sonore stéréophonique que doivent présenter les systèmes sonores de télédiffusion et de radiodiffusion sonore doivent être les plus hautes possibles, et généralement indiscernables par rapport à celles de la source;
- f) qu'il faudrait tenir compte de l'interopérabilité et de l'exploitation des réseaux assurant des liaisons radiophoniques, par exemple de contribution ou de distribution;
- g) qu'il faudrait tenir compte de l'interopérabilité avec les équipements audio multivoies grand public existants, tels que décodeurs d'ambiophonie matricielle et décodeurs multicanaux discrets;
- h) que, lors de la mise en application de la Recommandation UIT-R BS.1115 d'un système sonore multicanal dans un service de radiodiffusion existant, il faut tenir compte de la compatibilité nécessaire avec les récepteurs existants pour pouvoir continuer d'assurer le service;
- j) que, plus généralement, étant donné les nombreuses applications auxquelles se prêtent ces systèmes, il faudrait en déterminer clairement toutes les exigences techniques, qualitatives et opérationnelles;
- k) que la qualité de fonctionnement des systèmes de codage audio multicanaux répondant à la Recommandation UIT-R BS.1196 est largement tributaire de la configuration suivant laquelle est exploité le système (débit binaire, utilisation d'un prématriçage, utilisation d'un codage composite, etc.);
- l) que plusieurs services de radiodiffusion font déjà usage ou ont déjà spécifié l'utilisation des systèmes préconisés dans la Recommandation UIT-R BS.1196;
- m) qu'en conséquence les radiodiffuseurs ont un besoin urgent de disposer de toutes les informations qui sont nécessaires à l'établissement de tous les paramètres de codage disponibles pour les systèmes préconisés dans la Recommandation UIT-R BS.1196;
- n) que la mise en service de systèmes présentant des caractéristiques de qualité de fonctionnement analogues, mais incompatibles, n'est pas du tout souhaitable;

o) que les radiodiffuseurs qui n'ont pas encore démarré de services doivent être à même de choisir le système le mieux adapté à leur application et présentant le meilleur rapport coût/efficacité,

recommande

1 que les systèmes de codage audio destinés à la télédiffusion et à la radiodiffusion sonore numériques pour les applications de contribution et de distribution soient conformes aux spécifications énoncées dans l'Annexe 1;

2 que les systèmes de codage audio destinés à la télédiffusion et à la radiodiffusion sonore numériques pour des applications d'émission soient conformes aux spécifications énoncées dans l'Annexe 2.

ANNEXE 1

Spécifications pour la contribution et la distribution

Les systèmes de codage audio destinés à la télédiffusion et à la radiodiffusion sonore numériques pour les applications de contribution et de distribution doivent être conformes aux spécifications énoncées ci-dessous.

1 Spécifications de service

1.1 Configurations des canaux

Pour les services audio, il faut pouvoir disposer, selon les besoins des applications visées dans la Recommandation UIT-R BS.775 – Système de son stéréophonique multicanal avec ou sans image associée, des configurations de canaux suivantes:

| | | |
|----------|------------|---|
| 1 canal | 1/0 | Monophonique |
| 2 canaux | 2/0 | Gauche, droit |
| 3 canaux | 3/0 2/1 | Gauche, droit, centre Gauche, droit/ambiophonique |
| 4 canaux | 3/1 2/2 | Gauche, droit, centre/ambiophonique Gauche, droit/ambiophonique gauche, ambiophonique droit |
| 5 canaux | 3/2 | Gauche, droit, centre/ambiophonique, gauche, ambiophonique droit |

ainsi qu'éventuellement d'un canal d'amélioration basse fréquence (CABF).

Pour la contribution, il pourrait être en outre nécessaire d'acheminer des programmes produits dans d'autres formats que ceux visés ci-dessus, par exemple 3/4, de sorte que le système de codage devrait permettre de prévoir des canaux supplémentaires de haute qualité.

1.2 Attribution souple des canaux

Un train de bits doit fournir des données d'identification pour la signalisation et la commande des configurations des signaux sonores. Il doit être possible dans le système de transmission de passer dynamiquement de l'une à l'autre des configurations de canaux mentionnées à la section 1.1.

1.3 Données auxiliaires

Le système de codage audio doit permettre de transmettre des données auxiliaires, lesquelles peuvent véhiculer différents types d'information, tels que commande de la dynamique, réglage du volume sonore et données d'utilisateur, ainsi que toutes métadonnées nécessaires au codeur à l'émission qui codera le signal audio final destiné aux auditeurs.

2 Spécifications de qualité de fonctionnement

2.1 Qualité audio

2.1.1 Qualité audio de base

La qualité du son reproduit au terme d'une cascade de contribution/distribution de référence (cinq codecs de contribution et trois codecs de distribution en tandem) ne devrait pas pouvoir être subjectivement distinguée de celle de la source pour la plupart des types de programme audio. Pour ce faire, il faut obtenir, en utilisant la méthode en double aveugle à triple stimulus et référence dissimulée, exposée dans la Recommandation UIT-R BS.1116 – Méthodes d'évaluation subjective des dégradations faibles dans les systèmes audio y compris les systèmes sonores multivoies, des notes moyennes généralement supérieures à 4,5 sur l'échelle de dégradation à 5 notes, pour des auditeurs occupant la position d'écoute de référence. La note la plus basse ne devrait pas être inférieure à 4 (Recommandation UIT-R BS.775).

NOTE 1 – L'intervalle de confiance (barre d'erreur) associé à la note moyenne unique correspondant à un codec et un élément détermine la marge au-dessus et au-dessous de la note moyenne en question à l'intérieur de laquelle peut se situer la note réelle, avec un taux de certitude donné, habituellement de 95%. La note réelle peut ici atteindre la limite inférieure de l'intervalle de confiance correspondant à la note indiquée. Pour procéder à une évaluation significative de la qualité de fonctionnement attendue des codecs en cascade, l'intervalle de confiance associé aux notes moyennes indiquées pour chacun des codecs doit être approximativement égal ou inférieur à la différence existant entre les notes qui sont comparées.

NOTE 2 – Les opérations de contribution et distribution en cascade ne devraient pas, lorsqu'elles sont associées en tandem au codec d'émission, entraîner de réduction importante de la qualité par rapport à la qualité audio de base du codec d'émission. La spécification précise devra faire l'objet d'un complément d'étude.

NOTE 3 – Les paramètres de qualité objective audio pour les applications de contribution/distribution pourront être ultérieurement intégrés, conformément à la Recommandation UIT-R BS.1387.

2.1.2 Résolution de quantification

La résolution exigée devrait être au moins de 18 bits pour la distribution et de 20 bits ou plus de préférence pour la contribution.

2.1.3 Fréquence d'échantillonnage

Conformément à la Recommandation UIT-R BS.646 – Codage à la source des signaux audio-numériques dans les studios de radiodiffusion, la fréquence d'échantillonnage doit être de 48 kHz.

2.1.4 Largeur de bande

Canaux audio principaux: 20-20 000 Hz.

CABF: 15-120 Hz.

2.1.5 Accentuation

Le système de codage audio devrait être dépourvu d'accentuation.

2.1.6 Nombre de codecs tolérables en tandem

Le nombre de codecs pouvant être utilisés en tandem dépend de l'application selon le tableau ci-dessous:

| | |
|--------------|---------------------|
| Distribution | 3 codecs en cascade |
| Contribution | 5 codecs en cascade |

Ces chiffres proviennent d'expériences précédentes réalisées pour évaluer des systèmes de radio-diffusion sonore à deux canaux (voir la Recommandation UIT-R BS.1115 – Codage audio à faible débit binaire) et peuvent ne pas être représentatifs des conditions d'exploitation pratiques des systèmes de radiodiffusion et de télédiffusion. Ce point devra être précisé à l'aide d'informations supplémentaires.

2.1.7 Possibilité de postproduction

Le traitement après production à réaliser est fortement tributaire de l'application. Pour la distribution, on peut appliquer des fondus-enchaînés ainsi que la commande de la dynamique. Il est plus difficile de déterminer le traitement des signaux qui peut intervenir entre des liaisons de contribution.

2.2 Retard dû au codage

Le retard dû au codage doit être identique pour tous les canaux d'un programme. Il doit être aussi court que possible, compte tenu de la qualité du codage (c'est-à-dire quantité de réduction du débit binaire) exigée. Dans le cas du son télévisuel, le retard dû au codage des signaux audio doit correspondre à celui de la vidéo. Il est souhaitable que le codeur audio produise des trames audio codées (unités d'accès) correspondant exactement à la durée de la trame vidéo correspondante.

2.3 Protection contre les erreurs

Il faut prévoir dans le train des bits audio un mécanisme permettant au décodeur de repérer des erreurs résiduelles dans les canaux et de les masquer.

2.4 Délai de récupération

Le délai de récupération devrait être aussi court que possible. Dans le cas où on utilise les unités d'accès audio (AAU, *audio access unit*), il devrait être de quelques AAU au maximum, et le son devrait être rétabli de préférence dès la réception de la première AAU ne comportant pas d'erreur.

3 Spécifications fonctionnelles et opérationnelles

3.1 Débit binaire et système de codage

Pour les liaisons de distribution et de contribution, la Recommandation UIT-R BS.1115 préconise d'utiliser le système de codage de Couche II de l'Organisation internationale de normalisation (ISO)/Commission électrotechnique internationale (CEI) IS 11172-3 avec un débit binaire d'au moins 180 kbit/s par canal ou plus. Pour plusieurs raisons, il est possible d'utiliser le système à un débit binaire plus élevé ou d'employer d'autres systèmes.

Ces raisons sont, entre autres, les suivantes:

- il est possible d'insérer entre des générations de codage une marge supplémentaire pour permettre le traitement des signaux (cette possibilité n'a pas été testée ni vérifiée lors de l'élaboration de la Recommandation UIT-R BS.1115);
- pour obtenir un débit binaire inférieur dans la liaison de distribution et de contribution;
- pour obtenir une meilleure qualité;
- possibilité de synchronisation et de commutation avec des signaux vidéo associés.

3.2 Codage composite

Un élément de programme bicanal ou multicanal présente souvent une certaine corrélation statistique intercanal. Le codage composite peut être une bonne solution pour réduire la redondance intercanal ou pour éliminer les signaux qui n'ont aucun effet du point de vue de la discrimination entre les canaux et accroît donc l'efficacité du codage. Certains systèmes de codage utilisent des critères perceptuels pour éliminer une partie de ces signaux en réunissant plusieurs canaux dans des régions fréquentielles où l'oreille humaine peut difficilement repérer d'où vient le son. L'inconvénient de cette technique est qu'il n'est en général pas possible de bien repositionner ultérieurement les informations sonores dans les canaux d'origine. Pour les applications de contribution et pour de nombreuses applications de distribution, il convient de ne pas utiliser pareils systèmes de codage composite.

ANNEXE 2

Spécifications pour l'émission

Les systèmes de codage audio destinés à la télédiffusion et à la radiodiffusion sonore numériques pour les applications d'émission doivent être conformes aux spécifications énoncées ci-dessous.

1 Spécifications de service

1.1 Configurations des canaux

Pour les services audio il faut disposer, selon les besoins, des applications visées dans la Recommandation UIT-R BS.775, des configurations de canaux suivantes:

| | | |
|----------|------------|---|
| 1 canal | 1/0 | Monophonique |
| 2 canaux | 2/0 | Gauche, droit |
| 3 canaux | 3/0 2/1 | Gauche, droit, centre Gauche, droit/ambiophonique |
| 4 canaux | 3/1 2/2 | Gauche, droit, centre/ambiophonique Gauche, droit/ambiophonique gauche, ambiophonique droit |
| 5 canaux | 3/2 | Gauche, droit, centre/ambiophonique gauche, ambiophonique droit |

ainsi qu'éventuellement d'un CABF.

1.2 Services audio

Parallèlement au service audio principal, il est possible que soient assurés en fonction des besoins d'applications particulières les autres services audio suivants: service multilingue, services audio pour les malentendants et les personnes atteintes d'un handicap visuel, plus une certaine capacité prévue pour les données auxiliaires (Recommandation UIT-R BS.775) [ATSC, 1992]. Comprenant un seul canal ou plusieurs canaux indépendants, le service multilingue peut être utilisé pour distribuer un programme avec un commentaire en plusieurs langues. Le service destiné aux personnes ayant un handicap visuel fournit généralement une description vocale du contenu de l'image, alors que le service pour les malentendants efface la musique et les effets spéciaux, ou bien en abaisse le niveau sonore, pour améliorer l'intelligibilité des dialogues. Les données auxiliaires peuvent véhiculer divers types d'information, tels que commande de la dynamique, réglage du volume sonore et données d'utilisateur.

Il est possible de classer les divers services comme suit:

- *Service principal* (chaque canal, même le canal CABF, y est assigné au même programme).
- *Service(s) étendu(s)*, qui peuvent être des:
 - *services indépendants* (pour des programmes additionnels, indépendants du service principal, tels que commentaires, ou d'autres services utilisant deux ou plusieurs canaux; la configuration des canaux peut être choisie parmi les possibilités indiquées au tableau sous le § 1.1);
 - *services de remplacement* (pour des programmes destinés à remplacer un ou plusieurs des canaux du service principal, par exemple multilingue ou pour les malentendants);
 - *services additionnels* (utilisant des canaux qui viennent s'ajouter à ceux du service principal, tels que le commentaire, ou des canaux additionnels pour des systèmes audio améliorés comme la télévision en trois dimensions).

Etant donné que tout système de transmission doit comporter une couche système capable de réaliser des opérations de multiplexage, il n'est pas nécessaire que tous les services audio mentionnés ci-dessus soient assurés par un seul train de bits.

1.3 Attribution souple des canaux

Un train de bits fournit des données d'identification pour la signalisation et la commande des configurations des signaux sonores. Le système de transmission doit permettre de passer dynamiquement de l'une à l'autre des configurations de canaux mentionnées au § 1.1.

1.4 Données auxiliaires

Le système de codage audio doit permettre de transmettre des données auxiliaires, lesquelles peuvent véhiculer différents types d'information, tels que commande de la dynamique, réglage du volume sonore et données d'utilisateur.

2 Spécifications de qualité de fonctionnement

2.1 Qualité audio

2.1.1 Qualité audio de base

Le radiodiffuseur doit normalement trouver un compromis entre la qualité audio et le débit binaire appliqué aux signaux sonores. En principe, la qualité du son reproduit après décodage sera subjectivement identique à celle du signal d'origine pour la plupart des types de programme audio.

Pour ce faire, il faut obtenir, en utilisant la méthode en double aveugle à triple stimulus et référence dissimulée, exposée dans la Recommandation UIT-R BS.1116, des notes moyennes généralement supérieures à 4 sur l'échelle de dégradation à 5 notes pour des auditeurs occupant la position d'écoute de référence. Dans la pratique, on utilise parfois pour des impératifs commerciaux des débits binaires inférieurs à celui qui est nécessaire pour obtenir ce niveau de qualité, que le système doit néanmoins permettre au radiodiffuseur d'atteindre.

NOTE 1 – Les paramètres de qualité objective audio pour les applications de contribution/distribution pourront être ultérieurement intégrés, conformément à la Recommandation UIT-R BS.1387.

2.1.2 Qualité de l'image sonore stéréophonique

Dans le cas de configurations stéréophoniques (bicanal ou multicanal), la qualité de l'image sonore des signaux sources devrait être préservée. Dans les configurations comprenant un canal central (3/0, 3/1, 3/2), la stabilité directionnelle de l'image sonore frontale doit être maintenue à l'intérieur de limites raisonnables sur une plage d'écoute plus étendue que celle à laquelle donne lieu une stéréophonie bicanal classique. Dans les configurations faisant intervenir un effet d'ambiophonie (2/1, 2/2, 3/1, 3/2), la sensation de réalité spatiale (ambiance) doit être nettement améliorée par rapport à celle que fournit une stéréophonie bicanal classique (Recommandation UIT-R BS.775).

2.1.3 Résolution de quantification

La résolution exigée devrait être au moins de 16 bits.

2.1.4 Fréquence d'échantillonnage

Conformément à la Recommandation UIT-R BS.646, la fréquence d'échantillonnage doit être de 48 kHz.

2.1.5 Largeur de bande

Canaux audio principaux: 20-20 000 Hz.

CABF: 15-120 Hz.

2.1.6 Accentuation

Le système de codage audio ne devrait pas employer d'accentuation.

2.1.7 Possibilité de postproduction

Le traitement après production à réaliser est fortement tributaire de l'application. Dans le cas des liaisons d'émission, il peut se limiter à une égalisation et à un ajustement de la dynamique (par exemple, pour faire correspondre la dynamique du programme à celle de l'environnement d'écoute).

2.2 Retard dû au codage

Le retard dû au codage doit être identique pour tous les canaux d'un programme. Dans le cas du son télévisuel, le retard dû au codage des signaux audio doit correspondre à celui des signaux vidéo.

2.3 Protection contre les erreurs

Il faut prévoir dans le train des bits audio un mécanisme permettant au décodeur de repérer des erreurs résiduelles dans les canaux et de les masquer.

2.4 Délai de récupération

Le délai de récupération devrait être aussi court que possible. Dans le cas où on utilise des AAU, il devrait être de quelques AAU au maximum, mais dans l'idéal d'une seule AAU.

3 Spécifications fonctionnelles et opérationnelles

3.1 Compatibilité

3.1.1 Compatibilité aval (Recommandation UIT-R BS.775)

Le format du train de bits multicanal doit lui permettre d'être décodé par des classes de décodeurs de complexité variable. Il doit être possible dans le décodeur de prévoir un nombre de canaux inférieur à celui des canaux émis, suivant les possibilités de reproduction dont dispose l'utilisateur, sans entraîner d'autre dégradation que la perte de l'effet de localisation stéréophonique ou multicanal.

Deux méthodes permettent d'assurer une compatibilité aval avec une faible complexité des récepteurs. La première utilise le procédé matriciel: un récepteur bon marché ne doit alors être pourvu que des voies A et B à l'instar du système 2/0, c'est-à-dire un système n'utilisant pas de matrice à compatibilité amont. La deuxième s'applique au système d'émission discret 3/2: les signaux émis sont numériquement combinés à l'aide des équations qui permettent de fournir le nombre de signaux voulus; dans le cas de signaux codés à la source de faible débit, il est possible de procéder à la combinaison aval des signaux 3/2 avant la partie de synthèse du processus du décodage (où se situe la plus grande partie de la complexité).

3.1.2 Compatibilité amont par rapport à la Recommandation UIT-R BS.1115

La spécification en question correspond à la situation où une application monophonique/stéréophonique existante, conforme aux dispositions de la Recommandation UIT-R BS.1115, doit être adaptée à un son multicanal, la desserte des récepteurs existants continuant d'être assurée. La Recommandation UIT-R BS.1115 préconise (pour les applications nécessitant uniquement un codage monophonique ou stéréophonique à faible débit) l'utilisation du système de codage de Couche II de l'ISO/CEI IS 11172-3. Dans le cas de systèmes employant déjà ce type de codage audio pour un son monophonique ou stéréophonique, un décodeur ISO/CEI IS 11172-3 pourra, grâce à cette compatibilité amont qui permet un codage multicanal à faible débit, décoder des informations en stéréophonie de base, constituées par une combinaison aval appropriée des informations audio provenant de tous les canaux source. Pour satisfaire à cette spécification, on peut appliquer soit la méthode de la transmission simultanée, soit la méthode du matriçage.

Méthode de la transmission simultanée

Une solution est de continuer d'assurer le service existant, correspondant aux canaux conformes à la Recommandation UIT-R BS.1115, et d'ajouter le nouveau service des canaux 3/2. Il s'agit de la méthode dite de transmission simultanée. L'avantage est que le service existant conforme à la Recommandation UIT-R BS.1115 pourra être interrompu le moment voulu et qu'il est possible d'optimiser de façon indépendante les combinaisons des programmes 2/0 et 3/2.

Méthode de matriçage

L'autre solution consiste à utiliser des matrices de compatibilité. Il est en effet possible de recourir aux équations matricielles pour assurer la compatibilité avec les récepteurs existants. Dans ce cas, les canaux d'émission gauche et droite existants sont utilisés pour transporter les signaux

matriciels A et B compatibles. On utilise d'autres canaux d'émission pour transporter les signaux matriciels T, Q₁ et Q₂. L'avantage de cette méthode est peut-être qu'il faut, pour ajouter le nouveau service, une moindre capacité additionnelle de données.

3.1.3 Compatibilité aval par rapport à la Recommandation UIT-R BS.1115

En ce qui concerne les applications où le nouveau système multicanal doit coexister avec le système monophonique/stéréophonique défini dans la Recommandation UIT-R BS.1115, il faut éventuellement que les décodeurs puissent décoder un train de bits audio conforme à la norme ISO/CEI IS 11172-3.

3.2 Débit binaire

Il est recommandé d'utiliser le système de codage de Couche II de l'ISO/CEI IS 11172-3 [ISO/CEI, 1993] avec un débit binaire de 128 kbit/s par canal (Recommandation UIT-R BS.1115). Le débit de 2×128 kbit/s peut donc être considéré comme une limite supérieure pour le service bicanal et le débit de 5×128 kbit/s comme une limite supérieure pour le service multicanal dans le cas où la compatibilité amont (voir le § 3.1.2) n'est pas nécessaire. Etant donné le gain de codage additionnel que les techniques de codage composite devraient permettre d'obtenir, le nouveau système de codage multicanal devrait être capable d'atteindre un débit binaire de 512 kbit/s au maximum pour le service principal à 5 canaux.

Le système de codage devrait assurer la qualité audio définie au § 2.1 pour un débit binaire comparable ou inférieur à celui que nécessiterait le système préconisé dans la Recommandation UIT-R BS.1115.

3.3 Complexité du décodeur

Le décodeur pour le programme audio devrait être d'une complexité raisonnable pour que son prix puisse être bas. Dans l'hypothèse où un nombre de canaux moindre, M , doit être reproduit à partir d'un programme audio contenant N canaux, la complexité du décodeur devrait être inférieure à celle du décodeur correspondant à l'ensemble de N canaux.

Références Bibliographiques

- ATSC [1992] Services audio et de données auxiliaires numériques pour un service de télévision évolué. Document ATSC T3/186. Comité de systèmes de télévision évolués.
- ISO/CEI [1993] Coding of moving pictures and associated audio for digital storage media at up to 1,5 Mbit/s. Norme ISO/CEI IS 11172-3. Organisation internationale de normalisation/Commission électrotechnique internationale.
-