

**Zeitschrift:** Technische Mitteilungen / Schweizerische Post-, Telefon- und Telegrafienbetriebe = Bulletin technique / Entreprise des postes, téléphones et télégraphes suisses = Bollettino tecnico / Azienda delle poste, dei telefoni e dei telegrafi svizzeri

**Herausgeber:** Schweizerische Post-, Telefon- und Telegrafienbetriebe

**Band:** 69 (1991)

**Heft:** 8

**Artikel:** Système de codage MICDA-SB pour signaux sonores de haute qualité

**Autor:** Mury, Roland / Pexa, Günther / Vörös, Paul

**DOI:** <https://doi.org/10.5169/seals-876302>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 15.01.2026

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

# Systeme de codage MICDA-SB pour signaux sonores de haute qualité

Roland MURY, Berne, Günther PEXA, Munich, Paul VÖRÖS, Berne et Daniel WILD, Zurich

## 1 Introduction

La deuxième moitié des années 80 a vu la technique de transmission numérique s'imposer pour la distribution des programmes de radiodiffusion. Les systèmes MIC actuellement en service dans le réseau de l'Entreprise des PTT suisses permettent la transmission d'un canal radiophonique de haute qualité (bande passante de 15 kHz) à 384 kbit/s, conformément aux Recommandations J41 et J42 du CCITT. L'objet de cet article est de rendre compte du résultat de travaux effectués au Département de la recherche et du développement des PTT dans le domaine du codage de signaux audio de haute qualité. Le fruit de ces travaux est une méthode de codage MICDA-SB (Modulation par impulsions et codage différentiel adaptatif en sous-bandes) permettant la transmission d'un signal audio de 15 kHz à l'aide de 128 kbit/s, la qualité étant identique à celle des systèmes de codage MIC de 384 kbit/s (fig. 1). L'emploi de processeurs de signaux numériques a permis l'implantation d'un codeur MICDA-SB sur une seule carte de format 11 x 16 cm. Grâce au faible retard introduit par le codage MICDA-SB ( $\leq 7$  ms), ce système s'impose tout particulièrement dans les applications interactives, telles que le reportage.

## 2 Principe du système de codage MICDA-SB

### Réduction du débit binaire

La modulation par impulsions et codage (MIC) est la forme la plus simple de codage numérique, consistant à faire un échantillonnage périodique de la sortie d'une source analogique et à représenter chaque échantillon par un nombre fini de bits. Le débit binaire d'environ 700 kbit/s par signal monophonique qu'exigent les enregistrements MIC sur disque compact (le signal est échantillonné à 44,1 kHz et quantifié avec 16 bits par échantillon), peut être sensiblement réduit pour la transmission sans qu'il en résulte une dégradation subjective de la qualité.

Le débit peut être réduit:

- en éliminant la redondance du signal d'origine par la méthode de la prédiction ou par la répartition du spectre en sous-bandes

- en ne transmettant que l'information située à la portée spectrale et temporelle de l'oreille humaine. Cette information ou entropie perceptive se monte à moins de 2 bits par échantillon pendant 98 % du temps, aucune valeur supérieure à 2,23 bits par échantillon n'ayant été relevée à ce jour [1]
- par la compression du débit binaire à la sortie, en utilisant un procédé de codage découvert par D.H. Huffman [2].

Le codage fondé sur l'entropie perceptive englobe la division du spectre en sous-bandes dites «bandes critiques» d'environ 100 Hz à l'extrémité basse et s'élargissant à 4 kHz à l'extrémité haute du spectre audio [3]. Cela peut être effectué par codage, soit en sous-bandes, soit par transformation dans le domaine des fréquences [4, 5]. Cependant, cette technique est assez complexe, sensible aux erreurs binaires et hautement non-linéaire ce qui pourrait rendre problématique tout traitement ultérieur, tel que le mixage. De plus, le filtrage acoustique résultant des caractéristiques de la salle d'écoute pourrait modifier chez l'auditeur la structure du masquage et ainsi rendre la distorsion audible.

Comme il est mentionné ci-dessus, la prédiction et la répartition spectrale sont des techniques en concurrence pour la réduction des redondances et le modèle de l'oreille représentée comme une batterie de filtres met en évidence le fait qu'un rapport-signal-sur-bruit (RSB) élevé n'est pas suffisant pour garantir que le bruit est

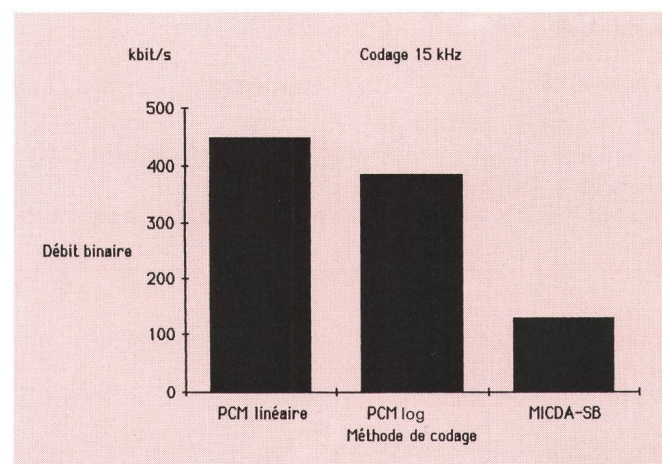


Fig. 1 Evolution du débit binaire requis pour coder un signal audio de haute qualité



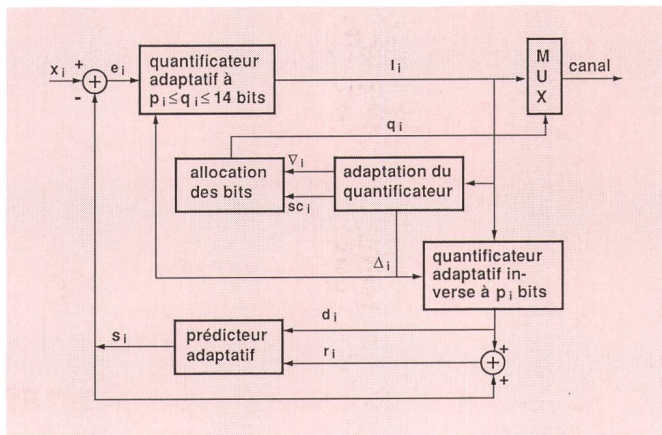


Fig. 2 Schéma de principe du codeur MICDA dans la SBi pour  $1 \leq i \leq 4$

correctement masqué. Dans la mesure où l'objectif est un codage à  $2 \times 64$  kbit/s par canal monophonique à 15 kHz, le débit binaire utilisé correspond à environ deux fois le minimum théorique, ce qui finalement n'implique pas une trop grande complexité. A cet effet, on a développé un algorithme basé sur un modèle à quatre sous-bandes combiné avec une quantification et une prédiction adaptatives (MICDA-SB).

Ce modèle constitue un bon compromis qualité/complexité avec effets de masquage prononcés et suffisamment de gain de prédiction dans chaque sous-bande pour la plupart des signaux.

#### Description de l'algorithme [6, 7]

Une batterie de filtres miroirs en quadrature (FMQ) du 32<sup>e</sup> ordre ordonnée en cascade à trois étages est utilisée pour obtenir la décomposition en quatre sous-bandes (SB) [0-2], [2-4], [4-8] et [8-16] kHz, numérotées SB1, SB2, SB3 et SB4 respectivement.

A partir des signaux sonores échantillonnés à 32 kHz, les groupes FMQ d'analyse et de synthèse produisent un retard d'environ 7 ms. Vu la grande diversité des signaux sonores, le problème de l'affectation des bits créé par la décomposition du spectre en sous-bandes ne peut être convenablement résolu que par une répartition adaptative des bits. De même, des quantificateurs et des prédicteurs adaptatifs sont utilisés en raison de la nature non-stationnaire de ces signaux.

#### Quantification et prédiction adaptatives

Le schéma de principe de la boucle du codeur est représenté par la figure 2. La corrélation entre les échantillons d'entrée peut être éliminée en calculant  $s(n)$ , la valeur probable ou «prédiction» du signal d'entrée  $x(n)$  à l'instant  $n$  à partir des échantillons précédents  $x(n-1)$ ,  $x(n-2)$ ,... L'erreur de prédiction  $|e(n)| = |x(n) - s(n)|$  est normalement plus petite que  $|x(n)|$  et pour les signaux typiques un gain de prédiction statistique de 5 à 15 dB peut être réalisé. Pour une tonalité, le gain peut aller jusqu'à 25 dB.

Le prédicteur à deux pôles et six zéros, normalisé dans la Recommandation G.722 du CCITT (Comité consultatif international télégraphique et téléphonique), est utilisé dans les quatre sous-bandes. Pour un fonctionnement

stable du prédicteur à des débits binaires variables dans les sous-bandes, un codage imbriqué est indiqué [8]: chaque mot du code à  $q$  bits consiste en un nombre fixe de bits de prédiction  $p = q_{\min}$  et une suite variable de bits pour une quantification plus fine. Le robuste quantificateur imbriqué à sept segments (fig. 3) exige peu de mémoire et peu d'opérations arithmétiques pour quantifier un échantillon, cependant il produit un rapport-signal-sur-bruit (RSB) dans les limites d'un décibel du quantificateur idéal de Laplace pour  $q$  supérieur ou égal à 3 bits par échantillon. Pour s'assurer que le quantificateur fonctionne dans sa plage RSB la plus favorable, le facteur d'échelle logarithmique du quantificateur  $\nabla$  à l'instant  $n$  est corrigé en fonction des coefficients multiplicateurs  $w[l(n-1)]$  qui dépendent de l'échantillon transmis à l'instant précédent  $l(n-1)$ :

$$\nabla(n) = \beta \cdot \nabla(n-1) + w[l(n-1)], \quad \beta = 127/128 \quad (1)$$

où  $\beta$  est le facteur de perte qui rend l'adaptation résistante aux erreurs de transmission. Le facteur d'échelle linéaire du quantificateur  $\Delta(n)$  est calculé de la façon suivante:

$$\Delta(n) = 2^{[\nabla(n)-15]}, \quad 0 \leq \nabla(n) \leq 12 \quad (2)$$

Le mot de code  $l(n)$  est obtenu par le multiplexage du signe, du code du segment  $sc$  à 2 bits, du reste du signal quantifié uniformément dans le même segment, en ne retenant que les  $q$  bits de poids fort (truncation).

Pour  $q = 8$ , par exemple, on obtient:

$$\begin{aligned} |e(n)| / \Delta(n) &= [\pm] 0101.01110010110 [\text{BIN}] \\ \rightarrow l(n) &= [\pm] 1010011 \end{aligned} \quad (3)$$

Le quantificateur inverse produit la valeur représentative  $d(n)$  en ajoutant 0,45 fois le bit de valeur minimale (BVM). Le signal reconstitué  $r(n)$ , qui est calculé de la même manière dans le décodeur, correspond au signal d'entrée dans les limites de l'erreur de quantification.

#### Affectation dynamique des bits

La répartition des bits est habituellement basée sur la puissance des signaux donnée par la variance  $\sigma^2$  dans les sous-bandes calculée pour des blocs d'échantillons

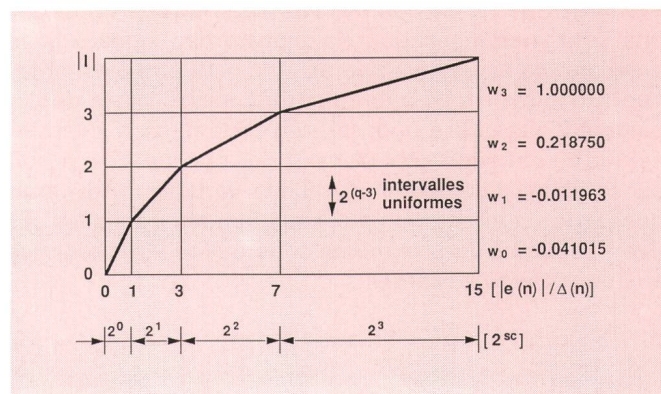


Fig. 3 Quantificateur à 7 segments pour  $q_i \geq 3$  bits et  $1 \leq i \leq 4$



fixes et transmise comme information accessoire au récepteur [9]. Le RSB augmente avec la diminution de la dimension du bloc, mais la longueur du bloc ne peut pas être arbitrairement courte.

Pour l'algorithme, la répartition des bits est basée sur le signal d'erreur  $e(n)$  quantifié approximativement en utilisant seulement les bits du segment  $sc$ . Cette valeur, tout comme le facteur d'échelle logarithmique  $\nabla$ , peut être calculée à la réception sans information accessoire et la dimension du bloc dans la sous-bande la plus étroite peut être réduite à un seul échantillon! Pour un signal d'erreur approchant les caractéristiques laplaciennes, dont le facteur de crête est de 12 dB, 90 % des échantillons d'erreur  $[e(n)/\Delta(n)]$  sont compris dans l'intervalle de 1/3 à 15 comme illustré dans la figure 3 et l'amplitude de chacun de ces échantillons est connue avec à peu près la même précision que le RSB du mot de code entier (c'est-à-dire dans les limites du BVM ou 6 dB). Pour le plus petit bloc possible de 8 échantillons d'entrée, il y a 1, 1,2 et quatre échantillons de sortie dans  $SB_1$ ,  $SB_2$ ,  $SB_3$  et  $SB_4$  respectivement. Pour une puissance du bruit de quantification uniforme, le nombre de bits  $q$  attribué à l' $i^e$  échantillon est donné par

$$q_i = (\nabla_i + sc_i) + \sum_{i=1}^8 q_i/8 - \sum_{i=1}^8 (\nabla_i + sc_i)/8 \text{ avec } 1 \leq i \leq 8 \quad (4)$$

où  $\sum q/8$ , la moyenne des bits par échantillon, est égal à 4.

Le nombre de bits de quantification est soumis à une double limitation:

- pour la prédiction et la quantification adaptatives, un minimum de 2 bits par échantillon peut être nécessaire (information d'amplitude et signe)
- le gain de prédiction augmente avec  $p$  (pour  $p$  supérieur à 3 seulement marginalement), c'est pourquoi  $p$  devrait être aussi grand que possible tout en gardant un certain nombre de bits en réserve.

Des simulations sur ordinateur montrent que, pour une large classe de signaux (à l'exception des passages silencieux) la demande de bit  $q$  en  $SB_1$ ,  $SB_2$  et  $SB_3$ , tombe rarement au-dessous de 6, 4 et 3 bits par échantillon, respectivement. Dans  $SB_4$  elle est inférieure à 1,5 bit 75 % du temps et elle n'excède que rarement 3 bits par échantillon.

Bien que les résultats objectifs démontrent que les avantages de l'inclusion de  $SB_4$  dans la répartition des bits sont marginaux pour la plupart des signaux, une amélioration subjective appréciable peut être ainsi obtenue. On la réalise en imbriquant le quantificateur de Laplace à 2 bits utilisé pour la prédiction dans le quantificateur à sept segments. Ainsi, pour chaque bloc à huit échantillons d'entrée, il y a 8 bits variables («dynamiques») dont l'affectation aux sous-bandes – dans les cas extrêmes, tous dans la même sous-bande – est soumise aux contraintes suivantes:

$$p_1 = 5 \leq q_1 \leq 14; \quad p_2 = 4 \leq q_2 \leq 12; \quad p_{31-32} = 3 \leq q_{31-32} \leq 9; \\ p_{41-44} = 2 \leq q_{41-44} \leq 5 \quad (5)$$

Bien qu'un minimum de 6 bits soit attribué à  $SB_1$ , pour des raisons indiquées plus loin, seulement 5 bits sont

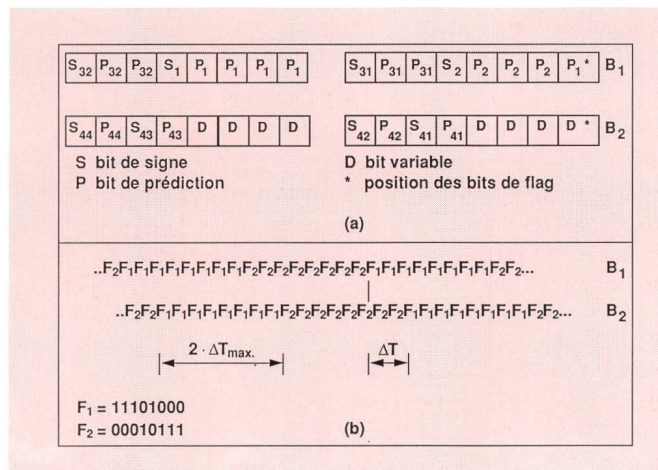


Fig. 4 Format de sortie des octets après multiplexage (a) et insertion des Flags  $F_1$  et  $F_2$  pour la synchronisation des canaux  $B_1$  et  $B_2$  (b)

employés pour la prédiction. En l'absence d'erreurs de transmission, la répartition des bits basée sur (4) et (5) livre les mêmes résultats dans le décodeur. Cependant, en présence d'erreurs la robustesse du décodeur dépend du multiplexage des bits variables; avec n'importe quel schéma de multiplexage la principale source de dégradation se révèle être la désadaptation du prédicteur.

#### Adaptation de l'algorithme à la transmission dans le RNIS

Le débit binaire produit par cet algorithme peut être transmis par n'importe quelle paire de canaux transparents à 64 kbit/s, mais les deux canaux  $B$  indépendants de l'accès de base d'un RNIS posent un problème de synchronisation. Le schéma de multiplexage proposé pour cette application est indiqué à la figure 4a. Le récepteur peut synchroniser les canaux et identifier la séquence correcte d'octets en utilisant les flags  $F_1$  et  $F_2$  (fig. 4b). Dans chaque 16<sup>e</sup> trame, le bit  $P^*$  est copié dans la position  $D^*$  et remplacé par un bit flag. Le même bit flag remplace le bit  $D^*$  dans le canal  $B_2$ , deux trames plus tard. Si  $\Delta T_{\max}$  est la différence de retard maximum attendu entre les deux canaux  $B$ , ces flags sont alors transmis alternativement, chacun pour une période de  $2\Delta T_{\max}$ . Selon l'état de synchronisation, le récepteur peut décoder le canal  $B_1$  seul, les deux canaux  $B_1$  et  $B_2$ , ou aucun des canaux.

Si le canal  $B_2$  devient indisponible (par exemple en raison d'un appel téléphonique), il n'est pas nécessaire d'en informer le codeur, mais la perte de ce canal devrait, de préférence, être signalée au décodeur par une salve de zéros assez longue pour entraîner une perte de synchronisme. Afin d'éliminer la distorsion due au recouvrement spectral qui peut être audible dans ce cas, le troisième étage FMQ de synthèse est remplacé par un autre filtre avec une fréquence de coupure plus basse, et il n'est pas nécessaire de modifier l'étage de sortie analogique. La séquence d'octets correcte est néanmoins maintenue dans le canal  $B_1$ , mais une petite fraction des bits  $P_1$  (BVM) est perdue. Dans le but d'assurer la stabilité de la boucle du décodeur  $SB_1$ , seuls 5 bits sur un minimum de 6 bits assignés à  $SB_1$  de façon permanente sont utilisés pour la prédiction.



## Qualité subjective

Le dernier chaînon – mais le plus important dans un système de communication – est l'utilisateur. Les mesures objectives habituelles de la qualité telles que RSB, la réponse fréquentielle, la distorsion d'intermodulation, etc., peuvent donner des indications fiables sur la qualité subjective du système, mais, par exemple, dans le cas d'un codec exploitant les propriétés de masquage de l'appareil auditif humain, les qualités objectives et subjectives sont à peine corrélées, et des tests subjectifs deviennent indispensables pour évaluer les codecs. Le test subjectif formel effectué à l'aide d'un modèle hardware utilisant l'algorithme MICDA-SB a été exécuté selon la Recommandation 562-2 [10] du CCIR (Comité consultatif international pour la radiodiffusion) à l'aide de l'échelle de comparaison à sept notes où la qualité «identique à la référence» (référence = MIC à 512 kbit/s) est notée 0 et les qualités «légèrement meilleures/plus mauvaises» correspondent à +1/−1.

Le résultat peut être résumé de la façon suivante:

- La note moyenne obtenue pour 13 signaux sélectionnés sur le disque compact EBU-SQAM, y compris des instruments en solo (gong électronique à 500 Hz, jeux de cloches, piano, trompette, guitare, piccolo et castagnettes) et plusieurs extraits de musique vocale et orchestrale n'indique aucune différence significative entre les versions originales et codées.
- Aucune dégradation notable n'a été constatée, si le taux d'erreurs des bits est inférieur à  $10^{-4}$ . Il s'ensuit qu'aucune protection contre les erreurs ne sera nécessaire pour la plupart des applications.

## 3 Réalisation: système de reportage SB15

L'algorithme MICDA-SB développé par le Département des recherches et du développement des PTT suisses pour la transmission de signaux audio de haute qualité

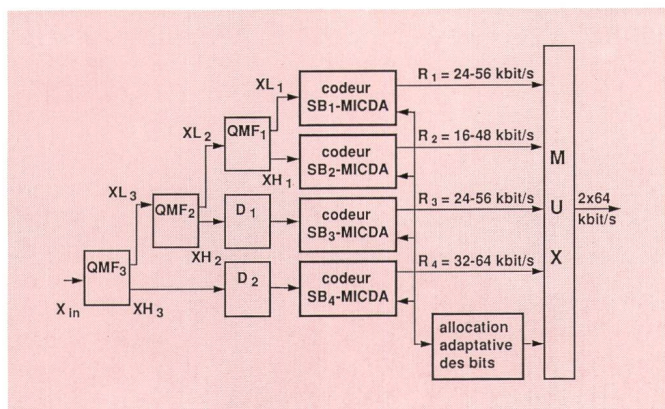


Fig. 5 Schéma de principe du codeur SB-MICDA

- $X_{in}$  Valeur d'entrée (représentation uniforme)
- QMF Filtre miroir en quadrature
- XL, XH Sortie de la sous-bande inférieure et supérieure respectivement
- $D_1$  Déphaseur à 16 échantillons (à 8 kHz)
- $D_2$  Déphaseur à 48 échantillons (à 16 kHz)
- $R_i$  Débit dans la SB, pour  $1 \leq i \leq 4$

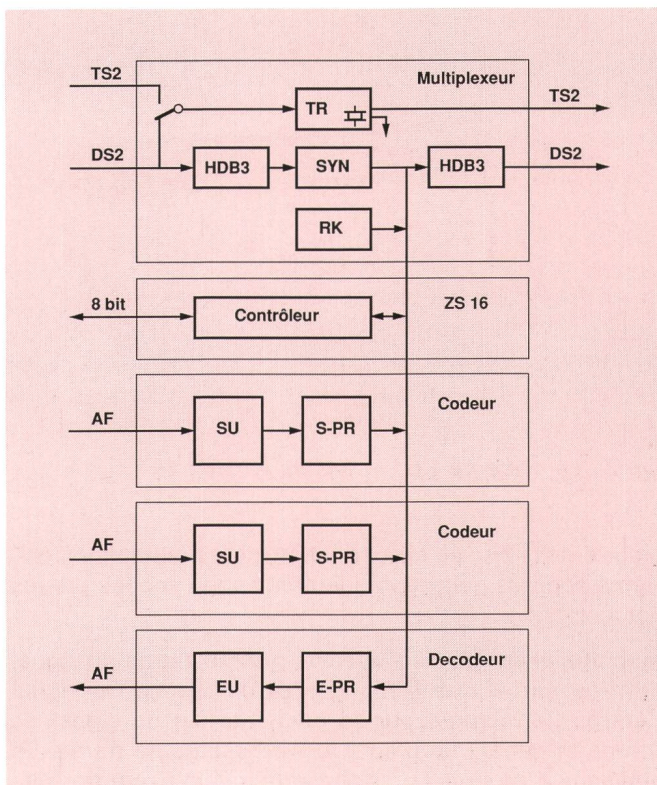


Fig. 6 Répartition des fonctions du système SB15

- DS2 Signal numérique 2048 kbit/s
- TS2 Rythme 2048 kbit/s
- HDB3 Codeur/Décodeur HDB3
- SYN Verrouillage de trame
- RK Générateur de mot de verrouillage de trame
- SU Convertisseur émission
- S-PR Processeur de signaux émission
- E-PR Processeur de signaux réception
- EU Convertisseur réception
- AF Signal audio

avec un débit numérique de 128 kbit/s est appliqué dans le système de reportage SB15 de la maison Siemens (fig. 5).

Le codec du système de reportage est équipé d'interfaces analogiques pour les signaux audio et d'interfaces CCITT G.703 côté transmission. Le système SB15 utilise la trame CCITT G.704 standardisée pour la transmission de signaux téléphoniques, ce qui permet la transmission simultanée de signaux audio, téléphoniques et de données. Un signal audio occupe la place de deux voies téléphoniques, de sorte qu'un signal à 2048 kbit/s peut contenir jusqu'à 15 voies audio monophoniques ou 7 voies stéréophoniques.

### Hardware

Le système SB15 permet la transmission point à point de 15 voies audio. L'attention a cependant été portée tout particulièrement sur le couplage et le découplage d'une ou de plusieurs voies, selon les besoins. L'appareil possède une entrée et une sortie numérique pour signaux à 2048 kbit/s. Le multiplexeur (fig. 6) se synchronise après décodage du signal HDB3 sur la trame entrante (SYN). A l'exception des intervalles de temps venant d'être couplés, le signal d'entrée se retrouve à la sortie du codeur HDB3 sous sa forme originale. Il est de



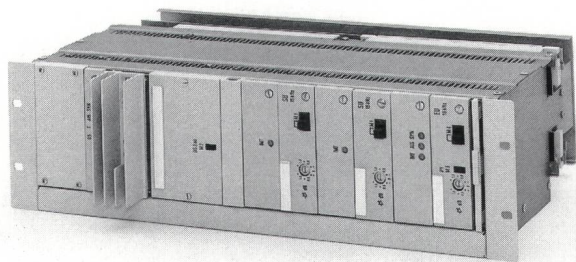


Fig. 7 Le système SB15 dans sa version 19"

la sorte possible, en cascadeant plusieurs multiplexeurs — même séparés géographiquement —, de «collecter» les signaux audio.

Le multiplexeur travaille avec un rythme extrait du signal d'entrée numérique (TR). En cas d'absence de signal d'entrée, le multiplexeur se connecte sur un signal de rythme et génère lui-même le verrouillage de trame. En l'absence d'un signal de rythme, le multiplexeur travaille avec le rythme interne généré par un oscillateur à quartz. Le rythme utilisé est disponible à une sortie «rythme» pour les synchronisations externes: la synchronisation de tous les signaux numériques est indispensable si les signaux doivent être commutés numériquement.

Les signaux audio (AF) sont échantillonnés à la fréquence de 32 kHz et quantifiés uniformément par le convertisseur émission SU. Le signal ainsi généré est transmis au processeur de signaux émission S-PR.

A la réception, le signal audio est reconstitué à l'aide de la configuration inverse: processeur de signaux réception E-PR et convertisseur réception EU.

Une unité SB15 peut être équipée pour trois voies audio au maximum.

La figure 7 représente le cas de deux codeurs et d'un décodeur. Les appareils sont disponibles en châssis de 19" pouvant être montés dans des bâtis BW72 avec un système d'alarme AK84.

#### Processeur de signaux

L'algorithme de codage décrit au paragraphe 2 est implanté dans des processeurs de signaux. La réalisation est basée sur le processeur *Motorola DSP 56001* avec une fréquence de quartz de 26,5 MHz (fig. 8).

A l'émission, le processeur de signaux réalise la division du spectre en quatre sous-bandes, le codage MICDA des différentes sous-bandes et l'allocation variable des bits aux différentes sous-bandes. A la réception, les quatre signaux MICDA sont décodés et les sous-bandes assemblées pour reformer le signal audio.

Le processeur de réception assume également des tâches de surveillance de voie telle que l'alarme de voie

(AIS) ou la mise en sourdine en cas de coupure de la ligne.

#### Caractéristiques du système

Les caractéristiques objectives du système correspondent à celles du système MIC utilisant 384 kbit/s par voie (système Siemens MStDF) et sont indiquées ci-dessous.

Les caractéristiques subjectives ont été l'objet de nombreux tests pratiques (voir paragraphe 2) qui ont confirmé les mesures en laboratoire.

Courbe de réponse,  
par rapport à 1 kHz/—12 dBmOs

40 Hz . . . . .	+0,25 dB/—0,5 dB
125 Hz à 14 kHz . . . . .	±0,25 dB
15 kHz . . . . .	+0,25 dB/—1 dB

Distorsion du temps de

propagation de groupe . . . . . < 3 ms

Bruit psophométrique par voie libre . . . . . < —56 dbqOs

Perturbation isolée (psophométrique) . . . . . < —73 dBmOps

Distorsions non linéaires

Facteur de distorsion

40 Hz à 125 Hz/ +9 dBmOs . . . . . <0,5 % (46 dB)

125 Hz à 2 kHz/ +9 dBmOs . . . . . <0,25 % (52 dB)

2 kHz à 4 kHz/ +6 dBmOs . . . . . <0,25 % (52 dB)

Intermodulation

Dans la bande 0 dBmOs . . . . . > 55 dB

Hors bande 0 dBmOs . . . . . > 65 dB

Atténuation de la diaphonie audible

40 Hz à 15 kHz . . . . . > 84 dB

#### Caractéristiques stéréo

Différences de niveau entre voies A et B

40 Hz . . . . . <0,5 dB

125 Hz à 10 kHz . . . . . <0,25 dB

14 kHz . . . . . <0,5 dB

15 kHz . . . . . <1 dB

Déphasage entre voies A et B

40 Hz . . . . . <10°

200 Hz à 4 kHz . . . . . <5°

14 kHz . . . . . <10°

15 kHz . . . . . <15°

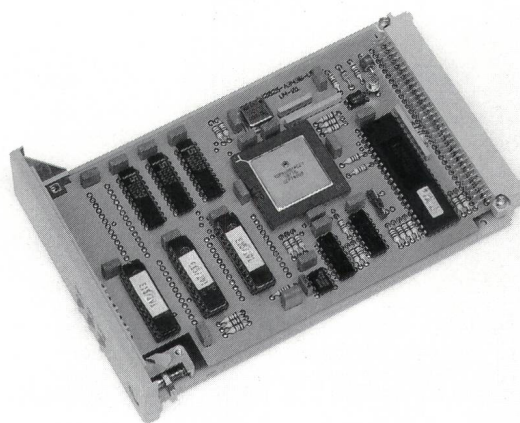


Fig. 8 L'unité «Processeur de signaux»



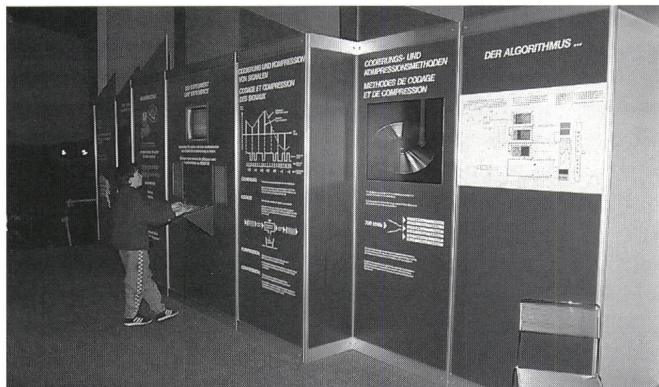


Fig. 9 Présentation du système à l'exposition COMM89

## 4 Présentations et applications du système de codage MICDA-SB

La méthode de codage MICDA-SB a été présentée pour la première fois à l'occasion des réunions interrégionales du CCIR en automne 1987 à Genève. La démonstration informelle organisée en marge des travaux des commissions d'études illustra pratiquement les résultats des tests objectifs et subjectifs obtenus avec cette nouvelle méthode de codage.

Au cours de l'Exposition «COMM89» marquant le 30<sup>e</sup> anniversaire du Musée suisse des transports de Lucerne, l'Entreprise des PTT a procédé à une démonstration pratique des possibilités d'application sur le RNIS du codage à bas débit des signaux sonores dans le domaine de l'audiovisuel.

Il s'agissait pour le visiteur de consulter, avec l'aide d'un PC, une banque de données sonores issue de disques compacts dont les signaux étaient codés à  $2 \times 64$  kbit/s pour la transmission (fig. 9). Cet exemple révéla l'intérêt d'un nouveau service permettant à l'utilisateur d'écouter à distance des extraits du disque compact qu'il désire acheter et cela en toute tranquillité, à son domicile.

Une application différente du codage à bas débit a été identifiée avec la vidéoconférence. En général, les codecs de vidéoconférence à 2 Mbit/s sont équipés d'au moins deux interfaces à 64 kbit/s pour la transmission de données. Il pourrait être intéressant de les utiliser pour la transmission d'une voie radiophonique de haute qualité à  $2 \times 64$  kbit/s. Cette nouvelle possibilité a été expérimentée à l'occasion de l'inauguration en 1990, de la salle de vidéoconférence de l'Entreprise des PTT à Berne, en réalisant une liaison avec un studio des PTT autrichiens (Vienne) pour un échange de séquences de programmes sonores simultanément avec la vidéoconférence. La séance s'est déroulée dans les meilleures conditions, sans qu'il en résulte d'inconvénients pour l'un ou l'autre des services.

Les performances remarquables du codage MICDA-SB ont créé de nouvelles conditions favorables à l'identification des futurs services sur le RNIS. Parmi ceux-ci, la téléaudiovidéothèque est déjà dans une phase d'expérimentation à l'Entreprise des PTT suisses sous la forme d'un modèle comportant plusieurs terminaux et serveurs raccordés au Swissnet. Ce développement, réalisé avec la collaboration d'Ascom et de Siemens devrait permet-

tre d'explorer les possibilités d'un tel service et de définir ses exigences avant d'envisager son expérimentation à une plus grande échelle.

C'est toutefois dans le domaine plus conventionnel des réseaux numériques de distribution des programmes radiophoniques que le codage MICDA-SB a déjà trouvé une application des plus prometteuses. Ce n'est donc plus une surprise que de voir arriver sur le marché un multiplexeur 2 Mbit/s à 15 voies sonores de haute qualité comme le SB15 de Siemens. A l'occasion de sa présentation officielle, une liaison stéréophonique à  $2 \times 128$  kbit/s par le biais d'un satellite de la génération Eutelsat-1 a été expérimentée en direct.

Des essais par comparaison avec des signaux issus directement d'un disque compact ont démontré les performances remarquables de l'algorithme dans des conditions de transmission réelles et aussi d'écoute optimale au studio d'acoustique de l'Entreprise des PTT à Berne.

Cette liaison à trajets multiples comportait en outre plusieurs tronçons sur câbles à basse fréquence et sur des systèmes numériques sur câbles et faisceaux hertziens.

Toutes ces démonstrations ont confirmé que l'algorithme de codage MICDA-SB est particulièrement bien adapté aux caractéristiques spécifiques des réseaux de distribution numérique et du RNIS à bande étroite.

### Champ d'application du codage MICDA-SB

Le système SB15 de Siemens représente, dans le domaine de la transmission numérique des programmes de radiodiffusion, un exemple type d'une concrétisation matérielle des progrès les plus récents en matière de technologie et de l'impact de la prodigieuse puissance du traitement numérique du signal appliqué au codage à bas débit des signaux sonores de haute qualité. La méthode de codage MICDA-SB suggère des perspectives d'applications qui, jusqu'à présent, ne pouvaient être raisonnablement envisagées. Aujourd'hui, le temps est venu d'entreprendre une démarche concrète vers l'utilisateur en expérimentant par le biais du RNIS les possibilités et les avantages des services audiovisuels interactifs à  $2 \times 64$  kbit/s.

Dans un autre domaine, les Administrations des télécommunications et les Organismes de radiodiffusion doivent continuer à assurer et à améliorer les liaisons de contribution, de transmission et de distribution répondant à des besoins spécifiques caractérisés par des exigences de qualité, de bande passante, de densité d'informations et de coûts d'exploitation très différents. Pour répondre à cette demande plus conventionnelle, des équipements comparables au SB15 offrent une solution équilibrée en termes de qualité intrinsèque de codage, de résistance aux erreurs de transmission, de flexibilité et de simplicité relative. Les principales applications de ce type d'équipements sont le multiplexage et la transmission de signaux de programmes radiophoniques de haute qualité sur des conduits à 2 Mbit/s des réseaux de base nationaux et des circuits internationaux. Plus concrètement, les systèmes basés sur des algorithmes de codage MIC différentiel ou en sous-bandes sont appelés à remplacer dans un proche avenir



les systèmes à 384 kbit/s conformes aux Recommandations J.41 et J.42 du CCITT. Il en résultera une économie importante du débit binaire et d'autres avantages décisifs du point de vue du transmetteur telle que la résistance aux inévitables erreurs de transmission. Toutefois, il faudra être très attentif à maintenir les délais de transmission dans des limites qui n'occasionnent pas d'inconvénients dans l'exploitation traditionnelle des circuits de reportage.

Les Administrations pourraient être également intéressées à utiliser les possibilités du codage MICDA-SB avec les multiplexeurs programmables. Cette application est très prometteuse dans la mesure où il s'agit d'augmenter la flexibilité et le rendement des circuits à 64 kbit/s grâce à la gestion informatisée des réseaux (Network Management).

C'est toutefois dans le domaine des applications destinées à l'utilisateur que la solution décrite dans cet article s'impose par son originalité en réalisant automatiquement la synchronisation des deux canaux B du RNIS. Cette contrainte liée à la procédure de commutation étant résolue, une barrière importante est levée à l'introduction de nouveaux services à  $2 \times 64$  kbit/s.

Parmi ceux-ci figurera vraisemblablement l'audiovidéothèque photographique, c'est-à-dire le vidéotex tel qu'il est connu aujourd'hui mais qui sera doté en plus de la parole, du son de haute qualité et de l'image photographique. Cependant, pour envisager l'introduction d'un tel service, il manquait impérativement des algorithmes de compression du signal sonore de haute qualité et vidéo suffisamment efficaces pour réduire sans perte de qualité subjective le débit binaire à une valeur égale, voire inférieure à 128 kbit/s, équipant par exemple, un PC à l'usage de l'audiovidéothèque. Toute une série de nouvelles prestations pourront être proposées à l'utilisateur dans de nombreux domaines de l'enseignement, de l'information, de la culture et des loisirs. En voici quelques exemples:

- le service de l'audiovidéothèque permettra de consulter le catalogue des nouveautés du disque et de procéder à l'écoute d'extraits musicaux. Il est facile d'imaginer quel pourrait être l'intérêt d'un tel service
- dans les domaines de l'information et de l'enseignement assisté, l'audiovidéothèque permettra à l'utilisateur de s'affranchir des contraintes de la programmation des organismes de radiodiffusion et d'avoir accès à des services interactifs
- la TDHF à la carte! L'audiovidéothèque sera en mesure d'élargir les services traditionnels de la télédiffusion en offrant à l'auditeur un large choix de programmes.

## 5 Conclusion

Jusqu'à présent l'application de la transmission numérique des signaux audio s'est limitée à la distribution des programmes de radiodiffusion. La possibilité offerte par le système de codage MICDA-SB de transmettre un signal de haute qualité sur une voie de 128 kbit/s à un prix raisonnable ouvre la porte à un nouveau champ d'applications. C'est ainsi que le système SB15 permet la création de réseaux de reportage utilisant le réseau téléphonique numérique existant. Dans un proche avenir, l'application du MICDA-SB au réseau numérique à intégration de services (RNIS) pourrait permettre d'offrir à l'utilisateur une gamme de services audio interactifs impossible à réaliser jusqu'alors.

## Bibliographie

- [1] Johnston J. D. «Estimation of Perceptual Entropy Using Noise Masking Criteria», Proc. ICASSP 88, New York, pp. 2524–2527.
- [2] Huffman D. A. «A Method for the Construction of Minimum Redundancy Codes», Proc. I.R.E., 40, 1098, Sept. 1952.
- [3] Zwicker E. «Psychoakustik», Springer Verlag, Berlin, Heidelberg, New York 1982.
- [4] Krahé D. «Ein Verfahren zur Datenreduktion bei digitalen Audiosignalen unter Ausnutzung psychoakustischer Phänomene», Rundfunktechn. Mitt., n° 3, Mai–Juni 1986, pp. 117–123.
- [5] Johnston J. D. «Transform Coding of Audio Signals Using Perceptual Noise Criteria», IEEE Journal on Selected Areas in Comm., Vol. 6, No. 2, February 1988.
- [6] Vörös P. «High Quality Sound Coding within  $2 \times 64$  kbit/s Using Instantaneous Dynamic Bit-Allocation», Proc. ICASSP 88, New York, pp. 2536–2539.
- [7] Rohrecker L. «Ein Tonsignalcodierer hoher Qualität mit einer Datenrate von  $2 \times 64$  kbit/s durch ein adaptatives 4-Subband-Verfahren», Rundfunktechn. Mitt., Jahrg. 33 (1989) H.4, pp. 145–148.
- [8] Goodman D. J. «Embedded DPCM for variable bit rate transmission», IEEE Trans., COM-28, pp. 1040–1046, 1980.
- [9] Ramstad T. A. «Considerations on Quantization and Dynamic Bit-Allocation in Subband Coders», Proc. ICASSP 86, Tokyo, pp. 841–844.
- [10] CCIR. «Subjective Assessment of Sound Quality», Recommendations and Reports of the CCIR 1986, Vol. X – Part 1, pp. 275–279.



## Zusammenfassung

*Codierungssystem MICDA-SB für Tonsignale hoher Qualität*

Nachdem einige Grundbegriffe über die differentielle adaptive Pulsmodulation in Unterbändern (MICDA-SB) erwähnt werden, wird ein Codierungsalgorithmus zu  $2 \times 64$  kbit/s näher betrachtet. Dieser bietet eine Übertragungsqualität für Rundfunkprogramm-Signale, die mit jener bestehender 384-kbit/s-Systeme vergleichbar ist. Ferner wird ein Multiplexer behandelt, der es erlaubt, 15 Tonkanäle zu 15 kHz so zu gruppieren, dass sie über einen 2-Mbit/s-Kanal übertragen werden können. Zum Schluss werden einige Betrachtungen über die Verwendungsmöglichkeiten der MICDA-SB-Codierung sowie über die heutigen Erfahrungen dargelegt.

## Résumé

*Système de codage MICDA-SB pour signaux sonores de haute qualité*

Après avoir rappelé quelques notions fondamentales sur la modulation par impulsions et codage différentiel adaptatif en sous-bandes (MICDA-SB), on examine plus en détail un algorithme de codage à  $2 \times 64$  kbit/s qui offre une qualité de transmission des signaux de programmes radiophoniques comparable à celle des systèmes existants à 384 kbit/s. Par la suite, on aborde la réalisation pratique d'un équipement sous la forme d'un multiplexeur qui permet de regrouper 15 voies sonores à 15 kHz dans un conduit à 2 Mbit/s. L'article s'achève par quelques considérations sur le champ d'application du codage MICDA-SB et sur des expériences réalisées à ce jour.

## Riassunto

*Sistema di codifica MICDA-SB per segnali vocali di elevata qualità*

Dopo aver presentato alcuni concetti fondamentali della modulazione a impulsi codificati differenziale e adattativa nelle sottobande (MICDA-SB), gli autori spiegano un po' più dettagliatamente un algoritmo di codifica a  $2 \times 64$  kbit/s. Questo algoritmo offre una qualità di trasmissione dei segnali di programmi radiofonici paragonabile a quella dei sistemi esistenti a 384 kbit/s. In seguito descrivono un multiplexer che permette di raggruppare 15 canali vocali di 15 kHz ciascuno e di trasmetterli attraverso un canale a 2 Mbit/s. Infine fanno alcune considerazioni sulle possibilità d'applicazione della codifica MICDA-SB e sulle esperienze acquisite fino ad oggi.

## Summary

*Coding System MICDA-SB for High Quality Sound Signals*

The explanation of some basics of the Differential Adaptive Pulse Code Modulation in Subbands (MICDA-SB) is followed by a closer look at a coding algorithm for  $2 \times 64$  kbit/s offering a grade of transmission performance for sound broadcasting signals comparable to existing systems at 384 kbit/s. Furthermore, a multiplexer is presented which is capable to group 15 sound channels of 15 kHz each, so that they can be transmitted in one 2 Mbit/s channel. Finally, some considerations on the practical use of the MICDA-SB coding are stated and actual experiments are mentioned.