



UNION INTERNATIONALE DES TÉLÉCOMMUNICATIONS

UIT-T

SECTEUR DE LA NORMALISATION
DES TÉLÉCOMMUNICATIONS
DE L'UIT

H.261

(03/93)

**UTILISATION DES LIGNES POUR LA
TRANSMISSION DES SIGNAUX AUTRES
QUE TÉLÉPHONIQUES**

**CODEC VIDÉO POUR SERVICES
AUDIOVISUELS À $p \times 64$ kbit/s**

Recommandation UIT-T H.261

(Antérieurement «Recommandation du CCITT»)

AVANT-PROPOS

L'UIT-T (Secteur de la normalisation des télécommunications) est un organe permanent de l'Union internationale des télécommunications (UIT). Il est chargé de l'étude des questions techniques, d'exploitation et de tarification, et émet à ce sujet des Recommandations en vue de la normalisation des télécommunications à l'échelle mondiale.

La Conférence mondiale de normalisation des télécommunications (CMNT), qui se réunit tous les quatre ans, détermine les thèmes que les Commissions d'études de l'UIT-T doivent examiner et à propos desquels elles doivent émettre des Recommandations.

La Recommandation révisée UIT-T H.261, élaborée par la Commission d'études XV (1988-1993) de l'UIT-T, a été approuvée par la CMNT (Helsinki, 1-12 mars 1993).

NOTES

1 Suite au processus de réforme entrepris au sein de l'Union internationale des télécommunications (UIT), le CCITT n'existe plus depuis le 28 février 1993. Il est remplacé par le Secteur de la normalisation des télécommunications de l'UIT (UIT-T) créé le 1^{er} mars 1993. De même, le CCIR et l'IFRB ont été remplacés par le Secteur des radiocommunications.

Afin de ne pas retarder la publication de la présente Recommandation, aucun changement n'a été apporté aux mentions contenant les sigles CCITT, CCIR et IFRB ou aux entités qui leur sont associées, comme «Assemblée plénière», «Secrétariat», etc. Les futures éditions de la présente Recommandation adopteront la terminologie appropriée reflétant la nouvelle structure de l'UIT.

2 Dans la présente Recommandation, le terme «Administration» désigne indifféremment une administration de télécommunication ou une exploitation reconnue.

© UIT 1994

Droits de reproduction réservés. Aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie et les microfilms, sans l'accord écrit de l'UIT.

TABLE DES MATIÈRES

		<i>Page</i>
1	Champ d'application.....	1
2	Brève spécification.....	1
	2.1 Entrée et sortie vidéo	2
	2.2 Sortie et entrée numériques.....	2
	2.3 Fréquence d'échantillonnage.....	2
	2.4 Algorithme de codage de la source.....	2
	2.5 Débit	2
	2.6 Symétrie de transmission.....	3
	2.7 Traitement des erreurs	3
	2.8 Exploitation multipoint.....	3
3	Codeur de source.....	3
	3.1 Format de la source.....	3
	3.2 Algorithme de codage de la source vidéo.....	3
	3.3 Contrôle du débit des données.....	6
	3.4 Mise à jour forcée	6
4	Codeur de multiplexage vidéo.....	7
	4.1 Structure des données	7
	4.2 Configuration du multiplex vidéo.....	7
	4.3 Configurations multipoint.....	18
5	Codeur d'émission	19
	5.1 Débit binaire	19
	5.2 Mise en mémoire des données vidéo	19
	5.3 Retard de codage vidéo.....	20
	5.4 Correction d'erreur sans voie de retour pour le signal vidéo codé.....	20
	Annexe A – Spécification de la précision de la transformée inverse	21
	Annexe B – Décodeur fictif de référence	22
	Annexe C – Méthode de mesure du retard introduit par un codec	23
	Annexe D – Transmission d'images fixes.....	24

CODEC VIDÉO POUR SERVICES AUDIOVISUELS À $p \times 64$ kbit/s

(Genève, 1990; révisée à Helsinki, 1993)

Le CCITT,

considérant

- (a) qu'il existe une forte demande de la part de la clientèle pour les services de visiophonie, de visioconférence et d'autres services audiovisuels;
- (b) que les circuits permettant de satisfaire à cette demande peuvent être assurés par transmission numérique en utilisant les débits B et H_0 ou leurs multiples allant jusqu'au débit primaire, ainsi que les débits H_{11}/H_{12} ;
- (c) qu'on peut disposer des services RNIS dans un certain nombre de pays qui fournissent un service de transmission commuté aux débits B, H_0 ou H_{11}/H_{12} ;
- (d) que l'existence de hiérarchies numériques différentes et de normes de télévision différentes en diverses parties du monde, complique les problèmes de spécification de codage et de normes de transmission pour les communications internationales;
- (e) qu'un certain nombre de services audiovisuels sont susceptibles d'apparaître utilisant les accès de base et à débit primaire RNIS et qu'une certaine intercommunication devrait être possible entre ces terminaux;
- (f) que le codec vidéo constitue un élément essentiel de l'infrastructure des services audiovisuels qui permet cette intercommunication dans le cadre de la Recommandation H.200;
- (g) que la Recommandation H.120 du service de visioconférence utilisant le débit numérique primaire de transmission est la première d'une série évolutive de Recommandations,

appréciant

que des progrès ont été réalisés dans la recherche et la mise au point de techniques de codage vidéo et de réduction de débit binaire, qui conduiront à l'utilisation de débits binaires plus faibles allant jusqu'à 64 kbit/s, de sorte qu'on peut considérer la présente Recommandation comme la deuxième de la série évolutive de Recommandations,

et notant

que l'objectif fondamental du CCITT est de recommander des solutions uniques en ce qui concerne les communications internationales,

recommande

qu'en plus des codecs satisfaisant à la Recommandation H.120, les codecs ayant les caractéristiques de traitement, de transmission et de codage de signaux décrites ci-dessous, soient utilisés pour les communications de visioconférence internationale.

NOTES

- 1 Des codecs de ce type conviennent également pour certains services de télévision lorsque la qualité de diffusion habituelle ne s'impose pas.
- 2 Un équipement pour le transcodage assurant l'interfonctionnement dans les deux sens avec les codecs décrits dans la Recommandation H.120 est à l'étude.

1 Champ d'application

La présente Recommandation décrit les méthodes de codage et de décodage pour la composante image animée des services audiovisuels aux débits de $p \times 64$ kbit/s, p étant un nombre entier allant de 1 à 30.

2 Brève spécification

Un diagramme synoptique du codec est représenté à la Figure 1.

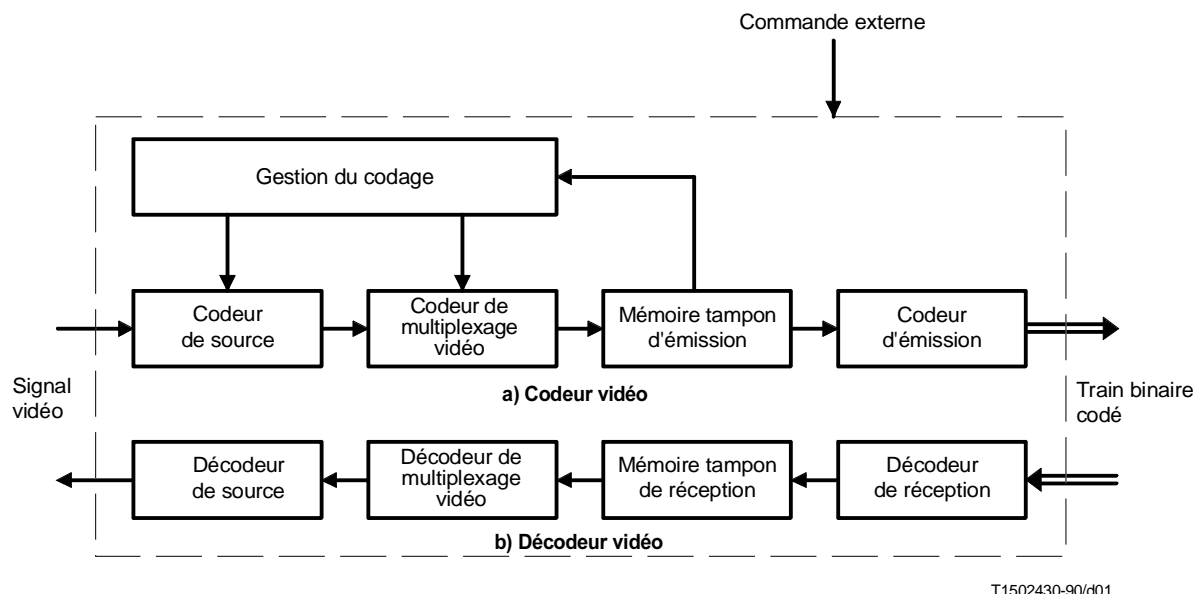


FIGURE 1/H.261
Diagramme synoptique du codec vidéo

2.1 Entrée et sortie vidéo

Pour permettre l'élaboration d'une Recommandation unique qui couvre l'utilisation des services audiovisuels dans et entre les régions appliquant des normes de télévision à 625 et à 525 lignes, le codeur de source travaille sur des images échantillonnées selon un format intermédiaire commun (CIF) (*common intermediate format*). Les normes des signaux de télévision d'entrée et de sortie qui peuvent, par exemple, être des signaux composites ou à plusieurs composantes, analogiques ou numériques et les méthodes employées pour effectuer toute conversion nécessaire au et du format de codage de source, ne font pas l'objet de recommandations.

2.2 Sortie et entrée numériques

Le codeur vidéo fournit un train binaire numérique autonome qui peut être combiné avec d'autres signaux multiservice (tels, par exemple, que ceux définis dans la Recommandation H.221). Le décodeur vidéo exécute l'opération inverse.

2.3 Fréquence d'échantillonnage

Les images sont échantillonnées suivant un multiple entier de la fréquence de ligne vidéo. Ce rythme d'horloge d'échantillonnage et le rythme d'horloge du réseau sont asynchrones.

2.4 Algorithme de codage de la source

On adopte un algorithme hybride de prédiction interimage qui utilise la redondance temporelle et de codage par transformée du signal restant qui réduit la redondance spatiale. Le décodeur est capable d'utiliser la compensation de mouvement, permettant l'introduction facultative de cette technique dans le codeur.

2.5 Débit

La présente Recommandation vise essentiellement l'utilisation de débits vidéo allant approximativement de 40 kbit/s à 2 Mbit/s.

2.6 Symétrie de transmission

On peut se servir du codec pour une communication audiovisuelle bidirectionnelle ou unidirectionnelle.

2.7 Traitement des erreurs

Le train de bits émis contient un code BCH (*Bose, Chaudhuri Hocquengham*) de correction d'erreur sans voie de retour (511,493) dont l'emploi par le décodeur est facultatif.

2.8 Exploitation multipoint

Les caractéristiques nécessaires pour une exploitation multipoint avec commutation sont incluses.

3 Codeur de source

3.1 Format de la source

Le codeur de source agit sur des images non entrelacées apparaissant 30 000/1001 (approximativement 29,97) fois par seconde. La tolérance sur la fréquence d'image est de $\pm 50 \times 10^{-6}$.

Les images sont codées sous forme d'une composante de luminance et de deux composantes de différence de couleur (Y , C_B et C_R). Ces composantes et les codes représentant leurs valeurs échantillonnées, sont ceux définis dans la Recommandation 601 du CCIR.

Noir = 16

Blanc = 235

Différence de couleur nulle = 128

Différence de couleur de crête = 16 et 240.

Ces valeurs sont des valeurs nominales et l'algorithme de codage fonctionne avec des valeurs d'entrée allant de 1 à 254.

Deux formats de balayage d'image sont spécifiés.

Dans le premier format (CIF), la structure d'échantillonnage de luminance est de 352 pixels par ligne, 288 lignes par image dans une disposition orthogonale. L'échantillonnage de chacune des deux composantes de différence de couleur est orthogonal à 176 pixels par ligne et 144 lignes par image. Les échantillons de différence de couleur sont situés de telle façon que leurs limites de bloc coïncident avec les limites de bloc de luminance comme indiqué à la Figure 2. La surface de l'image couverte par ces pixels et lignes a un format de 4:3 et correspond à la partie active du signal vidéo d'entrée conforme à la norme vidéo locale.

NOTE – Le nombre de pixels par ligne est compatible avec l'échantillonnage des parties actives des signaux de luminance et de différences de couleur respectivement à 6,75 et 3,375 MHz, pour des sources à 525 ou 625 lignes. Il existe un rapport simple entre ces fréquences et celles qui figurent dans la Recommandation 601 du CCIR.

Le second format vidéo quart de CIF (QCIF) (*quarter-CIF*) a un nombre de pixels et de lignes égal à la moitié des valeurs adoptées dans le premier format. Tous les codecs doivent pouvoir fonctionner au format QCIF; certains peuvent aussi fonctionner au format CIF.

Il devra être possible de limiter la cadence maximale du codeur, en choisissant 0, 1, 2 ou 3 images à ne pas transmettre entre les images transmises. La sélection de cette cadence et du format CIF ou QCIF se fera par des moyens externes (par exemple, ceux de la Recommandation H.221).

3.2 Algorithme de codage de la source vidéo

Le codeur de source est représenté sous forme synoptique à la Figure 3. Les principaux éléments sont la prédiction, la transformation de bloc et la quantification.

Pour le codage de l'erreur de prédiction (mode interimage ou INTER) ou de l'entrée vidéo (mode intra-image ou INTRA), l'image est subdivisée en blocs de 8 pixels \times 8 lignes qui sont classés en deux catégories: à transmettre, non transmis. De plus, quatre blocs de luminance et les deux blocs de différence de couleur qui leur correspondent spatialement sont combinés en un macrobloc (voir la Figure 10).

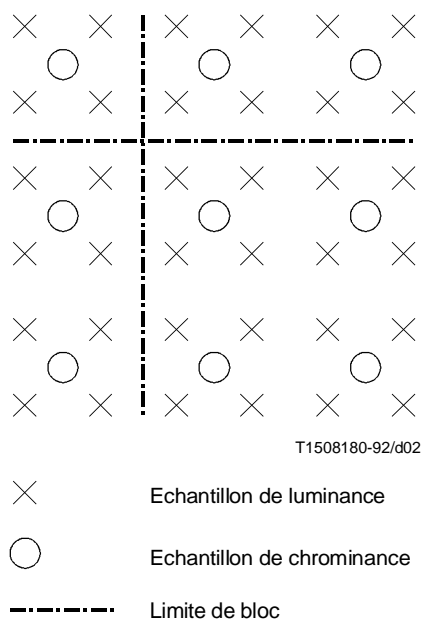


FIGURE 2/H.261

**Positionnement des échantillons de luminance
et de chrominance**

Les critères de choix du mode et de transmission d'un bloc ne font pas l'objet de recommandations et on peut les faire varier en tant que paramètres dynamiques intervenant dans la stratégie d'optimisation du codage. Les blocs transmis sont transformés et les coefficients résultants sont quantifiés, puis font l'objet d'un codage à longueur variable.

3.2.1 Prédiction

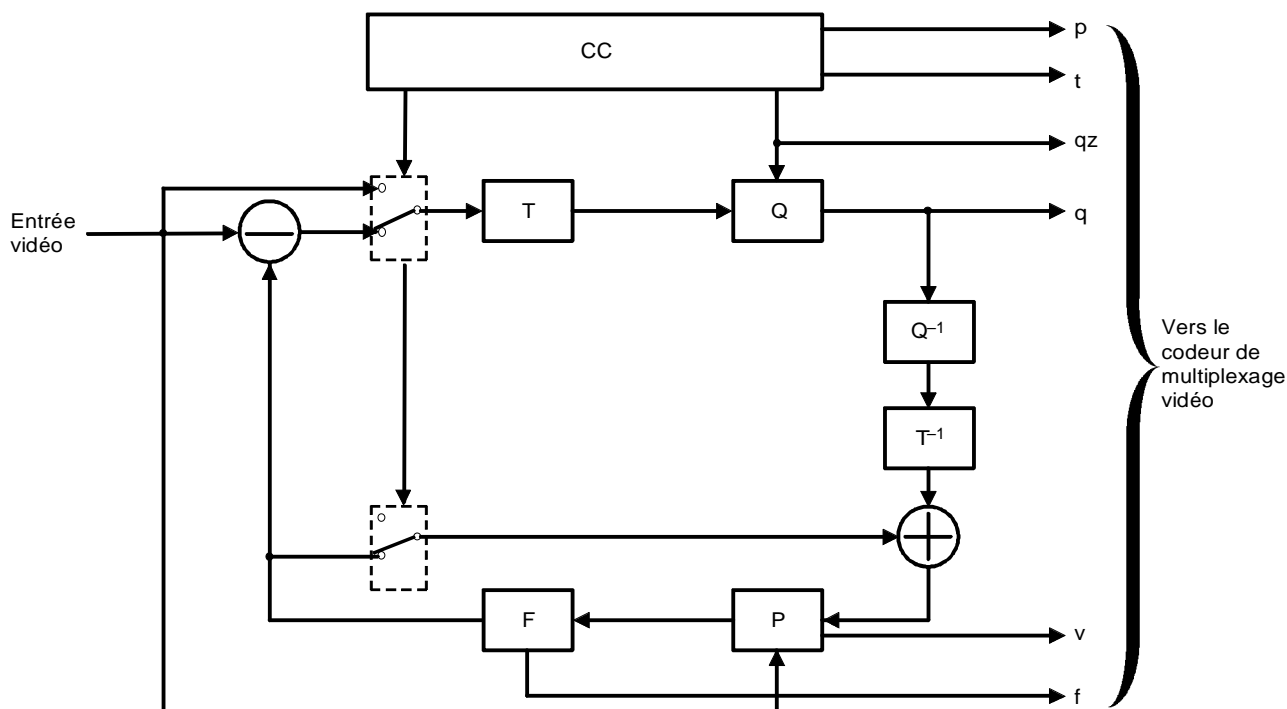
La prédiction est en interimage et peut être améliorée par compensation de mouvement (voir 3.2.2) et par un filtre spatial (voir 3.2.3).

3.2.2 Compensation de mouvement

La compensation de mouvement (MC) (*motion compensation*) est facultative dans le codeur. Le décodeur doit accepter un vecteur par macrobloc. Les composantes horizontale et verticale de ces vecteurs de mouvement sont des nombres entiers limités à l'intervalle ± 15 . Le vecteur est utilisé pour les quatre blocs de luminance du macrobloc considéré. Le vecteur de mouvement pour chacun des deux blocs de différence de couleur s'obtient en divisant par deux les valeurs des composantes du vecteur de macrobloc et en arrondissant par défaut les amplitudes à des valeurs entières.

Une valeur positive de la composante horizontale ou verticale du vecteur de mouvement indique que la prédiction repose sur des pixels de l'image précédente situés respectivement à droite ou au-dessous des pixels considérés.

Les vecteurs de mouvement sont bornés de telle sorte que tous les pixels auxquels ils font référence sont dans la zone d'image codée.



T1502441-90/d03

- T Transformation
- Q Quantificateur
- P Mémoire d'image avec retard variable de compensation de mouvement
- F Filtre de boucle
- CC Gestion du codage
- p Fanion pour INTRA/INTER
- t Fanion de transmission ou non d'un bloc
- qz Indication du quantificateur
- q Index de quantification pour les coefficients de la transformée
- v Vecteur de mouvement
- f Filtre activé ou non dans la boucle

FIGURE 3/H.261
Codeur de source

3.2.3 Filtre de boucle

Le traitement par prédiction peut être modifié par un filtre spatial (FIL) bidimensionnel qui agit sur les pixels d'un bloc de 8 par 8.

Ce filtre peut être factorisé en deux fonctions unidimensionnelles, l'une horizontale, l'autre verticale, non récursives, avec des coefficients de 1/4, 1/2, 1/4 sauf aux limites de blocs, où l'un des points sortirait du bloc considéré (en ce cas, les coefficients du filtre unidimensionnel sont 0, 1, 0). La précision arithmétique totale est conservée, avec arrondi à des nombres entiers de 8 bits à la sortie du filtre 2-D. Les valeurs dont la partie fractionnaire est égale à un demi sont arrondies par excès.

Le filtre est activé ou non pour les 6 blocs d'un macrobloc en fonction de la nature du bloc (voir 4.2.3 sous MTYPE).

3.2.4 Transformée

Les blocs transmis sont d'abord traités au moyen d'une transformée en cosinus discrète à 2 dimensions séparables (8 par 8). Les valeurs résultant de la transformation inverse vont de -256 à +255 après écrêtage permettant de les représenter sur 9 bits. La fonction de transfert de la transformation inverse répond à l'équation suivante:

$$f(x, y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u) C(v) F(u, v) \cos[\pi(2x + 1)u/16] \cos[\pi(2y + 1)v/16]$$

avec $u, v, x, y = 0, 1, 2, \dots, 7$

où x, y sont les coordonnées spatiales du domaine des pixels,

u, v sont les coordonnées du domaine transformé,

$$C(u) = 1/\sqrt{2} \text{ pour } u = 0, \text{ sinon } C(u) = 1,$$

$$C(v) = 1/\sqrt{2} \text{ pour } v = 0, \text{ sinon } C(v) = 1.$$

NOTE – Dans le bloc transformé, $x = 0$ et $y = 0$ se rapportent au point le plus proche des limites respectivement gauche et supérieure de l'image.

La méthode arithmétique de calcul des coefficients de transformation n'est pas définie, mais la transformée inverse doit répondre à la tolérance d'erreur spécifiée dans l'Annexe A.

3.2.5 Quantification

Le nombre de quantificateurs est de 1 pour la composante continue INTRA et de 31 pour tous les autres coefficients. Dans un macrobloc, le même quantificateur est utilisé pour tous les coefficients sauf la composante continue INTRA. Les seuils de décision ne sont pas définis. La composante continue INTRA est nominalement la valeur de la transformée quantifiée linéairement avec un pas de 8 sans plage interdite. Chacun des 31 autres quantificateurs est aussi nominalement linéaire, mais présente une plage interdite centrée autour du zéro et un pas de valeur paire comprise entre 2 et 62.

Les niveaux de reconstruction sont définis en 4.2.4.

NOTE – Pour les pas de quantification les plus petits, il n'est pas possible de représenter toute la gamme dynamique des coefficients de la transformée directe.

3.2.6 Ecrêtage de l'image reconstruite

Pour éviter que la distorsion d'amplitude des coefficients de la transformée due à la quantification ne provoque un débordement arithmétique dans les boucles du codeur et du décodeur, des fonctions d'écrêtage y sont insérées. Une telle fonction est appliquée à l'image reconstituée obtenue par la somme de la prédiction et de l'erreur de prédiction modifiées par le traitement de codage. L'écrêtage intervient sur les valeurs résultantes inférieures à 0 ou supérieures à 255, qui sont respectivement ramenées à 0 et à 255.

3.3 Contrôle du débit des données

Divers paramètres peuvent être modifiés pour gérer le débit de production de données vidéo codées, en agissant sur le traitement en amont du codeur de source, le quantificateur, le critère de sélection des blocs et le sous-échantillonnage temporel. L'importance relative de ces actions dans la stratégie globale de régulation ne fait l'objet d'aucune recommandation.

Lorsqu'il est mis en œuvre, le sous-échantillonnage temporel se fait par élimination d'images entières.

3.4 Mise à jour forcée

Cette fonction est obtenue en forçant l'algorithme de codage à revenir au mode INTRA. La séquence de mise à jour n'est pas définie. Pour limiter l'accumulation d'erreurs dans la transformation inverse, chaque macrobloc doit être mis à jour au moins une fois toutes les 132 fois où il est transmis.

4 Codeur de multiplexage vidéo

4.1 Structure des données

Sauf spécification contraire, le bit de plus fort poids est transmis d'abord. C'est le bit 1, le plus à gauche dans les tableaux de code de la présente Recommandation. Sauf spécification contraire, tous les bits inutilisés ou réservés sont mis à «1». Les bits en réserve ne doivent pas être utilisés tant que leur fonction n'aura pas été spécifiée par le CCITT.

4.2 Configuration du multiplex vidéo

Les signaux vidéo multiplexés sont organisés selon une structure hiérarchique en 4 couches, soit, de haut en bas:

- image,
- groupe de blocs (GOB) (*group of blocks*),
- macrobloc (MB) (*macroblock*),
- bloc.

La Figure 4 représente un diagramme syntaxique du codage de multiplexage vidéo; les abréviations y figurant sont explicitées plus loin.

4.2.1 Couche image

Pour chaque image, les données comportent un en-tête d'image suivi des données de groupes de blocs, selon la structure de la Figure 5. Les en-têtes des images éliminées ne sont pas transmis.

4.2.1.1 Code de début d'image (PSC) (20 bits)

Mot de 20 bits, de valeur: 0000 0000 0000 0001 0000.

4.2.1.2 Référence temporelle (TR) (5 bits)

Nombre comportant cinq bits, avec 32 valeurs possibles, formé par incrémentation d'une unité de sa valeur dans l'en-tête d'image précédemment transmis, à laquelle on ajoute le nombre d'images non transmises (à 29,97 Hz) dans l'intervalle. Le calcul ne porte que sur les cinq bits les moins significatifs.

4.2.1.3 Information de type (PTYPE) (6 bits)

Information concernant l'image entière:

- | | |
|-------|---|
| Bit 1 | Indicateur de partage d'écran: «0» inactif; «1» actif. |
| Bit 2 | Caméra de document: «0» inactif; «1» actif. |
| Bit 3 | Libération de gel d'image: «0» inactif; «1» actif. |
| Bit 4 | Indicateur de format: «0» QCIF; «1» CIF. |
| Bit 5 | Image fixe optionnelle en mode HI_RES défini en Annexe D; «0» inactif; «1» actif. |
| Bit 6 | En réserve. |

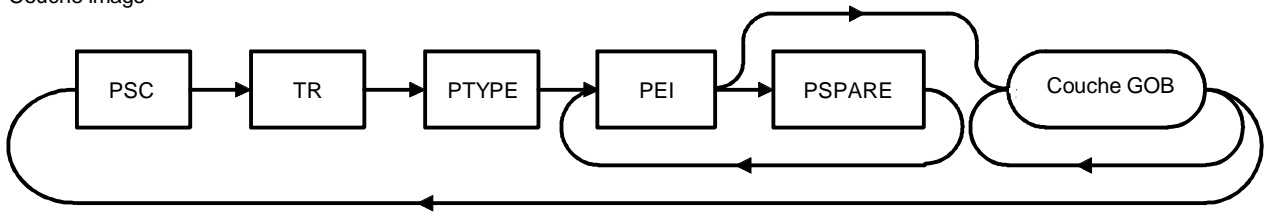
4.2.1.4 Information supplémentaire (PEI) (1 bit)

Un bit signalant lorsqu'il vaut 1 la présence du champ de données optionnel qui suit.

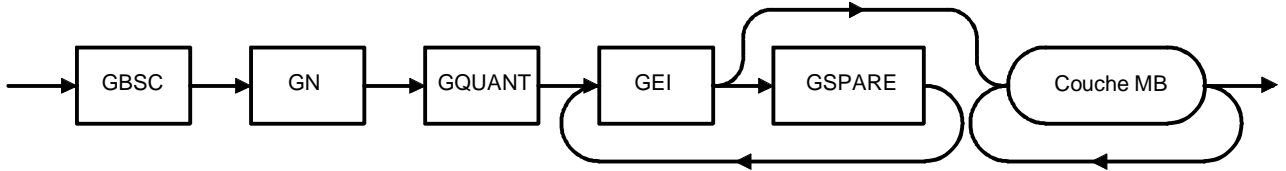
4.2.1.5 Informations réservées (PSPARE) (0/8/16 . . . bits)

Si PEI a la valeur «1», les 9 bits qui suivent sont 8 bits de données (PSPARE) et un autre bit PEI pour indiquer si 9 bits suivent encore, etc. Les codeurs ne doivent pas insérer PSPARE tant que cela n'est pas spécifié par le CCITT. Les décodeurs doivent être conçus de manière à éliminer PSPARE si PEI a la valeur 1. Le CCITT pourra ainsi spécifier de futures adjonctions compatibles avec les versions antérieures dans PSPARE.

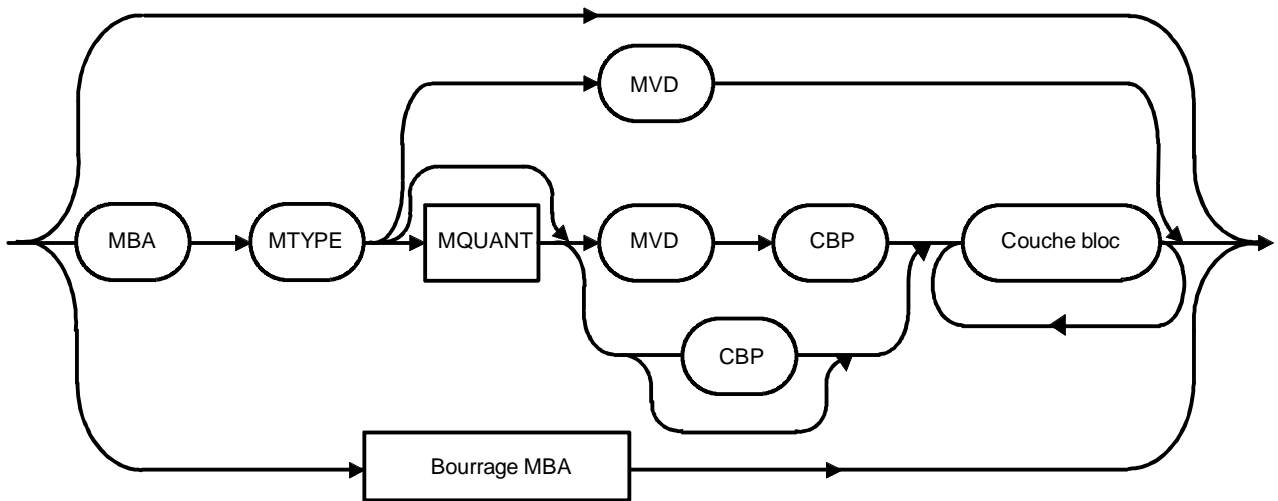
Couche image



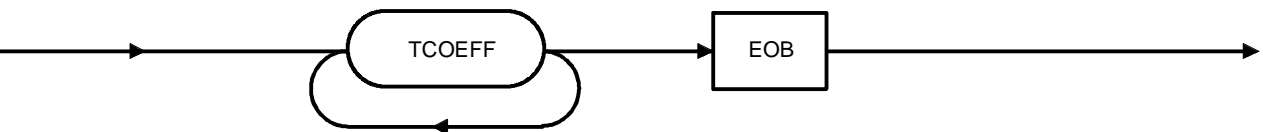
Couche GOB



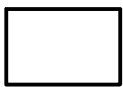
Couche MB



Couche bloc



T1502451-90/d04



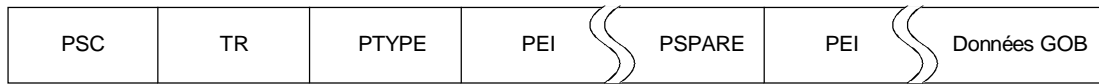
Longueur fixe



Longueur variable

FIGURE 4/H.261

Diagramme syntaxique du codage de multiplexage vidéo



T1514230-93/d05

FIGURE 5/H.261
Structure de la couche image

4.2.2 Couche groupe de blocs

Chaque image est subdivisée en groupes de blocs (GOB). Un groupe de blocs comprend un douzième de la zone d'image au format CIF ou un tiers de la zone d'image au format QCIF (voir la Figure 6). Un groupe de blocs (GOB) correspond à 176 pixels et 48 lignes de Y, et aux 88 pixels et 24 lignes de C_B ou C_R leur correspondant spatialement.

Les données de chaque groupe de blocs se composent d'un en-tête de groupe suivi des données correspondant aux macroblocs, selon la structure représentée à la Figure 7. Chaque en-tête de groupe de blocs est transmis une fois entre les codes de début d'image successifs, selon la séquence CIF ou QCIF de la Figure 6, même lorsque ce groupe de blocs ne comporte aucune donnée de macrobloc.

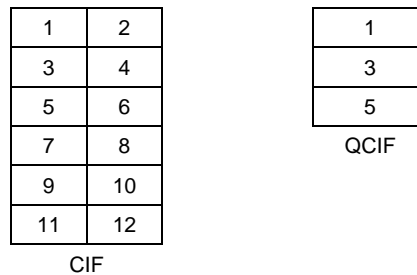
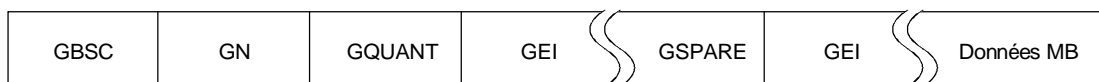


FIGURE 6/H.261

Disposition des groupes de blocs d'une image



T1514240-93/d06

FIGURE 7/H.261
Structure de la couche groupe de blocs

4.2.2.1 Code de début de groupe de blocs (GBSC) (16 bits)

Mot de 16 bits, de la forme 0000 0000 0000 0001.

4.2.2.2 Numéro de groupe (GN) (4 bits)

Quatre bits indiquant la position du groupe de blocs. Il s'agit de la représentation binaire des chiffres de la Figure 6. Les numéros de groupe 13, 14 et 15 sont réservés pour utilisation future. Le numéro de groupe 0 est utilisé dans le PSC.

4.2.2.3 Information de quantificateur (GQUANT) (5 bits)

Mot de code de longueur fixe de 5 bits, indiquant le quantificateur qui sera utilisé dans le groupe de blocs considéré jusqu'à infirmation par un code MQUANT. Les mots de code sont la représentation binaire naturelle des valeurs de QUANT (voir 4.2.4) qui, puisqu'ils correspondent à la moitié du pas, s'étagent de 1 à 31.

4.2.2.4 Information supplémentaire (GEI) (1 bit)

Un bit signalant lorsqu'il vaut 1 la présence du champ de données optionnel qui suit.

4.2.2.5 Information réservée (GSPARE) (0/8/16 . . . bits)

Si GEI a la valeur 1, les 9 bits suivants sont: 8 bits de données, (GSPARE) et un autre bit GEI pour indiquer si 9 autres bits suivront, etc. Les codeurs ne doivent pas insérer GSPARE en l'absence de spécification du CCITT. Les décodeurs doivent être conçus pour rejeter GSPARE si GEI a la valeur 1. Le CCITT sera ainsi en mesure de spécifier des adjonctions futures compatibles avec les versions antérieures dans GSPARE.

NOTE – L'émulation de codes de début peut se produire si la future spécification de GSPARE n'impose aucune restriction aux bits de données GSPARE.

4.2.3 Couche macrobloc

Chaque groupe de blocs est subdivisé en 33 macroblocs, selon la structure de la Figure 8. Un macrobloc correspond à 16 pixels par 16 lignes de Y, et aux 8 pixels par 8 lignes correspondants des signaux C_B et C_R .

Les données d'un macrobloc comprennent un en-tête de macrobloc suivi des données correspondant aux blocs (voir la Figure 9). La présence de MQUANT, MVD et CBP dépend de la valeur de MTYPE.

1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	21	22
23	24	25	26	27	28	29	30	31	32	33

FIGURE 8/H.261

Disposition des macroblocs d'un groupe de blocs

MBA	MTYPE	MQUANT	MVD	CBP	Données de bloc
-----	-------	--------	-----	-----	-----------------

FIGURE 9/H.261

Structure de la couche macrobloc

4.2.3.1 Adresse de macrobloc (MBA) (longueur variable)

Mot de code de longueur variable, indiquant la position d'un macrobloc dans un groupe de blocs. L'ordre de transmission est celui de la Figure 8. Dans le cas du premier macrobloc transmis dans un groupe de blocs, la MBA est l'adresse absolue de la Figure 8. Pour les macroblocs suivants, la MBA est la différence entre les adresses absolues du macrobloc considéré et du dernier macrobloc transmis. Le Tableau 1 donne la grille de codage des adresses MBA.

Dans cette grille, un mot de code supplémentaire est prévu pour le remplissage par des bits de bourrage après un en-tête de GOB ou un macrobloc codé (bourrage MBA). Ce mot de code doit être ignoré par le décodeur.

Le code de début est également indiqué dans le Tableau 1.

La MBA est toujours présente dans les macroblocs transmis.

Les macroblocs ne sont pas transmis lorsqu'ils ne comprennent aucune information concernant la partie de l'image considérée.

TABLEAU 1/H.261

Tableau des VLC pour l'adressage de macrobloc

MBA	Code	MBA	Code
1	1	17	0000 0101 10
2	011	18	0000 0101 01
3	010	19	0000 0101 00
4	0011	20	0000 0100 11
5	0010	21	0000 0100 10
6	0001 1	22	0000 0100 011
7	0001 0	23	0000 0100 010
8	0000 111	24	0000 0100 001
9	0000 110	25	0000 0100 000
10	0000 1011	26	0000 0011 111
11	0000 1010	27	0000 0011 110
12	0000 1001	28	0000 0011 101
13	0000 1000	29	0000 0011 100
14	0000 0111	30	0000 0011 011
15	0000 0110	31	0000 0011 010
16	0000 0101 11	32	0000 0011 001
		33	0000 0011 000
		Bourrage MBA	0000 0001 111
		Code de début	0000 0000 0000 0001

4.2.3.2 Information de type (MTYPE) (longueur variable)

Mots de code de longueur variable donnant des informations sur le macrobloc et les éléments de données présents. Les types de macroblocs, les éléments inclus et les mots de code de longueur variable sont indiqués au Tableau 2.

Le MTYPE est toujours présent dans les macroblocs transmis.

4.2.3.3 Quantificateur (MQUANT) (5 bits)

L'élément MQUANT n'est présent que lorsque le MTYPE l'indique.

Mot de code de 5 bits indiquant le quantificateur à utiliser dans ce macrobloc et les blocs suivants du groupe de blocs considéré jusqu'à remplacement par un élément MQUANT.

Les mots de code de l'élément MQUANT sont identiques à ceux de l'élément GQUANT.

TABLEAU 2/H.261

Tableau des VLC pour le MTYPE

Prévision	MQUANT	MVD	CBP	TCOEFF	VLC
Intra				x	0001
Intra	x			x	0000 001
Inter			x	x	1
Inter	x		x	x	0000 1
Inter + MC		x			0000 0000 1
Inter + MC		x	x	x	0000 0001
Inter + MC	x	x	x	x	0000 0000 01
Inter + MC + FIL		x			001
Inter + MC + FIL		x	x	x	01
Inter + MC + FIL	x	x	x	x	0000 01
NOTES					
1 «x» signifie que l'objet est présent dans le macrobloc.					
2 Il est possible d'appliquer le filtre dans un macrobloc compensé sans mouvement en le déclarant comme MC + FIL mais avec un vecteur nul.					

4.2.3.4 Données de vecteur de mouvement (MVD) (longueur variable)

Les données de vecteurs de mouvement (MVD) sont présentes dans tous les macroblocs à compensation de mouvement (MC). Les MVD sont obtenues à partir du vecteur du macrobloc en lui retranchant le vecteur du macrobloc précédent. Dans ce calcul, le vecteur du macrobloc précédent est considéré comme nul dans les trois cas suivants:

- 1) calcul des MVD pour les macroblocs 1, 12 et 23;
- 2) calcul des MVD pour les macroblocs d'adresses non consécutives;
- 3) lorsque le MTYPE du macrobloc précédent n'était pas MC.

Les données de vecteur de mouvement sont représentées par un mot de code de longueur variable pour la composante horizontale, suivi d'un mot de code de longueur variable pour la composante verticale. Les codes de longueur variable sont repris dans le Tableau 3.

Il est tiré parti du fait que la gamme des valeurs de vecteur de mouvement est limitée. Chaque mot de code de longueur variable représente deux valeurs de différence. Une seule de ces valeurs donne un vecteur de macrobloc compris dans la gamme autorisée.

4.2.3.5 Configuration de bloc codé (CBP) (longueur variable)

L'élément CBP est présent lorsque le MTYPE l'indique. Le mot de code fournit un numéro de séquence indiquant les blocs du macrobloc pour lesquels il est transmis au moins un coefficient de la transformée. Le numéro de séquence répond à la formule suivante:

$$32 \cdot P_1 + 16 \cdot P_2 + 8 \cdot P_3 + 4 \cdot P_4 + 2 \cdot P_5 + P_6$$

avec $P_n = 1$ lorsqu'un coefficient quelconque est présent pour le bloc n , et 0 dans le cas contraire. La numérotation des blocs est indiquée à la Figure 10.

Le tableau des codes de longueur variable pour CBP figure dans le Tableau 4.

TABLEAU 3/H.261

Tableau des codes de longueur variable VLC pour MVD

MVD	Code
-16 & 16	0000 0011 001
-15 & 17	0000 0011 011
-14 & 18	0000 0011 101
-13 & 19	0000 0011 111
-12 & 20	0000 0100 001
-11 & 21	0000 0100 011
-10 & 22	0000 0100 11
-9 & 23	0000 0101 01
-8 & 24	0000 0101 11
-7 & 25	0000 0111
-6 & 26	0000 1001
-5 & 27	0000 1011
-4 & 28	0000 111
-3 & 29	0001 1
-2 & 30	0011
-1	011
0	1
1	010
2 & -30	0010
3 & -29	0001 0
4 & -28	0000 110
5 & -27	0000 1010
6 & -26	0000 1000
7 & -25	0000 0110
8 & -24	0000 0101 10
9 & -23	0000 0101 00
10 & -22	0000 0100 10
11 & -21	0000 0100 010
12 & -20	0000 0100 000
13 & -19	0000 0011 110
14 & -18	0000 0011 100
15 & -17	0000 0011 010

4.2.4 Couche des blocs

Un macrobloc comprend 4 blocs de luminance et 2 blocs de différence de couleur (Figure 10).

Les données correspondant à un bloc se composent des mots de code des coefficients de la transformée suivis d'un marqueur de fin de bloc (EOB) (*end of block*) (Figure 11) transmis dans l'ordre indiqué à la Figure 10.

4.2.4.1 Coefficients de transformation (TCOEFF)

Les données de coefficients de la transformée sont toujours présentes pour la totalité des 6 blocs d'un macrobloc lorsque le MTYPE indique INTRA. Dans les autres cas, les indications «MTYPE» et «CBP» spécifient les blocs pour lesquels les coefficients sont transmis. Les coefficients quantifiés de la transformée sont transmis de façon séquentielle dans l'ordre indiqué à la Figure 12.

Le codage des combinaisons les plus courantes de séquences de zéros consécutifs (longueur) (RUN) et des valeurs qui suivent (amplitude) (LEVEL) est réalisé avec des codes de longueur variable. Les autres combinaisons (longueur, amplitude) (RUN, LEVEL) sont codées sur un mot de 20 bits composé de 6 bits d'échappement (ESCAPE), 6 bits de longueur (RUN) et 8 bits d'amplitude (LEVEL). Le codage de longueur variable se fait selon deux grilles de codage: la première est utilisée pour la première amplitude transmise dans les blocs INTER, INTER+CM et INTER+MC+FIL, la seconde pour toutes les autres amplitudes (LEVEL) à l'exception de la première des blocs INTRA; celle-ci a un code de longueur fixe de 8 bits.

TABLEAU 4/H.261

Tableau des VLC pour CBP

CBP	Code	CBP	Code
60	111	35	0001 1100
4	1101	13	0001 1011
8	1100	49	0001 1010
16	1011	21	0001 1001
32	1010	41	0001 1000
12	1001 1	14	0001 0111
48	1001 0	50	0001 0110
20	1000 1	22	0001 0101
40	1000 0	42	0001 0100
28	0111 1	15	0001 0011
44	0111 0	51	0001 0010
52	0110 1	23	0001 0001
56	0110 0	43	0001 0000
1	0101 1	25	0000 1111
61	0101 0	37	0000 1110
2	0100 1	26	0000 1101
62	0100 0	38	0000 1100
24	0011 11	29	0000 1011
36	0011 10	45	0000 1010
3	0011 01	53	0000 1001
63	0011 00	57	0000 1000
5	0010 111	30	0000 0111
9	0010 110	46	0000 0110
17	0010 101	54	0000 0101
33	0010 100	58	0000 0100
6	0010 011	31	0000 0011 1
10	0010 010	47	0000 0011 0
18	0010 001	55	0000 0010 1
34	0010 000	59	0000 0010 0
7	0001 1111	27	0000 0001 1
11	0001 1110	39	0000 0001 0
19	0001 1101		

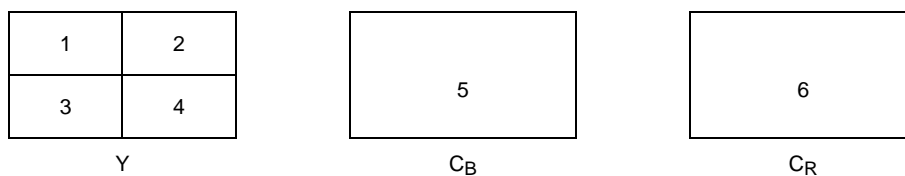


FIGURE 10/H.261

Disposition des blocs d'un macrobloc



FIGURE 11/H.261

Structure de la couche des blocs

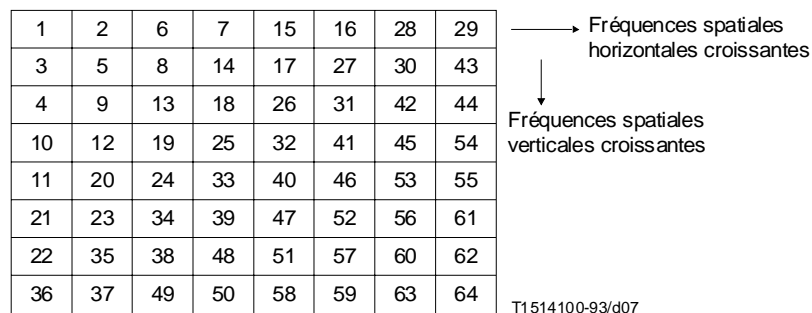


FIGURE 12/H.261

Ordre de transmission des coefficients de la transformée

Les codes sont indiqués dans le Tableau 5.

Les combinaisons les plus courantes de séquences de zéros et les valeurs qui suivent sont codées au moyen de codes de longueur variable donnés dans le Tableau 5. EOB (fin de bloc) y figure. CBP indiquant les blocs sans données de coefficient, EOB ne peut exister comme premier coefficient. EOB peut donc disparaître du tableau des VLC pour le premier coefficient.

Le dernier bit «s» indique le signe de l'amplitude, «0» pour positif et «1» pour négatif.

Les autres combinaisons (longueur, amplitude) sont codées au moyen d'un mot de 20 bits, composé de 6 bits échappement, 6 bits longueur et 8 bits amplitude. L'utilisation de ce mot de 20 bits pour le codage des combinaisons énumérées dans le tableau des VLC n'est pas interdite.

TABLEAU 5/H.261

Tableau des VLC pour TCOEFF

Longueur	Amplitude	Code
EOB		10
0	1	1s ^{a)} Si premier coefficient du bloc
0	1	11s Autre que 1 ^{er} coefficient du bloc
0	2	0100 s
0	3	0010 1s
0	4	0000 110s
0	5	0010 0110 s
0	6	0010 0001 s
0	7	0000 0010 10s
0	8	0000 0001 1101 s
0	9	0000 0001 1000 s
0	10	0000 0001 0011 s
0	11	0000 0001 0000 s
0	12	0000 0000 1101 0s
0	13	0000 0000 1100 1s
0	14	0000 0000 1100 0s
0	15	0000 0000 1011 1s
1	1	011s
1	2	0001 10s
1	3	0010 0101 s
1	4	0000 0011 00s
1	5	0000 0001 1011 s
1	6	0000 0000 1011 0s
1	7	0000 0000 1010 1s
2	1	0101 s
2	2	0000 100s
2	3	0000 0010 11s
2	4	0000 0001 0100 s
2	5	0000 0000 1010 0s
3	1	0011 1s
3	2	0010 0100 s
3	3	0000 0001 1100 s
3	4	0000 0000 1001 1s
4	1	0011 0s
4	2	0000 0011 11s
4	3	0000 0001 0010 s
5	1	0001 11s
5	2	0000 0010 01s
5	3	0000 0000 1001 0s
6	1	0001 01s
6	2	0000 0001 1110 s
7	1	0001 00s
7	2	0000 0001 0101 s
8	1	0000 111s
8	2	0000 0001 0001 s
9	1	0000 101s
9	2	0000 0000 1000 1s
10	1	0010 0111 s
10	2	0000 0000 1000 0s
11	1	0010 0011 s
12	1	0010 0010 s
13	1	0010 0000 s
14	1	0000 0011 10s
15	1	0000 0011 01s
16	1	0000 0010 00s
17	1	0000 0001 1111 s
18	1	0000 0001 1010 s
19	1	0000 0001 1001 s
20	1	0000 0001 0111 s
21	1	0000 0001 0110 s
22	1	0000 0000 1111 1s
23	1	0000 0000 1111 0s
24	1	0000 0000 1110 1s
25	1	0000 0000 1110 0s
26	1	0000 0000 1101 1s
Echappement		0000 01
a) Jamais utilisé dans les macroblochs INTRA.		

La longueur est codée en longueur fixe sur 6 bits

Longueur de séquence	Code
0	0000 00
1	0000 01
2	0000 10
.	.
.	.
63	1111 11

Amplitude est un code de 8 bits de longueur fixe

Amplitude	Code
-128	INTERDIT
-127	1000 0001
.	.
-2	1111 1110
-1	1111 1111
0	INTERDIT
1	0000 0001
2	0000 0010
.	.
127	0111 1111

Pour tous les coefficients autres que la composante continue INTRA, les amplitudes de reconstitution (REC) sont comprises entre -2048 et 2047 et sont obtenues par écrêtage des résultats donnés par les formules suivantes:

$$\left. \begin{array}{l} \text{REC} = \text{QUANT} \cdot (2 \cdot \text{amplitude} + 1); \text{amplitude} > 0 \\ \text{REC} = \text{QUANT} \cdot (2 \cdot \text{amplitude} - 1); \text{amplitude} < 0 \end{array} \right\} \text{QUANT} = \text{«impair»}$$

$$\left. \begin{array}{l} \text{REC} = \text{QUANT} \cdot (2 \cdot \text{amplitude} + 1) - 1; \text{amplitude} > 0 \\ \text{REC} = \text{QUANT} \cdot (2 \cdot \text{amplitude} - 1) + 1; \text{amplitude} < 0 \end{array} \right\} \text{QUANT} = \text{«pair»}$$

$$\text{REC} = 0; \text{amplitude} = 0$$

NOTE – QUANT est compris entre 1 et 31, il est transmis comme GQUANT ou MQUANT.

Amplitude de reconstruction (REC)

Amplitude	QUANT												
	1	2	3	4	.	8	9	.	17	18	.	30	31
-127	-255	-509	-765	-1019	.	-2039	-2048	.	-2048	-2048	.	-2048	-2048
-126	-253	-505	-759	-1011	.	-2023	-2048	.	-2048	-2048	.	-2048	-2048
.
-2	-5	-9	-15	-19	.	-39	-45	.	-85	-89	.	-149	-155
-1	-3	-5	-9	-11	.	-23	-27	.	-51	-53	.	-89	-93
0	0	0	0	0	.	0	0	.	0	0	.	0	0
1	3	5	9	11	.	23	27	.	51	53	.	89	93
2	5	9	15	19	.	39	45	.	85	89	.	149	155
3	7	13	21	27	.	55	63	.	119	125	.	209	217
4	9	17	27	35	.	71	81	.	153	161	.	269	279
5	11	21	33	43	.	87	99	.	187	197	.	329	341
.
56	113	225	339	451	.	903	1017	.	1921	2033	.	2047	2047
57	115	229	345	459	.	919	1035	.	1955	2047	.	2047	2047
58	117	233	351	467	.	935	1053	.	1989	2047	.	2047	2047
59	119	237	357	475	.	951	1071	.	2023	2047	.	2047	2047
60	121	241	363	483	.	967	1089	.	2047	2047	.	2047	2047
.
125	251	501	753	1003	.	2007	2047	.	2047	2047	.	2047	2047
126	253	505	759	1011	.	2023	2047	.	2047	2047	.	2047	2047
127	255	509	765	1019	.	2039	2047	.	2047	2047	.	2047	2047

NOTE – Les amplitudes de reconstruction sont symétriques par rapport au signe de l'amplitude sauf pour 2047/-2048.

Pour les blocs INTRA, le premier coefficient est nominalement la composante continue de la transformée, quantifiée linéairement avec un pas de 8, occupant la totalité de la zone d'information prévue à cet effet. Les valeurs résultantes sont représentées par 8 bits. Un bloc noir nominal est représenté par 0001 0000, le bloc blanc nominal par 1110 1011. Le code 0000 0000 n'est pas utilisé, non plus que le code 1000 0000, le niveau de reconstitution de 1024 étant codé sous la forme 1111 1111 (voir le Tableau 6).

Les coefficients suivant le dernier coefficient non nul ne sont pas transmis. L'indication EOB (code de fin de bloc) est toujours le dernier élément des blocs pour lesquels les coefficients sont transmis.

4.3 Configurations multipoint

Les services supplémentaires suivants sont fournis pour le fonctionnement multipoint en mode commuté.

4.3.1 Demande de gel d'image

Le décodeur gèle l'image affichée; la libération a lieu à la réception du signal de fin de gel ou à l'expiration d'une période de temporisation de 6 secondes au moins. La transmission de ce signal emprunte des moyens extérieurs (par exemple la Recommandation H.221).

4.3.2 Demande de rafraîchissement rapide

La demande de rafraîchissement rapide fait que le codeur code l'image suivante en mode INTRA avec les paramètres de codage prévenant tout débordement de la mémoire tampon. La transmission de ce signal emprunte des moyens extérieurs (par exemple la Recommandation H.221).

TABLEAU 6/H.261

Amplitude de reconstruction pour la composante continue du mode INTRA

FLC	Amplitude de reconstruction en transformée inverse
0000 0001 (1)	8
0000 0010 (2)	16
0000 0011 (3)	24
.	.
.	.
0111 1111 (127)	1016
1111 1111 (255)	1024
1000 0001 (129)	1032
.	.
.	.
1111 1101 (253)	2024
1111 1110 (254)	2032

NOTE – La valeur décodée correspondant au FLC «n» est 8 n, à l'exception du FLC 255 qui donne 1024.

4.3.3 Fin de gel d'image

Signal émis par le codeur en réponse à une demande de rafraîchissement rapide; le décodeur passe du mode gel au mode affichage normal. Le signal est transmis par le bit 3 de PTYPE (voir 4.2.1) dans l'en-tête d'image de la première image codée en réponse à la demande de mise à jour rapide.

5 Codeur d'émission

5.1 Débit binaire

Le rythme d'horloge d'émission est fourni de l'extérieur (par exemple, par une interface I.420).

5.2 Mise en mémoire des données vidéo

Le codeur doit réguler son train binaire sortant pour répondre aux spécifications du décodeur fictif de référence défini dans l'Annexe B.

En mode CIF, le nombre de bits créés par le codage d'une image quelconque ne doit pas dépasser 256 kbit/s.

En mode QCIF, le nombre de bits créés par le codage d'une image quelconque ne doit pas dépasser 64 kbit/s.

Dans ces deux cas, les valeurs recommandées comprennent le code de début d'image et toutes les données concernant l'image considérée y compris PSPARE, GSPARE et le bourrage MBA. Le décompte des bits ne comprend pas les bits de verrouillage de trame, de correction d'erreur, l'indicateur de bourrage (Fi), les bits de remplissage ou l'information de parité pour correction d'erreur décrite en 5.4.

Des données vidéo doivent être fournies à chaque cycle d'horloge: cela peut être assuré en recourant à l'indicateur de bits de bourrage (Fi) avec remplissage ultérieur de «1» dans le verrouillage de trame de bloc correcteur d'erreur (voir la Figure 13), au bourrage MBA (voir 4.2.3) ou aux deux procédés conjointement.

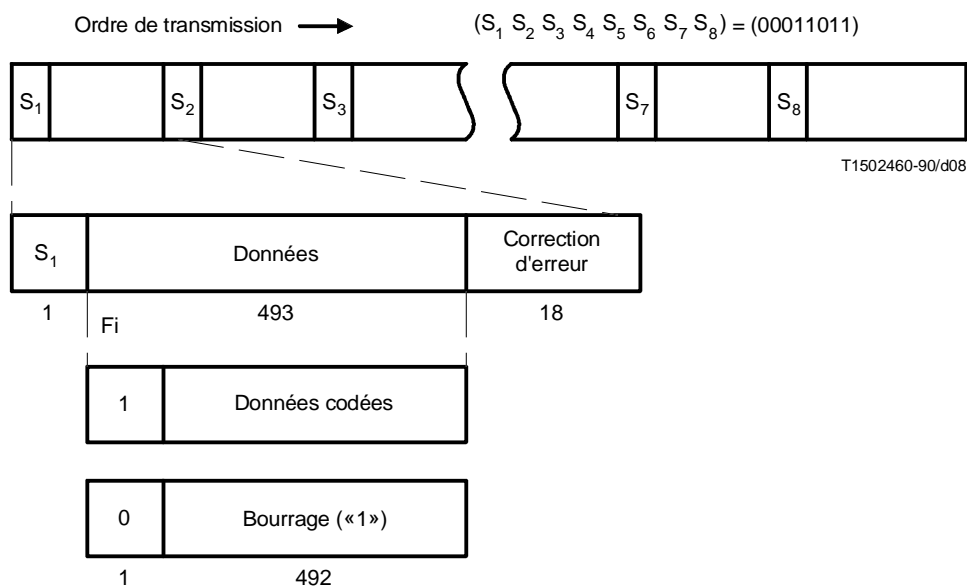


FIGURE 13/H.261
Trame de correction d'erreur

5.3 Retard de codage vidéo

Cet élément est inclus dans la présente Recommandation car les retards dus au codeur et au décodeur vidéo doivent être connus pour permettre d'appliquer des retards compensateurs sur l'audio lorsque la procédure H.261 est appliquée dans un service du type conversationnel. On peut ainsi conserver le synchronisme labial. L'Annexe C recommande une méthode de mesure des retards. D'autres méthodes peuvent être appliquées, mais elles doivent donner des résultats semblables.

5.4 Correction d'erreur sans voie de retour pour le signal vidéo codé

5.4.1 Code de correction d'erreur

Le train de bits émis contient un code de correction d'erreur sans voie de retour BCH (511,493). Son utilisation par le décodeur est facultative.

5.4.2 Polynôme générateur

$$g(x) = (x^9 + x^4 + 1)(x^9 + x^6 + x^4 + x^3 + 1)$$

Exemple: Pour l'entrée «01111 . . . 11» (493 bits), le polynôme générera la séquence de vérification «011011010100011011» (18 bits).

5.4.3 Verrouillage de trame de correction d'erreur

Pour que les données vidéo et l'information de correction d'erreur puissent être identifiées par un décodeur, un schéma de verrouillage de trame de correction d'erreur est inclus: il consiste en une multitrame de 8 trames comportant 1 bit de verrouillage de trame, 1 bit indicateur de bourrage (Fi), 492 bits de données codées (ou un remplissage de «1») et 18 bits de correction d'erreur. Le schéma de verrouillage de trame est:

$$(S_1S_2S_3S_4S_5S_6D_7S_8) = (00011011).$$

La disposition de la trame est indiquée à la Figure 13. La parité est calculée par rapport aux 493 bits, indicateur de bourrage (Fi) compris.

L'indicateur de bourrage (Fi) peut être mis à zéro par le codeur. Dans ce cas, il n'est émis que 492 bits de remplissage consécutifs (uniquement des «1»), ainsi que les bits de parité, aucune donnée codée n'étant émise. Cette méthode peut être utilisée pour répondre aux exigences du paragraphe 5.2 concernant l'émission de données vidéo dans chaque cycle d'horloge.

5.4.4 Temps de rétablissement du verrouillage de trame de correction d'erreur

Trois séquences consécutives (24 bits) de verrouillage de trame de correction d'erreur doivent être reçues avant que le verrouillage de trame soit considéré comme établi. Le décodeur doit être conçu de telle manière que le verrouillage de trame soit rétabli moins de 34 000 bits après un changement de phase de verrouillage de trame de correction d'erreur.

NOTE – A cette fin, on admet que les données vidéo ne contiennent pas 3 émulations de séquence de verrouillage de trame de correction d'erreur en phase entre elles pendant la période de rétablissement du verrouillage de trame.

Annexe A

Spécification de la précision de la transformée inverse

(Cette annexe fait partie intégrante de la présente Recommandation)

A.1 Générer des valeurs de pixels entières aléatoires allant de $-L$ à $+H$, selon le générateur de nombres aléatoires ci-dessous (version C). Disposer en blocs de 8×8 . Des ensembles de données de 10 000 blocs chacun doivent être générés pour $(L = 256, H = 255)$, $(L = H = 5)$ et $(L = H = 300)$.

A.2 Pour chaque bloc de 8×8 , effectuer une transformée en cosinus discrète directe, orthonormale séparable, avec multiplication matricielle, en virgule flottante sur au moins 64 bits conformément à la fonction de transfert suivante:

$$F(u, v) = \frac{1}{4} C(u) C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos[\pi(2x + 1)u/16] \cos[\pi(2y + 1)v/16]$$

avec $u, v, x, y = 0, 1, 2, \dots, 7$

où x, y sont les coordonnées spatiales dans le domaine des pixels,

u, v sont les coordonnées dans le domaine de la transformée,

$C(u) = 1/\sqrt{2}$ pour $u = 0$; sinon $C(u) = 1$,

$C(v) = 1/\sqrt{2}$ pour $v = 0$; sinon $C(v) = 1$.

A.3 Pour chaque bloc, arrondir les 64 coefficients de la transformée ainsi obtenus à la valeur entière la plus proche. Tronquer les valeurs à l'intervalle $[-2048$ à $+2047]$. On obtient ainsi des données d'entrée à 12 bits pour la transformée inverse.

A.4 Pour chaque bloc de 8×8 données à 12 bits produit en A.3, effectuer une transformée discrète inverse (IDCT) (*inverse discrete transform*) en cosinus, orthonormale séparable, avec multiplication matricielle, en virgule flottante sur au moins 64 bits. Arrondir les valeurs ainsi obtenues au nombre entier le plus proche, et les tronquer à l'intervalle $[-256$ à $+255]$. Ces blocs de 8×8 éléments constituent les données de sortie de référence de la transformée en cosinus discrète inverse.

A.5 Pour chaque bloc de 8×8 données produit en A.3, utiliser le dispositif à l'essai de transformation en cosinus discrète inverse, en tronquant la sortie à l'intervalle $[-256$ à $+255]$. Ces blocs de 8×8 pixels constituent les données de sortie d'essai de la transformée en cosinus discrète inverse.

A.6 Pour chacun des 64 pixels de sortie de la transformée en cosinus discrète inverse, et pour chacun des ensembles de 10 000 blocs générés, mesurer l'erreur maximale, l'erreur moyenne et l'erreur quadratique moyenne entre les données de référence et les données d'essai.

A.7 Pour tout pixel, l'erreur absolue maximale ne doit pas être supérieure à 1.

Pour tout pixel, l'erreur quadratique moyenne ne doit pas dépasser 0,06.

Sur l'ensemble, l'erreur quadratique moyenne ne doit pas dépasser 0,02.

Pour tout pixel, l'erreur absolue moyenne ne doit pas être supérieure à 0,015.

Sur l'ensemble, l'erreur absolue moyenne ne doit pas être supérieure à 0,0015.

A.8 Tous les zéros à l'entrée doivent donner des zéros à la sortie.

A.9 Procéder à une seconde série de mesures en utilisant exactement les mêmes valeurs de données qu'en A.1, mais en inversant le signe de chaque pixelc.

Programme «C» de génération de nombres aléatoires

```

/* L et H sont des entiers longs de 32 bits */
long rand    (L,H)
long        L,H;
{
    static long randx = 1;                /* long est à 32 bits */
    static double z = (double) 0x7fffffff;

    long i,j;
    double x;                            /* double est à 64 bits */

    randx = (randx * 1103515245) + 12345;
    i = randx & 0x7fffffff;               /* conserver 30 bits */
    x = ((double)i) / z;                  /* valeur de 0 à 0,99999 ... */
    x * = (L+H+1);                         /* valeur de 0 à < L+H+1 */
    j = x;                                 /* partie entière */
    return( j - L );                       /* valeur de -L à H */
}

```

Annexe B

Décodeur fictif de référence

(Cette annexe fait partie intégrante de la présente Recommandation)

Le décodeur fictif de référence (HRD) (*hypothetical reference decoder*) se définit comme suit:

B.1 Le HRD et le codeur ont la même fréquence d'horloge ainsi que la même cadence d'images CIF et ils travaillent en synchronisme.

B.2 La capacité de la mémoire tampon de réception du HRD est ($B + 256$ kbit/s). La valeur de B est définie comme suit:

$$B = 4R_{max}/29,97, R_{max} \text{ étant le débit vidéo maximal à utiliser sur la connexion.}$$

B.3 Initialement, la mémoire tampon du HRD est vide.

B.4 La mémoire tampon du HRD est examinée à intervalles CIF (≈ 33 ms). Si une image codée complète au moins se trouve dans cette mémoire, toutes les données correspondant à l'image la plus ancienne sont instantanément supprimées (par exemple à t_{n+1} sur la Figure B.1). Aussitôt après la suppression de ces données, l'occupation de la mémoire tampon doit être inférieure à B . C'est là une condition imposée au train binaire sortant du codeur, comprenant les données d'images codées et le bourrage, MBA, mais non les bits de verrouillage de trame, l'indicateur de bourrage (Fi), les bits de remplissage ou l'information de correction d'erreur selon la description du paragraphe 5.4.

Pour respecter cette condition, le nombre de bits de la $(n + 1)^{\text{e}}$ image codée d_{n+1} doit satisfaire l'inégalité:

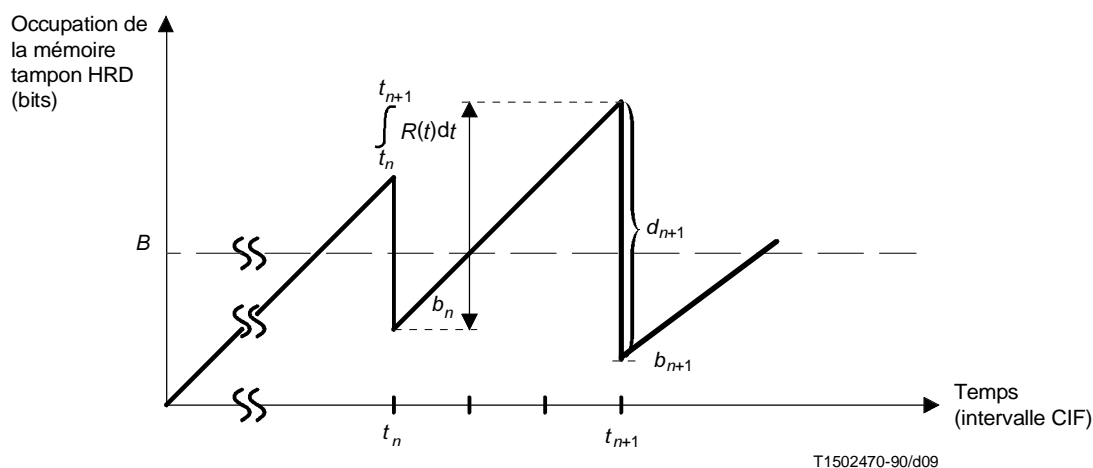
$$d_{n+1} \geq b_n + \int_{t_n}^{t_{n+1}} R(t) dt - B$$

dans laquelle

b_n est l'occupation de la mémoire tampon juste après l'instant t_n ;

t_n est l'instant où la $n^{\text{ième}}$ image codée est supprimée de la mémoire tampon du HRD,

$R(t)$ est le débit binaire vidéo à l'instant t .



NOTE – L'intervalle $(t_{n+1} - t_n)$ est un nombre entier de périodes d'images (1/29,97, 2/29,97, 3/29,97, ...).

FIGURE B.1/H.261
Occupation de la mémoire tampon HRD

Annexe C

Méthode de mesure du retard introduit par un codec

(Cette annexe fait partie intégrante de la présente Recommandation)

Les retards introduits par les codeurs et les décodeurs vidéo varient selon la réalisation ainsi que selon le format de l'image (QCIF, CIF) et le débit utilisé. On trouvera dans cette annexe une spécification de la méthode permettant d'établir la valeur de ce retard pour un modèle donné. Pour permettre une compensation correcte par un retard du signal audio, il faut déterminer le retard total du signal vidéo tel qu'il est perçu par l'utilisateur dans des conditions typiques d'observations.

Le point A est l'entrée du codeur vidéo. Le point B est la sortie du canal à partir du terminal vidéo (c'est-à-dire y compris la correction d'erreur, verrouillage de trame, etc.). Le point C est la sortie vidéo du décodeur.

Une séquence vidéo de plus de 100 secondes est introduite à l'entrée du codeur vidéo (point A) de la Figure C.1. Cette séquence vidéo doit avoir les caractéristiques suivantes:

- elle doit contenir des scènes animées typiques correspondant à l'utilisation prévue pour le codec vidéo;
- elle doit produire une cadence d'images codées de 7,5 Hz au débit utilisé;
- elle doit contenir un repère d'identification visible périodiquement tout au long de la séquence. Cette identification visible doit changer toutes les 97 trames vidéo entrantes et être située dans la zone d'image représentée par le premier groupe de blocs de l'image. Par exemple, le premier bloc de l'image peut passer du noir au blanc à intervalles de 97 périodes de trame vidéo. Ce repère d'identification doit être choisi de manière à pouvoir être détecté au point B et ne doit pas notablement influencer la qualité globale du codage.

Le codec et la séquence vidéo doivent être tels que le train binaire contienne moins de 10% de bits de remplissage (bourrage MBA + bits de remplissage de code correcteur d'erreur).

Le retard dû au codeur s'obtient en mesurant la durée qui sépare le changement du repère visible en A de la détection de ce changement en B. De même, le retard dû au décodeur s'obtient par des mesures faites en B et en C.

Plusieurs mesures doivent être faites pendant la durée de la séquence afin d'en établir la valeur moyenne. Plusieurs essais doivent être faits pour garantir l'obtention d'une valeur moyenne cohérente pour les retards dus au codeur et au décodeur.

Des valeurs moyennes doivent ainsi être établies pour chaque combinaison de format d'image et de débit dans le cadre des possibilités d'un type particulier de codec.

NOTE – En raison du prétraitement et du post-traitement, il peut être nécessaire de prendre un niveau moyen pour déterminer la transition du repère d'identification aux points B et C.

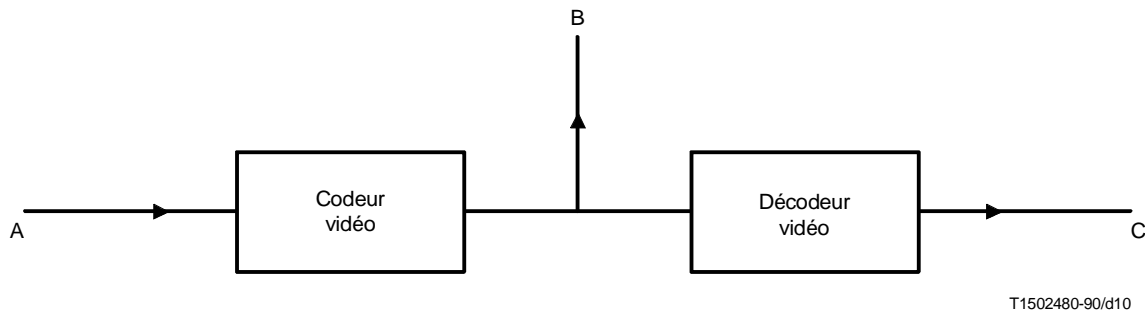


FIGURE C.1/H.261

Points de mesure

Annexe D

Transmission d'images fixes

(Cette annexe fait partie intégrante de la présente Recommandation)

D.1 Introduction

Cette annexe décrit la procédure de transmission d'images fixes dans le cadre de la présente Recommandation. Elle permet à un codeur vidéo conforme à la norme H.261 de transmettre des images fixes avec une résolution quatre fois supérieure à celle obtenue par l'arrêt sur image vidéo animée. Les Administrations pourront utiliser cette procédure optionnelle comme méthode simple et bon marché de transmission d'images fixes. Toutefois, il est préférable de suivre à cette fin la procédure définie dans la Recommandation T.81 (JPEG) lorsque cette procédure sera normalisée pour les systèmes audiovisuels.

La procédure de transmission d'images fixes peut assurer une transmission d'images de haute qualité avec des effets similaires à ceux obtenus avec les procédés progressifs et hiérarchiques. Un minimum de modifications par rapport à la procédure H.261 de base (faible coût), une compatibilité rétroactive avec les terminaux existants, une souplesse dans le choix de la qualité d'images en fonction de la vitesse de transmission ont constitué les points clés pris en considération pour le développement de cette procédure.

NOTE – Lorsqu'il transmet des images fixes, le codeur doit mettre à 0 un des bits de PTYPE (information de type) précédemment non utilisé (les bits non utilisés doivent être mis à 1). S'il ne tient pas compte de ce bit, un décodeur recevra cette image comme une image vidéo normale. Un décodeur qui passe à l'état d'erreur lorsque ce bit est à 0 gèlera vraisemblablement la trame vidéo précédente, et reprendra le décodage lorsque ce bit repassera à 1. Un décodeur disposant de cette nouvelle fonction affichera l'image avec une meilleure résolution, transférera l'image vers un écran graphique distinct et maintiendra l'image après la reprise de la transmission vidéo, imprimera ou enregistrera l'image, etc.

D.2 Format d'image fixe

Le format d'image fixe est quatre fois celui de l'image vidéo qui était en cours de transmission. Si le format vidéo était le QCIF (quart de CIF), alors l'image fixe est une trame CIF (format intermédiaire commun). Si le format vidéo était le CIF avec 352×288 échantillons de luminance, l'image fixe contiendra 704×576 échantillons de luminance et un accroissement correspondant du nombre d'échantillons de chrominance (trame CCIR 601).

Pour la transmettre avec la procédure H.261, l'image fixe est sous-échantillonnée au rapport 2:1 horizontalement et verticalement en quatre sous-images dans le format vidéo en cours de transmission. La Figure D.1 montre le dispositif de sous-échantillonnage de l'image fixe. Les échantillons étiquetés 0, 1, 2 et 3 forment respectivement les quatre sous-images 0, 1, 2 et 3.

0	3	0	3	0	3	0	3
1	2	1	2	1	2	1	2
0	3	0	3	0	3	0	3
1	2	1	2	1	2	1	2
0	3	0	3	0	3	0	3
1	2	1	2	1	2	1	2

FIGURE D.1/H.261

Disposition de sous-échantillonnage

D.3 Multiplexage de couche image

Lorsque le bit HI_RES est à «0», les deux bits de poids le plus faible de la référence temporelle (TR) (*temporal reference*) identifient l'une des quatre sous-images 0, 1, 2 ou 3. Les trois bits de poids le plus fort de TR seront mis à «0».

Pour transmettre une image fixe, le codeur met le bit HI_RES à «0» puis transmet séquentiellement les quatre sous-images 0, 1, 2 et 3. Il est permis de transmettre plus d'une trame pour chaque sous-image, mais une fois la transmission de la sous-image suivante commencée, aucun retour en arrière n'est possible. Le codeur peut reprendre à tout moment la transmission vidéo animée en remettant le bit HI_RES à «1».

NOTE – La référence en mémoire de la trame en cours est toujours la trame précédente, que cette trame corresponde à de la vidéo animée ou à une image fixe.

D.4 Considérations multipoints

Une image fixe transmise dans un train binaire vidéo peut être diffusée sur une liaison multipoints par diffusion de la vidéo. Les commandes de diffusion de la vidéo (MCV) (*multipoint command visualization-forcing*) et annuler-MCV (annulation de diffusion de la vidéo) définies dans la Recommandation H.230 offrent cette possibilité. Un terminal peut ainsi forcer un pont de conférence à diffuser ses images vidéo en envoyant la commande MCV, puis revenir au mode de fonctionnement précédent en envoyant la commande annuler-MCV. La prise en charge de ces commandes est imposée pour les ponts de conférence, mais facultative pour les terminaux.

D.5 Autres considérations

- Tous les modes de codage vidéo sont autorisés (intra-trame, inter-trames, compensation de mouvement, etc.);
- les dispositions de multiplexage en dessous de la couche image restent les mêmes (groupes de blocs, macroblocs, etc.);
- le nombre maximal autorisé de bits par trame (sous-image) ne doit pas être dépassé (256 kbit/s pour le format CIF et 64 kbit/s pour le format QCIF);
- la correction d'erreur sans boucle de retour n'est pas affectée.