

## Dynamique des métriques dans l'algorithme de VITERBI

D.DUPONTEIL  
P.TORTELIER

CNET/PAB/ETR  
38-40 RUE DU GENERAL LECLERC  
92131 ISSY LES MOULINEAUX

### RESUME

L'algorithme de VITERBI se trouve au cœur d'un nombre de plus en plus grand de réalisations mettant en œuvre la détection ou l'estimation optimales, en raison des progrès continus dans le domaine des circuits intégrés. Or, dans le cadre d'une réalisation de ce type, il convient de dimensionner au mieux le nombre de bits nécessaires à la représentation des grandeurs intervenant dans le calcul, en particulier des métriques de nœud dont la comparaison est à la base de la détermination du message le plus vraisemblable. Cet article donne une méthode pour calculer la borne supérieure atteinte par ces quantités.

### SUMMARY

The VITERBI Algorithm is more and more implemented in systems based on optimal detection due to growing improvements in VLSI circuits. Nevertheless, for such implementation it is necessary to know the exact number of bits needed for the representation of the values computed by the algorithm. This paper describes a method to determine the true upperbound of the node metrics whose comparison gives the most likely message.

#### Accroissement de métrique sur un chemin

Soit  $R_1, R_2 \dots R_L$  une suite de  $L$  symboles reçus en sortie du démodulateur, chacun d'eux étant un point de  $R^2$  de la forme  $R_k = (X_k, Y_k)$ . On considère un chemin  $C$  constitué de  $L$  branches consécutives du treillis, portant respectivement les symboles  $Z_1, Z_2 \dots Z_L$ . L'accroissement de métrique sur ce chemin est égal à :

$$\Delta_C = \sum_{k=1}^L \|R_k - Z_k\|^2 = \sum_{k=1}^L \|R_k\|^2 + \|Z_k\|^2 - 2 R_k \cdot Z_k$$

On suppose que la modulation utilisée est une MDP4 (QPSK) de sorte que les symboles émis sont de la forme  $Z_k = (2A_k - 1, 2B_k - 1)$  où  $A_k$  et  $B_k$  sont des éléments de  $GF(2)$ . L'accroissement de métrique prend alors la forme :

$$\Delta_C = 2L + \sum_{k=1}^L \|R_k\|^2 - 2 \sum_{k=1}^L X_k (2A_k - 1) + Y_k (2B_k - 1)$$

Pour un ensemble de symboles reçus donné, cet accroissement prend une valeur minimum atteinte pour une suite de bits  $\alpha_k$  et  $\beta_k$  telle que, pour tout indice  $k$  :

$$(2\alpha_k - 1) X_k \geq 0 \text{ et } (2\beta_k - 1) Y_k \geq 0$$

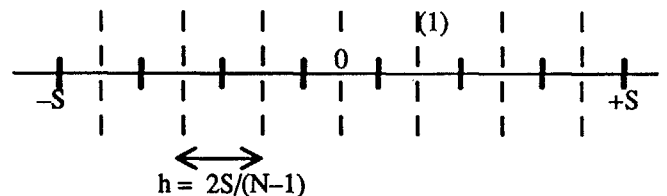
Comme cette valeur minimum est indépendante des symboles de branche  $Z_k$ , il est naturel de considérer, pour tous les chemins  $C$ , l'écart entre l'accroissement de métrique et sa valeur minimum :

$$\delta_C = 4 \sum_{k=1}^L X_k (\alpha_k - A_k) + Y_k (\beta_k - B_k)$$

Il est facile de vérifier que la quantité ainsi définie est non-

négative, et qu'elle correspond à l'opération qui consiste, à chaque étape, à retrancher la plus petite métrique de branche à toutes les métriques de branches de sorte que la distance du symbole reçu au symbole le plus proche soit égale à zéro.

D'autre part nous pouvons introduire l'opération de quantification effectuée sur les signaux en sortie du démodulateur, faite de manière identique sur les deux axes: Le nombre d'intervalles  $N$  est choisi pair, la valeur crête est  $S$  et  $h$  le pas de quantification.



les valeurs quantifiées seront des multiples impairs de la quantité  $h/2$  de la forme

$$X_k = M_k h/2 \text{ et } Y_k = N_k h/2$$

où  $M_k$  et  $N_k$  sont des nombres entiers impairs appartenant à l'ensemble :

$$\{ -(N-1) \dots -7, -5, -3, -1, 1, 3, 5, 7, \dots (N-1) \}$$

A un facteur multiplicatif près qu'on peut négliger parce qu'il n'affecte pas le résultat des comparaisons de métriques de nœud, on peut prendre comme accroissement de métrique sur un chemin  $C$  la quantité

$$\delta_C = \sum_{k=1}^L M_k (\alpha_k - A_k) + N_k (\beta_k - B_k)$$

Enfin, en remarquant que  $M_k (\alpha_k - A_k) = |M_k| (\alpha_k \oplus A_k)$ , où  $\oplus$  désigne l'addition modulo 2, l'équation précédente se met sous la forme suivante: (eq.1)

Expression qui montre que les accroissements de métriques pos-



$$\delta_c = \sum_{k=1}^L |M_k| (\alpha_k \oplus A_k) + |N_k| (\beta_k \oplus B_k)$$

sibles à chaque itération sont:

0, |M<sub>k</sub>|, |N<sub>k</sub>|, |M<sub>k</sub>| + |N<sub>k</sub>|. Toutes les opérations sur les métriques mises en œuvre par l’algorithme de Viterbi (addition, comparaison, choix) porteront donc sur des nombres entiers.

### Rappels sur le fonctionnement du codeur

Soient G(D) et H(D) les deux polynômes générateurs du codeur, de même degré K :

$$G(D) = \sum_{k=0}^K g_k D^k \quad \text{et} \quad H(D) = \sum_{k=0}^K h_k D^k$$

Le motif élémentaire du treillis correspondant à ce codeur est un "papillon" ( voir figure 1 ).

Les coefficients de degré 0 et ceux de degré K des deux polynômes générateurs du codeur sont toujours choisis égaux à 1 de sorte que les symboles binaires qu’on peut trouver sur les deux branches issues d’un même nœud ou convergeant en un même nœud soient complémentaires et les symboles MDP4 correspondants opposés. Les accroissements de métriques sur les branches (métriques de branches) seront respectivement égaux à:

$\delta_1 = |M| (\alpha \oplus A) + |N| (\beta \oplus B)$ , sur une branche de symbole (A,B)

$\delta_2 = |M| (\alpha \oplus A \oplus 1) + |N| (\beta \oplus B \oplus 1)$ , sur une branche de symbole complémentaire (A⊕1,B⊕1)

Il y a 4 cas possibles selon l’emplacement du symbole reçu :

$\alpha$	$\beta$	$\delta_1$	$\delta_2$
A	B	0	M  +  N
A	B ⊕ 1	N	M
A ⊕ 1	B	M	N
A ⊕ 1	B ⊕ 1	M  +  N	0

Les nouvelles métriques de nœud Γ3 et Γ4 sont calculées par:

$$\Gamma_3 := \min(\Gamma_1 + \delta_1, \Gamma_2 + \delta_2)$$

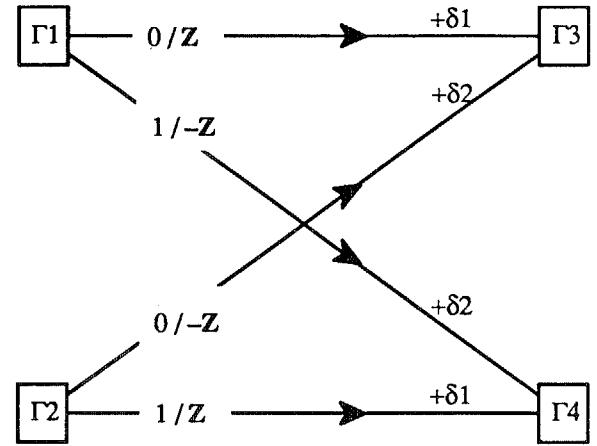
$$\Gamma_4 := \min(\Gamma_1 + \delta_2, \Gamma_2 + \delta_1)$$

D’autre part, si on se donne deux états quelconques (e<sub>-1</sub>, e<sub>-2</sub>, ..., e<sub>-K</sub>) et (e<sub>K-1</sub>, ..., e<sub>1</sub>, e<sub>0</sub>), il existe un chemin, et un seul, de K branches dans le treillis qui permet de passer du premier au second.

Cette propriété ne fait que traduire le fait que la mémoire du codeur se présente sous la forme d’un registre à décalage de longueur K: Pour passer du premier état au second il faut entrer successivement dans le codeur les bits e<sub>0</sub>, e<sub>1</sub>, ..., e<sub>K-1</sub>.

### Evolution de la plus petite métrique

La quantité δ<sub>c</sub> définie par eq.1 dépend de la suite de symboles reçus R<sub>k</sub>, de la longueur L du chemin considéré et de la suite des



Γ<sub>i</sub> : métriques de nœud . Z = symbole MDP4

fig.1 : "papillon" élémentaire du treillis d’un codeur 1/2

bits A<sub>k</sub> et B<sub>k</sub> correspondants à ce chemin . On notera m(L) la valeur minimum de δ<sub>c</sub> sur tous les chemins de longueur L; c’est donc une quantité qui dépend uniquement de la suite des symboles reçus R<sub>k</sub> ( et de L ) . On a l’encadrement suivant:

$$m(L-1) \leq m(L) \leq m(L-1) + (N-1)$$

En effet, considérons un "papillon" du treillis dont l’un des états initiaux (noté N\*) a une métrique égale à m(L-1) . On peut supposer sans perte de généralité que Γ<sub>1</sub>=m(L-1) . D’après les cas recensés plus haut, si le symbole reçu est dans le même quadrat que le symbole (A<sub>L</sub>, B<sub>L</sub>) ou son complémentaire, une des deux métriques de branche δ<sub>1</sub> ou δ<sub>2</sub> est nulle et un des successeurs de ce nœud aura pour métrique la quantité min(Γ<sub>1</sub>+0, Γ<sub>2</sub>+|M<sub>L</sub>|-|N<sub>L</sub>|) toujours égale à Γ<sub>1</sub> puisque Γ<sub>2</sub> ≥ Γ<sub>1</sub> .(fig.2, cas (1)) . Il y a donc continuité de la plus petite métrique.

Dans le cas contraire les métriques de branches sont égales à |M<sub>L</sub>| et |N<sub>L</sub>| et les successeurs du nœud N\* auront une métrique non supérieure à Γ<sub>1</sub> + max(|M<sub>L</sub>|, |N<sub>L</sub>|), quantité inférieure à m(L-1) - (N-1) car les valeurs quantifiées sont toujours inférieures ou égales à (N-1) en valeur absolue ( fig.2, cas(2) ). Dans le but de montrer qu’un accroissement maximum de la plus petite métrique se fait toujours à partir d’un nœud de plus petite métrique, on peut vérifier qu’un nœud de plus petite métrique m(L-1)+(N-1) ne peut pas provenir d’un nœud de métrique m(L-1)+u, avec u < N-1 .

Si u est impair, il faudrait une métrique de branche égale à N-1-u, quantité paire ( N a été choisi pair ) qui ne peut être obtenue que comme la somme de deux métriques de branches élémentaire impaires |M<sub>L</sub>| + |N<sub>L</sub>|, chacune d’entre elles étant donc < N-1. Comme on est dans le cas(2) de la figure 2, les deux successeurs du nœud N\* auraient une métrique inférieure ou égale respectivement à Γ<sub>1</sub> + |M<sub>L</sub>| et Γ<sub>1</sub> + |N<sub>L</sub>|, toutes deux inférieures à Γ<sub>1</sub> + N-1. Si u est pair, il faudrait une métrique de branche égale à N-1-u qui cette fois est une quantité impaire et est donc égale à une métrique de branche élémentaire |M<sub>L</sub>| ou |N<sub>L</sub>| inférieure à N-1. Mais dans ce cas un des deux successeurs (pas forcément les deux) de N\* aurait une métrique inférieure à N-1. Reste le cas u=N-1.



faut avoir  $|M_L| = |N_L| = N-1$ , sinon  $\Gamma_3$  et  $\Gamma_4$  seraient inférieurs à  $\Gamma_1 + N - 1$ . Mais alors  $\Gamma_3 = \Gamma_4 = \Gamma_1 + N - 1 = m(L-1) + (N-1)$  qui provient donc aussi d'un nœud de métrique  $m(L-1)$ .

Donc dans tous les cas un accroissement maximum de la plus petite métrique se fait toujours à partir d'un nœud de plus petite métrique et implique que  $|M_L| = |N_L| = N-1$  .i.e: Ses deux successeurs ont la même métrique  $m(L-1) + N - 1$ . S'il arrivait que la plus petite métrique augmente K fois de suite d'un accroissement maximum  $N-1$ , les  $2^K$  nœuds du treillis auraient tous la même métrique. Comme les branches du treillis portent tous les symboles possibles, quelque soit le symbole reçu une des métriques de branches serait nulle à l'étape suivante et la plus petite métrique n'augmenterait pas. On a donc:

La plus petite métrique est une fonction non décroissante du temps ; au mieux, elle ne peut augmenter que K fois de suite de son incrément maximum.

On s'intéresse maintenant à l'écart entre la plus grande et la plus petite métriques (la dynamique).

### Comparaison entre accroissements de métriques

On considère deux chemins  $C_1$  et  $C_2$  de même longueur  $L$  et portant respectivement les suites de symboles  $Z_{i,1}, Z_{i,2}, \dots, Z_{i,L}$  avec  $i=1,2$ . Les accroissements respectifs de métriques sur les deux chemins sont donc donnés par:

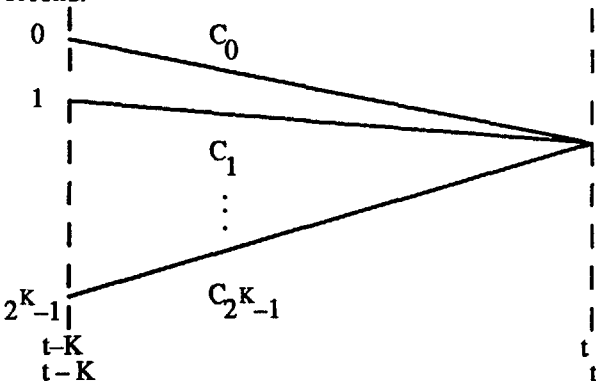
$$\delta_C(i) = \sum_{k=1}^L |M_k| (\alpha_k \oplus A_{i,k}) + |N_k| (\beta_k \oplus B_{i,k})$$

avec  $i=1,2$ . La soustraction de ces deux accroissements fait intervenir des quantités de la forme  $(a \oplus b) - (a \oplus c)$  où  $a, b, c$  sont dans  $GF(2)$ . On vérifie facilement (il suffit de passer en revue les 8 cas possibles) que  $|(a \oplus b) - (a \oplus c)| = b \oplus c$  et donc:

$$|\delta_C(1) - \delta_C(2)| \leq \sum_{k=1}^L |M_k| (A_{1,k} \oplus A_{2,k}) + |N_k| (B_{1,k} \oplus B_{2,k})$$

Le maximum étant atteint lorsque l'un des deux termes  $\delta_C(1)$  ou  $\delta_C(2)$  est nul.

**Application:** On considère les  $2^K$  nœuds du treillis à l'instant  $t-K$  et un nœud quelconque à l'instant  $t$ . On sait qu'il existe un et un seul chemin permettant de joindre un quelconque des premiers au second.



Appelons  $A$  la quantité définie par

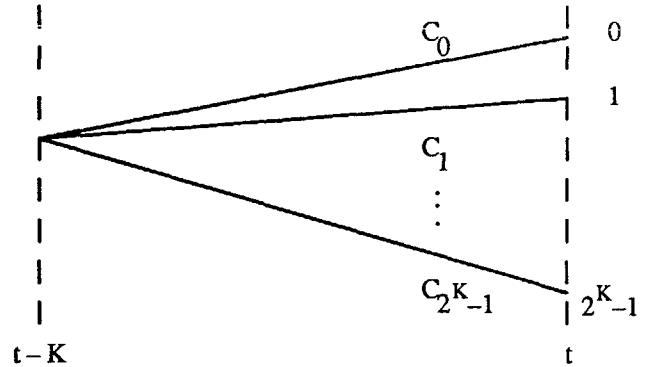
$$A = \sup_{i,j} \left\{ \sum_{k=1}^L (A_{i,k} \oplus A_{j,k}) + (B_{i,k} \oplus B_{j,k}) \right\}$$

On peut montrer que cette quantité ne dépend pas du nœud choisi à l'instant  $t$ . La différence maximum entre les accroissements de métriques sur deux chemins  $C_i$  et  $C_j$  est alors égale à  $A(N-1)$ . On peut en tirer une conclusion sur la dynamique des métriques à l'instant  $t-K$ . En effet soit  $n_1$  le nœud de plus petite métrique (notée  $m$ ) à l'instant  $t-K$ ,  $n_2$  le nœud de plus grande métrique (notée  $M$ ) au même instant.  $\delta_1$  et  $\delta_2$  les accroissements respectifs de métriques sur les chemins reliant ces nœuds à un même nœud  $n$  à l'instant  $t$ . Supposons que  $M-m > A(N-1)$ . Alors

$$M + \delta_2 - (m + \delta_1) = (M-m) + (\delta_2 \pm \delta_1) > A(N-1) + (\delta_2 \pm \delta_1)$$

quantité qui, à cause de la définition de  $A$ , est non-négative. Donc, lors d'une comparaison de métrique, le nœud  $n_2$  serait systématiquement préféré au nœud  $n_1$ . D'où: Les seuls nœuds à l'instant  $t-K$  qui participeront à des opérations de comparaison de métrique à l'instant  $t$  sont ceux dont l'écart de métrique avec le nœud de plus petite métrique est inférieur ou égal à  $A(N-1)$ .

De manière analogue considérons un nœud quelconque à l'instant  $t-K$  et les  $2^K$  chemins, tous de longueur  $K$ , permettant de joindre ce nœud à un nœud quelconque à l'instant  $t$ :



La quantité  $B$  définie de manière analogue à  $A$  par:

$$B = \sup_{i,j} \left\{ \sum_{k=1}^L (A_{i,k} \oplus A_{j,k}) + (B_{i,k} \oplus B_{j,k}) \right\}$$

ne dépend pas du nœud de départ choisi. Considérons à l'instant  $t$  deux nœuds  $n_1$  et  $n_2$  de métriques respectives  $m_1$  et  $m_2$ . Appelons  $n_0$  l'ancêtre de  $n_1$  à l'instant  $t-K$ , de métrique  $m_0$ , et  $\delta m_1$ ,  $\delta m_2$  les accroissements respectifs de métrique sur les chemins permettant d'aller de  $n_0$  à  $n_1$  et  $n_2$ . On peut donc écrire  $m_1 = m_0 + \delta m_1$ . Supposons que  $m_2 - m_1 > B(N-1)$ . Alors  $m_2 - (m_0 + \delta m_2) = m_2 - m_1 + \delta m_1 - \delta m_2 > B(N-1) + \delta m_1 - \delta m_2$  quantité non négative par définition de  $B$ . Le chemin provenant de  $n_0$  aurait une métrique inférieure à  $m_2$ , il serait donc préférable. A un instant  $t$  donné, l'écart entre deux métriques de nœud quelconques est borné par  $B(N-1)$

On a dit que les quantités  $A$  et  $B$  ne dépendaient pas du nœud choisi aux instants  $t-K$  et  $t$ . Il suffit de choisir le nœud  $0$  d'où part

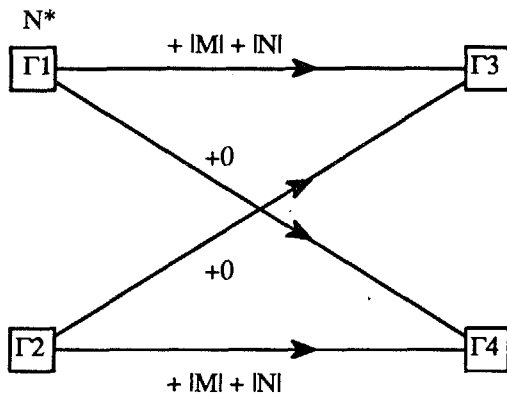


ou arrive un chemin de poids de Hamming nul. A est alors égal au poids maximum d'un chemin de K branches issu du nœud 0, B au poids maximum d'un chemin de K branches incident au nœud 0. Les deux quantités sont obtenues à l'aide d'une exploration du treillis sur une profondeur de K étapes par une méthode analogue à celle du calcul de la fonction de transfert d'un codeur convolutif.

Le résultat important est que la dynamique des métriques de nœuds est bornée, ce qui, combiné à la renormalisation de ces mêmes métriques, permet de les représenter sur un nombre fini de bits.

**Exemples**

Soit le codeur à 4 états de polynômes générateurs



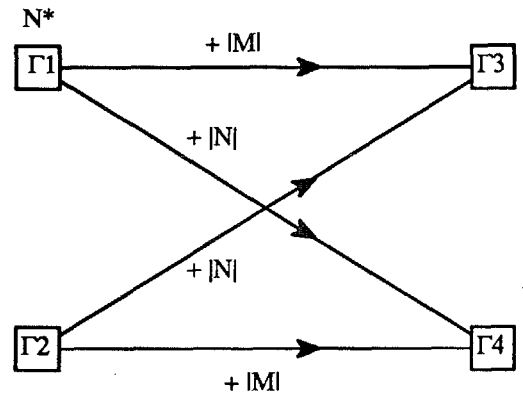
cas (1)

$$G(D)=1+D+D^2 \text{ et } H(D)=1+D^2$$

On trouve  $A=B=3$

Avec une quantification sur 3 bits ( $N=8$ ) cela donne une dynamique égale à 21.

On sait que ce codeur possède une distance libre égale à 5; on pourrait penser que A (et B) sont liés de manière simple à la distance libre. Afin de montrer que ce n'est pas nécessairement le cas, on considère un codeur à 64 états de polynômes générateurs  $G(D)=1+D^2+D^4+D^6$  et  $H(D)=1+D+D^2+D^5+D^6$  dont la fonction de transfert est  $X^8+X^9+3X^{10}+\dots$  pour ce codeur on trouve  $A=B=7$  Mais  $7=9-2$  et le premier chemin qui reconverge vers le chemin tout-zéro est bien de poids égal à 9. Avec une quantification sur 3 bits, la dynamique est donc : 49, à comparer avec la valeur  $6 \times 14 = 84$  obtenue avec la majoration usuelle et qui nécessite un bit de plus pour être représentée en binaire.



cas (2)

fig.2

**REFERENCES**

[1] J-K OMURA  
Optimal Receiver Design for Convolutional Codes and Channels with Memory Via Control Theoretical Concepts  
Information Sciences, Vol. 3, 1971, pp 243-266

[2] A-J VITERBI, J-K OMURA

Principles of Digital Communications and Coding.  
Mc Graw-Hill, 1979

[3] G-D FORNEY  
The Viterbi Algorithm  
Proc. IEEE, mars 1973, pp 268-278