Cuantizarea vectorială
Partea a I-a. Prezentarea cuantizării vectoriale

Ş.l. ing Cornel Balint
prof. dr. ing. Miranda Nafornită
Universitatea Politehnica Timişoara

ABSTRACT. During last years, vector quantization becomes a powerful technique for data compression, in particular for image and speech signals. This paper, structured as two part series, present a tutorial on vectorial coding, introducing the fundamentals vectorial cuantizer and of adaptive vectorial coding as a method for coding nonstationary sources. Part two presents a general mathematical definition of adaptive vector quantization and an original classification.

1. Introducere

Codarea vectorială este o generalizare a codării scalare, fiind o extindere a acesteia la codarea unui vector privit ca un set de numere reale. Un vector poate reprezenta de exemplu un segment de semnal vocal sau de imagine.

În literatură sunt prezentate numeroase metode de cuantizare vectorială, fiecare dintre acestea fiind dezvoltate şi optimizate pentru un anumit tip de semnal particular, avându-se în vedere anumite criterii de performanţă impuse.

Un număr mare din aceste metode de cuantizare se numesc metode de cuantizare adaptivă, acest calificativ fiind utilizat în diverse sensuri, pentru a indica faptul că algoritmul este adaptat într-un fel sau altul semnalului de intrare, în vederea realizării unei cuantizări optime într-un sens precizat, pentru semnalul dat [GG93]. În cele ce urmează se vor prezenta principiile cuantizării vectoriale şi se va realiza o clasificare a metodelor de cuantizare adaptivă.

2. Cuantizarea vectorială

Un cuantizor vectorial VQ [GG93], [Ger82], k - dimensional şi de mărime N este o aplicaţie a unui vector sau a unui punct din spaţiul euclidian k-dimensional, P^k pe o mulţime finită X numită dicţionar (codebook) şi conţinând N elemente, numite cuvinte de cod:

$$Q : \mathrm{R}^k \rightarrow \mathrm{C}, \quad \mathrm{C} \in \mathrm{R}^k \tag{2.1.}$$

$$\mathrm{C} = (\mathbf{y}_{1,}\mathbf{y}_2 ....\mathbf{y}_N) \text{ şi } \mathbf{y}_i \in \mathrm{R}^k \text{, pentru orice } i \in \mathbf{I} \; = \{1,2,3,...N\}, \tag{2.2.}$$

unde I este o mulţime de numere întregi, numite index.

Cuantizorul vectorial divide spaţiul R^k în regiuni:

$$R_i = \left\{ x \in \mathrm{R}^k : Q(x) = \mathbf{y}_i \right\}, \quad i \in \mathbf{I} \tag{2.3.}$$

care formează o partiţie a spaţiului R^k : $\bigcup_i R_i = \mathrm{R}^k$ şi $R_i \cap R_j = \varnothing$ (2.4.)

şi asociază fiecărei regiuni un cuvânt de cod din C.

Se defineşte un cuantizor vectorial ca fiind regulat, dacă partiţiile P_i ale spaţiului vectorial de intrare formează un set convex, adică dacă pentru pentru orice $a_i, b_i \in \mathrm{R}_i$ şi orice $0 < \lambda < 1$, avem îndeplinită condiţia: $(1-\lambda)a_i + \lambda b_i \in \mathrm{R}_i$

Un cuantizor vectorial este politopal dacă este regulat şi dacă partiţiile P_i ale spaţiului P_k sunt limitate de hiperplane în spaţiul P_k.

Un cuantizor vectorial constă din două părţi: un codor şi un decodor, reprezentate în fig. 2.1.

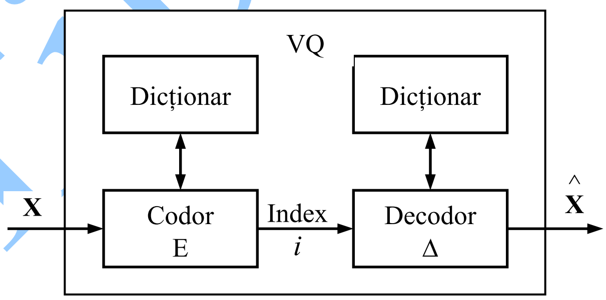

Fig. 2.1. Codorul şi decodorul vectorial

Codorul E realizează maparea spaţiului R^k spre indexul I, iar decodorul Δ realizează maparea inversă:

$$\mathrm{E} : \mathrm{P}^k \rightarrow \mathrm{I} \quad \text{şi} \quad \Delta : \mathrm{I} \rightarrow \mathrm{P}^k \tag{2.5.}$$

şi prin urmare operaţia de cuantizare vectorială se poate reprezenta ca o succesiune a celor două operaţii de codare şi decodare:

$$\hat{\mathbf{x}} = Q(\mathbf{x}) = \mathrm{D}(\mathrm{E}(\mathbf{x})) \qquad (2.6.)$$

Atât codorul cât şi decodorul dispun de câte o copie a dicţionarului.

3. Cuantizarea optimală a unui vector aleator

Fie **X** un vector aleator, de dimensiune *k*, continuu distribuit în spaţiul P^k, adică având funcţia densitate de probabilitate $f_{\mathbf{X}}(\mathbf{x})$ finită şi continuă pe P^k.

În timp ce **X** este continuu distribuit, $\hat{\mathbf{X}}$ are o distribuţie discretă, densitatea sa de probabilitate fiind [FOW96]

$$f_{\hat{\mathbf{X}}}(\mathbf{x}) = \sum_{i=1}^{N} p_i \delta(\mathbf{x} - \mathbf{y}_i)\text{, unde } p_i \text{ este probabilitatea ca } \hat{\mathbf{X}} = \mathbf{y}_i . \qquad (3.1.)$$

Se defineşte rata de cuantizare, în biţi pe simbol prin:

$$R_Q = \frac{1}{k} \max_{f_{\mathbf{X}}} I(\mathbf{X}; \hat{\mathbf{X}}), \qquad (3.2.)$$

unde $I(\mathbf{X}; \hat{\mathbf{X}})$ este informaţia mutuală între vectorii **X** şi $\hat{\mathbf{X}}$. Se arată [FOW96] că pentru un cuantizor vectorial aceasta este

$$R_Q = \frac{1}{k} \log_2 N\text{, unde } N = |\mathrm{C}| \text{ este mărimea dicţionarului.} \qquad (3.3.)$$

Performanţele unui codor vectorial se pot estima prin măsura distorsiunii medii pe care o introduce cuantizorul, definită prin relaţia:

$$D_Q = E\left[d(\mathbf{X}, \hat{\mathbf{X}}) \right] = \int_{\mathrm{R}^k} d(\mathbf{X}, \hat{\mathbf{X}}) f_{\mathbf{X}}(\mathbf{X}) d\mathbf{x} = \int_{\mathrm{R}^k} d(\mathbf{X}, Q(\mathbf{X})) f_{\mathbf{X}}(\mathbf{X}) d\mathbf{x} \qquad (3.4.)$$

în care *d* este o funcţie de cost.

Cuantizorul optimal pentru un semnal cu o densitate de probabilitate dată este acel cuantizor care va minimiza distorsiunea exprimată prin relaţia (10). Un cuantizor este optimal dacă îndeplineşte două condiţii:

a) condiţia de cel mai apropiat vecin: cea mai bună cuantizare este aceea care atribuie unui vector cuvântul de cod din dicţionar care este cel mai apropiat de vectorul dat, în sensul măsurii distorsiunii, adică $\hat{\mathbf{x}} = Q(\mathbf{x})$ dacă şi numai dacă $d(\mathbf{x}, \mathbf{y}_i) \le d(\mathbf{x}, \mathbf{y}_j), \forall j$. (3.5.)

b) condiţia de centroid: pentru o partiţie dată a spaţiului P^k, cuvântul de cod asociat cu fiecare partiţie trebuie astfel ales, încât să minimizeze eroarea medie faţă de toţi vectorii din acea regiune:

$$\mathbf{y}_i = \arg\min_{\mathbf{y}} E\left[d(\mathbf{X}, \mathbf{y}) \middle| \mathbf{X} \in \mathrm{R}_i \right] \qquad (3.6.)$$

care corespunde definiției centroidului în cazul în care d este eroarea medie pătratică $\mathbf{y}_i = E\left[\mathbf{X} \middle| \mathbf{X} \in \mathrm{R}_i\right]$

Funcția de cost nenegativă, $d(\mathbf{x}, \hat{\mathbf{x}})$ asociată vectorului de intrare $\mathbf{x}$ şi vectorului reconstituit $\hat{\mathbf{x}}$, care defineşte eroarea sau distorsiunea de cuantizare şi permite aprecierea performanțelor cuantizorului, poate avea diferite forme, dintre care câteva cazuri particulare, care prezintă interes practic sunt [GG93], [Fow96], [GK82], [Gra90], [ZBL94]:

- eroarea pătratică sau distanța euclidiană, definită prin relația:

$$d(\mathbf{X}, \hat{\mathbf{X}}) = \left\| \mathbf{X} - \hat{\mathbf{X}} \right\|^2 = \sum_{i=1}^{k} (X_i - \hat{X}_i)^2 \qquad (3.7.)$$

- distorsiunea definită de o funcție de tip $d(\mathbf{X}, \hat{\mathbf{X}}) = \sum_{i=1}^{k} d_m(X_i, \hat{X}_i)$ în care $d_m(x, \hat{x})$ este o funcție aditivă, în cele mai multe cazuri de forma

$$d_m(x, \hat{x}) = \left| x - \hat{x} \right|^m \qquad (3.8.)$$

cu m pozitiv (pentru m = 1 se regăseşte norma l_1 iar pentru m = 2 se regăseşte distrosiunea medie pătratică).

- eroarea pătratică ponderată:

$$d(\mathbf{x}, \hat{\mathbf{x}}) = (\mathbf{x} - \hat{\mathbf{x}})^t \mathbf{W} (\mathbf{x} - \hat{\mathbf{x}}) \qquad (3.9.)$$

unde $\mathbf{W}$ este o matrice pătrată de ponderare, simetrică şi pozitiv definită. Un caz particular este acela în care matricea $\mathbf{W}$ este o matrice diagonală, cu elementele $w_{ii} > 0$:

$$d(\mathbf{x}, \hat{\mathbf{x}}) = \sum_{i=1}^{k} w_{ii} (x_i - \hat{x}_i)^2 , \qquad (3.10.)$$

care permite atribuirea de ponderi diferite diverselor componente ale vectorului. Eroarea poderată poate fi interpretată ca o eroare neponderată între vectorii $\mathbf{x'} = \mathbf{Ax}$ şi $\hat{\mathbf{x}}' = \mathbf{A}\hat{\mathbf{x}}$ unde matricea $\mathbf{A}$ se obține prin factorizarea $\mathbf{W} = \mathbf{A}^t\mathbf{A}$. Măsurile $d(\mathbf{x}, \mathbf{y})$ ale distorsiunii prezentate sunt simetrice față de argumentele x şi y. În unele cazuri, se poate alege o matrice $\mathbf{W}$ nesimetrică, ca de exemplu $\mathbf{W}(\mathbf{x}) = \|x\|^{-2}\,\mathbf{I}$, cu $\mathbf{I}$ matricea identitate (unitate): $d(\mathbf{x}, \hat{\mathbf{x}}) = \left\| \mathbf{x} - \hat{\mathbf{x}} \right\|^2 / \|\mathbf{x}\|^2$, care atribuie o valoare cu atât mai mare distrosiunii, cu cât semnalul este mai mic [GG93].

- eroarea maximă sau norma l_∞, definită prin:

$$d_{\max}(\mathbf{x},\hat{\mathbf{x}}) = \max_i \left| x_i - \hat{x}_i \right|. \tag{3.11.}$$

4. Proiectarea unui cuantizor vectorial

Obținerea unui cuantizor vectorial optim presupune existența unui set de vectori de antrenare, T, care conține un număr suficient de mare de vectori cu o distribuție apropiată de a vectorilor care urmează a fi codați. Condițiile necesare de cuantizare optimală prezentate anterior nu sunt însă şi condiții suficiente, dar stau la baza celui mai cunoscut algoritm de construcție a dicționarului unui cuantizor vectorial, algoritmul LBG (Linde – Buzo – Gray) [LBG80], care presupune următorii paşi:

Pasul 1. Fiind dat un dicționar inițial, de dimensiune N, se partiționează spațiul P^k folosind regula celui mai apropiat vecin.

Pasul 2. Se calculează centriodul pentru fiecare partiție a P^k şi se creează un nou dicționar, care conține valorile astfel calculate pentru centroizi.

Pasul 3. Se calculează distorsiunea medie. Dacă distorsiunea medie s-a modificat cu mai puțin decât o valoare de prag impusă algoritmul se opreşte; dacă nu, se revine la pasul 1 şi se reiau calculele.

Se arată că acest algoritm conduce întotdeauna la un cuantizor local optimal.

O altă metodă de a construi dicționarul este bazată pe rețele neuronale artificiale, fiind cunoscută sub numele de învățare Kohonen [Koh88]. Pentru fiecare vector **X**(n) din setul de învățare, se calculează un nou cuvânt de cod definit prin:

$$y_i(n) = y_i(n-1) + \alpha_i(n) N_i(j,n)\left[X(n) - y_i(n-1)\right] \tag{4.1.}$$

unde $\alpha_i(n)$ este rata de învățare, o funcție care scade de la 0 la 1 pe măsură ce învățarea progresează şi $N_i(j,n)$ este o funcție de vecinătate, care exprimă relația de vecinătate a vectorilor din dicționar cu noul vector **X**(n).

În literatură se prezintă şi alte metode de construcție a unui dicționar pentru un cuantizor vectorial: inițializarea aleatoare, relaxarea stimulată [FMF89], despicarea dicționarului [GG93].Cuantizarea vectorială după regula celui mai apropiat vecin, ilustrată în fig. 4.1, este definită prin relația:

$$R_i = \left\{ \mathbf{x} : d(\mathbf{x},\mathbf{y}_i) \le d(\mathbf{x},\mathbf{y}_j), \quad \forall j \in \mathbf{I} \right\} \tag{4.2.}$$

Cuantizorul este complet determinată dacă se cunoaşte dicționarul şi se defineşte măsura distorsiunii. Partițiile spațiului P^k se numesc în acest caz regiuni Voronoi, dacă satisfac condițiile de regulat şi politopal [GG93].

Algoritmul de codare se poate rezuma astfel:
Pasul 1. Se fac inițializările: $d = d_0$, $j = 1$ *şi* $i = 1$.
Pasul 2. Se calculează $D_j = d(x, y_j)$.
Pasul 3. Dacă $D_j<d$, *se pune* $Dj \rightarrow d$ *şi* $j \rightarrow i$.
Pasul 4. Dacă $j<N$ *se pune* $j + 1 \rightarrow j$ *şi se revine la pasul 2.*
Pasul 5. Dacă $j = N$, *stop. Rezultatul este indexul i.*

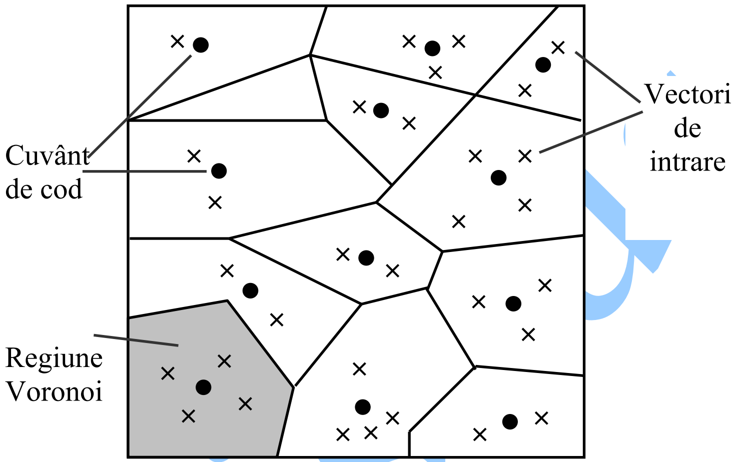

Fig. 4.1. Cuantizarea vectorială

Pentru un dicționar dat, cuantizorul se numeşte optimal în sens local, dacă o perturbație suficient de mică a unui vector de cod nu conduce la creşterea distorsiunii D. Un cuantizor este optimal în sens global, dacă nu există un alt cuantizor, cu un dicționar diferit de cel considerat, care să asigure o distorsiune mai mică decât D.

Bibliografie

[FMF89] **J. K. Flanagan, D. R. Morell, R. L. Frost, C. J. Read, B. E. Nelson** – *Vector Quantization Codebook Design generation Using Simulated Annealing*, Proceedings of ICASSP 89, vol. 3, August 1989.

[Fow96] **J. E. Fowler** – *Adaptive Vector Quantization for the Coding of Nonstationary Sources*, Ph. D. Dissertation, Ohio State University, 1996.

[Ger82] **A. Gersho** – *On the Structure on Vector Quantizers,* IEEE Trans. Inf. Theory, IT vol. 28, march 1982.

[GG93] **A. Gersho, R. M. Gray** – *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1993.

[GK82] **R. M. Gray, E. Karnin** – *Multiple Local Optima in Vector Quantizer*, IEEE Trans on Inform. Theory IT vol. 28, march 1982.

[Gra90] **R. M. Gray** – *Source Coding Theory*, Kluwer Academic Press, Boston, 1990.

[Koh88] **T. Kohonen** – *Self-Organization and Associative Memory,* Springer Verlag, Berlin, 1988.

[LBG80] **Z. Linde, A. Buzo, R. M. Gray** – *An Algorithm for Vector Quantizer Design*, IEEE Transaction On Communications, vol. COM-28, January, 1980.

[ZBL94] **K. Zeger, A. Bist, T. Linder** – *Universal Source Coding with Codebook Transmission,* IEEE Transaction on Communications, vol. 42, Feb.,March, April 1994.