

⑫

DEMANDE DE BREVET D'INVENTION

A1

②2 Date de dépôt : 23.12.91.

③0 Priorité :

④3 Date de la mise à disposition du public de la  
demande : 25.06.93 Bulletin 93/25.

⑤6 Liste des documents cités dans le rapport de  
recherche : *Se reporter à la fin du présent fascicule.*

⑥0 Références à d'autres documents nationaux  
apparentés :

⑦1 Demandeur(s) : LE GOFF Jean-Marc — FR.

⑦2 Inventeur(s) : LE GOFF Jean-Marc.

⑦3 Titulaire(s) :

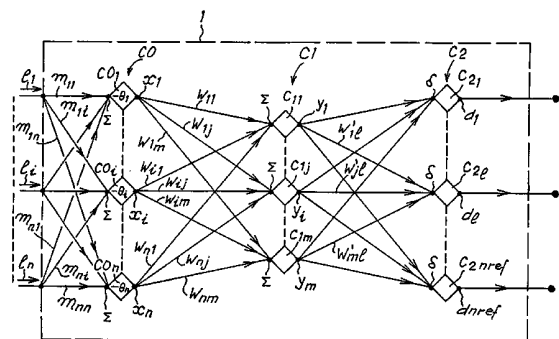
⑦4 Mandataire : Cabinet Beau de Loménie.

⑤4 Dispositif de traitement de signaux et son utilisation.

⑤7 La présente invention est relative à un dispositif de traitement de signaux et son utilisation pour déterminer notamment la complexité d'un signal à partir d'une estimation de la dimension fractale d'attracteurs construits à partir du signal.

Un dispositif comporte un réseau (1) de neurones comportant une première couche ( $C_1$ ) de premiers neurones de poids synaptiques respectifs ( $w_{ij}$ ) qui constituent un nombre ( $m$ ) de premiers vecteurs ( $W_j$ ); le réseau de neurones comporte une deuxième couche ( $C_2$ ) de deuxième neurones de poids synaptiques respectifs ( $w'_{k\tau}$ ) qui constituent un nombre  $nref$  de deuxième vecteurs ( $W'_\tau$ ); un état de sortie ( $s_j$ ) d'un des premiers neurones d'indice ( $j$ ) est déterminé par le produit scalaire entre le vecteur d'entrée ( $E$ ) et le premier vecteur ( $W_j$ ) d'indice  $j$ ; un état ( $D_\tau$ ) d'un des deuxième neurones d'indice  $\tau$  est relatif à une distance  $d(W'_\tau, Y)$ ; les premiers neurones sont dotés d'une première loi d'apprentissage qui est une loi de HEBB étendue, et les deuxième neurones sont dotés d'une deuxième loi d'apprentissage par compétition.

Le domaine technique de l'invention est celui des procédés et dispositifs de traitement de signaux.



Dispositif de traitement de signaux et son utilisation.

---

DESCRIPTION

La présente invention est relative à un dispositif de traitement  
5 de signaux et son utilisation pour déterminer notamment la complexité  
d'un signal à partir d'une estimation de la dimension fractale  
d'attracteurs construits à partir des signaux à analyser.

Le domaine technique de l'invention est celui des procédés et  
dispositifs de traitement de signaux.

10 Parmi les nombreuses applications possibles de l'invention, on  
peut citer notamment l'utilisation de dispositifs selon l'invention  
pour le traitement de signaux d'électroencéphalogrammes (EEG) et  
d'électrocardiogrammes (ECG), ainsi que l'estimation de la rugosité de  
surfaces par l'analyse de figure d'interférence se formant sur ces  
15 surfaces.

Le traitement envisagé opère sur une image multidimensionnelle  
(attracteur constitué à partir d'un signal au moins et qui consiste en  
une suite de points ou vecteurs) obtenue en prétraitant le signal par  
des procédures de "fenêtrage" et de "discrétisation". De par sa  
20 conception, le traitement mis en oeuvre dans l'invention permet donc  
aussi de traiter directement des images telles que des figures  
d'interférences ou des hologrammes, pour autant que ces images se  
présentent sous forme d'une série discrète de points ou vecteurs  
(image scannée numérisée selon un prétraitement d'image classique). Le  
25 procédé et dispositif selon l'invention permet alors de caractériser  
la complexité d'une image à partir d'une estimation de sa dimension  
fractale.

Très généralement, le procédé et dispositif qui font l'objet de  
l'invention permettent de caractériser les signaux ou images qui sont  
30 difficiles à traiter par des méthodes classiques d'analyse de Fourier,  
du fait d'une interaction non linéaire des sources qui engendrent ces  
signaux ou images. Les méthodes classiques d'analyse de Fourier  
supposent en effet qu'un signal est une combinaison linéaire de  
sinusoïdes élémentaires. L'analyse de FOURIER fait apparaître les  
35 fréquences fondamentales d'un nombre limité d'oscillateurs théoriques  
qui, par combinaison linéaire de leurs effets, engendreraient un  
signal comparable au signal expérimental analysé. Ce type d'analyse

s'applique bien à la caractérisation de signaux qui sont engendrés par des sources indépendantes, à condition que ces dernières interagissent entre elles en combinant linéairement leurs effets ; ce type de signaux est généralement caractérisé par des spectres de raies ou spectres discontinus.

De nombreux signaux expérimentaux échappent à l'analyse de FOURIER alors même qu'ils sont engendrés par des sources indépendantes en nombre fini et ce parce qu'il existe une interaction non linéaire des sources qui les engendrent. Pour ces signaux complexes pour lesquels l'analyse de FOURIER donne des spectres difficilement interprétables (spectres continus ou plats, ou décroissants par exemple en  $1/f$ ), on peut envisager une analyse dynamique utilisée pour la caractérisation des systèmes non linéaires et pour l'étude du chaos.

A la base de cette analyse on construit à partir du signal dans un espace multidimensionnel dit "de plongement" ("embedding space" en anglais) un objet géométrique, appelé attracteur, qui caractérise la dynamique du système qui engendre le signal. Différentes méthodes peuvent être employées pour construire cet attracteur, par exemple la méthode des retards décrite dans le chapitre VI "attracteurs étranges" du livre "L'ordre dans le chaos" de BERGE et al (1988, Hermann éditeur, Paris).

La signification du terme "signal" utilisé dans le présent document est très générale et suppose simplement l'enregistrement au cours du temps d'au moins une grandeur caractéristique d'un système dynamique, cette grandeur pouvant être sujette à des variations. Il existe des variantes pour la construction de l'attracteur d'un signal.

Dans le cas où on possède différents enregistrements d'un même signal de même nature en  $n$  points différents du système étudié (exemple : enregistrement d'un même électrocardiogramme par différentes électrodes placées en des points différents du corps du patient), on peut définir un point de l'attracteur comme le  $n$ -uplet  $s(t) = [s_1(t), s_2(t), \dots, s_n(t)]$  des valeurs prises par chacun desdits signaux en ces  $n$  points à un instant donné  $t$ . En faisant varier  $t$  (par échantillonnage desdits signaux) on construit l'attracteur dans un espace de dimension  $n$  parfois appelé espace de phase. L'attracteur est par exemple l'ensemble des points  $x_i = [s(t + i\pi)]$  où  $\pi$  définit la fréquence d'échantillonnage des enregistrements.

Cette méthode est utilisée par exemple dans le document : Physical Review letters, Vol. 50, n°5, 1983, "Characterization of strange Attractors", GRASSBERGER & PROCACCIA, p. 343 à 349.

5 Dans un autre cas où on ne possède qu'un enregistrement du signal en un point unique du système étudié, on peut commencer par engendrer à partir dudit signal une série temporelle, c'est-à-dire une suite des valeurs prises par le signal à différents instants. Cette suite peut être obtenue par simple discrétisation, à une fréquence préétablie (par exemple d'échantillonnage), du signal. Elle peut résulter aussi  
10 d'un prétraitement du signal. Par exemple la détection des pics dans un électrocardiogramme permet d'accéder à la suite des intervalles entre 2 complexes QRS qui constitue la série temporelle soumise à l'analyse. On utilise alors cette série temporelle pour créer l'attracteur dans un espace de plongement par la méthode suivante :

15 Soit  $X_1, X_2, \dots, X_i, \dots$  la série temporelle engendrée à partir du signal. Un point de l'attracteur sera le nuplet  $x_i = [X_i, X_{i+L}, \dots, X_{i+(n-1)L}]$  où  $L$  définit une fréquence d'échantillonnage de la série temporelle. Le point suivant de l'attracteur sera  $x_{i+1} = [X_{i+J}, X_{i+L+J}, \dots, X_{i+(n-1)L+J}]$ , où  $J$  définit un délai dans la construction  
20 des points successifs de l'attracteur. Cette méthode est détaillée dans le document Physical Review A, Vol. 38 n°6, 1988, "Singular-value Decomposition and the GRASSBERGER-PROCACCIA Algorithm", ALBANO et Al., p. 3017 à 3026.

Dans le cas de signaux complexes, l'attracteur du signal possède  
25 fréquemment une structure en feuillets, dite fractale, avec des repliements caractéristiques et une auto-similarité entre le tout et les parties. On parle alors d'attracteurs étranges.

Le problème général à résoudre est celui de la caractérisation d'un signal et plus particulièrement de la caractérisation de  
30 l'attracteur d'un signal reconstruit dans son espace englobant, notamment par l'estimation de la dimension fractale de cet attracteur ; un problème similaire à résoudre est celui de la reconnaissance de forme par la caractérisation d'image multidimensionnelle se présentant sous forme de  $p$  points de dimension  $n$ , par exemple une image  
35 d'interférence produite à la surface d'une couche mince par interaction photon/matière, permettant d'estimer la rugosité de cette surface par une estimation de la dimension fractale de l'image.

Dans le paragraphe "Mesure de la dimension des attracteurs étranges" du chapitre du livre de BERGE et al précédemment cité, on trouve une description des méthodes traditionnelles de définition et de calcul de la dimension fractale d'attracteurs construits à partir de signaux.

Dans le cas d'objets fractals de petite dimension, on peut chercher à calculer la dimension de HAUSDORFF-BESICOVITCH en utilisant le "Box Counting Algorithm".

On divise l'espace englobant en hypercubes de côté  $c$ , on compte le nombre  $N(c)$  d'hypercubes qui contiennent au moins un point de l'objet fractal. On recommence en faisant tendre  $c$  vers 0. On calcule la dimension DF par :

$$DF = - [\ln [N(c)] / \ln (c)]$$

La convergence de cette méthode est mauvaise dès que la dimension de l'espace englobant est supérieure à 2. On lui préfère deux méthodes de calcul de la dimension fractale : l'algorithme de GRASSBERGER & PROCACCIA pour l'évaluation de la dimension de corrélation et l'algorithme de FARMER et al. pour l'évaluation de la dimension ponctuelle moyenne d'un attracteur.

Le document "an approach to error estimation in the application of dimension algorithms" par J. HOLZFUSS et G. MAYER-KRESS du livre "Dimensions & Entropies in chaotic systems" (1989, ed. by G. MAYER-KRESS, Springer Verlag, Berlin) résume ces méthodes de calcul de la dimension fractale applicables à la caractérisation d'attracteurs construits à partir de séries temporelles de valeurs numérisées de signaux. Ces méthodes qui sont faciles à mettre en oeuvre et possèdent la meilleure convergence, reposent sur le calcul d'un coefficient  $C_{nref,p}(r)$  qui caractérise la densité relative moyenne de distribution des  $p$  points constituant l'objet fractal à l'intérieur de  $n_{ref}$  hypersphères de rayon  $r$  choisies dans l'espace de plongement.

Dans l'algorithme de FARMER ou dans l'algorithme de GRASSBERGER & PROCACCIA, les hypersphères sont centrées sur  $n_{ref}$  points de l'espace de plongement qui sont choisis parmi les points de l'attracteur. Dans l'algorithme de GRASSBERGER & PROCACCIA, on prend souvent tous les points de l'attracteur soit  $n_{ref}=p$ . Pour chaque point de référence  $x$ , on considère l'hypersphère de rayon  $r$  centrée sur  $x$  et on compte parmi les  $p$  points de l'attracteur combien sont contenus dans l'hypersphère.

Soit  $N_{x,p}(r)$  ce nombre et  $C_{x,p}(r) = N_{x,p}(r)/p$  la densité relative qui lui correspond.

Selon le principe de l'algorithme de FARMER, on calcule pour  $n_{ref}$  points de références choisis au hasard sur l'attracteur une "moyenne logarithmique"  $C_{nref,p}(r)$  des  $C_{x,p}(r)$  selon la formule :

$$\text{Log } C_{nref,p}(r) = \sum_{(i=1 \text{ à } n_{ref})} \text{Log } C_{xi,p}(r) / n_{ref}$$

On utilise le fait que pour  $r$  suffisamment petit

$$\text{Log } [C_{nref,p}(r)] = D_{fF} \text{Log}(r)$$

pour déduire la dimension fractale notée  $D_{fF}$  qui est la dimension "ponctuelle moyenne" au sens de FARMER.

Selon le principe de l'algorithme de GRASSBERGER & PROCACCIA, on calcule pour  $n_{ref}$  points de références choisis au hasard sur l'attracteur le logarithme de la moyenne  $C'_{nref,p}(r)$  des  $C_{x,p}(r)$  selon la formule

$$\text{Log}[C'_{nref,p}(r)] = \text{Log}[\sum_{(i=1 \text{ à } n_{ref})} C_{xi,p}(r) / n_{ref}]$$

On utilise le fait que pour  $r$  suffisamment petit

$$\text{Log}[C'_{nref,p}(r)] = D_{fGP} \text{Log}(r) + \text{constante}$$

pour déduire par régression linéaire à partir du calcul des  $\text{Log } C'_{nref,p}(r)$  pour différents  $r$  petits, la pente  $D_{fGP}$  de la droite de corrélation en fonction des  $\text{Log } r$ , qui est la dimension fractale au sens de GRASSBERGER et PROCACCIA.

L'inconvénient de ces différentes méthodes est la longueur du temps de calcul mis en jeu pour déterminer aussi bien  $D_{fF}$  que  $D_{fGP}$ , qui ne permet pas en général une estimation suffisamment rapide de la dimension fractale pour envisager le suivi en régime dynamique (ou "en temps réel") d'un processus dynamique complexe, dès lors que les dimensions fractales sont élevées, ou les processus rapides.

La demande de brevet WO 90/10278 (E.I. DU PONT DE NEMOURS AND COMPANY) décrit par exemple une méthode de calcul d'une valeur approchée de la dimension fractale d'images tridimensionnelles dans un domaine de FOURIER, du fait que le calcul dans le domaine réel (ou spatial) nécessite des ressources de calculs très importantes et trop lourdes à mettre en oeuvre.

Dans la demande de brevet WO 90/10278, on envisage le calcul de la dimension fractale d'un objet dans un espace de dimension 3 selon une procédure qui diffère totalement des procédures de FARMER ou de GRASSBERGER & PROCACCIA. On calcule la surface  $A(r)$  de la surface

tridimensionnelle à différentes échelles de résolution  $r$  pour déduire

$$D(r) = 2 - d[\text{LogA}(r)]/d[\text{Log } r]$$

qui est une dimension qui varie avec  $r$  et possède un maximum  $D(r_{\text{max}})$  qui peut servir comme estimation de la dimension de l'objet fractal. Comme dans le cas du "Box-Counting Algorithm" pour le calcul de la dimension de HAUSDORFF-BESICOVITCH, cette procédure ne peut être généralisée dans le cas où la dimension de l'espace de plongement dans lequel l'objet fractal est construit est supérieur à 3, du fait d'un manque de convergence de la méthode. En particulier cette procédure ne peut souvent pas être utilisée pour la caractérisation d'un signal (ou image) dont la dimension fractale  $D_f$  est supérieure à 2 (ce qui est le cas le plus général, un signal ECG possédant par exemple une dimension fractale  $D_f \sim 4$ ).

Si l'on considère un objet fractal (attracteur ou image) constitué de  $p$  points dans un espace (de plongement dans le cas d'un signal) de dimension  $n$ , supposons que l'on désire calculer sa dimension fractale selon l'algorithme de GRASSBERGER & PROCACCIA en estimant la valeur des coefficients  $C^{n_{\text{ref}},p}(r)$  pour différentes valeurs de  $r$  fixées à l'avance et pour  $n_{\text{ref}}$  points de références choisis sur l'attracteur. Le temps de calcul d'un coefficient augmente en fonction de  $p$ ,  $n$  et  $n_{\text{ref}}$ .

Or, le choix de ces nombres et/ou dimensions doit respecter certaines règles :

- ladite dimension ( $n$ ) de l'espace de plongement doit être au moins supérieure à la dimension fractale notée  $D_f$  de l'objet à analyser ; pour des objets de structure complexe, un théorème dû à F.TAKENS montre qu'il peut être nécessaire, pour ne pas perdre d'informations, de construire l'attracteur du signal dans un espace de plongement dont la dimension peut atteindre  $n = 2 D_f + 1$ , soit  $n = 10$  pour l'attracteur d'un ECG de dimension fractale  $d = 4,5$ ,

- le nombre  $p$  de points choisi pour représenter l'objet doit être suffisant pour représenter l'objet et non pas seulement une partie de l'objet : dans la plupart des cas, 10 000 à 100 000 points au moins sont nécessaires,

- le nombre  $n_{\text{ref}}$  de points de référence choisis dans l'espace de phase doit être suffisant (1000 à 10 000 points par exemple) pour moyenner les variations de la densité relative partielle qui dépend de

la topologie de l'attracteur .

Pour augmenter la rapidité de calcul, il a été proposé de réduire la dimension (n) de l'espace de plongement en opérant une analyse en composantes principales (ACP) de l'objet (ou signal ou image) avant de  
5 calculer la dimension fractale, par exemple dans le document : ALBANO et Al. cité plus haut.

Cette méthode diminue le temps de calcul proprement dit de la dimension fractale par les algorithmes conventionnels en réduisant à  $m < n$  la dimension de l'espace de construction de l'objet. Elle limite  
10 par ailleurs l'influence du bruit dans l'analyse. L'emploi de méthodes matricielles pour la réalisation de l'ACP impose cependant des limites à la rapidité de calcul de la dimension fractale, la durée de l'ACP augmentant considérablement lorsque la dimension n de l'espace de plongement (aussi appelé espace de phase ou espace englobant) initial  
15 et le nombre p de points de l'objet soumis à l'analyse, sont élevés.

On connaît par ailleurs des procédés et dispositifs de traitement d'images ou de vecteurs qui utilisent des réseaux de neurones.

Les réseaux de neurones désignent des techniques inspirées du fonctionnement du cerveau pour résoudre des problèmes de  
20 reconnaissance, classification ou optimisation, et dans le domaine des circuits électroniques, désignent un certain type de circuit capable de réaliser des transformations sur des vecteurs ayant un nombre élevé de composantes ; ces techniques de réseaux de neurones sont décrites d'une façon générale dans "IEEE ASSP MAGAZINE" Avril 87, R. LIPPMANN,  
25 pages 4 à 22.

La demande de brevet FR 2639739 (LABORATOIRES D'ELECTRONIQUE ET DE PHYSIQUE APPLIQUEE) décrit un procédé de compression d'images dans lequel un réseau de neurones effectue un apprentissage (non supervisé ou adaptatif), puis à l'aide de vecteurs tests issus de l'image génère  
30 un dictionnaire adaptatif ; cette adaptation utilise des attracteurs du réseau de neurones qui subissent une procédure d'optimisation ; le procédé décrit dans cette demande de brevet est destiné à traiter en temps réel des images de type télévision ; on connaît également d'autres structures à réseau de neurones par exemple par la demande de  
35 brevet FR 2655444 (THOMSON CSF) qui décrit un procédé d'apprentissage de réseau neuronal qui consiste à effectuer la correction des coefficients synaptiques à l'aide de circuits neuronaux



supplémentaires, lorsque la base d'apprentissage a été présentée à l'entrée du réseau et que tous les vecteurs de sortie ont été calculés par ce réseau.

Les réseaux de neurones utilisés dans le document FR 2 655 444  
5 sont des réseaux à apprentissage supervisé. Ceux décrits dans FR 2 639 739 sont à apprentissage non supervisé, c'est-à-dire que la modification des poids synaptiques des réseaux de neurones au cours de l'apprentissage, ne tient compte que des entrées qui sont présentées auxdits réseaux de neurones et des poids synaptiques calculés lors des  
10 itérations précédentes ; ce dernier document est relatif à la compression d'images pour leur stockage en mémoire et leur rappel.

L'objet de la présente invention est de remédier aux inconvénients des procédés et dispositifs connus de traitement de signaux (ou d'images), afin de permettre de déterminer la dimension fractale  
15 d'attracteurs construits à partir des signaux analysés (ou la dimension d'images fractales) ; un objectif de l'invention est notamment de diminuer très sensiblement les moyens de calcul nécessaires et les temps de calcul correspondant.

La solution au problème posé consiste à procurer un dispositif de  
20 traitement d'au moins un signal numérisé sous forme de  $(p)$  vecteurs d'entrée  $(E)$  ou  $(X)$  de dimension  $(n)$ , qui comporte au moins un moyen de réseau  $(1)$  de neurones comportant au moins une première couche  $(C1)$  ayant un nombre  $(m)$ , inférieur ou égal audit nombre  $(n)$ , de premiers neurones de poids synaptiques respectifs  $(w_{ij})$ , lesquels poids  
25 synaptiques  $(w_{ij})$  peuvent constituer ledit nombre  $(m)$  de premiers vecteurs  $(W_j)$  qui ont pour dimension ledit nombre  $(n)$ , lequel dispositif comporte au moins une deuxième couche  $(C2)$  ayant un nombre  $(n_{ref})$  de deuxièmes neurones de poids synaptiques respectifs  $(w'_{jl})$ , lesquels poids synaptiques  $(w'_{jl})$  peuvent constituer ledit nombre  
30  $(n_{ref})$  de deuxièmes vecteurs  $(W'_l)$  qui ont pour dimension ledit nombre  $(m)$ , et par ailleurs tel qu'un état de sortie  $(y_j)$  d'un desdits premiers neurones d'indice  $(j)$  est déterminé par le produit scalaire entre ledit vecteur d'entrée  $(E)$  (ou  $X$ ) et ledit premier vecteur  $(W_j)$  d'indice  $(j)$ , l'ensemble desdits états de sortie pouvant constituer un  
35 vecteur intermédiaire  $(Y)$ , et tel qu'un état  $(d_l)$  d'un desdits deuxièmes neurones d'indice  $(l)$  est relatif, de préférence est proportionnel, à une distance  $d(W'_l, Y)$ , et en ce que lesdits poids

synaptiques desdits premiers neurones sont dotés d'une première loi d'apprentissage non supervisé qui est une loi de HEBB étendue, et en ce que lesdits poids synaptiques desdits deuxièmes neurones sont dotés d'une deuxième loi d'apprentissage non supervisé par compétition.

5 Dans ce document, on désigne par réseau de neurones ou moyen de réseau de neurones, des circuits électroniques analogiques et/ou numériques comportant une pluralité de neurones, tout ou partie du réseau de neurones pouvant être simulé dans un ordinateur (tel que décrit par exemple dans "implémentation de l'algorithme de cartes  
10 topologiques de KOHONEN sur transputers" dans la "lettre du transputer", DORIZZI et al. p 17 et suivantes) sous forme de circuits électroniques de préférence doté de plusieurs processeurs ; des réseaux de neurones sont par exemple décrits dans les demandes de brevets WO 90/16039 (GENERAL ELECTRIC COMPANY), EP 377908  
15 (INTERNATIONAL BUSINESS MACHINES CORPORATION), ou US 4254474 (COOPER et al).

Dans le présent document, le terme "premier" différencie les liens (ou poids) synaptiques, les vecteurs de poids synaptiques et les neurones de ladite première couche  $C_1$  ; le terme "deuxième"  
20 différencie les liens (ou poids) synaptiques, les vecteurs de poids synaptiques et les neurones de ladite deuxième couche  $C_2$ ; le terme "troisième" différencie des liens (ou poids synaptiques), les vecteurs de poids synaptiques et les neurones d'une troisième couche  $C_3$ .

Ledit nombre  $n$  de composantes d'un vecteur d'entrée  $E$  et/ou d'un  
25 vecteur d'entrée centré  $X$  représente, dans le cas d'un signal, la dimension de l'espace de plongement dans lequel on construit l'attracteur de ce signal et représente alternativement la dimension de l'espace qui contient l'image fractale à analyser.

Ledit nombre  $m$  de composantes desdits vecteurs intermédiaires  $Y$   
30 qui constituent des vecteurs de sortie de ladite première couche desdits premiers neurones, lesquels vecteurs intermédiaires sont présentés à l'entrée de ladite deuxième couche desdits deuxièmes neurones, correspond à la dimension du sous espace de projection pour l'analyse en composantes principales (ACP).

35 Dans le présent document, on désigne par le terme "Loi de HEBB étendue", une Loi d'apprentissage définie par :

$$dW_j/dt = a_1(t) \langle W_j | X \rangle (X - \sum_k \langle W_k | X \rangle W_k)$$

$$+ b_1(t) X \langle W_j | (X - \sum_k \langle W_k | X \rangle W_k) \rangle \quad , k \text{ variant de } 1 \text{ à } j,$$

relation (1)

5 dans laquelle :

$W_j$  = premier vecteur d'indice  $j$  desdits premiers poids synaptiques desdits premiers neurones, aussi noté  $W_j(t)$ ,

$t$  = "temps calcul",

$dt$  = incrément de temps  $t$ ,

10  $dW_j/dt = W_j(t+dt) - W_j(t) / dt$

ce qui peut aussi s'écrire :

$$dW_j(t) = W_j(t+1) - W_j(t) \text{ aussi noté } dW_j$$

$a_1(t)$  et  $b_1(t)$  = fonctions décroissantes du temps, tendant vers 0, dont l'intégrale diverge lorsque  $t$  tend vers  $l^\infty$  ; par ex :  $a_1(t) = 1/t$

15  $X$  = vecteur d'entrée

$\langle W_j | X \rangle$  : produit scalaire de  $W_j$  par  $X$  (égale à la composante  $y_j$  d'indice  $j$  dudit vecteur intermédiaire  $Y$  :  $y_j = \langle W_j | X \rangle$ ) ; ce qui peut également s'écrire  $y_j(t) = \langle W_j(t) | X(t) \rangle = \sum_i w_{ij}(t) x_i(t)$ ,  $i=1$  à  $n$ .

Ce type de Loi est caractéristique d'un apprentissage non supervisé puisque l'évolution (dans le temps) des poids synaptiques est effectué sans relation avec des valeurs "désirées" ou prédéterminées des états de sortie des neurones concernés ; ce type de Loi est généralement qualifié de "feedforward" du fait que les variations ( $dW_j/dt$ ) d'un premier vecteur d'indice  $j$  dépendent des premiers vecteurs d'indice (k)  $<$  ou égal à  $j$ , ainsi bien sûr que des vecteurs d'entrée ( $X$ ).

25 On peut noter que dans le cas particulier où la fonction  $b_1(t)$  est nulle pour toutes valeurs de  $(t)$ , la relation (1) peut s'écrire :

$$W_j(t+1) = W_j(t) + a_1(t) [y_j X - y_j \sum_k y_k W_k],$$

vecteurs dont les composantes vérifient la relation :

30  $w_{ij}(t+1) = w_{ij}(t) + a_1(t) (y_j(t)x_i(t) - y_j(t) \sum_k y_k(t) w_{ik}(t)) \quad , k \leq j$

ce qui correspond à la Loi de HEBB dite "généralisée" décrite dans le document Neural Networks Vol. 2, 1989, "Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network", SANGER, p. 459 à 473.

35 Il est également intéressant de noter que la Loi de HEBB étendue, définie par la relation (1) ci-dessus, peut, grâce à son deuxième terme provoquer une convergence plus rapide desdits poids synaptiques.

Ladite deuxième Loi d'apprentissage non supervisée par compétition est du type :

$$dW'_l(t) = W'_l(t+1) - W'_l(t) = a_2(t) f[z_l(t)] [Y(t) - W'_l(t)] \quad , \text{ relation (2)}$$

5 dans laquelle :

$W'_l$  = deuxième vecteur desdits deuxièmes poids synaptiques dudit deuxième neurone d'indice  $l$ , aussi noté  $W'_l(t)$ ,

$Y$  = vecteur intermédiaire qui constitue un vecteur d'entrée de ladite deuxième couche desdits deuxième neurones, aussi noté  $Y(t)$ ,

10  $a_2(t)$  : fonction décroissante du temps, telle que quant  $t$  tend vers l'infini  $a_2(t)$  tend vers 0, et dont l'intégrale  $(\sum_t a_2(t))$  tend vers l'infini ;  $a_2(t)$  permet d'améliorer la convergence desdits deuxième poids synaptiques,

$f(x)$  : fonction de  $R^+ \rightarrow [0,1]$ , où

15  $R^+$  ensemble des réels positifs,

$[0, 1]$  intervalle des réels compris entre 0 et 1 ;  $f$  est une fonction de pondération pour l'adaptation des poids, et est telle que :

- $f(0) = 1$ ,
- 20 -  $f$  est décroissante,
- $\lim_{x \rightarrow \infty} f(x) = 0$  lorsque  $x$  tend vers l'infini.

On peut choisir par exemple :  $f(x) = \{1 - H[x - b_2(t)]\}$ , où

$H$  fonction de Heaviside  $H(x) = 1$  si  $x > 0$ ,  $H(x) = 0$  sinon,

$b_2(t)$  : fonction décroissante du temps qui tend vers 0 (ou  $b_2(t) = 0$ )

25 Nota : sauf indication contraire, dans le présent document,  $H()$  désigne la fonction de Heaviside et  $d(x,y)$  désigne une distance entre  $x$  et  $y$ .

Selon l'invention,  $n_{ref}$  états de sortie  $d_1 \dots d_l \dots d_{n_{ref}}$  desdits deuxièmes neurones sont susceptibles de constituer des vecteurs  $D$  de distance aux références (constituées par lesdits deuxièmes vecteurs  $W'$  de poids synaptique desdits deuxièmes neurones), et sont définis par :

$$d_l(t) = d(W'_l(t), Y(t))$$

dans laquelle :

$d(W'_l, Y)$  = distance, euclidienne par exemple, entre  $W'_l$  et  $Y$ ,

35 Avantageusement, des vecteurs  $D^*$  de distance modifiée peuvent être définis par :

$$D^*(t) = U(t).D(t) \quad , \text{ ce qui peut s'écrire :}$$

$$\begin{aligned}
d^*_l(t) &= u_l(t) \cdot d_l(t) \quad , \quad 1 \leq l \leq n_{\text{ref}}, \\
\text{avec : } u_l(1) &= 1 \quad , \quad 1 \leq l \leq n_{\text{ref}} \\
\text{et : } du_l(t) &= u_l(t+1) - u_l(t) \\
&= a'_2(t) g(z_l(t)) \quad , \quad \text{relation (3)}
\end{aligned}$$

5 où  $a'_2(t)$  et  $g(x)$  sont des fonctions ayant respectivement les mêmes caractéristiques que celles des fonctions  $a_2(t)$  et  $f(x)$  définies ci-avant,

$$\text{et avec } z_l(t) = d^*_l(t) - \text{Min}[d^*_k(t), k=1 \text{ à } n_{\text{ref}}] \quad , \quad \text{relation (4)}.$$

Avantageusement, ledit état  $d_l$  dudit deuxième neurone d'indice  $l$  est un état de sortie analogique dudit deuxième neurone, et ceci pour chacun desdits deuxième neurones.

Selon un mode préférentiel de réalisation de l'invention, ladite deuxième couche (C2) dudit réseau de neurones comporte ledit nombre (nref), de préférence très inférieur audit nombre (p), desdits  
15 deuxièmes neurones (C21...C2nref) dotés chacun dudit nombre (m) de deuxièmes liens synaptiques de poids respectifs ( $w'_{kl}$ ,  $k$  variant de 1 à  $m$ ,  $l$  variant de 1 à  $n_{\text{ref}}$ ).

Avantageusement, ledit nombre  $n_{\text{ref}}$  est supérieur à 100 et inférieur à 100 000, et par exemple ledit nombre  $n_{\text{ref}}$  est compris  
20 entre 1000 et 10 000.

Avantageusement, ledit nombre  $m$  est inférieur à 100, par exemple voisin de 10.

Avantageusement, ledit dispositif comporte des moyens (2) de comparaison de la distance  $d(W'_l, Y)$  entre chacun desdits deuxièmes  
25 vecteurs ( $W'_l$ ) et ledit vecteur intermédiaire ( $Y$ ) avec un nombre ( $q$ ) de valeurs  $r_h$  ( $h$  variant de 1 à  $q$ ) prédéterminées, lequel nombre  $q$  est par exemple de l'ordre de 10.

Préférentiellement, lesdits moyens (2) de comparaison de distance comportent un moyen de calcul, à chaque itération (et/ou à chaque  
30 présentation d'un vecteur d'entrée), d'un nombre noté  $N_{n_{\text{ref}}, r_h}$ , correspondant à une moyenne dite "spatiale", lequel nombre peut de préférence être défini par :

$$\begin{aligned}
N_{n_{\text{ref}}, r_h}(t) &= 1/n_{\text{ref}} \sum_{l=1}^{n_{\text{ref}}} H[r_h - d(W'_l(t) - y(t))] \\
h &= 1 \text{ à } q, \quad , \quad \text{relation (5)}
\end{aligned}$$

35 Avantageusement, ledit dispositif comporte des moyens (3, DF) de calcul d'une moyenne temporelle dudit nombre  $N_{n_{\text{ref}}, r_h}$  pour chacune desdites  $q$  valeurs  $r_h$ .

A un instant  $t$ , ladite moyenne temporelle peut être définie par un coefficient  $K_{nref}(r_h)$  tel que :

$$K_{nref}(r_h) = 1/p' \sum_{(t=1 \text{ à } p')} [N_{ref}, r_h(t)]$$

, relation (6)

5        pour un nombre ( $p'$ ) desdits vecteurs d'entrée qui ont été présentés à l'entrée dudit réseau de neurones, audit instant  $t$ .

Ledit coefficient  $K_{nref}(r_h)$  est utilisé comme une estimation des coefficients notés  $C_{nref,p'}(r_h)$  de l'algorithme de GRASSBERGER et PROCACCIA.

10        Lesdits moyens (3, DF) comportent de préférence des moyens de calcul de la pente (notée  $DF(p')$ ) de la droite de regression linéaire de logarithme desdits coefficients  $K_{nref}(r_h)$  en fonction de logarithmes desdites valeurs  $r_h$ , laquelle pente est appelée dimension fractale estimée.

15        Avantageusement, le dispositif comporte des moyens de visualisation, de préférence en permanence, de l'ensemble desdits deuxièmes vecteurs ( $W'_j$ ), permettant d'afficher une allure de l'attracteur dudit signal correspondant auxdits vecteurs d'entrée.

20        Avantageusement, il comporte des moyens de visualisation, de préférence en permanence, de ladite dimension fractale.

La solution au problème posé consiste également à utiliser un dispositif de traitement de signaux selon l'invention de la manière suivante :

dans une phase d'apprentissage :

25        a) - on affecte auxdits premiers et deuxièmes vecteurs de poids synaptique, des valeurs initiales choisies par exemple de façon aléatoire ou pseudo-aléatoire, ou alternativement, égales à des valeurs obtenues précédemment,

30        b) - on présente successivement un nombre  $p$  de vecteurs choisis parmi lesdits vecteurs d'entrée, de préférence jusqu'à convergence desdits vecteurs de poids synaptiques des différentes couches dudit réseaux de neurones (ou stabilisation des composantes desdits vecteurs de poids synaptiques dans des intervalles de dimension prédéterminée), puis dans une phase d'exploitation :

35        c) - on fige les valeurs desdits poids synaptiques,

d) - on présente successivement un nombre  $p'$  de vecteurs choisis parmi lesdits vecteurs d'entrée, et on calcule des valeurs de ladite

5 pente  $DF(p')$  qui donne une estimation de la dimension fractale d'un attracteur construit à partir desdits vecteurs d'entrée, ledit attracteur étant projeté sur un sous espace de meilleure projection (au sens d'une analyse en composantes principales), lequel sous espace a pour dimension ledit nombre  $m$ , et simultanément on choisit  $n_{ref}$  points (ou vecteurs) dudit sous espace, qui déterminent une allure dudit attracteur, et peuvent ainsi être facilement utilisés pour visualiser une allure dudit attracteur.

10 Dans un exemple d'utilisation d'un dispositif selon l'invention, lesdits signaux numérisés sont des signaux d'électrocardiogramme et ledit nombre  $m$  est voisin de 10.

Selon l'invention, il est proposé un procédé et un dispositif original de calcul de la dimension fractale, qui utilise non pas des méthodes matricielles, mais des techniques de réseaux de neurones.

15 La caractéristique principale de la présente invention est la mise en oeuvre de moyens de réseaux de neurones à apprentissage non supervisé (adaptatif) pour le calcul desdits coefficients  $Knref(r_h)$  en vue de l'estimation de la dimension fractale par l'algorithme de GRASSBERGER & PROCACCIA ; une autre caractéristique réside dans un choix particulier des points (ou vecteurs) de référence utilisés pour ces calculs.

20 Le procédé et le dispositif associent une "squelettisation" de l'objet qui s'apparente à une ACP (analyse en composantes principales), qui est réalisée par ladite première couche desdits premiers neurones, à une recherche optimale des points de référence pour le calcul desdits coefficients de type  $Knref(r_h)$  qui sont utilisés pour calculer la dimension fractale selon une méthodologie qui s'apparente à l'algorithme de GRASSBERGER & PROCACCIA, cette recherche optimisée étant réalisée par ladite deuxième couche desdits 25 deuxièmes neurones, lesdits points ou vecteurs de référence étant constitués par lesdits points ou vecteurs de poids synaptiques desdits deuxièmes neurones.

30 De par leur conception, le procédé et le dispositif de traitement de signaux selon l'invention permettent un parallélisme massif dans les calculs. Ils mettent en jeu des algorithmes distribués qui peuvent être effectués comme tâches indépendantes par des structures électroniques de type transputer, conduisant à une très grande 35

rapidité d'exécution. Il opère de plus selon un processus adaptatif qui fait que le calcul de la dimension fractale d'un signal (ou d'une image) est d'autant plus rapide que ce signal (image) est proche de celui qui avait été traité auparavant.

5 Les nombreux avantages procurés par l'invention seront mieux compris au travers de la description suivante qui se réfère aux dessins annexés qui illustrent sans aucun caractère limitatif, des modes de réalisation et de mise en oeuvre de dispositifs et de procédés de traitement de signaux selon l'invention et leur  
10 utilisation pour la détermination d'estimations de la dimension fractale d'image ou d'attracteur.

Les figures 1 à 3 illustrent des étapes dans des procédés de traitement de signaux, appliqués à un signal sensiblement sinusoïdal.

15 Les figures 4 à 6 correspondent respectivement aux figures 1 à 3 pour une deuxième signal, donné à titre d'exemple, qui est un signal déterministe mais non linéaire.

La figure 7 illustre de façon schématique un mode de réalisation d'un moyen de réseau de neurones d'un dispositif selon l'invention.

20 La figure 8 illustre schématiquement les principaux composants d'un dispositif de traitement de signaux selon l'invention permettant de comprendre le fonctionnement du procédé de traitement correspondant.

25 La figure 9 illustre un mode particulier de réalisation d'une partie d'un dispositif selon l'invention, dans lequel lesdits moyens 2 de comparaison de distance, comportent une quatrième couche C4 de quatrièmes neurones.

30 La figure 10 illustre schématiquement un dispositif de traitement de signaux et son utilisation pour le traitement de signaux d'électrocardiogramme ou de signaux vidéo correspondants à une image filmée par une caméra.

Les figures 11 et 13 illustrent sous forme de schéma matriciel, des moyens de centrage des vecteurs d'entrée et respectivement des moyens de détermination des vecteurs de distance modifiés  $D^*$  à partir des vecteurs de distance  $D$  issus de ladite deuxième couche desdits  
35 deuxièmes neurones d'un dispositif selon l'invention.

La figure 12 illustre de façon très schématique l'architecture générale de moyens de traitement de signaux selon l'invention et leur



connexions.

Par référence à la figure 1, un signal  $S_1$  à analyser, sensiblement périodique et sinusoïdal, peut, en vue de la construction d'un attracteur correspondant à ce signal, faire l'objet d'une  
 5 décomposition en série temporelle comme illustré sur cette figure où l'axe des abscisse représente le temps et l'axe des ordonnées représente l'amplitude du signal  $S_1$ .

Ledit signal peut être échantillonné et numérisé en des instants séparés par un intervalle de temps  $dt$  constant correspondant à la  
 10 fréquence d'échantillonnage dudit signal, lesquelles valeurs peuvent être, par référence à la figure 10, échantillonnées et numérisées par un dispositif 13 puis stockées dans une mémoire tampon 14.

On voit que parmi tous les points échantillonnés à la fréquence d'échantillonnage, qui sont repérés par des points sur la courbe du  
 15 signal sinusoïdal  $S_1$ , certains points peuvent être choisis afin de constituer les séries temporelles permettant de construire l'attracteur à partir dudit signal ; on voit sur la figure 1 qu'un premier vecteur (ou point) de l'attracteur construit à partir du signal, peut être constitué à partir de points  $x^1_1, x^1_2, x^1_3$  qui  
 20 peuvent constituer un vecteur  $X^1$  illustré figure 2, chacune desdites valeurs constituant lesdites composantes dudit vecteur étant séparées dudit intervalle de temps  $dt$  ; de la même façon un deuxième vecteur dudit attracteur (qui sera alors noté  $X^2$ ) peut être constitué par les  
 25 trois composantes constituées par trois valeurs successives  $x^2_1, x^2_2, x^2_3$  dudit signal, comme illustré à la figure 1, qui sont séparées par ledit intervalle  $dt$  ; on voit que lesdites composantes dudit premier vecteur et dudit deuxième vecteur ainsi constitués afin de construire ledit attracteur, seront espacés d'un intervalle de temps  $TAU$  supérieur audit intervalle de temps  $dt$ .

On voit sur la figure 2 un attracteur  $A_1$  qui peut être constitué desdits vecteurs  $X^1, \dots, X^9$  construits à partir dudit signal  $S_1$  de la  
 30 figure 1 ; ledit attracteur  $A_1$  est construit dans un espace englobant  $EE_1$  de dimension 3, d'origine  $O_1$ , où l'on voit que ledit vecteur  $X^1$  a pour composantes lesdites composantes  $x^1_1, x^1_2, x^1_3$  dans ledit repère  
 35 dudit espace englobant.

La figure 3 illustre la projection  $A'_1$  dudit attracteur repéré  $A_1$  sur la figure 2, dans un espace  $EACP_1$  de dimension 2, d'origine  $O'_1$ ,

ledit espace EACP<sub>1</sub> étant généré par deux vecteurs W<sub>1</sub> et W<sub>2</sub>, ledit espace EACP<sub>1</sub> constituant un espace de meilleure projection au sens d'une analyse en composantes principales dudit attracteur A<sub>1</sub> ; par référence aux repères utilisés dans le dispositif de traitement de signaux selon l'invention, la projection A'<sub>1</sub> dudit attracteur peut être constituée par une courbe formée de points (ou vecteurs) Y (intermédiaires) qui sont obtenus en sortie de ladite première couche desdits moyens de réseaux de neurones du dispositif selon l'invention.

De la même façon qu'illustré respectivement figures 1, 2 et 3, on a illustré aux figures 4, 5 et 6, respectivement un signal S<sub>2</sub> déterministe non linéaire pouvant faire l'objet d'un traitement de signaux du même type.

Sur la figure 4, l'axe des abscisses représente le temps et l'axe des ordonnées représente l'amplitude dudit signal S<sub>2</sub>.

On voit à la figure 5 que ledit signal S<sub>2</sub> peut par exemple, après constitution d'une série temporelle constituée de vecteurs de dimension 5, être représenté dans un espace englobant EE<sub>2</sub> de dimension 5 et d'origine O<sub>2</sub>, espace englobant dans lequel ledit attracteur A<sub>2</sub> est constitué par lesdits vecteurs ou points de la série temporelle qui constitueront les points ou vecteurs X appliqués à l'entrée de ladite première couche dudit réseau de neurones selon l'invention, lequel attracteur est par exemple (dans l'exemple pris en compte et représenté sur la figure 5) de dimension inférieure à 5.

On voit à la figure 6 que ledit attracteur A<sub>2</sub> de la figure 5 a fait l'objet d'une projection repérée A'<sub>2</sub> sur un espace EACP<sub>2</sub> de meilleure projection au sens d'une analyse en composantes principales. On voit que ledit espace de meilleure projection EACP<sub>2</sub> de dimension 3 (donc inférieure à la dimension de l'espace englobant EE<sub>2</sub> de la figure 5) a une origine O'<sub>2</sub> et est généré par trois vecteurs W<sub>1</sub>, W<sub>2</sub>, W<sub>3</sub> ; la projection A'<sub>2</sub> dudit attracteur est constituée de points ou vecteurs Y de dimension 3, qui sont normalement en très grand nombre et pour lesquels on va chercher des vecteurs W' permettant une quantification vectorielle dudit attracteur, lesquels vecteurs W' sont repérés par des croix, afin de simplifier le calcul du fait du moindre nombre desdits vecteurs ou points W' par rapport auxdits points ou vecteurs Y constituant la projection A'<sub>2</sub> dudit attracteur.

La figure 7 illustre un mode préférentiel de réalisation d'un

moyen 1 de réseau de neurones qui comporte trois couches de neurones repérées successivement  $C_0$ ,  $C_1$ ,  $C_2$ , dans le sens des flux d'informations ou de données.

Ladite couche  $C_0$  appelée aussi couche d'entrée comporte un nombre  
 5 de neurones d'entrée égal à  $n$  qui est la dimension des vecteurs d'entrée qui peuvent être les vecteurs ou points de l'attracteur du signal considéré ou des vecteurs constitués à partir d'une image ;  
 chacun desdits neurones d'entrée repéré  $Co_i$  ( $i=1$  à  $n$ ) de ladite  
 couche d'entrée est doté de liens synaptiques repérés par leur poids  
 10 respectif notés  $m_{11}, \dots, m_{nn}$  par lesquels lesdits neurones d'entrée, reçoivent les composantes  $e_1, \dots, e_i, \dots, e_n$ , dudit vecteur d'entrée  
 chacun desdits neurones  $Co_i$  de ladite couche d'entrée étant doté par ailleurs d'un seuil  $\theta_i$ .

Sur les figures 7 et 9, lesdits neurones dudit dispositif desdits  
 15 moyens de réseau de neurones dudit dispositif, sont représentés sous forme de losange muni dans sa partie gauche d'un point ; lesdits neurones munis à leur extrémité gauche dudit point repéré par le symbol  $\Sigma$ , sont des neurones effectuant les opérations habituelles de sommation des produits des poids synaptiques par les valeurs  
 20 présentées en entrée des liens synaptiques correspondant ; ceci concerne comme illustré à la figure 7, les neurones de ladite couche d'entrée  $Co$ , lesdits premiers neurones de ladite première couche  $C_1$  et  
 comme illustré à la figure 9, lesdits troisièmes neurones de ladite troisième couche  $C_3$  et lesdits quatrièmes neurones de ladite quatrième  
 25 couche  $C_4$  ; par contre comme illustré à la figure 7, lesdits deuxièmes neurones de ladite deuxième couche  $C_2$  sont munis à leur extrémité gauche d'un point repéré par la lettre grecque delta, ce qui illustre que lesdits deuxièmes neurones effectuent des calculs de distance.

Comme illustré figure 7, ladite couche d'entrée  $Co$  a ledit nombre  
 30  $n$  desdits neurones d'entrée de poids synaptique  $m_{ki}$ , lesquels poids synaptiques desdits neurones d'entrée sont représentés à la figure 11  
 comme constituant une matrice carrée  $M$  d'ordre  $n$ , laquelle matrice  $M$  est appelée matrice de centrage spatial ; lesdits neurones d'entrée de  
 ladite couche d'entrée comportent par ailleurs, comme illustré à la  
 35 figure 7 à l'intérieur desdits losanges représentant lesdits neurones d'entrée, un seuil  $\theta_1 \dots \theta_i \dots \theta_n$ , l'ensemble desdits seuils desdits neurones d'entrée pouvant constituer comme illustré à la figure 11, un

vecteur  $\theta$  ayant pour dimension ledit nombre  $n$ , lequel vecteur  $\theta$  est appelé vecteur de centrage temporel.

Ainsi, préférentiellement dans de nombreuses applications de traitement de signaux, lesdits neurones de ladite couche d'entrée  
5 permettent de centrer les valeurs desdits états de sortie  $x_i$  de ladite couche d'entrée, par rapport à la moyenne temporelle et "spatiale" des composantes desdits vecteurs  $E$  qui se présentent à l'entrée dudit réseau de neurones.

On obtient ainsi, lesdits vecteurs d'entrée centrés  $X$  de  
10 composantes  $x_1 \dots x_n$  qui sont obtenus à partir desdits vecteurs d'entrée non centrés  $E$  de composantes  $e_1 \dots e_n$ , par la relation suivante :

$X = M \times E - \theta$ , relation matricielle illustrée à la figure 11, qui  
15 peut s'écrire par une relation entre les composantes des vecteurs sous la forme :

$$x_i = \sum_k m_{ki} e_k - \theta_i, 1 \leq k \leq n.$$

Avantageusement, lorsqu'on obtient lesdits vecteurs centrés  $X$  à partir desdits vecteurs d'entrée non centrés  $E$  par la relation matricielle définie ci-dessus, lesdites composantes de ladite matrice  
20  $M$  qui correspondent auxdits poids synaptiques de ladite couche d'entrée, sont invariants et ne sont donc pas munis d'une loi d'apprentissage ; par contre lesdits seuils  $\theta$  desdits neurones d'entrée de ladite couche  $C_0$  sont avantageusement munis d'une loi d'apprentissage de GROSSBERG, comme décrite par exemple dans le  
25 document Neural Networks, Vol. 3, pages 87-92, 1990, "A Convergence Theorem for GROSSBERG Learning", CLARK et al.

On voit que ledit réseau de neurones comporte successivement, en série avec ladite couche d'entrée  $C_0$ , ladite première couche  $C_1$  puis ladite deuxième couche  $C_2$  de neurones ; ladite première couche  $C_1$   
30 comporte un nombre  $m$  de premiers neurones  $C_{11}, \dots C_{1j}, \dots C_{1m}$ , lequel nombre  $m$  correspond à la dimension de l'espace sur lequel on réalise l'analyse en composantes principales des vecteurs  $X$  (ou  $E$ ) d'entrée dont la dimension  $n$  est supérieure à la dimension de ladite première couche (ou nombre  $m$  desdits premiers neurones).

35 Lesdits premiers neurones sont respectivement caractérisés par des liens synaptiques repérés par leur poids synaptique correspondant ; ainsi, ledit premier neurone  $C_{11}$  est doté de poids synaptiques

$w_{11}, \dots, w_{i1}, \dots, w_{n1}$ , qui relient respectivement ledit premier neurone  $C_{11}$  auxdites composantes de vecteurs d'entrée  $x_1, \dots, x_n$  ; ledit premier neurone  $C_{1j}$  est doté de poids synaptiques  $w_{1j}, \dots, w_{ij}, \dots, w_{nj}$  ; de la même façon un premier neurone repéré  $C_{1m}$  est doté de liens ou poids synaptiques repérés  $w_{1m}, \dots, w_{im}, \dots, w_{nm}$ .

Lesdits premiers neurones  $C_{11}, \dots, C_{1j}, \dots, C_{1m}$  sont également caractérisés par leur état de sortie respectif  $y_1, \dots, y_j, \dots, y_m$  dont les valeurs sont obtenues par les relations définies ci-dessus, de façon classique dans les utilisations de réseau de neurones.

Ladite deuxième couche  $C_2$  comporte lesdits deuxièmes neurones  $C_{21}, \dots, C_{2l}, \dots, C_{2nref}$ , qui sont au nombre de  $nref$ , ledit nombre  $nref$  étant choisi de manière à constituer un nombre desdits vecteurs  $W'$  suffisant pour le calcul de ladite dimension fractale.

Chacun desdits deuxièmes neurones est caractérisé par ses liens synaptiques avec lesdits neurones de ladite première couche, par un état interne éventuel, et un état de sortie respectivement  $d_1, \dots, d_l, \dots, d_{nref}$ .

Ledit deuxième neurone  $C_{21}$  est doté de poids ou liens synaptiques non repérés sur cette figure pour des raisons de clarté, ainsi que ledit neurone  $C_{2nref}$  ; ledit deuxième neurone  $C_{2l}$  est doté de poids ou liens synaptiques repérés  $w'_{1l}, \dots, w'_{jl}, \dots, w'_{ml}$  reliant lesdits premiers neurones  $C_{11}, \dots, C_{1j}, \dots, C_{1m}$  audit deuxième neurone  $C_{2l}$ .

Par référence à la figure 8, on voit que les poids synaptiques  $W_{ij}$  ( $i$  variant de 1 à  $n$ ,  $j$  variant de 1 à  $m$ ) constituent un nombre  $m$  de vecteur  $W_1, \dots, W_j, \dots, W_m$ , lesquels vecteurs caractérisent ladite première couche  $C_1$  de premiers neurones dudit moyen de réseau de neurones 1 du dispositif de traitement de signaux selon l'invention ; chacun desdits vecteurs  $W$  est de dimension  $n$ , comme lesdits vecteurs d'entrée centrés  $X$ , qui peuvent correspondre à la sortie de ladite couche d'entrée  $C_0$  à laquelle sont appliqués lesdits vecteurs d'entrée (non centrés)  $E$  de composantes  $e_1, \dots, e_n$ , ladite couche d'entrée  $C_0$  étant caractérisé par lesdits seuils desdits neurones d'entrée ; ainsi lorsqu'on présente un vecteur  $X$  de composante  $x_1, x_i, x_n$  à l'entrée de ladite première couche  $C_1$  dudit réseau de neurones, on obtient en sortie de ladite première couche dudit réseau de neurones un vecteur  $Y$ , de composantes  $y_1, \dots, y_j, \dots, y_m$  dont chacune des composantes  $y_j$  est obtenue par le produit scalaire dudit premier vecteur de poids

synaptique  $W_j$  correspondant, avec ledit vecteur d'entrée présenté à l'entrée de ladite première couche dudit réseau de neurones ; ainsi ledit vecteur  $y_j$  correspond au produit scalaire du vecteur  $W_j$  par ledit vecteur  $X$ .

5 Ledit vecteur  $Y$  intermédiaire peut être présenté à l'entrée de ladite deuxième couche  $C_2$  dudit réseau de neurones, laquelle deuxième couche est constituée par les deuxièmes vecteurs  $W'_1, \dots, W'_l, \dots, W'_{nref}$  qui sont de dimension  $m$ .

10 En sortie de ladite deuxième couche dudit réseau de neurones, on obtient comme également représenté à la figure 7, un état de sortie  $d_1 \dots d_l \dots d_{nref}$  respectif desdits deuxièmes neurones  $C_{21} \dots C_{2nref}$  qui est de préférence analogique, et qui correspond à une distance entre ledit deuxième vecteur de poids synaptiques de ladite deuxième couche correspondant et ledit vecteur présenté à l'entrée de ladite deuxième

15 couche, en l'occurrence ledit vecteur intermédiaire  $Y$  ; les composantes du vecteur  $D$  de sortie de ladite deuxième couche dudit réseau de neurones peuvent être par exemple calculés par la formule :

$$d_l = d(W'_l, Y)$$

soit par exemple :

20 
$$d_l = (\sum_k (w'_{kl} - y_k)^2)^{1/2}, \quad k \text{ variant de } 1 \text{ à } m$$

Lesdites données de distance  $d_1 \dots d_{nref}$  (ou états de sortie) desdits neurones de ladite deuxième couche  $C_2$  sont de préférence véhiculées par des moyens de liaison 6, d'une part vers des moyens de comparaison 2 et d'autre part vers des moyens de détermination 4 de

25 l'indice ou du rang correspondant à celui desdits deuxièmes vecteurs  $W'$  qui est le plus proche dudit vecteur intermédiaire  $Y$  qui vient d'être présenté à ladite deuxième couche dudit réseau de neurones, de manière à réaliser l'adaptation desdits deuxièmes vecteurs, par une loi d'adaptation compétitive desdits deuxièmes vecteurs  $W'$  desdits

30 deuxièmes poids synaptiques.

Des moyens de liaison 5 permettent la modification (l'apprentissage) desdits deuxièmes poids synaptiques, ledit apprentissage étant de préférence sensible à la fréquence de choix du vecteur le plus proche dudit vecteur intermédiaire  $Y$ , c'est-à-dire du

35 vecteur gagnant lors de chaque étape d'apprentissage (ou de plusieurs vecteurs gagnants -c'est-à-dire les plus proches-).

Lesdites données de distance  $d_1, \dots, d_{nref}$  peuvent être transmises à

des bornes d'entrée  $2a_1...2a_{nref}$  desdits moyens de comparaison 2 qui peuvent recevoir par des moyens de liaison 8, sur des bornes  $2b_1...2b_4$ , quatre valeurs  $r_h$  auxquelles lesdites distances peuvent être comparées de manière à fournir en sortie d'un comparateur, pour

5 chacune desdites données de distance  $d_1...d_{nref}$ , un quadruplet de données binaires et/ou numériques qui sont notées respectivement  $(d_{11}...d_{14})...(d_{nref1}...d_{nref4})$  ; lesdites données numériques et/ou binaires  $d_{ij}$  correspondent à un état indiquant si ladite distance  $d_i$  est comprise dans une hypersphère de rayon  $r_j$  (transmis par lesdits

10 moyens de liaison 8), centrée sur ledit vecteur intermédiaire Y, cette opération étant effectuée à chaque présentation d'un vecteur d'entrée audit réseau de neurones ; lesdites données numériques font l'objet d'une moyenne spatiale dont les résultats sont transmis par des moyens de liaison 7 à des moyens 3 de compteur ou de sommateur temporel qui

15 permettent d'effectuer les opérations de calcul subsidiaire, permettant d'obtenir (selon l'une des méthodes explicitées ci-dessus) par un module DF, une estimation de la dimension fractale de la projection de l'attracteur du signal représenté par lesdits vecteurs d'entrée E ou X qui ont été successivement présentés auxdits moyens de

20 réseau de neurones.

La figure 9 illustre des deuxièmes neurones  $C_{21}...C_{2l}...C_{2nref}$  d'un réseau de neurones d'un dispositif selon l'invention, lequel réseau de neurones comporte, dans ce mode de réalisation, une

quatrième couche C4 de quatrièmes neurones  $C_{41}...C_{4q}$  qui constituent

25 une partie au moins desdits moyens 2 de comparaison desdites données de distance  $d_1...d_l...d_{nref}$  par rapport à q valeurs prédéterminées  $r_1...r_q$  ; lesdits moyens de réseaux de neurones comportent une troisième couche C3 de troisièmes neurones  $C_{31}...C_{3l}...C_{3nref}$  qui constituent une partie au moins des moyens (4) de détermination de

30 l'indice (ou du rang) correspondant à un au moins desdits deuxièmes vecteurs, le plus proche dudit vecteur d'entrée présenté à ladite deuxième couche dudit réseau de neurones (c'est-à-dire dudit vecteur intermédiaire).

On voit que dans ce mode de réalisation, lesdits moyens 2 de

35 comparaison sont munis de  $(q \times nref)$  modules  $2_{11}...2_{nrefq}$  de comparaison entre lesdites données de distance  $d_1...d_l...d_{nref}$  calculées par lesdits deuxièmes neurones  $C_{21}...C_{2l}...C_{2nref}$  et

respectivement  $q$  valeurs prédéterminées  $r_1 \dots r_q$ , de manière à fournir en sortie desdits modules comparateurs des informations  $d_{11} \dots d_{nrefq}$  numériques et/ou binaires déterminant l'appartenance dudit deuxième vecteur  $W'_i$  à  $q$  hypersphère de rayons respectifs  $r_1 \dots r_q$ , centrées sur ledit vecteur intermédiaire présenté à l'entrée de ladite deuxième couche, lesquelles informations peuvent être sommées (ou moyennées) par lesdits quatrièmes neurones, puis délivrées auxdits moyens 3 (de sommation et/ou moyenne temporelle) et DF permettant le calcul de l'estimation de ladite dimension fractale (par référence à la figure 8).

Avantageusement, lesdites informations numériques et/ou binaires  $d_{11} \dots d_{i1q} \dots d_{nref}$  sont présentées à l'entrée de ladite quatrième couche desdits quatrièmes neurones  $C_{41} \dots C_{4q}$  qui sont munis de liens synaptiques tous identiques dont les poids sont par exemple tous égaux à la valeur  $1/nref$  de sorte que chacun desdits quatrièmes neurones de ladite quatrième couche effectue une moyenne de la quantité desdits deuxièmes vecteurs situés dans lesdites hypersphères centrées sur ledit vecteur intermédiaire (moyenne spatiale).

On voit également sur cette figure que dans ce mode de réalisation, des données de distance modifiée  $d^*_1 \dots d^*_i \dots d^*_{nref}$  sont fournies en sortie desdits troisièmes neurones de ladite troisième couche, lesquels troisièmes neurones sont très simples puisqu'ils sont respectivement munis d'un seul lien synaptique  $U_1 \dots U_{nref}$  qui relie chacun desdits troisièmes neurones avec ledit deuxième neurone de ladite deuxième couche  $C_2$  correspondant, de manière à effectuer le calcul de ladite distance modifiée selon les méthodes indiquées ci-avant ; lesdites données de distance modifiée sont transmises auxdits moyens 4 comportant des moyens de calcul selon les relations (4), (3) et (2) définies ci-avant, des modifications à apporter auxdits deuxièmes et troisièmes vecteurs de poids synaptique.

La détermination de l'indice (ou du rang) correspondant à un au moins desdits deuxièmes vecteurs le plus proche dudit vecteur intermédiaire peut être effectué par une cinquième couche de neurones (non représentée) opérant sur les sorties desdits troisièmes neurones de ladite troisième couche comme indiqué sur la figure 3 de Neural Networks Vol. 3, 1990, "Competitive Learning Algorithms for Vector Quantization", AHALT et al. page 282.



Le fonctionnement des neurones de ladite cinquième couche est décrit par exemple dans Neural Networks, Vol. 2, 1989, "Minimum Distance Automata in Parallel Networks for Optimum Classification", WINTERS et al., pages 127 à 132.

5        La figure 10 illustre un mode de réalisation d'un dispositif de traitement de signaux selon l'invention qui peut être connecté soit à des moyens de capteurs 10, par exemple des capteurs d'électrocardiogramme d'un patient 9, ou alternativement comme indiqué par la flèche en pointillé, à des moyens 11 telle qu'une caméra vidéo.

10       Les signaux issus dudit capteur peuvent être prétraités par un module 10bis de filtrage ou de détection de pic (un tel module peut donner en sortie directement les variations de l'intervalle R-R entre deux complexes QRS de l'électrocardiogramme, lesdites variations constituant le signal S<sub>1</sub> à analyser). Les signaux ainsi prétraités  
15       sont appliqués à l'entrée d'un module 13 d'acquisition desdits signaux qui peut échantillonner et numériser lesdits signaux, lesquelles valeurs échantillonnées et numérisées peuvent être stockées dans une mémoire tampon 14.

      Ledit dispositif de traitement de signaux peut être principalement  
20       architecturé autour (ou à l'intérieur) d'un ordinateur 12 par exemple un mini ordinateur ou micro ordinateur de type "compatible", comportant des moyens spécifiques d'acquisition de données et de mémorisation désignée par les repères 13 et 14 et comportant avantageusement plusieurs processeurs afin d'accélérer les temps de  
25       calcul.

      Ledit ordinateur 12 comporte selon l'invention, ledit moyen de réseau de neurones 1, lesdits moyens 2 de comparaison de données de distance, lesdits moyens 3 de sommation et ledit module DF de calcul de l'estimation d'une dimension fractale de l'attracteur construit à  
30       partir desdits signaux ou de l'image correspondant auxdits signaux vidéo transmis par ladite caméra.

      Ledit ordinateur comporte également un écran 16 permettant de visualiser en permanence lesdits deuxièmes vecteurs W' de poids synaptiques desdits deuxièmes neurones, lesquels vecteurs sont  
35       représentatifs de la forme de la projection de l'attracteur ; avantageusement, ledit ordinateur permet également d'afficher à tout moment la valeur estimée de ladite dimension fractale dans une fenêtre

16a dudit écran 16 ; ledit calculateur comporte de façon connue des moyens 15 de gestion desdits modules comportant au moins un micro processeur, lesquels moyens 15 sont reliés auxdits moyens 13, 14, 1, 2, 3, DF, 16 par des moyens de liaison tels qu'un bus ; ledit  
5 ordinateur 12 peut être par ailleurs relié à une imprimante 17 permettant d'éditer les résultats du procédé de traitement de signaux selon l'invention.

La figure 12 illustre les principaux composants d'un dispositif selon l'invention ; par référence à cette figure on voit que lorsqu'on  
10 présente lesdits vecteurs d'entrée à l'entrée 20 d'un dispositif de traitement de signaux, lesdits vecteurs d'entrée E sont présentés à l'entrée de ladite couche d'entrée définie par lesdits poids synaptiques d'entrée constituant ladite matrice M, permettant le centrage spatial desdits vecteurs d'entrée et également caractérisée  
15 par lesdits n seuils formant ledit vecteur de seuil  $\theta$  permettant le centrage temporel desdits vecteurs d'entrée, ladite couche d'entrée délivrant des états de sortie formant ledit vecteur d'entrée centré X.

La figure 11 illustre sous forme de représentation matricielle le fonctionnement général de ladite couche d'entrée.

20 Ladite matrice M est une matrice carrée d'ordre n ; dans le cas particulier où cette matrice est la matrice identité, ladite couche d'entrée permet la moyenne temporelle desdits vecteurs d'entrée E présentée au réseau de neurones, les composantes dudit vecteur de seuil  $\theta$  étant adaptées selon une loi de GROSSBERG ; il est à noter que  
25 de tels moyens peuvent également être utilisés pour constituer lesdits moyens repérés 3 à la figure 10 permettant d'obtenir ladite moyenne temporelle (ledit coefficient  $K_{nref}(rh)$ ).

Par référence à la figure 12 ledit vecteur d'entrée centré X est présenté à l'entrée de ladite première couche desdits premiers  
30 neurones dont les poids synaptiques définissent lesdits premiers vecteurs W, ladite première couche délivrant des états de sortie formant lesdits vecteurs intermédiaires Y.

Ledit dispositif comporte lesdits moyens 23 permettant de calculer les modifications ( $dW$ ) à apporter auxdits poids synaptiques de ladite  
35 première couche, ceci à partir dudit vecteur d'entrée centré et dudit vecteur intermédiaire correspondant et selon ladite première loi d'apprentissage.

Lesdits vecteurs intermédiaires Y sont présentés à l'entrée de ladite deuxième couche desdits deuxièmes neurones caractérisée notamment par lesdits vecteurs W' desdits deuxièmes poids synaptiques, laquelle deuxième couche desdits deuxièmes neurones fournit en sortie  
5 ledit vecteur D dont les composantes correspondent aux données de distance entre chacun desdits vecteurs W' (vecteurs de référence) et ledit vecteur intermédiaire qui a été présenté à l'entrée de ladite deuxième couche.

Lesdites données de distance formant ledit vecteur D sont  
10 utilisées d'une part par lesdits moyens 2 de comparaison desdites données de distance avec des valeurs  $r_1 \dots r_q$  prédéterminées, lesquels moyens 2 fournissent les informations nécessaires auxdits moyens 3 et DF de sommation et de pondération permettant le calcul d'une estimation de ladite dimension fractale fournie à une sortie 21 dudit  
15 dispositif.

Par ailleurs, lesdites données de distance constituant ledit vecteur D font l'objet d'une transformation par des moyens 4 dont une partie 4<sub>1</sub> peut être constituée par exemple par ladite troisième couche desdits troisièmes neurones dont lesdits troisièmes vecteurs de poids  
20 synaptiques constituent ladite matrice U diagonale (selon la figure 13) ; lesdites données de distance modifiée constituant ledit vecteur D\* sont appliquées à l'entrée d'un module 22 permettant le calcul desdites données Z, lesquelles valeurs Z (selon la relation (4) ci-avant) sont utilisées par deux modules (constituant une deuxième  
25 partie 4<sub>2</sub> desdits moyens 4) effectuant les opérations de calcul des fonctions f et g respectivement selon les relations (2) et relation (3) ci-avant, lesquelles opérations permettent respectivement le calcul des modifications dW' à apporter auxdits vecteurs de ladite matrice W' desdits poids synaptiques desdits deuxièmes neurones, lors  
30 de la phase d'apprentissage, et permettant d'autre part le calcul de données dU correspondant aux modifications apportées auxdits troisièmes poids synaptiques de ladite troisième couche permettant l'apprentissage de ladite troisième couche dans le cas où celle-ci est utilisée.

35 Un mode préférentiel d'utilisation d'un dispositif selon l'invention est le suivant :

On suppose que l'objet (image ou attracteur construit à partir du

signal à analyser) est préalablement numérisé et se présente sous forme d'une suite discontinue de valeurs mesurées à partir de l'objet (image "scannée" ou attracteur d'un signal échantillonné avec ou sans prétraitement) qui constituent lesdits vecteurs d'entrée de dimension

5 n. La première étape consiste à reconstruire l'objet dans un espace de phase englobant ("embedding space" en anglais) de dimension suffisante (dans le cas de l'attracteur d'un signal, on utilisera par exemple une des méthodes décrites ci-dessus).

Ledit signal (ou l'image) à analyser, préalablement numérisé(e),

10 avec ou sans prétraitement, se présente sous forme d'une suite de p points (vecteurs d'entrée E) de dimension n. Lesdits p points sont stockés en mémoire et correspondent à une "fenêtre de mesure" dudit signal ou de l'image. La largeur de la fenêtre de mesure et la nature de l'échantillonnage et du prétraitement utilisé pour la numérisation

15 peuvent varier et sont à définir en fonction de la complexité dudit signal (de l'image) à analyser.

Selon un mode particulier de réalisation, un dispositif de traitement des signaux et images complexes, objet de l'invention, comporte un réseau de neurones comprenant au moins trois couches.

20 Une couche d'entrée repérée  $C_0$  compte n neurones  $C_{01}, \dots, C_{0n}$ , c'est-à-dire autant que ladite dimension n desdits vecteurs d'entrée E qui seront présentés successivement audit réseau.

Une première couche  $C_1$  comporte un nombre  $m < n$  neurones  $C_{11}, \dots, C_{1m}$ , ledit nombre m correspondant à la dimension du sous-espace sur

25 lequel on désire réaliser l'ACP.

Une deuxième couche  $C_2$  comprend un nombre  $n_{ref}$  neurones  $C_{21}, \dots, C_{2n_{ref}}$ , c'est-à-dire autant que de vecteurs de référence choisis pour calculer ledit coefficient  $K_{n_{ref}}(r_h)$ .

La connectivité (ensemble de liens synaptiques) entre lesdits

30 neurones de ladite couche d'entrée et ladite première couche est celle d'un réseau "feedforward". Elle est caractérisée par m vecteurs  $W_j$  ( $j=1$  à m) correspondant à des premiers vecteurs des poids synaptiques desdits neurones de ladite première couche, un premier vecteur  $W_j$  étant constitué par le n-uplet de poids synaptiques  $w_{ji}$  entre un

35 neurone  $C_{0i}$  de ladite couche d'entrée ( $i=1$  à n) et un neurone  $C_{1j}$  de ladite première couche. Il en est de même pour la connectivité entre ladite première et ladite deuxième couche dudit réseau que l'on peut

résumer par la donnée de nref deuxièmes vecteurs  $W'_l$  ( $l=1$  à nref) de poids synaptiques desdits neurones de ladite deuxième couche. On définit enfin pour chaque neurone  $C_{2l}$  de ladite deuxième couche, un paramètre  $U_l$  qui peut se modifier à chaque présentation d'un desdits points (ou vecteurs) d'entrée E audit réseau.

Au départ, les valeurs (ou composantes) desdits vecteurs  $W_j$  ( $j=1$  à  $m$ ) et deuxièmes vecteurs  $W'_l$  ( $l=1$  à nref) sont par exemple tirées au hasard et les valeurs initiales desdits paramètres  $U_l(1)$  sont fixés à 1 pour la première présentation d'un desdits vecteurs d'entrée. Lesdites valeurs (ou poids synaptiques) desdits premiers et deuxièmes vecteurs de poids synaptiques sont modifiées par la suite respectivement selon lesdites première et deuxième loi d'apprentissage, pendant ladite phase d'apprentissage durant laquelle on présente audit réseau de neurones successivement chacun desdits  $p$  points de l'objet (signal) à analyser.

Après la présentation d'un desdits vecteurs d'entrée audit réseau, lesdits premiers vecteurs  $W_j$  et deuxièmes vecteurs  $W'_l$ , ainsi que lesdits paramètres  $U_l(n)$  sont modifiés. L'ordre de présentation desdits  $p$  points peut être celui de leur stockage ou de leur apparition (ou mesure), ou bien faire l'objet d'un tirage au sort (procédure de Monte Carlo par exemple).

La série desdits  $p$  points peut être présentée plusieurs fois audit réseau jusqu'à la convergence de ce dernier. On appelle convergence dudit réseau le fait que lesdits vecteurs  $W_j$  et  $W'_l$  ne soient plus significativement modifiés lors de la présentation de nouveaux points d'entrée. L'état de convergence peut être testé par un "critère de convergence".

Lorsque la convergence est atteinte, les valeurs desdits vecteurs  $W_j$  et  $W'_l$  sont fixées et ledit réseau ainsi stabilisé est utilisé dans une procédure de calcul proprement dite (aussi appelée phase d'exploitation).

Ainsi, dans la phase d'apprentissage d'un mode particulier d'utilisation d'un dispositif selon l'invention, on réalise les opérations suivantes :

Lors de la présentation d'un point E de coordonnées  $e_i$  ( $i=1$  à  $n$ ), l'état de sortie de chaque neurone  $C_{0i}$  de ladite couche d'entrée prend la valeur  $x_i = \sum_m k_{im} e_i - \theta_i$  en fonction des coordonnées successives dudit

point ou vecteur d'entrée E. Les différentes tâches suivantes sont exécutées en parallèle :

- l'état de sortie de chaque neurone  $C_{1j}$  de ladite première couche prend une valeur égale à  $y_j = \langle W_j | X \rangle$  (produit scalaire de  $W_j$  par  $X$ ). On appelle  $Y$  le vecteur de coordonnées  $y_j$  ainsi créé sur ladite première couche du réseau, appelé vecteur intermédiaire, et simultanément, on utilise  $X$  pour adapter  $\theta$  selon une loi de GROSSBERG,
  - au fur et à mesure du calcul des composantes  $y_j$  dudit vecteur intermédiaire, le vecteur  $W_j$  correspondant au vecteur desdits poids synaptiques dudit neurone  $C_{1j}$  de ladite première couche est modifié selon ladite première loi d'apprentissage d'après la relation (1) définie ci-avant,
  - au fur et à mesure du calcul desdites composantes  $y_j$ , l'état de sortie  $d_l$  de chaque neurone  $C_{2l}$  de ladite deuxième couche prend une valeur proportionnelle à  $d(W'_l, Y)$ ,
  - au fur et à mesure du calcul des  $d_l$ , l'état de sortie de chacun desdits troisièmes neurones  $C_{3l}$  de ladite troisième couche prend une valeur égale à  $d_l^* = u_l \cdot d_l$  ; on détermine alors au moins un indice ou rang  $l^*$  d'un neurone  $C_{2l^*}$  de ladite deuxième couche qui correspond à un neurone  $C_{3l^*}$  de ladite troisième couche ayant en sortie une valeur correspondant à la valeur minimale desdites données de distance  $d_l$ , ledit deuxième neurone et/ou ledit troisième neurone correspondant et/ou ledit deuxième vecteur correspondant, étant appelé neurone et/ou vecteur "gagnant" c'est-à-dire le plus proche dudit vecteur intermédiaire  $Y$ . Pour ledit neurone  $C_{2l^*}$ , et éventuellement pour d'autres neurones de ladite deuxième couche, on calcule alors  $z_l(t)$  selon la relation (4) définie ci-avant, et on modifie lesdits deuxièmes vecteurs de poids synaptique selon la relation (2) définie ci-avant.
- Cette procédure est reprise lors de la présentation du point d'entrée suivant.
- Au fur et à mesure de la présentation desdits points d'entrée E, on assiste à une modification progressive des  $W_j$  qui tendent à s'orthogonaliser et à se normer. Lesdits premiers vecteurs  $W_j$  se stabilisent sur  $m$  premiers vecteurs propres de la matrice d'autocorrélation desdits vecteurs d'entrée, classés par ordre décroissant des valeurs propres de ladite matrice (on peut consulter à

ce sujet Neural Networks Vol.2 pages 459-473, 1989, "Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network". SANGER, 1989). Autrement dit ledit premier vecteur  $W_1$  définit un vecteur directeur normé de la droite de meilleur projection desdits vecteurs d'entrée, lesdits premiers vecteurs  $W_1$  et  $W_2$  définissent un plan de meilleur projection desdits vecteurs d'entrée, et lesdits premiers vecteurs de poids synaptiques  $W_1, W_2, \dots, W_m$  définissent un sous-espace de l'espace de dimension  $n$  desdits vecteurs d'entrée, lequel sous-espace est de dimension  $m$  et est sensiblement un sous-espace de meilleur projection au sens d'une analyse en composantes principales desdits vecteurs d'entrée.

L'intérêt desdites lois d'adaptation "feedforward" est de permettre des implémentations locales (cf. SANGER, p465), l'adaptation d'un neurone de ladite deuxième couche ne dépendant que des valeurs prises par lesdits neurones situés avant ce neurone sur ladite première ou ladite deuxième couche.

Parallèlement à cette adaptation, lesdits  $n_{ref}$  deuxièmes vecteurs  $W'_i$  opèrent par compétition, une quantification vectorielle de l'objet fractal à analyser ; lesdits  $n_{ref}$  deuxièmes vecteurs se répartissent de façon optimale sur ledit objet en respectant sa topologie, grâce à la sensibilité de ladite deuxième loi d'adaptation desdits deuxièmes neurones à la fréquence de choix des vecteurs "gagnants" correspondant auxdits deuxièmes neurones (les plus proches dudit vecteur intermédiaire  $Y$ ), les zones à forte densité dudit objet (attracteur) attirant plus de vecteurs de référence que celles à faible densité.

L'algorithme envisagé ici présente l'avantage de pouvoir subir une implémentation sous forme VLSI (cf AHALT et al).

Il est à noter que grâce au dispositif selon l'invention et son utilisation, lorsqu'on a présenté lesdits  $p$  vecteurs d'entrée  $E$ , on a créé ledit nombre  $p$  desdits vecteurs intermédiaires  $Y$  qui constituent l'attracteur des signaux correspondant auxdits vecteurs d'entrée, et qui correspondent plus particulièrement à la projection par analyse en composante principale dudit attracteur constitué à partir desdits signaux ; grâce au dispositif selon l'invention et son utilisation, on conserve lesdits  $n_{ref}$  vecteurs  $W'$ , qui constituent les vecteurs ou point de référence et qui ont été attirés par lesdits vecteurs intermédiaires  $Y$  lors de ladite phase d'apprentissage desdits poids

synaptiques desdits deuxièmes neurones, et non pas choisis parmi lesdits vecteurs intermédiaires comme il est généralement effectué dans les procédés et dispositifs connus de calcul de dimension fractale.

5 Lorsque lesdits premiers et deuxièmes vecteurs de poids synaptiques  $W_j$  et  $W'_j$  sont stabilisés on opère de la façon suivante :

Pour  $q$  valeurs prédéfinies d'un "rayon" noté  $r_h$  ( $h=1$  à  $q$ ), on pose  $K_{nref}(r_h)=0$ .

10  $p'$  points sont choisis parmi lesdits points d'entrée  $E$  et sont présentés successivement audit réseau.

Ledit réseau calcule lesdits vecteurs intermédiaires  $Y$  correspondants, en effectuant en parallèle lesdits produits scalaires  $\langle W_j | X \rangle$ . Pour lesdites  $q$  valeurs prédéfinies dudit rayon  $r_h$ , ledit réseau opère en parallèle le calcul desdits coefficients  $N_{nref,rh}$  selon la relation (5) ci-avant.

Au fur et à mesure de la présentation desdits  $p'$  points d'entrée, ledit réseau effectue pour chaque  $r_h$  en parallèle le calcul de ladite moyenne temporelle définie par la relation (6) ci-avant.

20 Une régression linéaire donne la pente de  $\log[K_{nref}(r_h)]$  en fonction de  $\log(r_h)$ , laquelle pente constitue une estimation de la dimension fractale recherchée, qui est en fait la dimension de la projection de l'attracteur.

Dans la procédure décrite ci-dessus, le calcul s'effectue sur les mêmes points d'entrée que ceux qui ont servi à l'apprentissage. Lors du traitement adaptatif d'un signal non linéaire, on peut envisager d'opérer le calcul sur lesdits  $p'$  points qui suivent lesdits  $p$  points qui servent à l'apprentissage dudit réseau. Ces mêmes  $p'$  points ou vecteurs d'entrée peuvent servir à l'apprentissage pour le calcul suivant.

30 Il est à remarquer que plus ledit signal d'entrée est stable, plus l'adaptation dudit réseau sera rapide à partir d'une configuration initiale ; de la sorte, après la phase de convergence dudit réseau sur des vecteurs d'entrée définis à partir d'un signal d'entrée relativement stable, une seule présentation desdits  $p$  points ou  
35 vecteurs d'entrée peut alors être suffisante pour réadapter lesdits points synaptiques dudit réseau et peut ainsi permettre de calculer et d'afficher en temps réel éventuellement l'attracteur dudit signal et



au moins la dimension fractale calculée par les procédés et dispositifs selon l'invention.

## REVENDECATIONS

1. Dispositif de traitement d'au moins un signal numérisé sous forme de vecteurs d'entrée (E) de dimension (n), caractérisé en ce qu'il comporte au moins un réseau (1) de neurones comportant au moins  
 5 une première couche (C<sub>1</sub>) ayant un nombre (m), inférieur ou égal audit nombre (n), de premiers neurones de poids synaptiques respectifs (w<sub>ij</sub>), lesquels poids synaptiques (w<sub>ij</sub>) peuvent constituer ledit nombre (m) de premiers vecteurs (W<sub>j</sub>) qui ont pour dimension ledit nombre (n), et ledit réseau de neurones comporte au moins une deuxième  
 10 couche (C<sub>2</sub>) ayant un nombre nref de deuxième neurones de poids synaptiques respectifs (w'<sub>k<sub>l</sub></sub>), lesquels poids synaptiques (w'<sub>k<sub>l</sub></sub>) peuvent constituer ledit nombre nref de deuxième vecteurs (W'<sub>l</sub>) qui ont pour dimension ledit nombre (m), et caractérisé en ce que un état de sortie (s<sub>j</sub>) d'un desdits premiers neurones d'indice (j) est  
 15 déterminé par le produit scalaire entre ledit vecteur d'entrée (E) et ledit premier vecteur (W<sub>j</sub>) d'indice j, l'ensemble desdits états de sortie pouvant constituer un vecteur intermédiaire (Y), et en ce que un état (D<sub>l</sub>) d'un desdits deuxième neurones d'indice l est relatif à une distance d(W'<sub>l</sub>,Y), et en ce que lesdits premiers neurones sont  
 20 dotés d'une première loi d'apprentissage qui est une loi de HEBB étendue, et en ce que lesdits deuxième neurones sont dotés d'une deuxième loi d'apprentissage par compétition.

2. Dispositif selon la revendication 1 caractérisé en ce que ladite deuxième loi d'apprentissage est sensible à la fréquence de  
 25 choix desdits deuxième vecteurs gagnants.

3. Dispositif selon l'une quelconque des revendications 1 à 2 caractérisé en ce qu'il comporte en outre au moins une couche d'entrée Co comportant n neurones d'entrée qui permet d'obtenir des vecteurs d'entrée (X) centrés spatialement et/ou temporellement.

30 4. Dispositif selon l'une quelconque des revendications 1 à 3 caractérisé en ce que ledit nombre nref est supérieur à 100 et inférieur à 100 000.

5. Dispositif selon l'une quelconque des revendications 1 à 4 caractérisé en ce que ledit nombre (m) est inférieur à 100.

35 6. Dispositif selon l'une quelconque des revendications 1 à 5 caractérisé en ce qu'il comporte des moyens (2) de comparaison de ladite distance d (W'<sub>l</sub>,Y) entre chacun desdits deuxième vecteurs

( $W'_1$ ) et ledit vecteur intermédiaire (Y), avec un nombre (q) de valeurs  $r_h$  prédéterminées.

5 7. Dispositif selon l'une quelconque des revendications 1 à 6 caractérisé en ce qu'il comporte des moyens (2, 3, DF) de calcul d'au moins une moyenne spatiale ( $N_{nref, r_h}$ ), au moins une moyenne temporelle ( $K_{nref}(r_h)$ ), et des moyens de calculs de regression de ladite moyenne temporelle en fonction de valeurs ( $r_h$ ).

10 8. Dispositif selon l'une quelconque des revendications 1 à 7 caractérisé en ce qu'il comporte des moyens (16) de visualisation de l'ensemble desdits deuxièmes vecteurs ( $W'_j$ ) permettant d'afficher une allure de l'attracteur dudit signal.

9. Dispositif selon l'une quelconque des revendications 1 à 8 caractérisé en ce qu'il comporte des moyens de visualisation de ladite dimension fractale.

15 10. Utilisation d'un dispositif de traitement de signaux selon l'une quelconque des revendications 1 à 9 caractérisé en ce que :

a) - on affecte auxdits premiers et deuxièmes vecteurs de poids synaptique, des valeurs initiales choisies (soit de façon aléatoire ou pseudo-aléatoire, soit les valeurs précédentes obtenues).

20 b) - on présente successivement p vecteurs choisis parmi lesdits vecteurs d'entrée, de préférence jusqu'à convergence desdits vecteurs de poids synaptiques,

c) - on fige les valeurs desdits poids synaptiques,

25 d) - on présente successivement p' vecteurs choisis parmi lesdits vecteurs d'entrée, et on calcule des valeurs d'une pente (DF (p')) qui donne une estimation de la dimension d'un attracteur construit à partir desdits vecteurs d'entrée, et on choisit  $n_{ref}$  points qui déterminent une allure dudit attracteur.

1/8

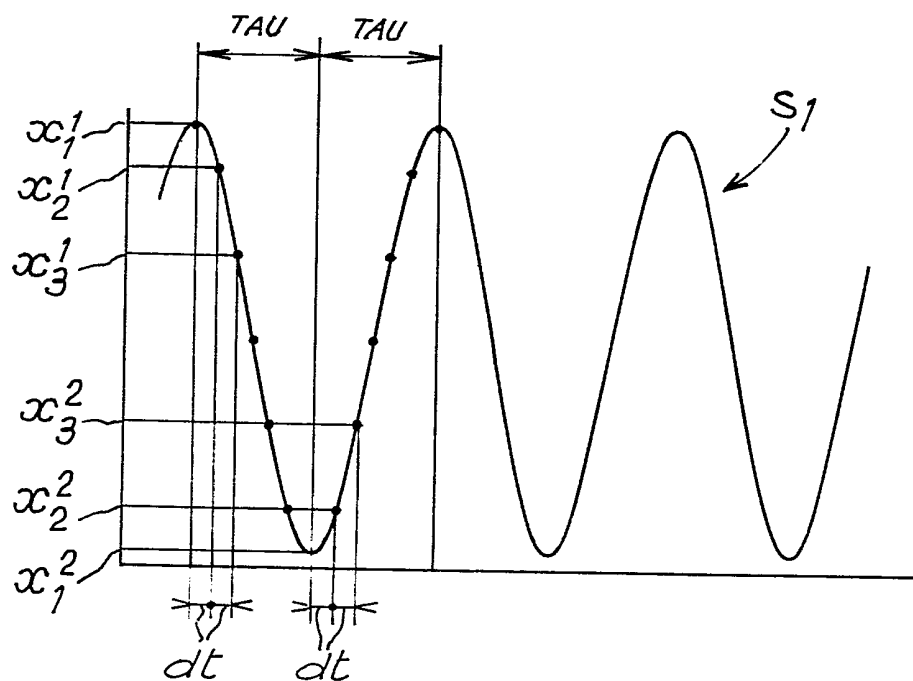


FIG. 1

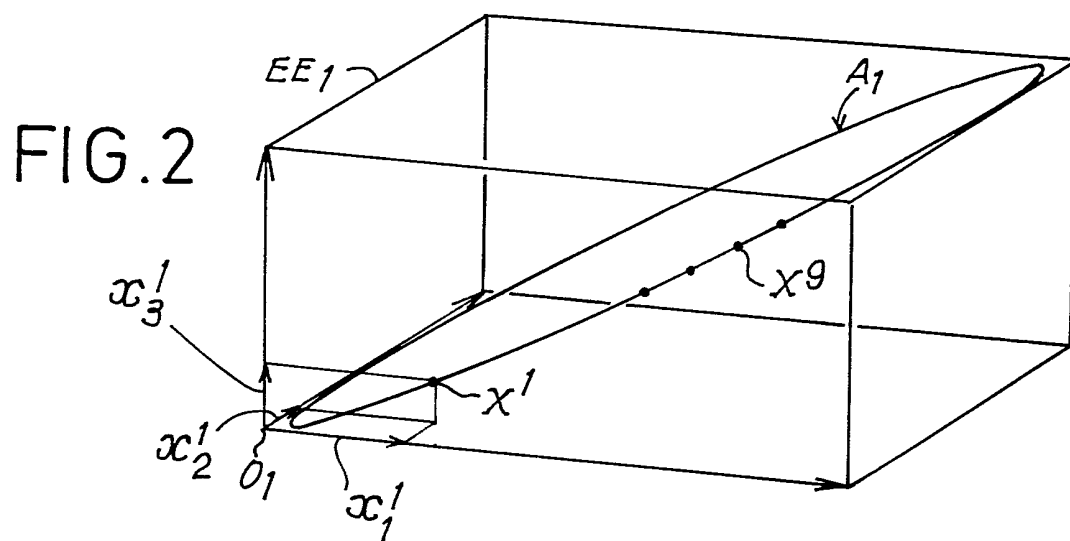


FIG. 2

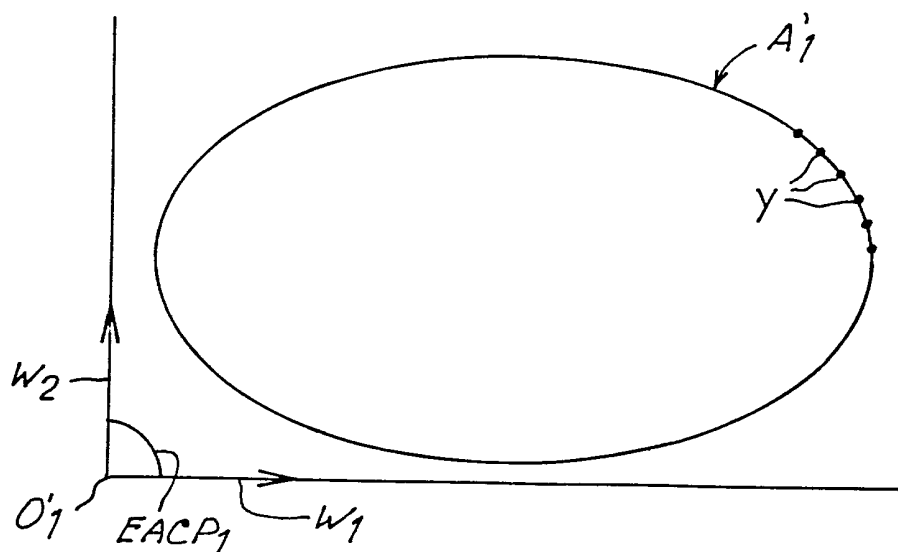


FIG. 3

2/8

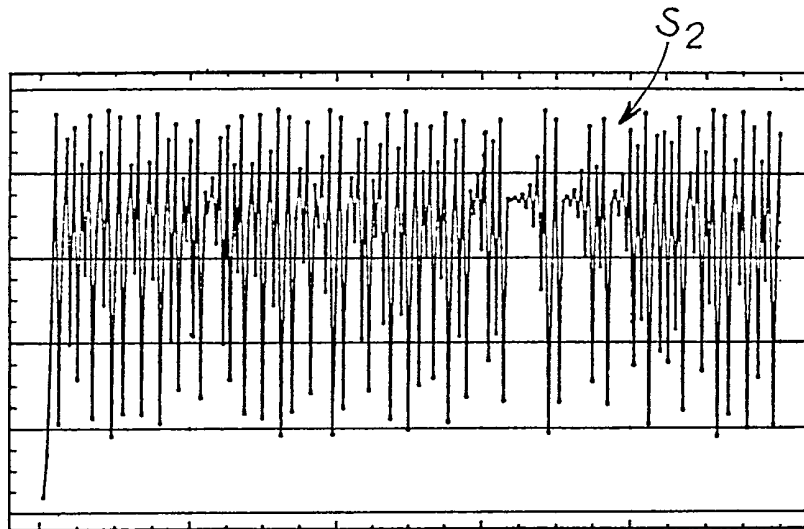


FIG. 4

FIG. 5

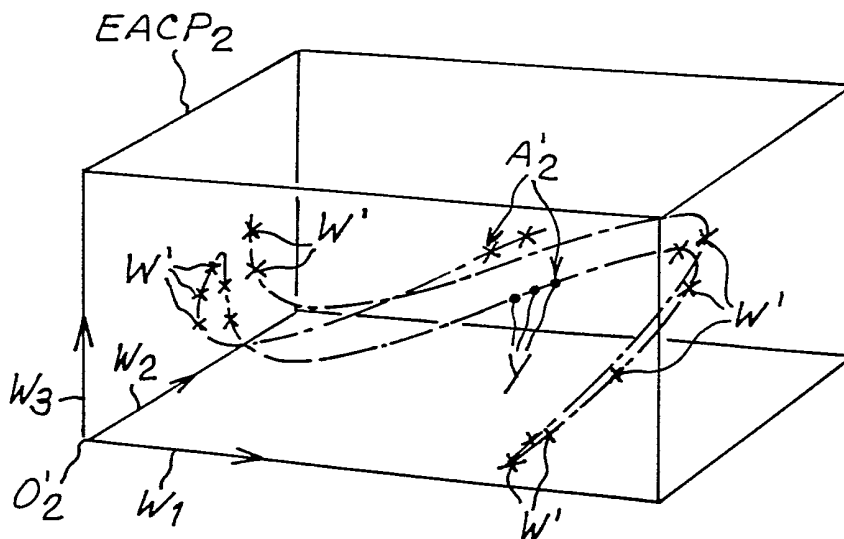
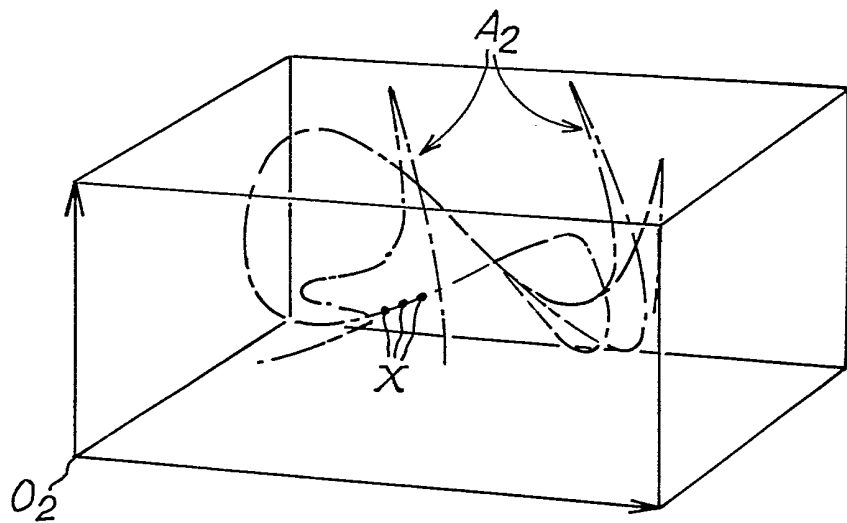
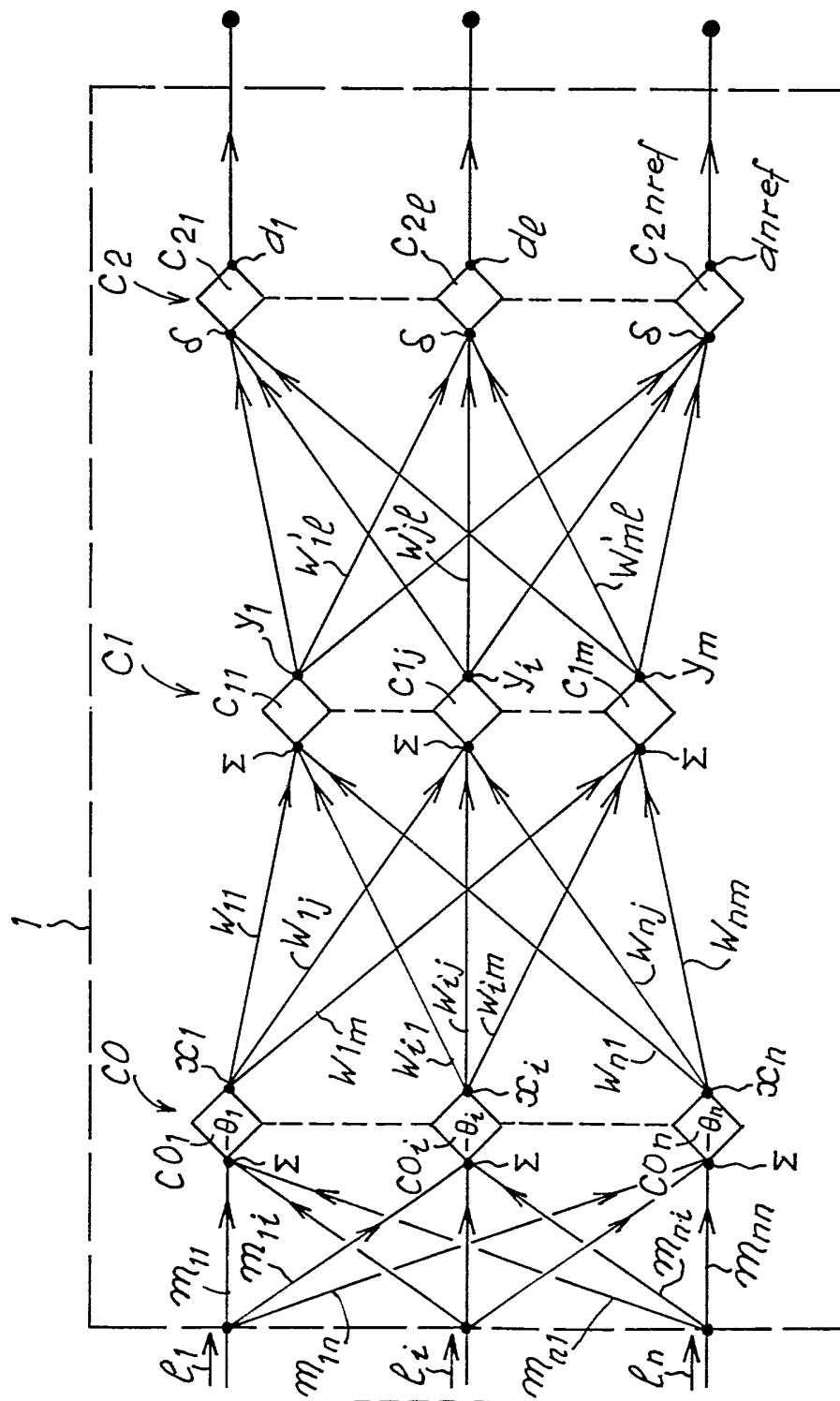
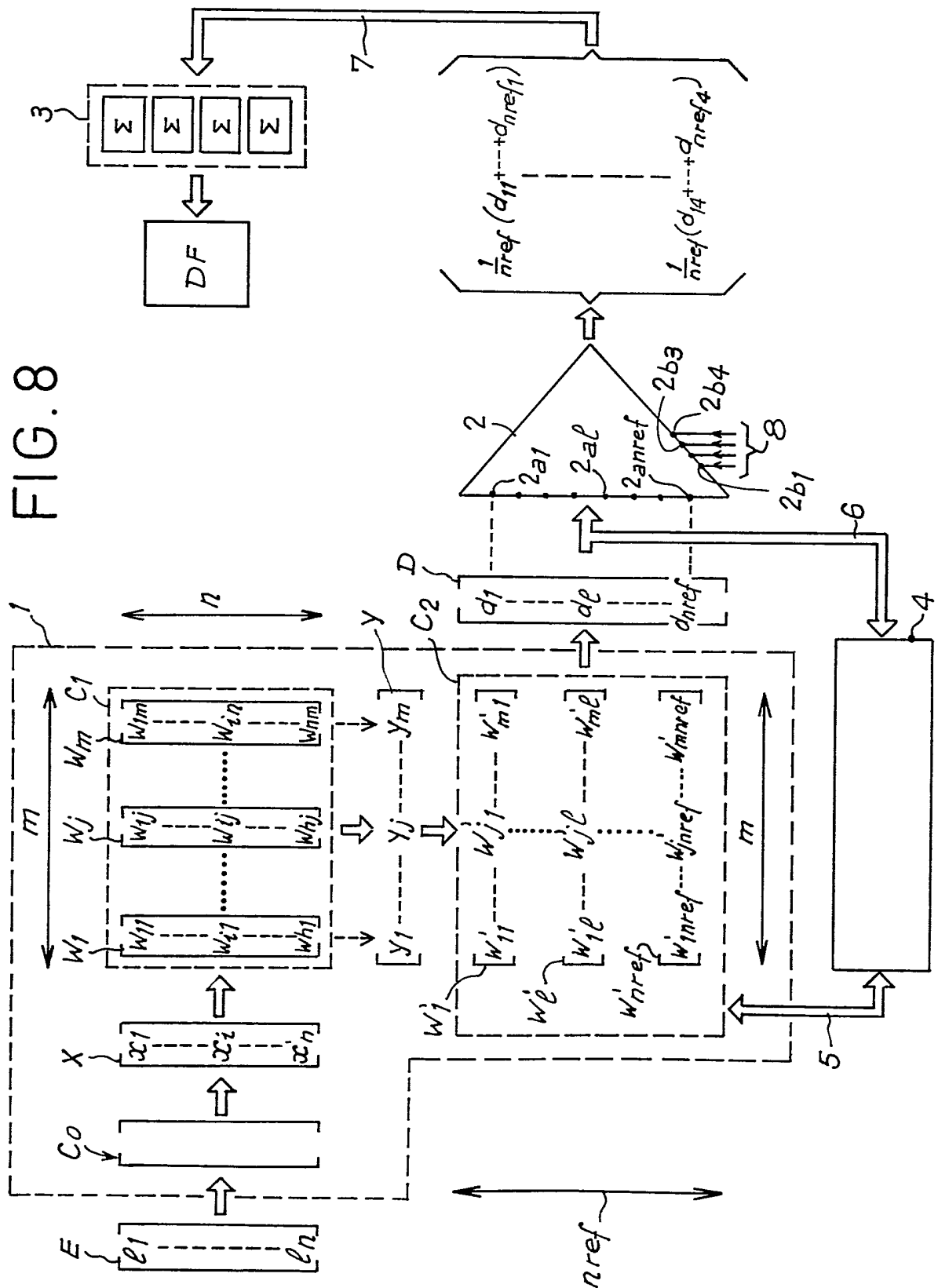


FIG. 6

FIG. 7





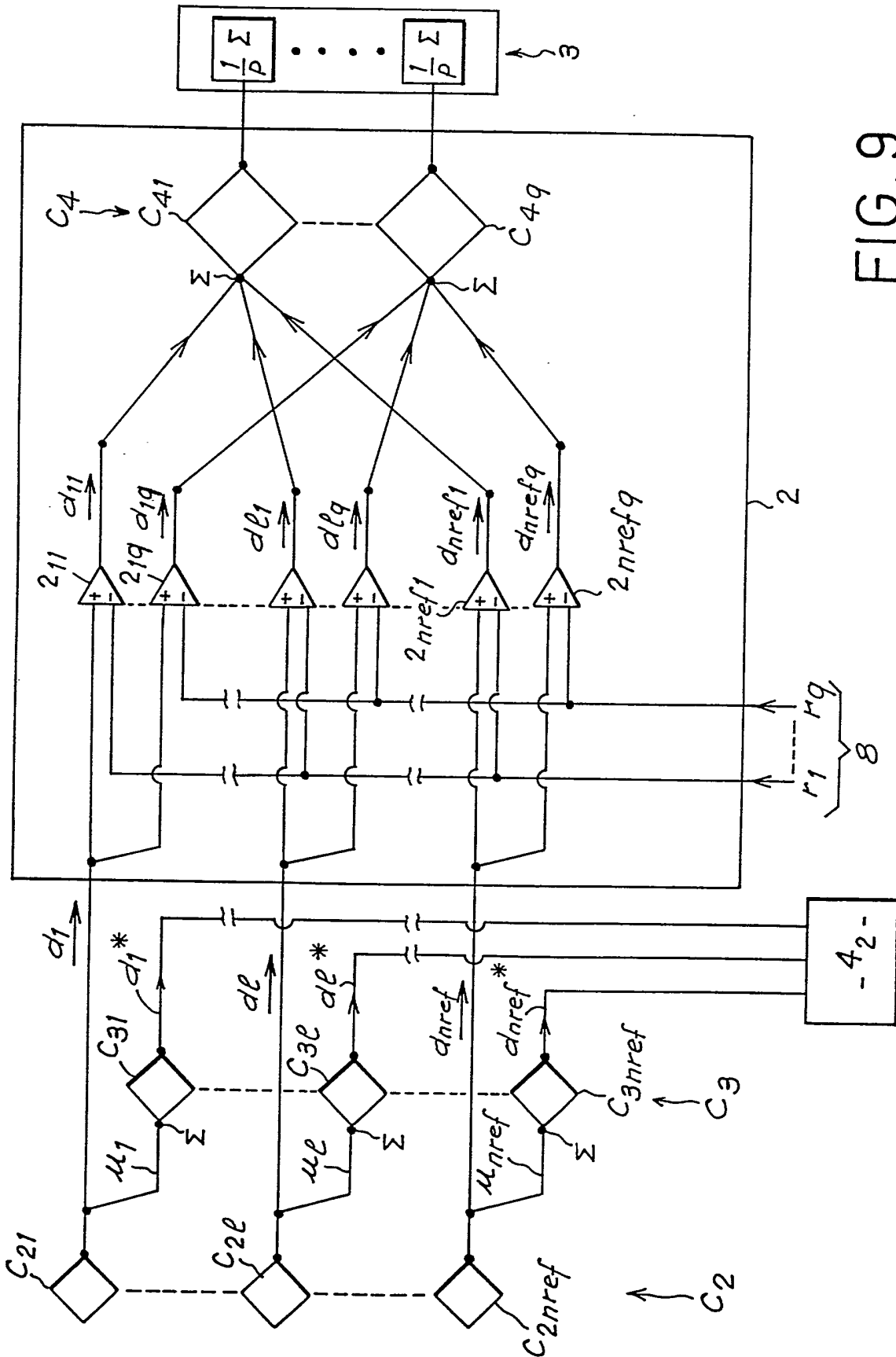
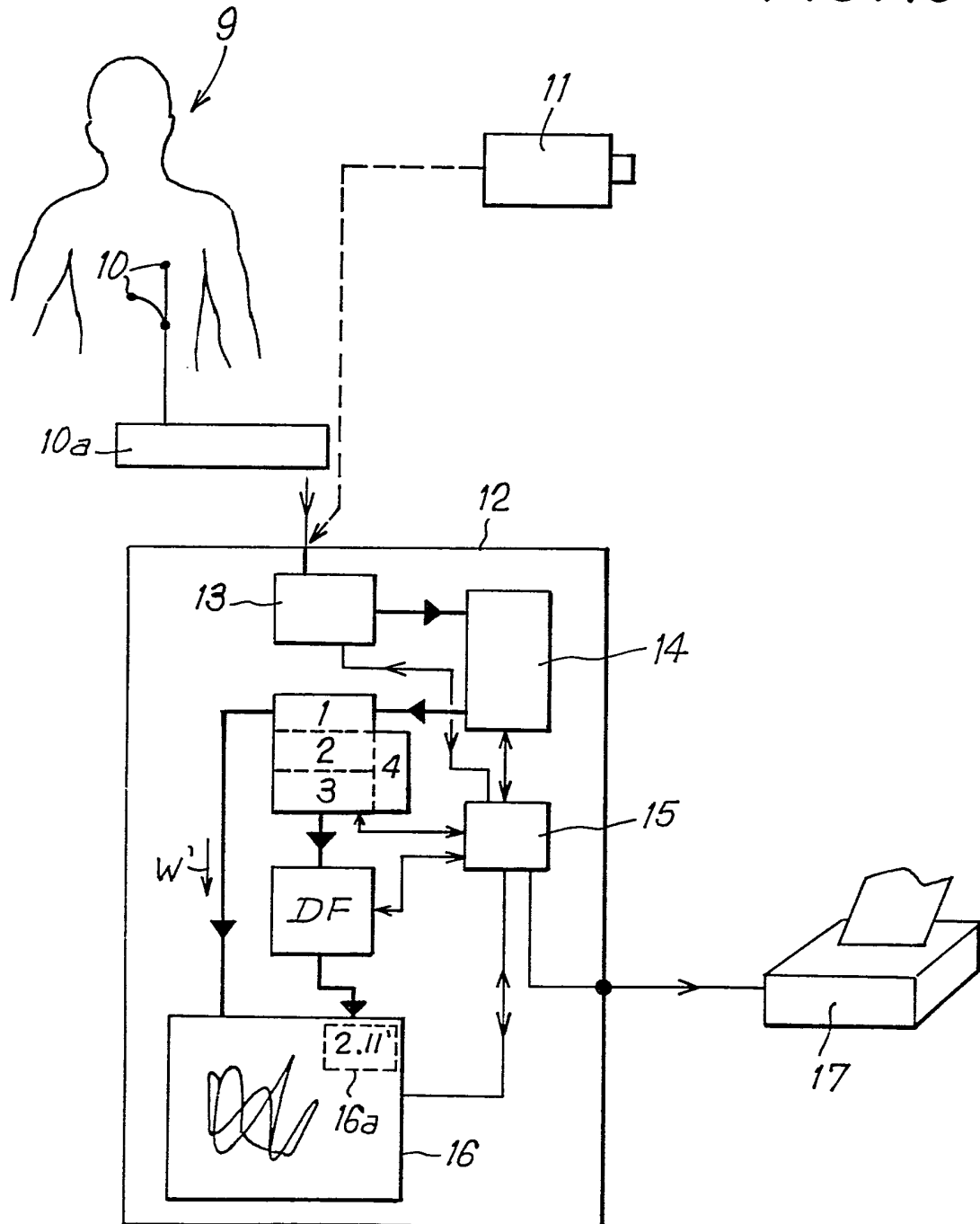


FIG. 9



6/8

FIG. 10



7/8

$$\begin{matrix} M \\ \left\{ \begin{matrix} m_{11} & \dots & m_{n1} \\ & \ddots & \\ m_{1n} & \dots & m_{nn} \end{matrix} \right\} \\ \downarrow n \\ \leftarrow n \rightarrow \end{matrix} \times \begin{matrix} E \\ \left\{ \begin{matrix} e_1 \\ \vdots \\ e_n \end{matrix} \right\} \end{matrix} - \begin{matrix} \theta \\ \left\{ \begin{matrix} \theta_1 \\ \vdots \\ \theta_n \end{matrix} \right\} \end{matrix} = \begin{matrix} x \\ \left\{ \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix} \right\} \end{matrix}$$

$\begin{matrix} \uparrow \\ \begin{matrix} \frac{d\theta}{dt} \\ \left\{ \begin{matrix} \frac{d\theta_1}{dt} \\ \vdots \\ \frac{d\theta_n}{dt} \end{matrix} \right\} \end{matrix} \end{matrix}$

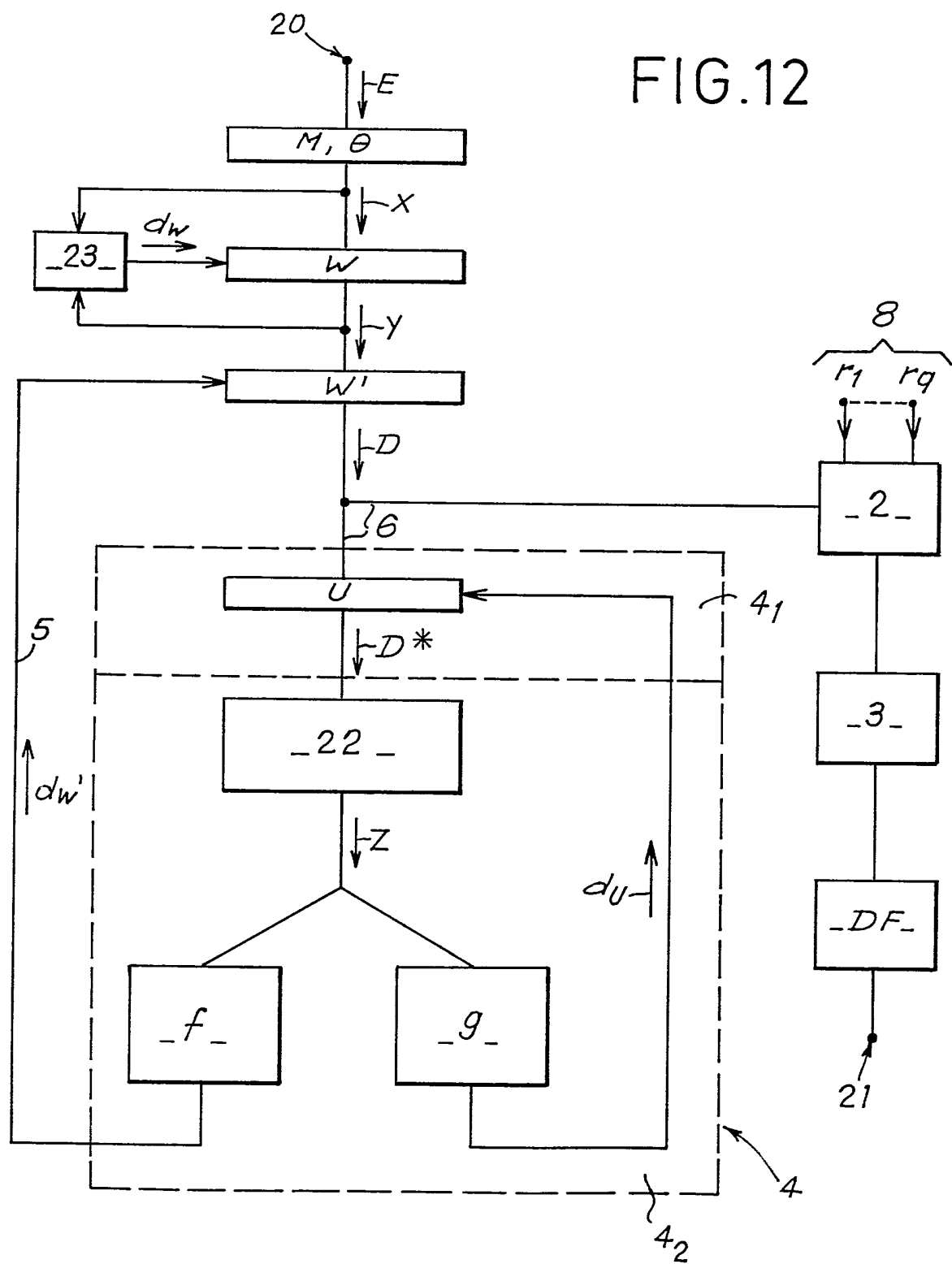
FIG. 11

FIG. 13

$$\begin{matrix} U \\ \left\{ \begin{matrix} u_1 & \dots & u_{nref} \end{matrix} \right\} \\ \downarrow nref \\ \leftarrow nref \rightarrow \end{matrix} \times \begin{matrix} D \\ \left\{ \begin{matrix} d_1 \\ \vdots \\ d_{nref} \end{matrix} \right\} \end{matrix} = \begin{matrix} D^* \\ \left\{ \begin{matrix} d_1^* \\ \vdots \\ d_{nref}^* \end{matrix} \right\} \end{matrix}$$

8/8

FIG.12



INSTITUT NATIONAL  
de la  
PROPRIETE INDUSTRIELLE

RAPPORT DE RECHERCHE  
établi sur la base des dernières revendications  
déposées avant le commencement de la recherche

FR 9116433  
FA 468959

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
D,Y	NEURAL NETWORKS vol. 3, no. 3, 1990, NEW YORK pages 277 - 290; STANLEY C. AHALT ET AL: 'Competitive Learning Algorithms for Vector Quantization' * page 279, colonne de droite, alinéa 3 - page 283, colonne de droite, alinéa 1; figure 3 *	1
Y	NETWORK - COMPUTATION IN NEURAL SYSTEMS vol. 2, no. 1, Février 1991, BRISTOL pages 63 - 84; CAROLA FASSNACHT ET AL: 'Recognition and categorization in a structured neural network with attractor dynamics' * page 65, alinéa 3 - page 66, dernier alinéa *	1
A	NEURAL NETWORKS vol. 4, no. 3, 1991, NEW YORK pages 337 - 347; WILLIAM R. SOFTKY ET AL: 'Correlations in High Dimensional or Asymmetric Data Sets: Hebbian Neuronal Processing' * page 338, colonne de droite, alinéa 3 - page 341, colonne de droite, alinéa 4; figure 2; exemple 3 *	1
D,A	NEURAL NETWORKS vol. 2, 1989, NEW YORK pages 459 - 473; TERENCE D. SANGER: 'Optimal unsupervised learning in a single-layer feedforward neural network' * le document en entier *	1
		DOMAINES TECHNIQUES RECHERCHES (Int. Cl.5)
		G06F
Date d'achèvement de la recherche 28 AOUT 1992		Examineur NICHOLLS J.
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons ..... &amp; : membre de la même famille, document correspondant</p>		