

# Analogie entre la perception bio-inspirée et la théorie des Topos

**Séminaire : Quelles mathématiques pour une IA fiable, frugale, moins artificielle ?**

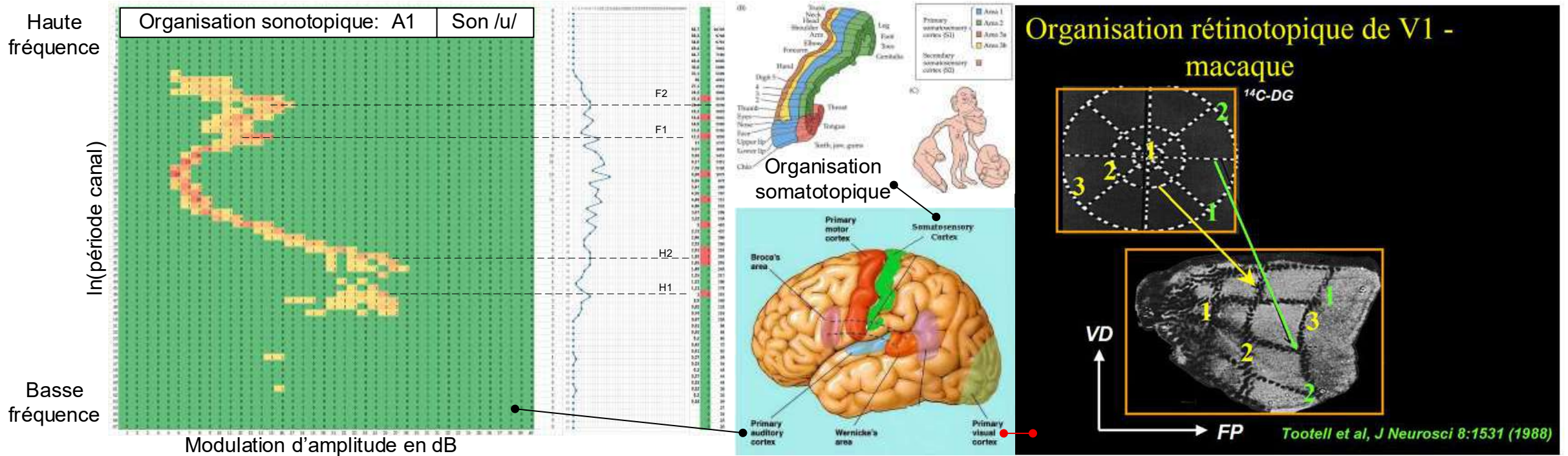
Patrick Pirim, 21 Mai 2026

# séparation claire en **4 niveaux** :

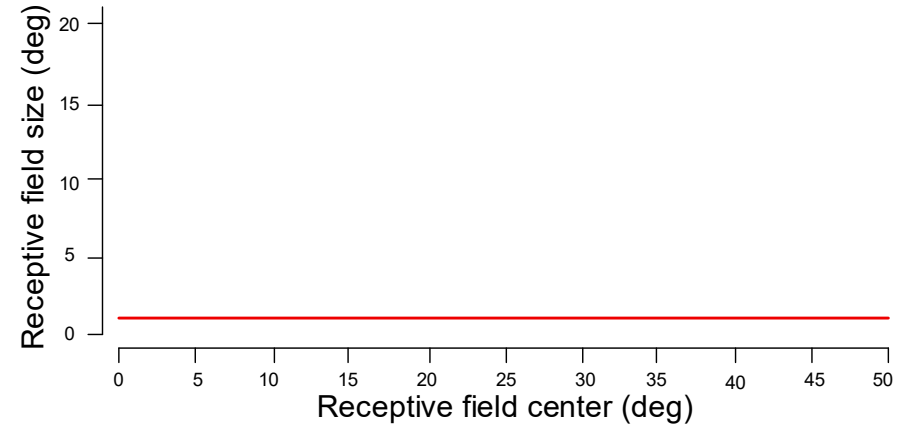
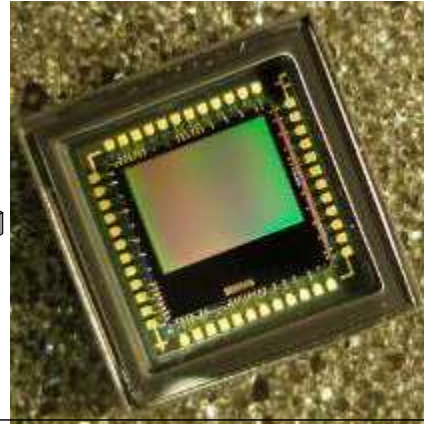
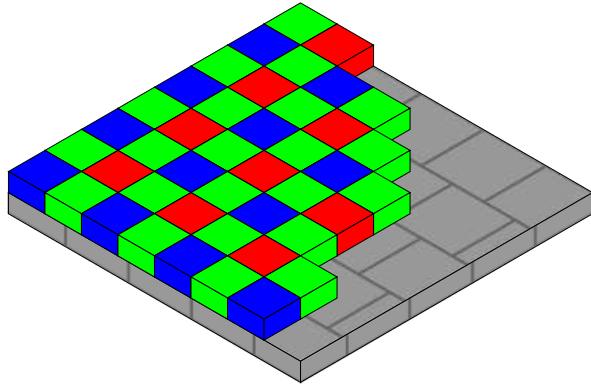
- perception physique
- intégration contextuelle
- structuration logique
- modélisation du monde
  
- Perspectives

perception physique

# Topicité des entrées perceptives



# Rétinotopie industrielle



# Rétinotopie biologique

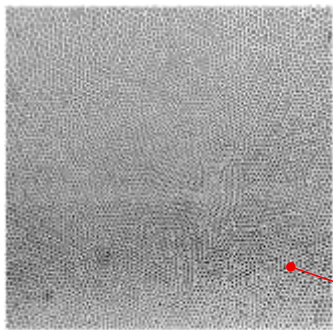
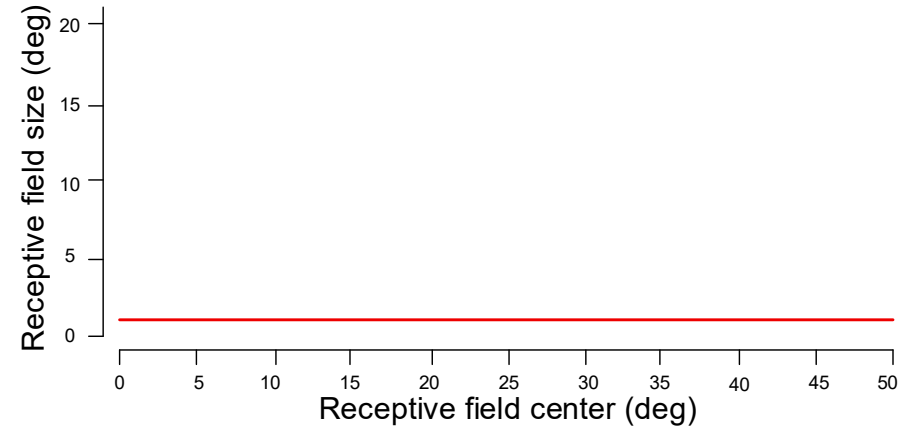
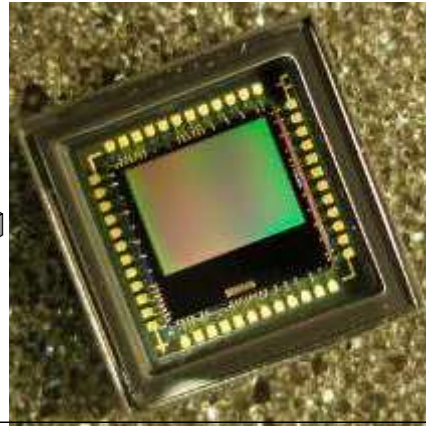
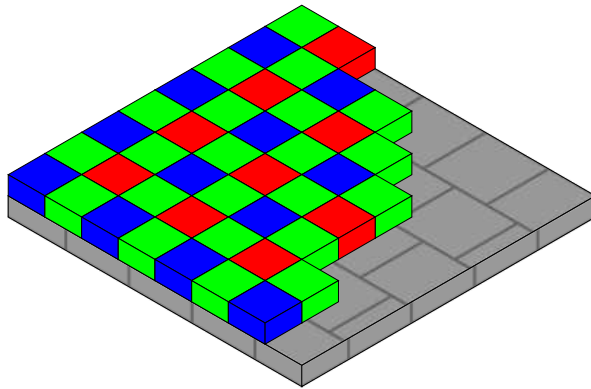
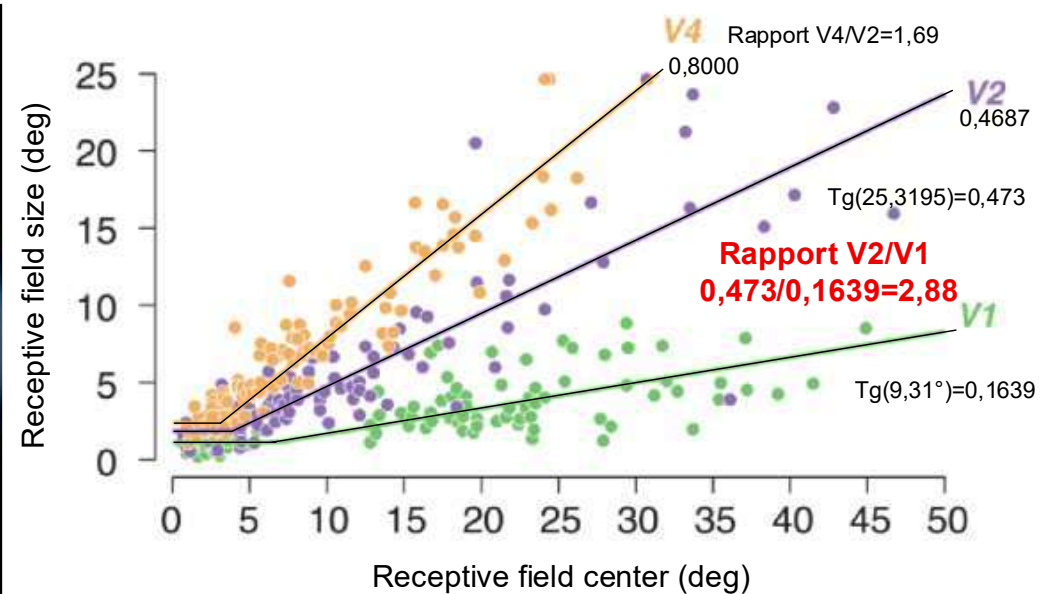
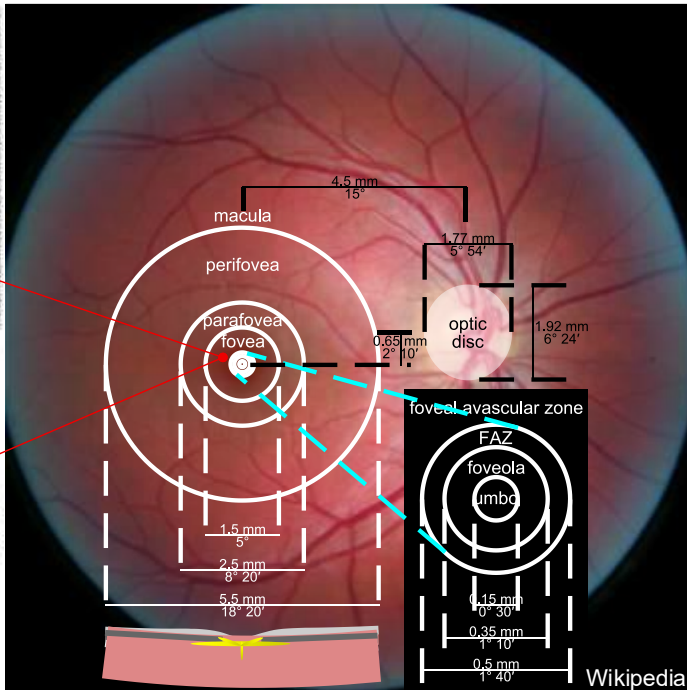
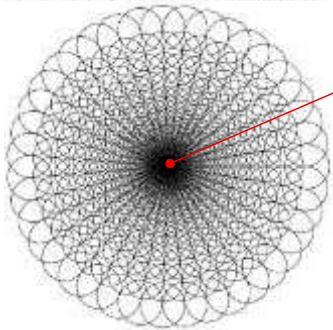
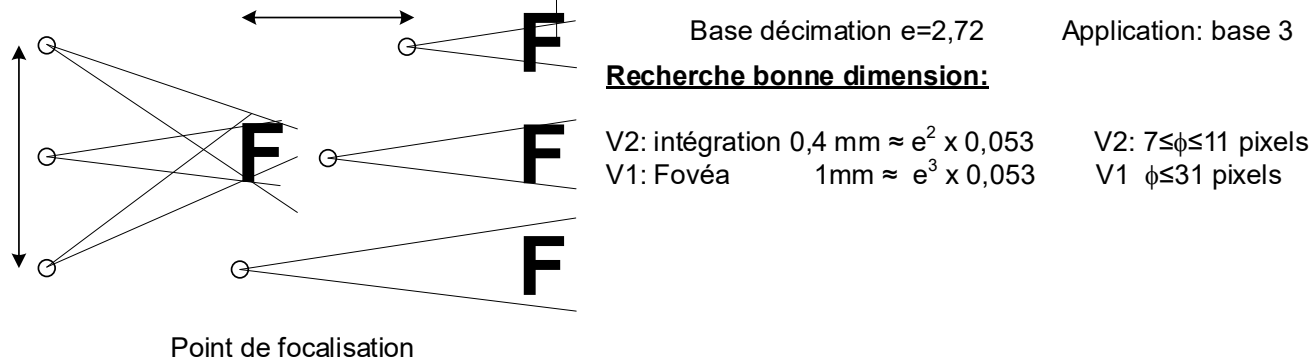
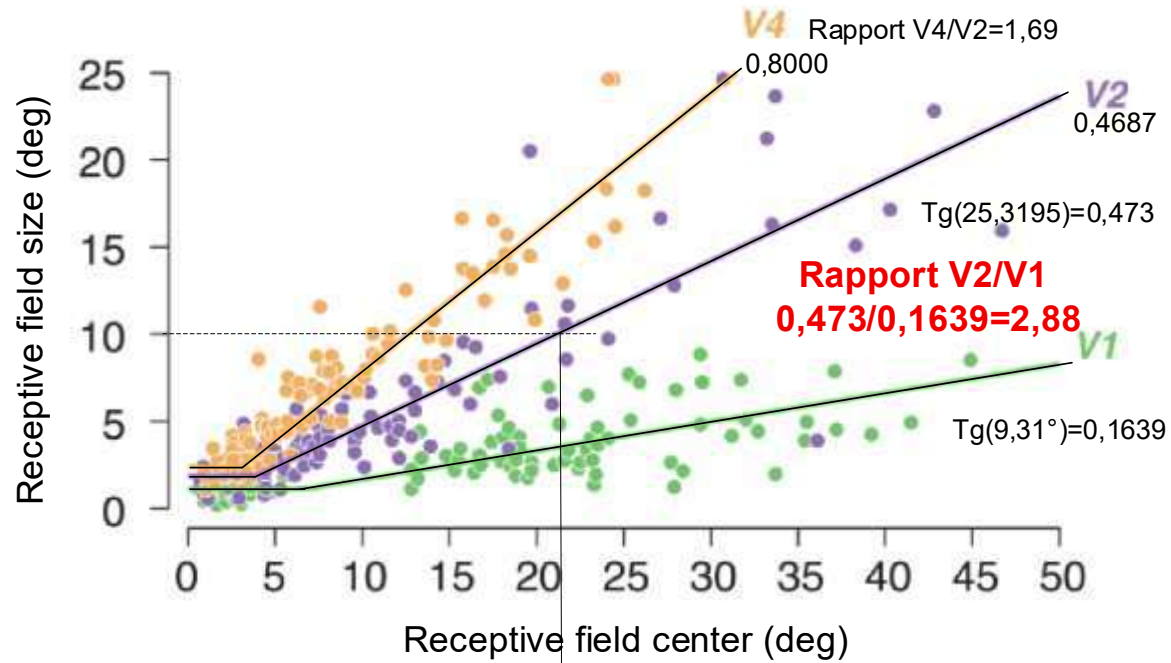


Figure 12.1. A horizontally sectioned and stained human retina at the foveal pit and mid foveal area. From Atiles, 2014.

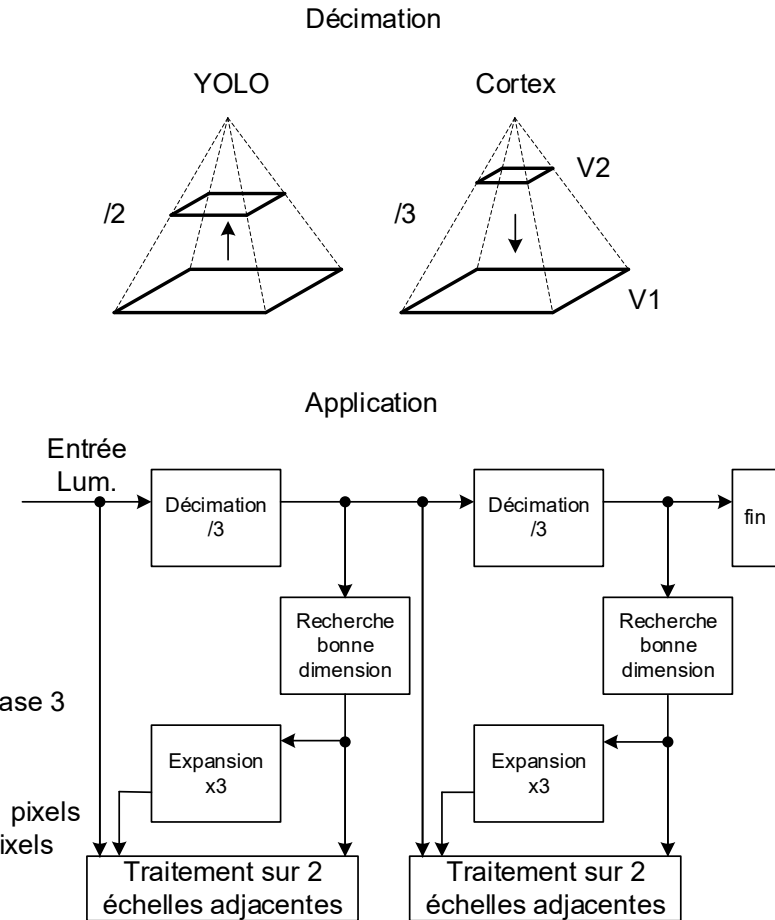
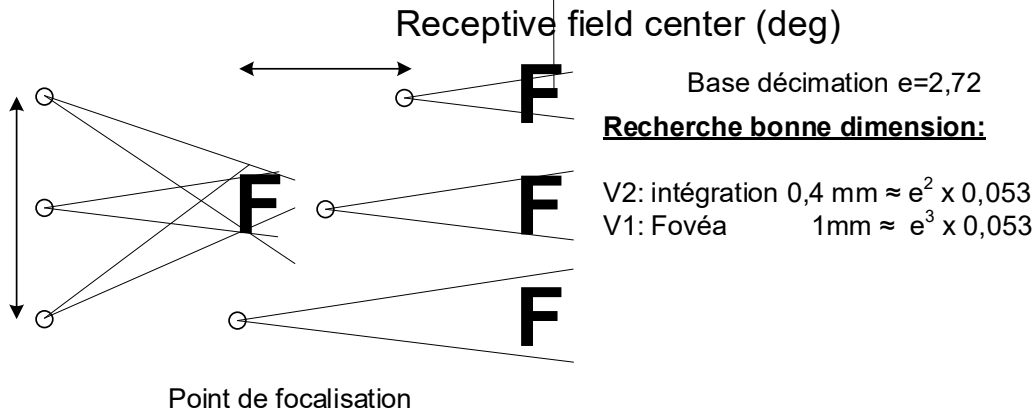
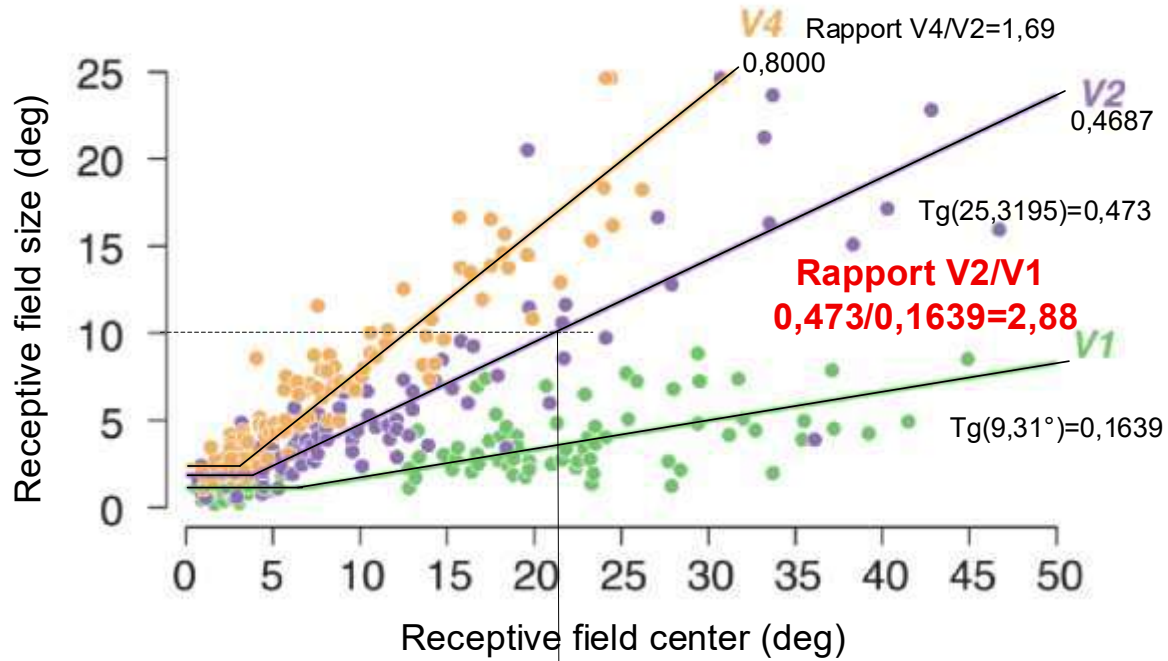


Cortical magnification factors in V1, V2 and V4 adapted from (Gattass et al, 1981) and (Gattass et al, 1988) by (Freeman and Simoncelli, 2011).

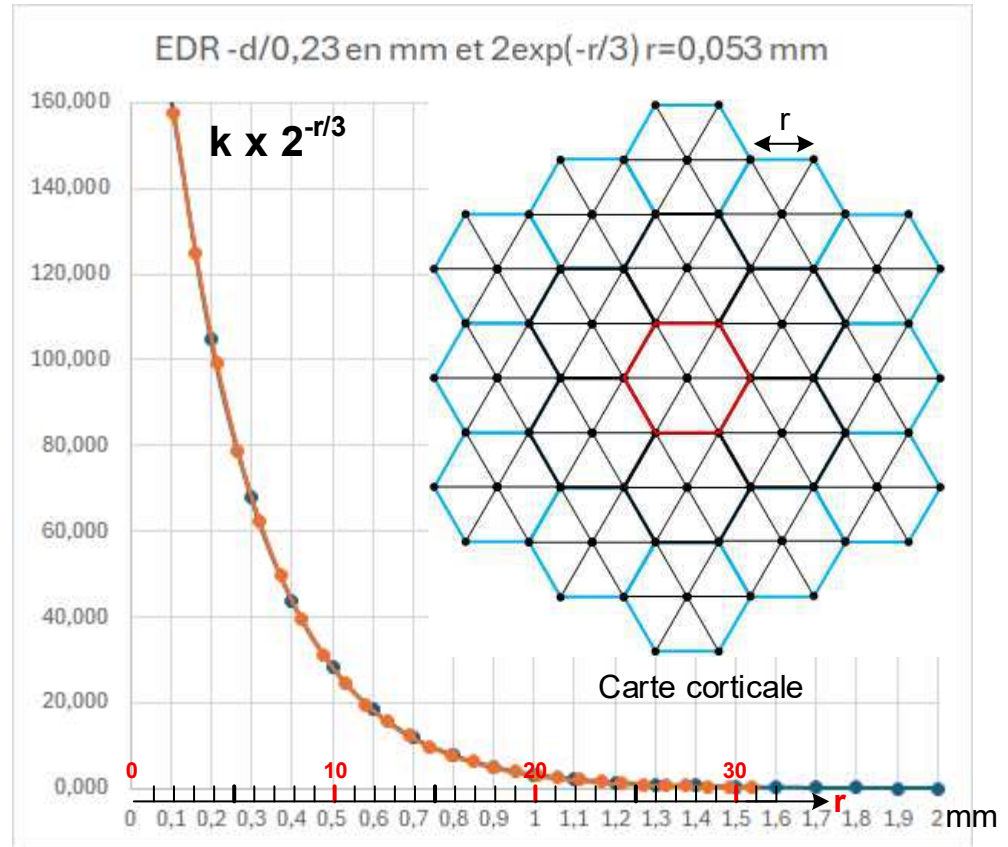
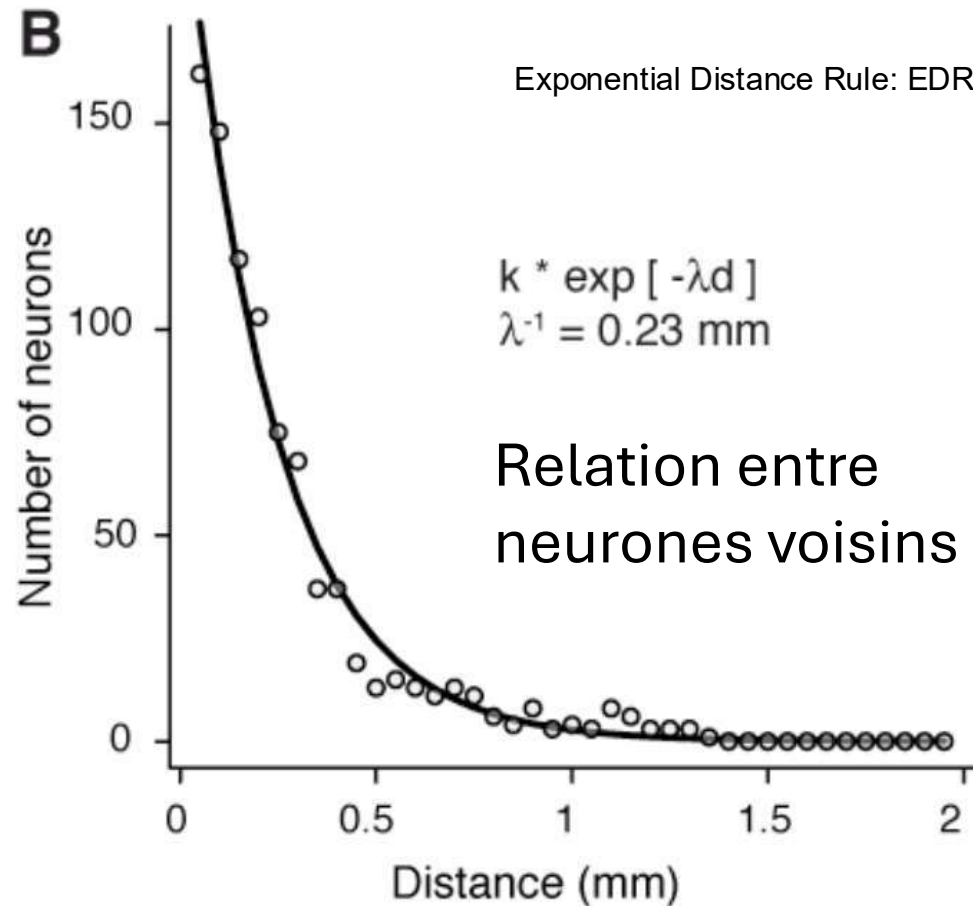
# Avantage de la rétinotopie biologique



# Adaptation de la rétinotopie bio-inspirée



# Topologie générique du cortex cerebral



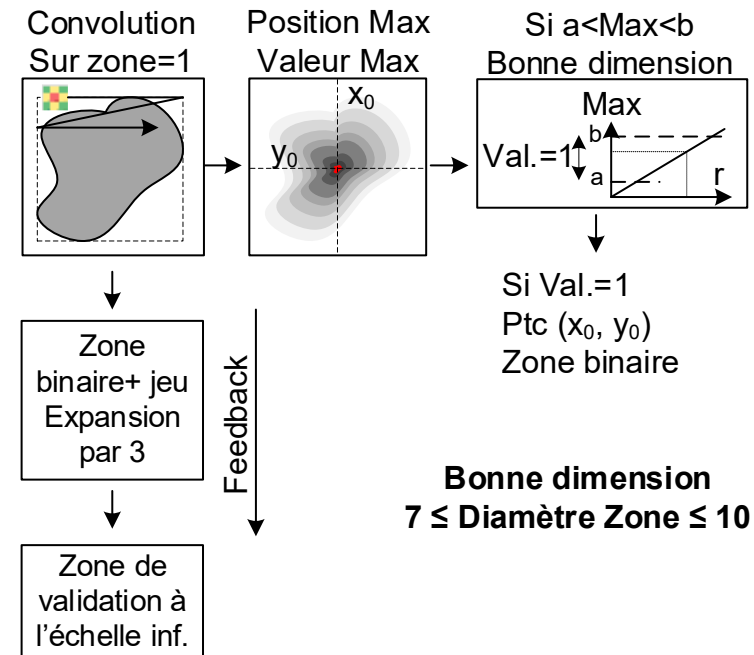
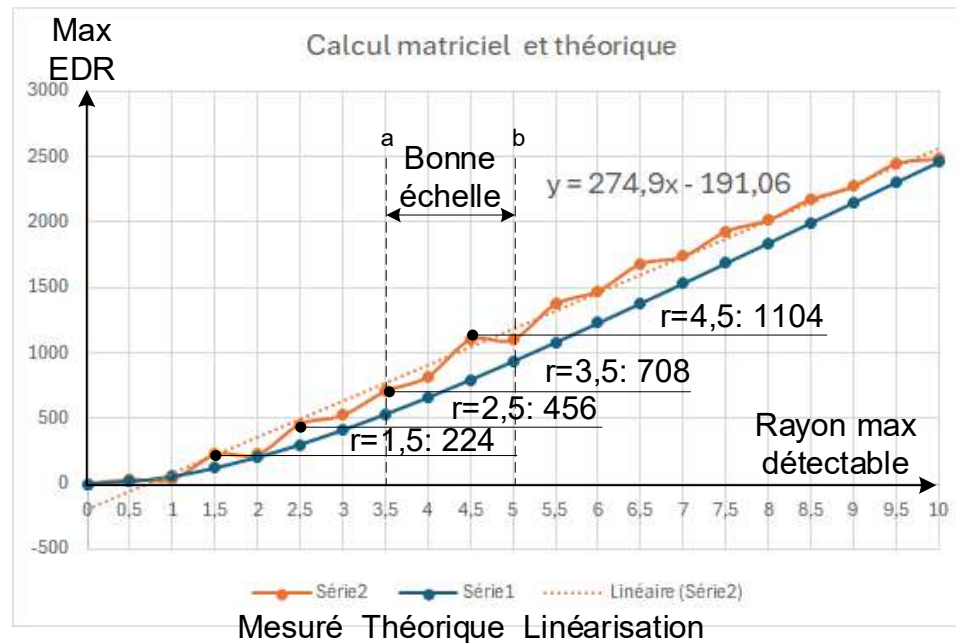
Les trois échelles biologiques en progression du nombre d'Euler 2,72:

Diamètre:  
 $e^3=20,1$  soit  $20,1 \times 0,053=1,07$  mm:  
 $e^2 \times 0,053=0,39$  mm  
 $e^1 \times 0,053=0,14$  mm:

This intrinsic connectivity is highly local, with an exponential decrease in the density of labeled neurons along increasing distance from injection site (Fig 1.2B)

**Exploration of the inter-areal cortico-cortical network of the macaque monkey**  
 Nikola Markov Th2010

# Point d'attention sur la bonne échelle de traitement



Permet l'invariance d'échelle de traitement: une seule mémorisation avec catégorisation de l'échelle

# En résumé:

**\*\*Le Catégorie de Base (Le Site) C \*\*** : Représente l'espace spatio-temporel des capteurs (rétine, cochlée). Les objets sont les zones réceptives (par exemple, les fenêtres locales 31x31) et les morphismes sont les inclusions ou les translations spatio-temporelles.

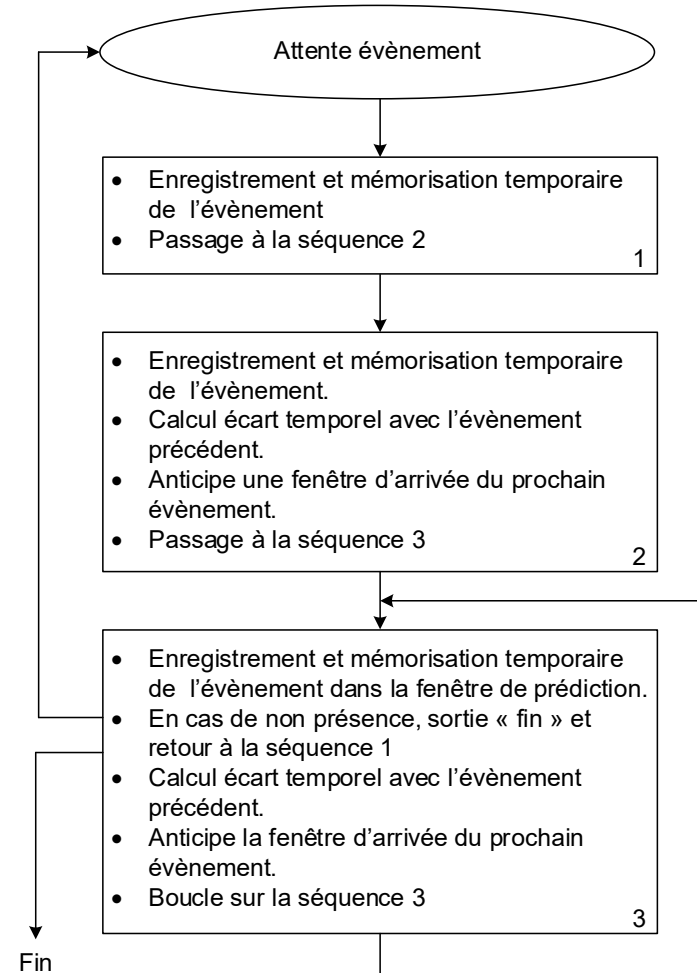
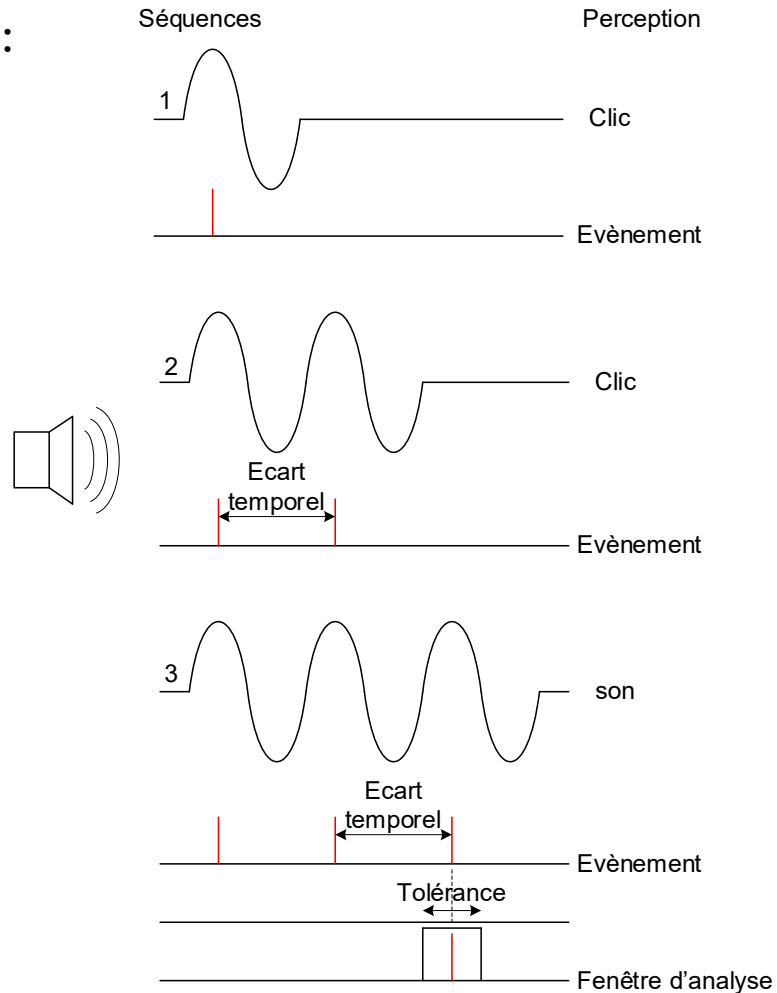
**\*\*Faisceaux (Sheaves) multi-échelles F \*\*** : Le système "multi-échelles" (Échelles 0 à 4 avec décimation par 3) correspond à un faisceau sur le site C . Une donnée (luminance L , variation) à une échelle fine se restreint (via la fonction de décimation/filtrage) à une échelle plus grossière. La condition de recollement du faisceau équivaut à la cohérence spatio-temporelle de la colonne corticale.

**Invariance perceptive (formes, sons, échelles, rotation)** : Modélisée par des faisceaux équivariants sous l'action du groupe de similarités (translations, rotations, scalings). L'invariance est obtenue par passage au quotient ou par colimites sur les orbites. Le texte note que le contexte est représenté en premier, puis l'invariance assure l'unicité — cela correspond au fait que le topos classifie d'abord les contextes (faisceaux sur le site de base) puis les objets invariants.

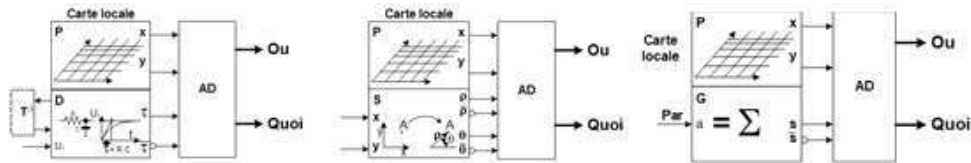
intégration contextuelle

# Prédiction Généralisée

Exemple sonore:



# L'élément de base: la colonne corticale



**Modalité DYNAMIQUE**

**Modalité STRUCTURALE**

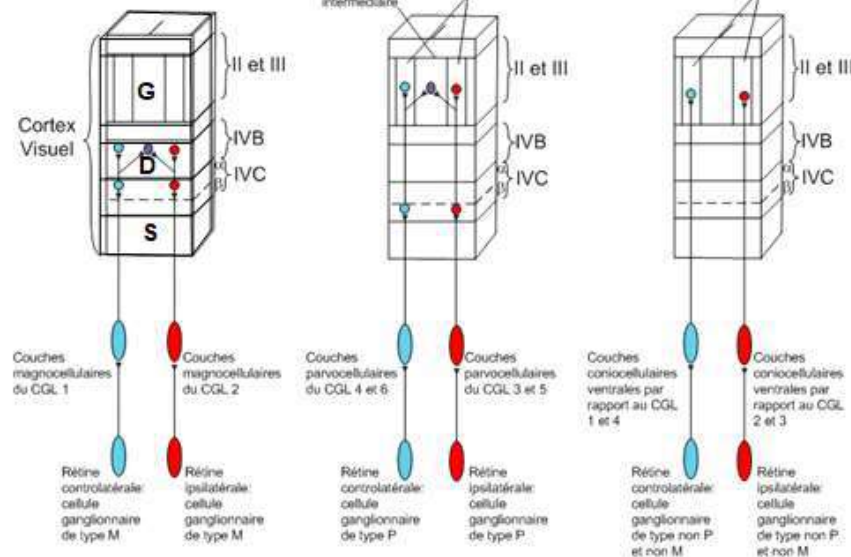
**Modalité GLOBALE**

Détection du mouvement

Détection de la forme

Détection des couleurs

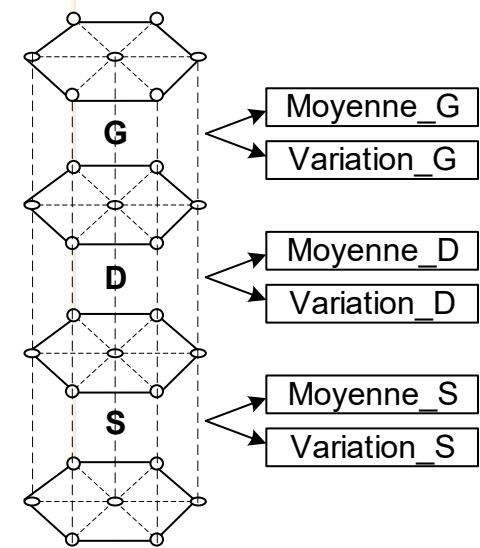
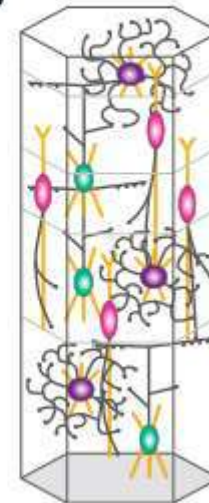
Microcircuit anatomy: (a) Microcircuit dimensions, (b) m-type distribution, and morphology selection



a



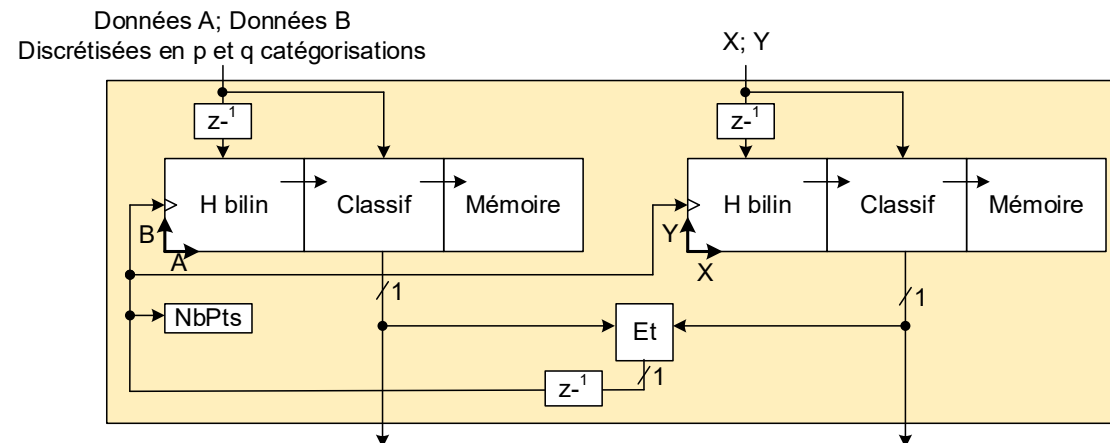
b



► David Hubel and Torsten Wiesel (Nobel Price medicine 1981)

# Règles de Hebb: Plasticité, association, mémorisation

Règles de Hebb		
<p>Plasticité cérébrale</p> <p>La plasticité cérébrale implique deux processus principaux : La <b>croissance</b> neuronale et l'<b>élagage</b> neuronal.</p>	<p>Association spatio-temporelle de deux entités</p> <p>les neurones qui s'<b>activent</b> ensemble, se <b>connectent</b> ensemble</p>	<p>Mémorisation</p> <p>La co-activation d'une assemblée de cellules crée une trace temporaire dans le cerveau qui représente un souvenir en tant qu'unité de mémoire appelé <b>engramme</b>.</p>
Attracteur dynamique à histogrammes bilinéaires		
<p>Plasticité cérébrale</p> <p>Deux groupes distincts de données connectés à un nouveau AD assure la <b>croissance</b> du réseau. L'histogramme bilinéaire de ces deux groupes distincts fait émerger une population locale de plus grande énergie qui définit une nouvelle catégorie, le reste des autres données est <b>élagué</b>.</p>	<p>Association spatio-temporelle de deux entités</p> <p><b>co-activation spatio-temporelle</b> d'une assemblée composée de deux groupes distincts de données catégorisées dont le résultat est l'émergence d'une population locale de plus grande énergie qui définit une <b>nouvelle catégorie</b></p>	<p>Mémorisation</p> <p>La classification du résultat de l'histogramme bilinéaire des deux groupes de données présentés fait émerger une population locale de plus grande énergie qui définit une nouvelle catégorie et représente un souvenir en tant qu'unité de mémoire appelé <b>engramme</b>.</p>



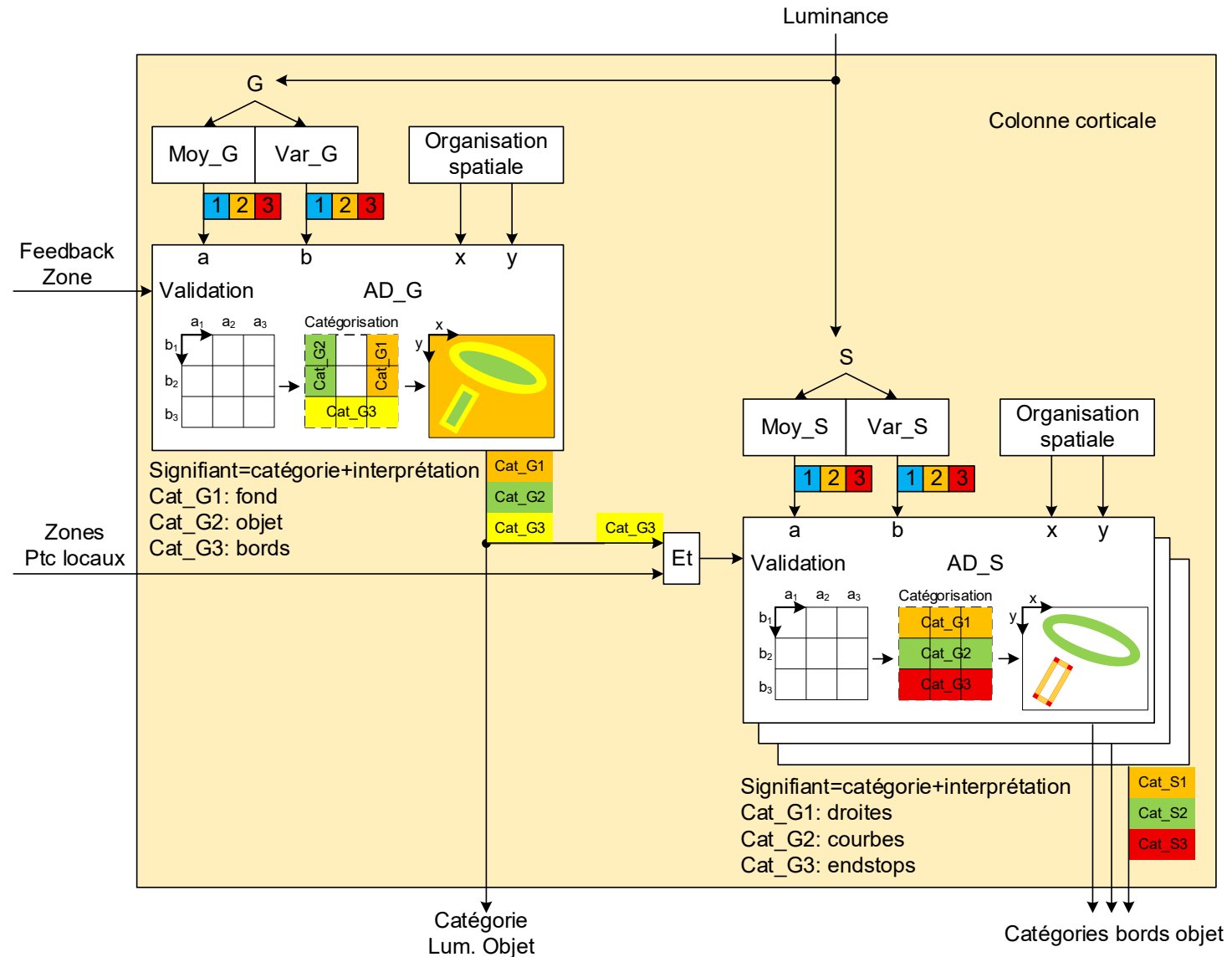
# Organisation partielle de la colonne corticale

1- Période critique:  
application du principe d'ergodicité sur de nombreuses données issues d'une large zone d'analyse.

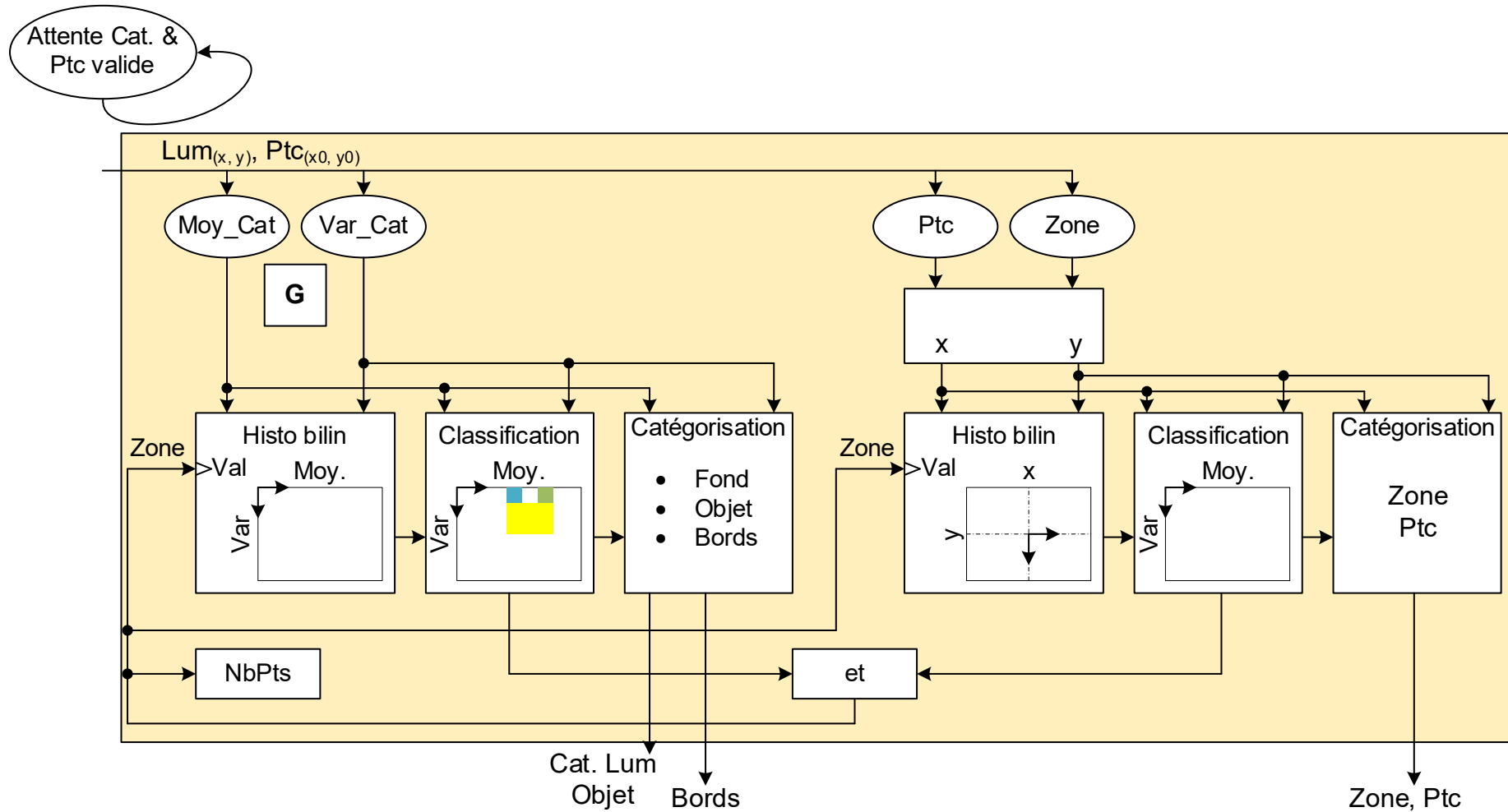
Données perceptives représentées en une carte corticale topique. Démarrage par un nombre réduit de catégories (3) par donnée.

2- Verrouillage apprentissage réseau et passage au traitement suivant.

Le nombre de catégories par données augmente avec l'accroissement de la précision des entrées



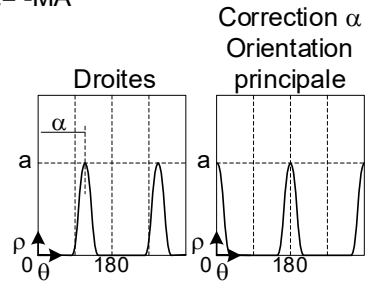
# Détail AD\_G



# Détail AD\_S: catégorisation générique des courbes

## Analyse catégories Droites:

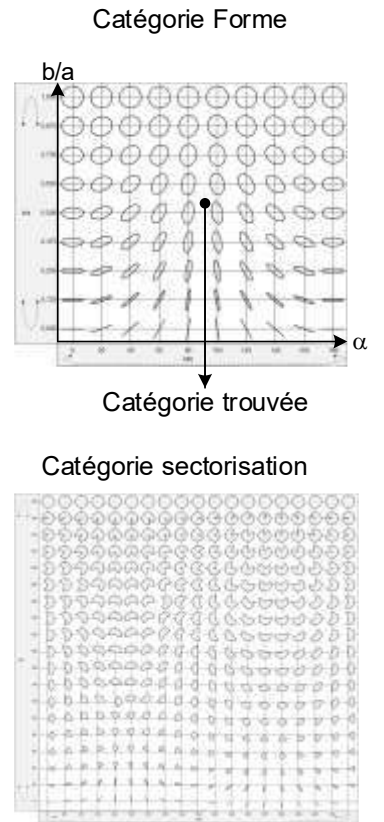
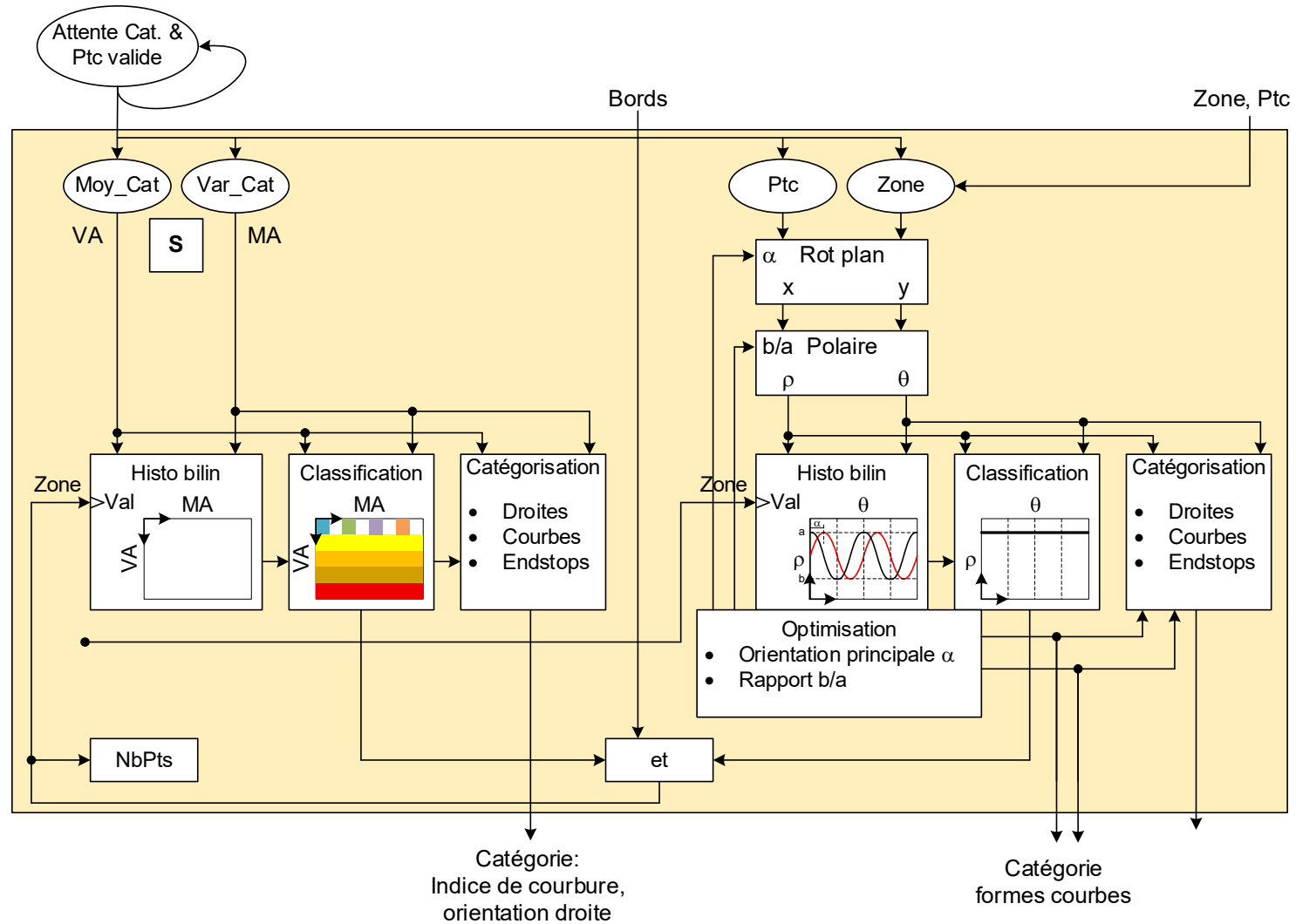
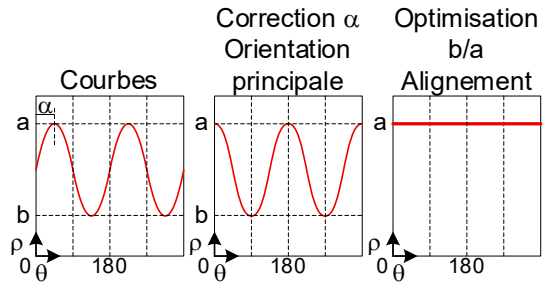
La valeur max MA issue du calcul d'histogramme bilinéaire correspond à l'orientation de la droite. Le module rotation plan exécute une rotation plan de  $\alpha = -MA$



Le module conversion polaire est bypassé. La zone englobe la droite.

## Analyse catégories Courbes:

Adaptation de la valeur  $\alpha$  et  $b/a$  afin d'obtenir une droite dans le plan  $\rho, \theta$ .



# En résumé:

**Foncteur d'Attraction Dynamique (AD)** : L'attracteur dynamique est un foncteur  $F:C \rightarrow H$  qui transforme l'espace topologique local en un espace d'états (les histogrammes bilinéaires  $H_{\text{bilin}}$  cartographiant les relations  $r, q$ ).

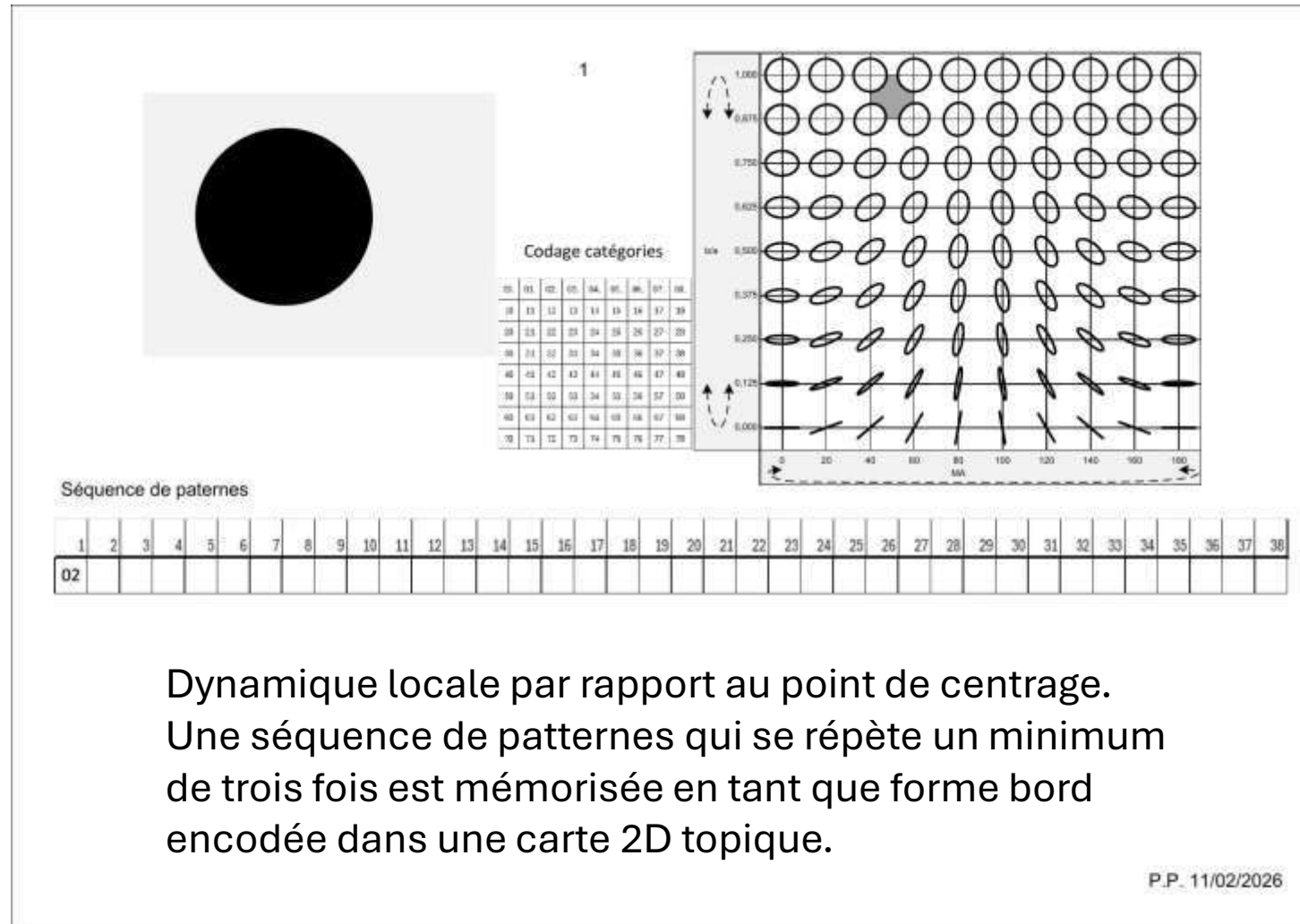
**Classificateur de Sous-objets  $\Omega$  (Mémoire Bijective et Catégorisation)** : La catégorisation mono-coup et la mémoire bijective correspondent au classificateur de sous-objets  $\Omega$  du topos. Chaque nouvelle forme (disque, ellipse, phonème) émerge comme un point dans  $\Omega$ , définissant une valeur de vérité/sémantique univoque ("Représentation sémantique OU et QUOI").

**Mémoire bijective incrémentale, apprentissage non supervisé mono-coup cumulatif** : La mémoire est l'objet classifiant des sous-objets  $\Omega$  du topos (ou un espace étalé). Bijectivité signifie que chaque catégorie correspond à un isomorphisme (ou une application étale) entre descripteur sémantique et instance. L'apprentissage incrémental est la construction inductive (colimite filtrante) : quand une nouvelle configuration ne se recolle pas dans les catégories existantes, on ajoute une nouvelle « carte de classification » (extension du site/topologie). La catégorisation est une flèche caractéristique vers  $\Omega$ . L'abstraction sémantique est le passage aux sections globales invariantes.

**Colonne corticale et modèle unifié** : La colonne corticale (avec plusieurs AD pour Moyenne/Variation par modalité, intégration des cartes  $r/q$ , feedback inter-échelles) est une catégorie interne au topos (ou un diagramme interne). Ses algèbres sont les perceptions cohérentes. Le modèle unifié multimodale est une

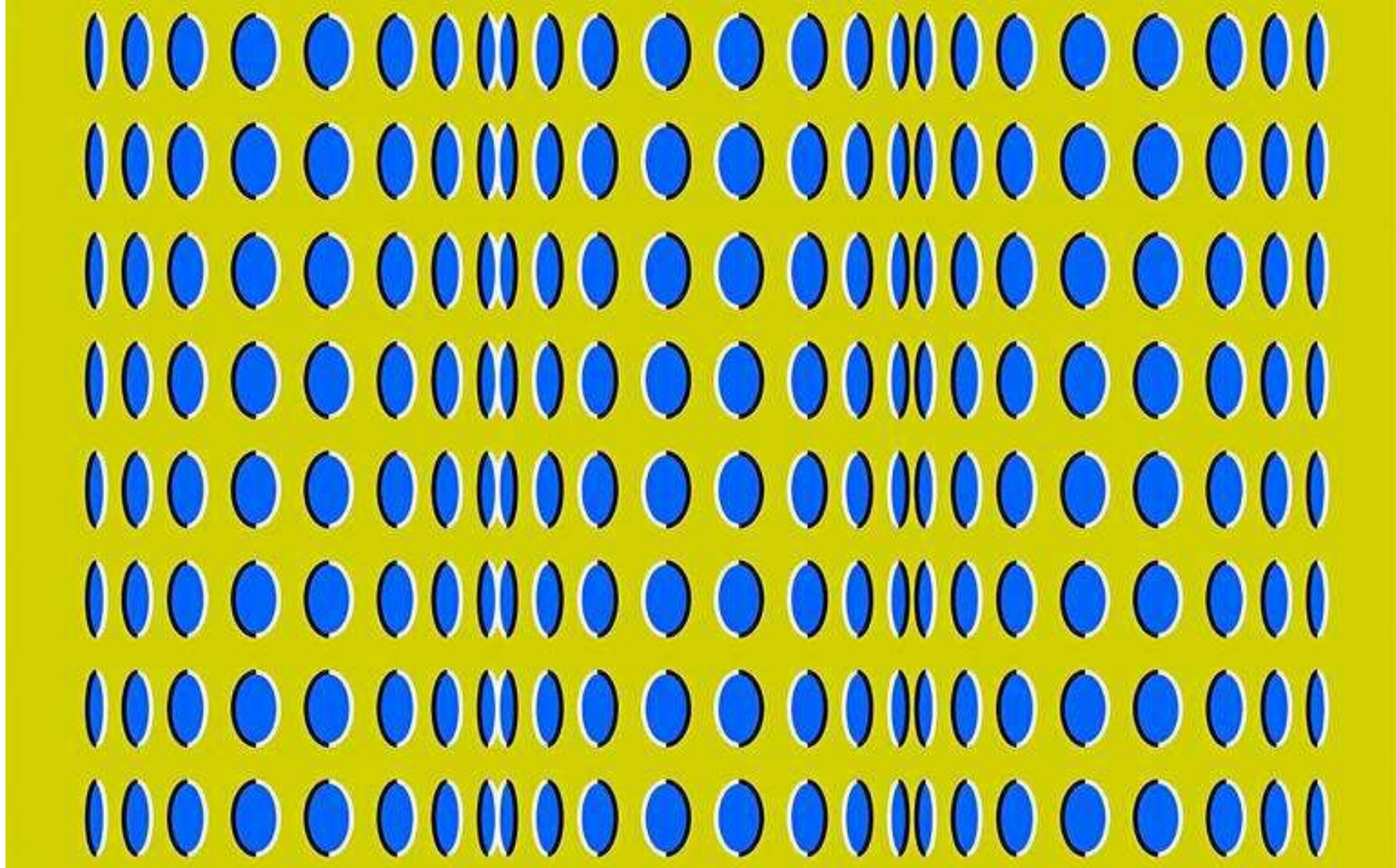
# structuration logique

# Catégorisation dynamique d'une forme simple



Carte mémoire catégorielle des formes simples

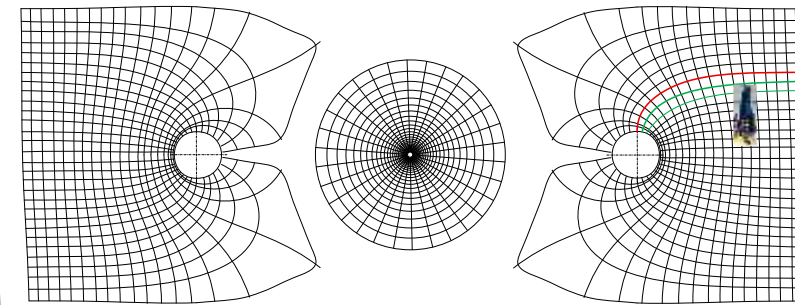
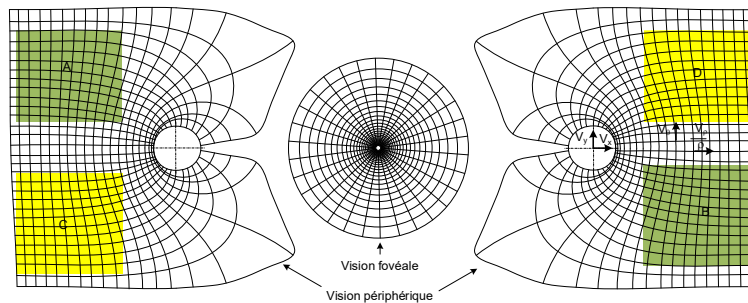
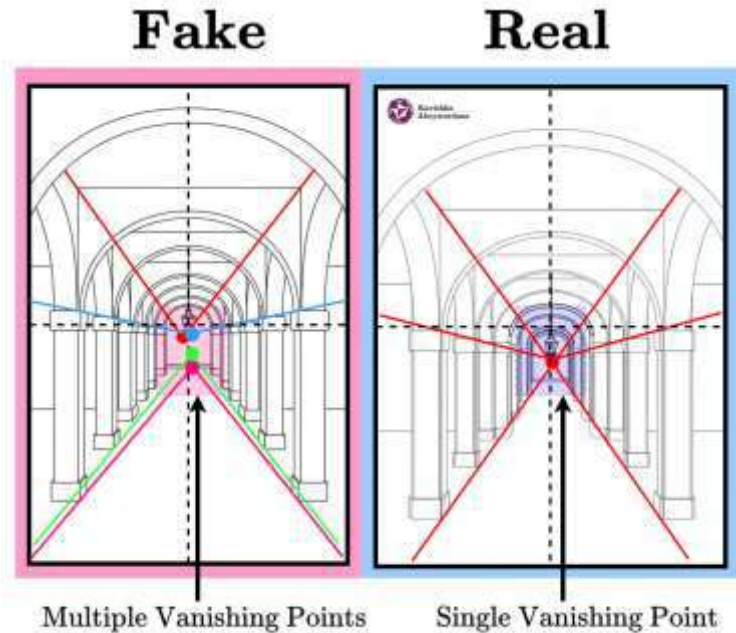
# Rotation induite



# Structuration de la vision fovéale

En vision fovéale le mouvement est constant quelque soit l'excentricité.

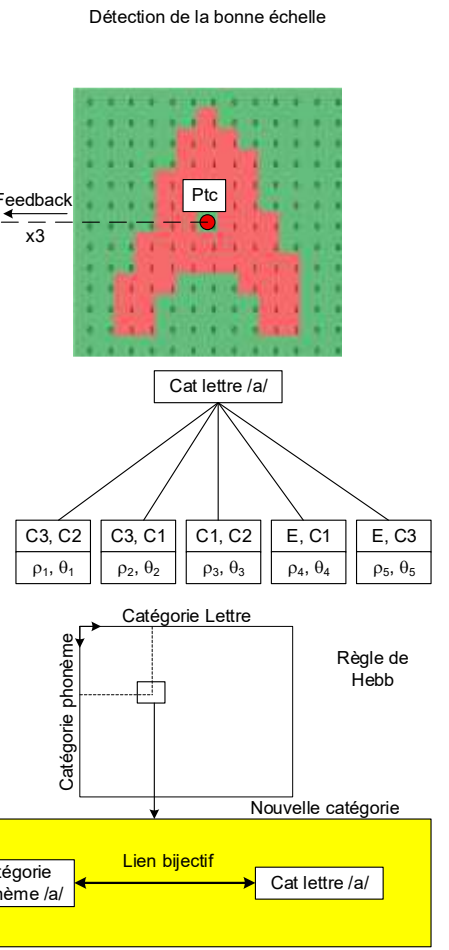
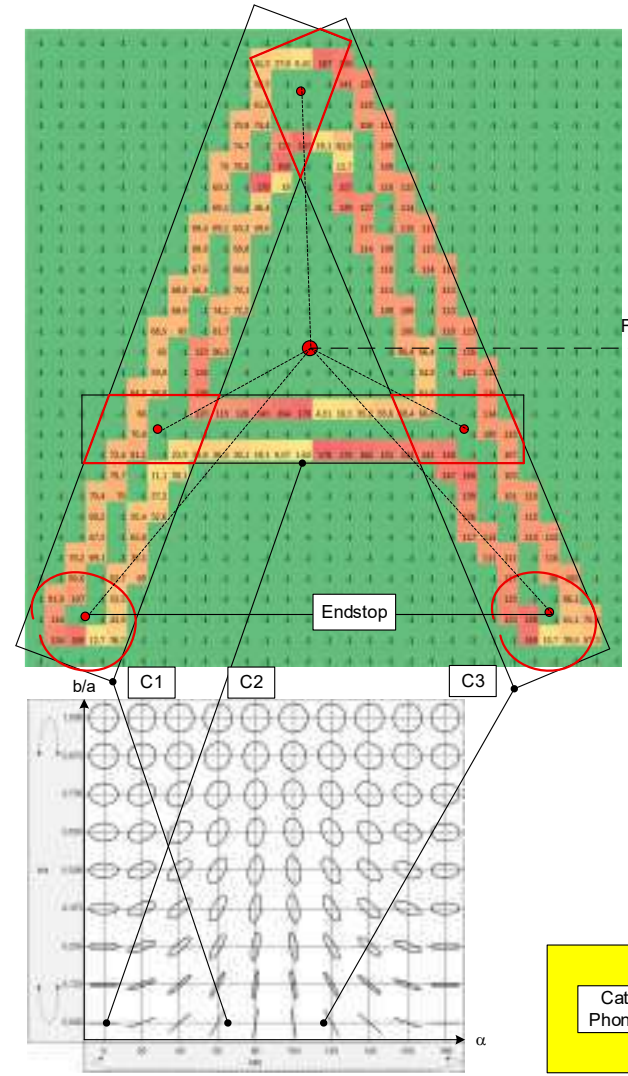
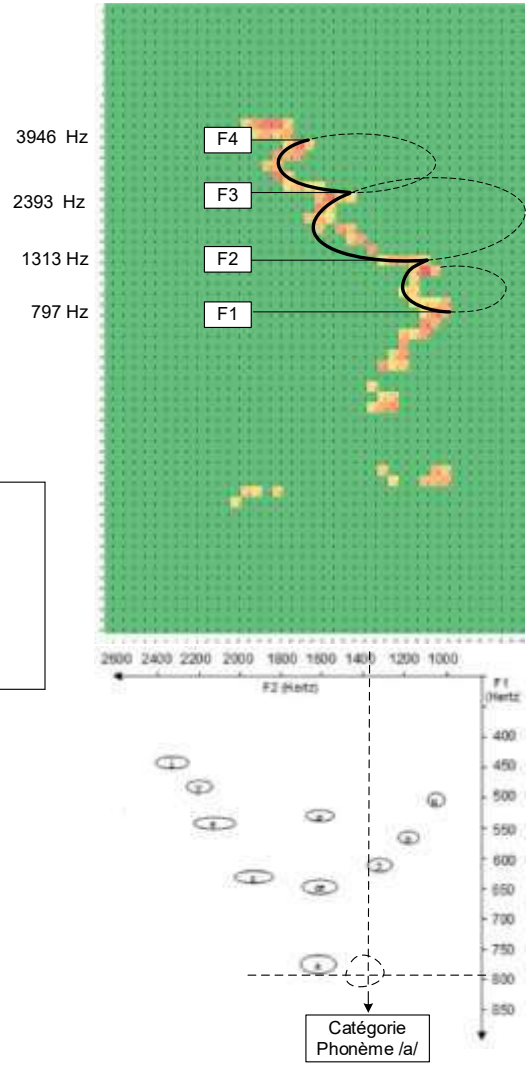
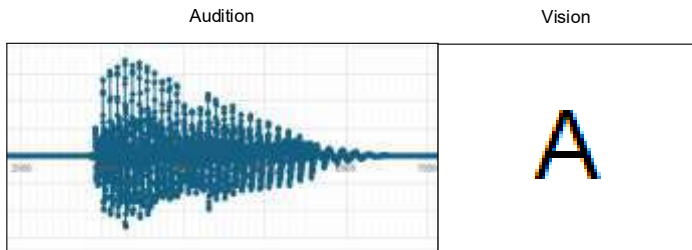
La différence de mouvement entre chaque coté fourni la position du point d'émergence du flot optique.



Lors du croisement d'une personne, son déplacement est perçu sur un rayon horizontal en vert sur la figure. Une position sur le rayon rouge est instantanément perçue comme un contact imminent.

En robotique, c'est un moyen simple d'évitement d'obstacles.

# Association multimodale



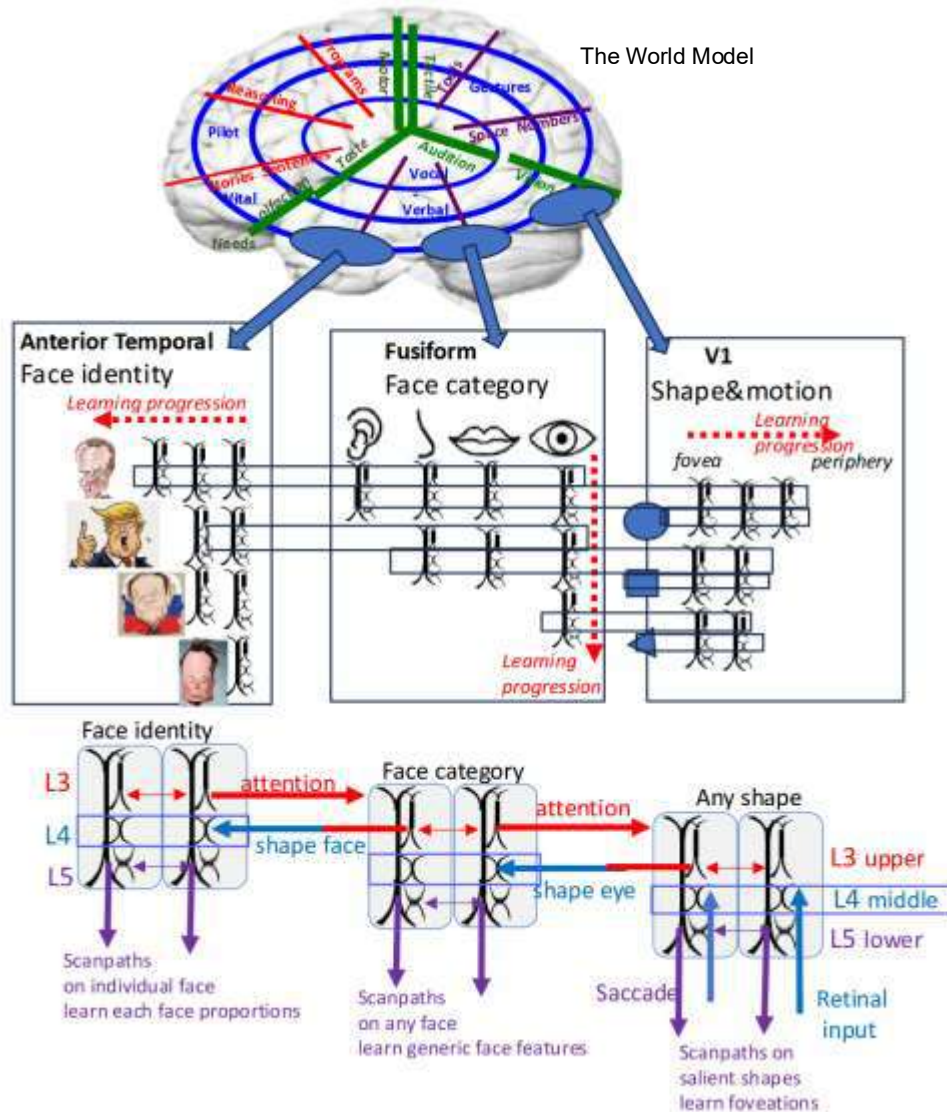
# En résumé:

**Topos Multimodal** : La vision et l'audition forment deux catégories différentes, mais elles sont mappées via des foncteurs vers un Topos Sémantique unifié où les concepts (le "quoi") sont abstraits de leur modalité d'origine. La "Règle de Distance Exponentielle" (EDR) agit comme un foncteur d'attention qui tire en arrière (pullback) l'information la plus saillante vers le centre du faisceau.

**Conscience et émergence** : Le sens émerge du mouvement à travers une variété de faible dimension (texte final), stabilisée par la tâche/contexte. En théorie des topos, cela correspond à des trajectoires dans l'espace des sections globales (ou dans un topos classifiant d'états internes), avec le Yoneda lemma assurant que les relations (contexte) déterminent l'objet (contenu perceptif). Des travaux récents sur les topos de la conscience (catégories de niveaux/contenus, faisceaux pour le recollement multi-agent, topos pour LLM/IA générative) s'alignent parfaitement.<sup>4</sup>

# modélisation du monde

# Identification d'un visage



Bio-model of columns	Similar AI methods
Neural population coding of match and mismatch in L3 Regular distribution of tuning curves	Mismatch information model Prediction Error Orthogonalization
Learn shapes in L4 + saccades L5 under the control of L3 attention	Attention-oriented learning Goal-directed learning
Incremental learning of new faces by progressive classification through successive shapes+saccades	Neural Network Growth Cascade classifiers Face Identity Recognition

# En résumé:

Le système est donc un **topos de faisceaux équivariants multi-échelles et multimodaux**, avec un générateur dynamique de topologies (via apprentissage), un foncteur d'attention (EDR), et des attracteurs dynamiques comme sélecteurs de sections globales sémantiques. Il est frugal, explicable (catégories explicites), et implémentable en hardware (NOC).

# Perspectives

# Implémentation Technique 3D en perspective

**FVLLMONTI : Ferroelectric Vertical Low energy Low latency low volume Modules fOr Neural network Transformers In 3D**

**Compact and low EDP 1-bit adder**

**N<sup>2</sup>C<sup>2</sup> concept**

**Scalable and versatile 3D architectural model leveraging reconfigurable 3D interconnect framework**

**Improving energy efficiency by 10<sup>4</sup>!**

- Physical regularity
- Functional versatility
- In-memory vector processing

**Non-volatile logic cells**

**Versatile VNWFET logic cell based 3D neural network compute cube (N<sup>2</sup>C<sup>2</sup>) for NN-based architecture design**

**Dedicated library of 3D logic cells leveraging VNWFET devices**

**Legend:**

- At the contact: PZT, GAA, 20nm, PZT
- 1-bit adder: PZT, GAA, 20nm, PZT
- Non-volatile logic cells: PZT, GAA, 20nm, PZT
- VNWFET: PZT, GAA, 20nm, PZT

**Grant n°101016776**

<https://fvllmonti.eu/>

<https://www.linkedin.com/in/fvllmonti-project/>

**ATELIER IA POUR LES SCIENCES DE L'INGENIERIE 28-29 JUN 2022 P 4**

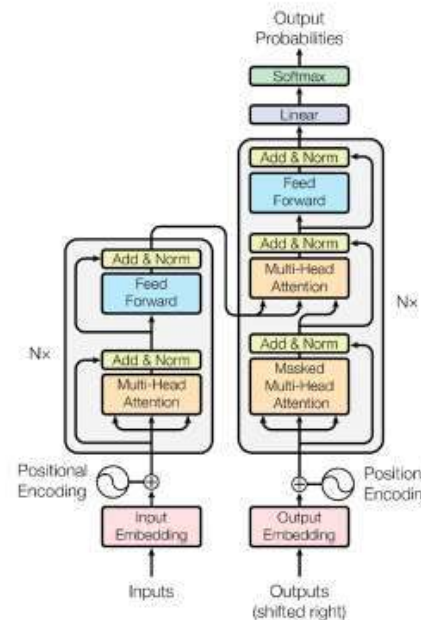
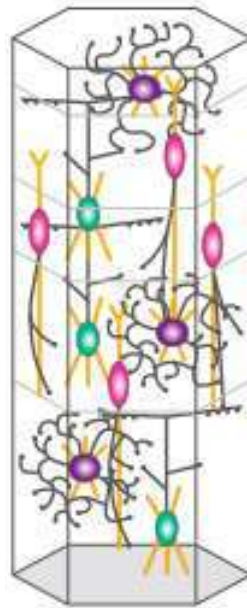
**cnrs**

# Deux paradigmes face à face

## A. L'approche Pirim (Attracteur Dynamique / Topos / Neuromorphique)

Cette approche est **géométrique, topologique et déterministe**. Elle s'inspire du fonctionnement bas niveau de la biologie (rétine, cochlée, aires V1/V2 du cortex).

- **Mécanisme de base** : Extraction mathématique d'invariants (règle de distance exponentielle - EDR, histogrammes bilinéaires). L'information est compressée immédiatement à la source.
- **Apprentissage** : Incrémental, **mono-coup (one-shot)** et non supervisé. Une nouvelle forme crée instantanément une nouvelle entrée dans une "mémoire bijective".
- **Frugalité** : Extrême. Le système a été conçu pour tourner sur des microprocesseurs de très faible puissance (< 30 MHz, 16 KB de mémoire).
- **Explicabilité** : "Boîte blanche" (White box). La mémoire étant bijective et géométrique, on peut tracer mathématiquement pourquoi une forme a été catégorisée ainsi.



## B. L'approche LeCun (IA Physique / V-JEPA / Apprentissage Profond)

L'approche de LeCun vise à donner du "sens commun" aux machines via l'**apprentissage auto-supervisé** et la prédiction dans des espaces latents abstraits.

- **Mécanisme de base** : Les réseaux de neurones (ex: Transformers, ConvNets) encodent l'état du monde. Le système apprend en prédisant ce qui va se passer (ou les parties manquantes d'une image/vidéo) dans un **espace latent abstrait**, en ignorant les détails non pertinents.
- **Apprentissage** : Décence de gradient sur des quantités massives de données (vidéos, images). L'apprentissage est continu mais lent et nécessite de voir de très nombreux exemples.
- **Frugalité** : Faible lors de l'entraînement (nécessite des clusters de GPU massifs), bien que l'inférence puisse être optimisée.
- **Explicabilité** : "Boîte noire" (Black box). Les représentations latentes (vecteurs de milliers de dimensions) sont très difficiles à interpréter pour un humain.

# Tableau comparatif



<b>Caractéristique</b>	<b>Approche Pirim (Attracteur Dynamique)</b>	<b>Approche LeCun (JEPA / IA Physique)</b>
<b>Philosophie</b>	Extraction d'invariants mathématiques.	Apprentissage de représentations par prédiction.
<b>Traitement</b>	Local, multi-échelles, histogrammes.	Global/Hiérarchique, attention, plongements latents.
<b>Besoins en calcul</b>	Minimes (Edge computing, puces dédiées).	Massifs (Datacenters, GPU).
<b>Apprentissage</b>	Mono-coup (One-shot), cumulatif sans oubli.	Statistique, nécessite des millions d'exemples.
<b>Explicabilité</b>	Totale (bijective, mathématiquement traçable).	Faible (intrication des poids neuronaux).
<b>Adaptabilité</b>	Rigide (dépend des filtres codés en dur).	Très flexible (le réseau découvre ses propres filtres).

# Quelle est la Meilleure approche ?

Il n'y a pas de "meilleure" approche absolue ; la réponse dépend entièrement du **cas d'usage**.

**L'approche Pirim est la meilleure pour : L'IA Embarquée Critique (Edge AI)** Si vous concevez le système de vision d'un micro-drone, d'un pacemaker intelligent, d'un missile, ou d'un capteur IoT industriel, l'approche bio-inspirée est supérieure.

**Pourquoi ?** Parce qu'elle ne consomme presque pas d'énergie (frugalité), elle réagit en temps réel strict sans latence, elle apprend sur le terrain en un seul coup d'œil, et elle est explicable (crucial pour les certifications de sécurité).

**L'approche LeCun est la meilleure pour : L'IA Généraliste et la Robotique Complexe** Si vous concevez un robot humanoïde (comme ceux de Boston Dynamics ou Tesla) ou un système de conduite autonome devant naviguer dans des villes chaotiques, l'approche de LeCun est supérieure.

**Pourquoi ?** Le monde réel est trop bruité, ambigu et complexe pour être entièrement modélisé par des équations géométriques codées à la main (les limites de l'IA symbolique/topologique). Les modèles de LeCun, en digérant des années de vidéos, construisent des "modèles du monde" capables de comprendre la physique naïve (ex: la gravité, l'occlusion) et de généraliser face à des situations jamais vues.

## Conclusion

Aujourd'hui, l'industrie et la recherche ont massivement validé **l'approche de LeCun (Deep Learning)** pour sa capacité impressionnante à passer à l'échelle (*scaling*) et à résoudre des problèmes très complexes, malgré son coût énergétique.

Cependant, l'approche de **Pirim (Neuromorphique/Frugale)** représente l'avenir de ce que l'on appelle l'*Edge AI* (l'IA à la périphérie).

La véritable convergence (et le Saint Graal de l'IA) sera très probablement **hybride** : des capteurs neuromorphiques extrêmement rapides et économes en énergie (façon Pirim) qui envoient des concepts prétraités à un cerveau prédictif de haut niveau (façon LeCun) pour la planification complexe.

# Questions?