



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

Explorando Adaptabilidad de Modelos Internos en
Agentes Artificiales

T E S I S

QUE PARA OBTENER EL GRADO DE :

Licenciatura en Ciencias

(Computación)

PRESENTA:

Javier Vargas Martínez

DIRECTOR DE TESIS:

Dr. Bruno Lara Guzmán

SINODALES:

Jorge Hermosillo Valadez (U.A.E.M.)

Alejandra Ciria Fernández-Varela (U.N.A.M.)

Cuernavaca, Morelos a 1 de marzo de 2017

Índice general

CAPÍTULO

1. Introducción	1
1.1. Objetivo	4
1.2. Hipótesis	4
1.3. Robótica cognitiva	4
1.3.1. Breve historia de la Inteligencia Artificial	4
1.3.2. Inteligencia Artificial moderna	9
1.4. Neuronas	10
1.5. Redes Neuronales Artificiales	11
2. Atención visual	17
2.1. Atención artificial	19
3. Campos Neuronales Dinámicos	23
3.1. Categorías de Amari	25
3.2. Modelo CND	26
4. Experimentos y resultados	29
4.1. Instrumentos	29
4.2. Metodología	30
5. Conclusiones	41
Glosario de términos y abreviaciones	45

Índice de figuras

1.1.	Esquema de una neurona	11
1.2.	Red (unidimensional)	12
1.3.	Vecindades de redes (2D)	13
1.4.	Esquema entrenamiento SOM	15
1.5.	Dominios	16
3.1.	Convolución (<i>sin</i> , <i>cos</i>)	27
4.1.	Robot Pioneer 3-DX	29
4.2.	Componentes del sensor visual	30
4.3.	Representación HSV	32
4.4.	Imágenes panorámicas (360°)	32
4.5.	Esquema de un toro	33
4.6.	Kernel gaussiano	34
4.7.	Experimento 1	36
4.8.	Experimento 2	38
4.9.	Experimento 3 (a)	39
4.10.	Experimento 3 (b)	40

RESUMEN

Explorando Adaptabilidad de Modelos Internos en Agentes Artificiales

Javier Vargas Martínez

Los procesos cognitivos se desarrollan continuamente en el tiempo además de poder adaptarse a través del desarrollo. La teoría de campos dinámicos es un marco teórico con una base neuronal —que ha sido útil— para describir el cómo emergen estos procesos cognitivos en un sistema corporeo —también llamados sistemas corporizados— añadase que uno de los métodos para demostrar cómo estos procesos se desarrollan es a través de la implementación de Campos Neuronales Dinámicos (CND) (Dynamic Neural Fields (DNF) por sus siglas en inglés) en robots. En este trabajo se investigó el potencial de los CND para procesar información visual en un agente artificial autónomo. En particular, el interés se centro en la modelación del proceso de rastreo como parte del fenómeno de atención visual. Importa añadir que los ambientes son cambiantes y el modelado de adaptabilidad a estos cambios es un problema abierto en robótica cognitiva.

Como plataforma de prueba se utilizó un robot Pioneer 3-DX con una cámara omnidireccional montada sobre este robot. Esta cámara logra obtener imágenes panorámicas de 360° con la ayuda de un espejo con forma esferoide. Se implementó un CND para procesar información de un canal de las imágenes de 360° obtenidas por el agente en su medio ambiente. La información visual fue procesada con distintos parámetros que definen a estos campos, con la finalidad de obtener valores de estos parámetros, que podrían facultar la capacidad del proceso de rastreo.

CAPÍTULO 1

Introducción

El trabajo aquí presentado pertenece al área de Inteligencia Artificial (IA) y se encuentra enfocado en el vínculo de dos áreas importantes. La primera es visión por computadora, área que intenta interpretar imágenes a través de extraer información de éstas (por ejemplo, identificar y ubicar objetos de interés en el ambiente). Extraer información de las imágenes se hace por dos principales objetivos. El primero es que una computadora “entienda” los elementos en la escena, el segundo objetivo es que las imágenes procesadas sean asequibles para un humano. La segunda área es Redes Neuronales Artificiales (RNA), la cual contiene modelos matemáticos inspirados en el funcionamiento de las neuronas biológicas, hay que mencionar además que estos modelos se utilizan como una herramienta de cómputo en distintas áreas y no sólo en IA. Entre las propiedades de estos modelos se encuentran la habilidad de aprender, clasificar y representar información además de cualidades tales como organizar y agrupar datos.

Los modelos de descripción de la actividad neuronal tanto por unidad así como los que describen regiones corticales, han sido abordados desde dos enfoques:

- Dinámica de interacción entre poblaciones de neuronas, también conocidos como Campos Neuronales (e.g. CND también conocido como modelo Wilson-Cowan, *Conditional Neural Fields* (Peng et al., 2009))
- Relación eléctrica (e.g. *integrate and fire* (Brette y Gerstner, 2005; Burkitt, 2006a), *spiking neurons* (Izhikevich et al., 2003; Brunel, 2000), *bursting neurons* (Zhao et al., 2014), modelo Hodgkin-Huxley —capacitancia— (Hodgkin y Huxley, 1952b,a))

Los Campos Neuronales Dinámicos (CND) se encuentran dentro de los modelos que intentan describir actividad neuronal (Wilson y Cowan, 1972, 1973). La teoría de campos dinámicos ha sido útil en describir el cómo emergen procesos cognitivos en un sistema corporeo, además es necesario recalcar que los procesos cognitivos se desarrollan continuamente en el tiempo. Una de las características principales de los CND es su alta

capacidad de respuesta a cambios en el ambiente, lo cual sigue siendo un problema abierto en el área de robótica.

Es conocido que los ambientes no son siempre estáticos y se considera importante modelar un sistema con facultad de adaptarse a condiciones dinámicas. Los cambios en el ambiente, percibidos por un agente son producto de dos fuentes:

- Causas exógenas que modifican el ambiente
- Acciones propias del agente sobre ese ambiente

Uno de los grandes retos dentro de la robótica cognitiva es simular, en ambientes dinámicos, el fenómeno de atención visual, este fenómeno involucra un conjunto de procesos importantes para su funcionamiento. El estudio de atención se ha dividido en distintas clases, por ejemplo, en correspondencia a la orientación del procesamiento de información. La dicotomía del fenómeno de atención más conocida es:

- *top-down* está relacionada al procesamiento de la información que involucra hacer inferencias sobre el mundo en base a la experiencia previa
- *bottom-up* es el procesamiento directo de la información captada por los sensores

En particular, el fenómeno de atención visual se puede separar en dos mecanismos diferentes, el primer mecanismo necesita de movimiento ocular para mantener la ubicación de los objetos relevantes en escena y es conocido por el nombre de *overt attention*, el segundo mecanismo permite detectar la ubicación de un estímulo relevante sin utilizar movimiento ocular y es conocido por el nombre de *covert attention*.

Algunas de las investigaciones que se realizaron con la intención de crear un modelo de atención visual (Hou y Zhang, 2007; Itti y Koch, 2000, 2001) utilizan procesos de adquisición de características relevantes derivados de la IA tradicional e inclusive cuando algunas de estas investigaciones toman en cuenta algunos aspectos neuronales, se emplean modelos neuronales que describen actividad por unidad (Izhikevich et al., 2003) pero es necesario tener en cuenta que para definir la localización relevante en una escena visual es importante valorar la relación que tienen las distintas posiciones en la escena además de no ocupar procesamientos que calculen características relevantes. Es por lo que se lleva dicho hasta aquí que el emplear el marco teórico de Campos Neuronales Dinámicos se considera podría resultar de

gran utilidad para investigar el fenómeno de atención visual, en robótica cognitiva, a través de utilizar esta teoría en un robot como plataforma de prueba. Se consideró que un excelente punto de partida es modelar *covert attention* por ser uno de los mecanismos más sencillos además de que este mecanismo podría proveer de la base para mecanismos más complejos (e.g. *overt attention*).

En particular, el interés se centró en la modelación del proceso de rastreo, proceso fundamental en el fenómeno de atención visual. En este trabajo la palabra rastreo está referida a la cualidad de mantener una región neuronal activa, correspondiente a los objetos (salientes) en escena, sea que estén estáticos e inclusive cuando estos objetos esten en movimiento. Agregando a lo anterior la posible presencia de ruido (es la contaminación de la señal de la imagen, por ejemplo información que no pertenece a la entrada), ya sea por consecuencia del ambiente o del sensor visual. Este trabajo se centra en los cambios percibidos que son producto de motivos exógenos.

El robot utilizado para realizar este trabajo es un agente artificial autónomo Pioneer 3-DX dotado de visión panorámica de 360° (ver figura 4.1).

En el capítulo 1 se encontrará una sinopsis sobre neuronas, robótica cognitiva, además de una exposición de algunos métodos para modelar propiedades del cerebro de los mamíferos.

En el capítulo 2 se hablará sobre el fenómeno de atención en el caso específico de la visión. Se describirán algunos procesos fundamentales de este fenómeno asimismo de peculiaridades de estos procesos y su relevancia en la robótica cognitiva. También se mencionarán algunos modelos de atención visual artificial. Estos modelos pueden tener fundamento biológico, ya sea a través de observaciones de la dinámica de activación neuronal tanto como de observaciones de la respuesta cognitiva.

En el capítulo 3 se dará una descripción del marco teórico que define a los CND. También se podrá observar algunos de los estudios que precedieron a los CND. Asimismo se explicaran categorías de estos campos además de los elementos que componen a los CND.

En el capítulo 4 se mostrarán los elementos —tanto de instrumentación como teóricos— requeridos para realizar este trabajo, asimismo de los experimentos realizados. A esto se añaden los resultados obtenidos junto con sus respectivos detalles, también se pormenorizará la metodología utilizada para realizar los experimentos descritos en esta investigación, además de cómo se utilizó la teoría de CND para la implementación de un sistema de procesamiento de información visual.

Finalmente en el capítulo 5 se presentarán las conclusiones a las que se

llegaron después de analizar los datos obtenidos en los resultados. Asimismo se expondrán algunas de las posibles repercusiones en el área de Robótica Cognitiva. Al final del capítulo se darán algunas sugerencias para posible trabajo futuro.

1.1. Objetivo

Implementar un CND en un robot Pioneer 3-DX que procese información visual con la intención de simular la propiedad de rastreo del fenómeno de atención visual.

Objetivos Particulares

- Programar un CND homogéneo unicapa de dos dimensiones
- Implementar un método de adquisición de imágenes 360° para el robot Pioneer 3-DX
- Observar el comportamiento de los parámetros que definen a los CND
- Analizar los resultados de las simulaciones obtenidas

1.2. Hipótesis

El marco teórico de CND será de utilidad para simular el proceso de rastreo del fenómeno de atención visual.

1.3. Robótica cognitiva

1.3.1. Breve historia de la Inteligencia Artificial

En la tentativa de simular el pensamiento humano, las matemáticas a través de lógica, probabilidad, ecuaciones diferenciales y otras ramas de ésta, han logrado imitar capacidades que requieren de “destreza mental” (e.g. jugar ajedrez, *backgammon*, *go*). A estas aproximaciones se les llamó Inteligencia Artificial alrededor de los años 50.

Los primeros acercamientos que hubo hacia definición y desarrollo de la IA, fueron explicados por Turing (1950) que propuso un método para “reconocer” pensamiento —mediante un ejercicio mental— al que llamó el *Juego de la Imitación*. En este método Turing (1950) propone que si un humano no pudiese distinguir las respuestas de una máquina de las de un humano entonces se habría encontrado *pensamiento* en máquinas.

En el juego de la imitación participan tres personas, un hombre, una mujer y un interrogador —que puede ser de cualquier sexo— que permanece separado de las primeras dos personas. Este interrogador puede hacer preguntas a estas personas con el objetivo de determinar quién es el hombre

y quién la mujer —tal que exclusivamente sabe que conversa con ‘X’ o bien con ‘Y’. El objetivo de la mujer es ayudar al interrogador a reconocerle como la mujer, mientras que el del hombre es engañarlo para no ser reconocido. La forma de comunicación es exclusivamente a través de escritura impresa (para evitar pistas tales como voz, forma de escritura u otras). Ciertamente una de las características de los humanos es el de poder “engañar”, por esta razón es probable que el interrogador pueda otorgar un juicio incorrecto sobre la identidad de las personas (Turing, 1950).

Turing (1950) propuso que considerar la pregunta, ¿Pueden las máquinas pensar?, tenía la limitación de tener que definir “máquina” y “pensamiento”, por lo tanto cambió su pregunta por otras dos. La primera, ¿Qué pasaría si una máquina tomara el lugar del hombre (en el juego de la imitación)?. La segunda, ¿Se equivocará el interrogador en el juego cuando consiste de un hombre y una mujer?. Conviene subrayar que las máquinas permitidas en el juego son sólo computadoras digitales. Es significativo que si se quisiera crear una máquina capaz de imitar el comportamiento de la “computadora humana”, hemos de realizar hacer un programa que describa las instrucciones en el orden necesario que explique cómo ejecutarse (Turing, 1950).

Hemos de realzar que pensar que la creación de programas competentes en la ejecución de tareas cognitivas (e.g. procesamiento de lenguaje) es suficiente para lograr construir sistemas “inteligentes”, tuvo repercusión en parte de la investigación sobre IA de la época. Una de estas repercusiones fue que algunas investigaciones se concentraron en identificar tareas que requieren de “inteligencia” para después solucionarles a través de construir un programa que resolvieran estas tareas (Newell y Simon, 1976). A esta forma de analizar la cognición se le conoce como Teoría Computacionalista.

En principio (Newell y Simon, 1976) consideraban que los símbolos se encontraban en la raíz de las “acciones inteligentes” además que uno de los requerimientos para estas acciones era la capacidad de almacenar y manipular símbolos. Estos autores tomaron parte del trabajo de Turing (1936) sobre la Máquina Universal (hoy en día es conocida como Máquina de Turing) para definir un posible modelo de procesamiento de símbolos. Por una parte, la teoría que describe a esta máquina generó una significativa importancia en Ciencias de la Computación además de un gran repercusión en la Teoría de la Computación, de aquí que resultó en lo que ahora conocemos como computadoras digitales. Por otra parte, esta teoría describe y define aquello que esta máquina —igualmente que su homóloga moderna (computadora digital)— puede realizar además de lo que no (Turing, 1936). Habría que decir también que Turing (1936) y Church (1936) llegaron a

resultados semejantes en sus respectivos trabajos por lo tanto hoy en día esta teoría es conocida y mencionada la Tesis Church-Turing (Sipser, 1997).

Newell y Simon (1976) llamaron a su sistema de procesamiento de símbolos, “sistema físico de símbolos” (*Physical Symbol System*) —el cual es una instancia de la Máquina universal— y lo complementaron con búsqueda heurística. Por un lado estos autores habían supuesto que resolver problemas era un indicador de inteligencia en los sistemas, también que el sistema físico de símbolos proporcionaba la base de “acción inteligente” pero no describía cómo. Por otro lado, pensaron que los sistemas de símbolos resuelven problemas con el proceso de búsqueda heurística, por lo tanto esa era la forma de describir el cómo de la “acción inteligente”. La búsqueda heurística permite encontrar solución a un problema sin la necesidad de buscar en todo el espacio de soluciones. La restringida capacidad de procesamiento de los sistemas físicos de símbolos —deriva de de la necesidad de responder en un intervalo de tiempo restringido además de tener sólo una cantidad finita de pasos para entregar una respuesta— causo que consideraran fundamentalmente complementar este sistema con búsqueda heurística (Newell y Simon, 1976).

A las aproximaciones de Inteligencia Artificial con un enfoque computacionalista ahora son denominadas Good Old-Fashioned Artificial Intelligence (GOFAI) (También llamada IA tradicional) (Haugeland, 1989). Para ilustrar mejor a que es referida la IA tradicional tomaremos por ejemplo un evento ocurrido en 1997. En este año hubo un encuentro de ajedrez entre una máquina y un humano. La empresa IBM creó un máquina para jugar ajedrez llamada *Deep Blue*. Esta máquina derrotó al campeón mundial de ajedrez Garry Kasparov en un encuentro de 6 juegos. Cabe mencionar que en esta máquina estaban programadas estrategias y reglas del juego —además tenía la capacidad de probar cientos de miles posibles soluciones (e.g. permutaciones factibles del tablero) en un periodo corto de tiempo (Campbell et al., 2002). Haciendo una analogía maquina-humano, la máquina *Deep Blue* tiene un procesador-cerebro para procesar la información del juego además de un programa-mente con las reglas y estrategias. Bien, pareciera por todo lo anterior que la máquina es un sistema “inteligente”, lo cual es cierto para la IA tradicional. Esta máquina sí es un sistema “inteligente” de acuerdo con el hecho que simula capacidades cognitivas (e.g. planificación) al manipular símbolos (e.g. posiciones factibles del tablero) para encontrar soluciones. Un análisis sobre los factores que pudieron contribuir a la victoria de *Deep Blue* sobre Kasparov se encuentra revisada por Campbell et al. (2002).

En el transcurso del tiempo algunos investigadores en IA se dieron cuenta

de problemáticas relacionadas a la solución de problemas (e.g. jugar ajedrez) a través de la manipulación de símbolos. Por ejemplo, aunque a través de manipulación de símbolos se puedan simular capacidades cognitivas del pensamiento humano, no son condición necesaria para demostrar que saben aquello que manipulan.

Una de las primeras críticas a esta IA tradicional es la de Searle (1980), en esa se contrastan las afirmaciones hechas por la IA tradicional —por ejemplo, un programa (computacional) puede explicar el pensamiento— a través del experimento mental llamado *Cuarto chino*. Este experimento muestra que los sistemas basados únicamente en procesamiento de símbolos carecen de Intencionalidad pese a que puedan imitar capacidades del pensamiento. Por intencionalidad nos referimos a la propiedad de los estados mentales para referirse o dar representación a cosas, propiedades y estados en el mundo —sean existentes o no— (Searle, 1983; Houdé et al., 2003) y no debe ser confundido con el significado de la palabra intención. En otras palabras es la capacidad de la mente de abstraer conceptos, por ejemplo un perro, casa, o un personaje ficticio (Jacob, 2014).

En el experimento del *cuarto chino* (Searle, 1980) se describe a una persona que se encuentra encerrada en un cuarto. Esta persona conoce el idioma inglés pero no conoce nada del idioma chino (ya sea hablado o escrito). A esta persona se le entregan textos en chino —guión (historia) y un cuestionario. Dentro del cuarto se encuentran dos textos en inglés, el primero contiene reglas que permiten relacionar los textos en chino (historia y cuestionario), mientras que el segundo, reglas necesarias para facultarle la posibilidad de devolver símbolos en chino cómo respuesta a los símbolos que recibió. Al conjunto de reglas en inglés se le llama “programa”, mientras que a los símbolos en chino que devuelve “respuestas”. En pocas palabras, la persona recibe texto y preguntas en chino, además debe responder en chino. En otro caso, se le entrega un guión y preguntas sobre esa historia —ambas en inglés— y debe responder en inglés. En efecto, sus respuestas en inglés serían indistinguibles para una persona que sabe inglés puesto que la persona dentro del cuarto sabe inglés. Además supone que la persona dentro del cuarto después de un tiempo se vuelve hábil en el procesamiento de símbolos en chino —al hacer uso de las reglas de manipulación de símbolos inglés-chino para responder símbolos en chino— tanto como los programadores para diseñar las reglas (“programa”), tal que las respuestas entregadas por la persona dentro del cuarto son indistinguibles para una persona que sabe chino (por ejemplo, una persona tal que su lengua materna sea chino). Es decir, las respuestas —ya sea en inglés o en chino— de la persona dentro del cuarto son “buenas” tanto en inglés como en chino

(Searle, 1980).

En conclusión, cuando la persona dentro del cuarto recibe y responde en inglés sabe de lo que trata el texto, además sabe lo que se le pregunta, mientras que en el caso cuando recibe texto y preguntas en chino, no tiene conocimiento de nada de lo que significa así tampoco sobre las respuestas que entrega.

Aplicando el análisis de Searle (1980) en el ejemplo sobre Deep Blue vs Kasparov nos lleva a pensar que la máquina Deep Blue no conoce el juego de ajedrez, ni cómo jugarlo. En otras palabras, es producto de sus capacidades de computo tal que es eficiente para calcular las posibles permutaciones en el tablero de juego —esto es, utilizar las reglas definidas por el programa— y entregar un resultado.

Este ejemplo y otros más permitieron esclarecer que asemejar al pensamiento con la habilidad de manipular símbolos (e.g. letras, palabras, 0's y 1's) aun cuando se puedan simular capacidades cognitivas humanas, no se logra Intencionalidad.

En (Harnad, 1990) se destaca que uno de los problemas de pensar que la mente es un sistema de símbolos y la cognición es manipulación de símbolos, reside —cómo había descrito (Searle, 1980)— en la carencia de significado intrínseco.

Los sistemas formales de símbolos (Newell y Simon, 1976) basados en manipulación de símbolos y no en semántica, no pueden ser facultados de tener significado. Si el significado de estos es externo entonces no pueden considerarse como un modelo factible para intencionalidad (i.e. la cognición no puede ser sólo manipulación de símbolos) (Searle, 1980; Harnad, 1990).

En la crítica de Harnad (1990) describe al “Problema de Cimentación de Símbolos” (*Symbol Grounding Problem*) a través de un ejemplo que consta de dos situaciones. En la primera, una persona —cómo en el cuarto chino de Searle (1980), la persona sabe inglés pero nada de chino— debe aprender el idioma chino, con sólo un diccionario de chino-chino. En la segunda, debe aprender chino como primera lengua —esto es, no conoce idioma alguno previamente— con sólo este diccionario. Ambas situaciones se caracterizan por no poder dar un significado a los símbolos en el diccionario chino-chino, por la razón de no poder darles un significado en el mundo (lo que se describe como un bucle entre símbolos), esto deja la pregunta sobre cómo dar representación a los símbolos del diccionario. Para la primera situación analiza que los criptólogos (al carecer de reglas) de decifrar códigos además de lenguas antiguas, es posible por la razón de poder cimentar estos símbolos en un idioma previo además de tener experiencia y conocimiento en el mundo. El problema de la cimentación de símbolos es en otras palabras,

la improbable posibilidad de salir del bucle entre símbolos (que no tienen sentido) (Harnad, 1990).

1.3.2. Inteligencia Artificial moderna

La Robótica Cognitiva surge como una IA moderna, por una parte en respuesta a las críticas hechas hacia la IA tradicional, mientras que por otra parte toma en cuenta teorías de las Ciencias Cognitivas —estudio interdisciplinario que le conciernen la Psicología, Filosofía, Inteligencia Artificial, Neurociencias, Lingüística y la Antropología (Thagard, 2014)— acogiendo un enfoque interdisciplinario. Ciertamente uno de los objetivos de la robótica cognitiva es construir agentes artificiales que, a través de interactuar en el ambiente puedan mostrar un comportamiento coherente.

Es importante mencionar que la robótica cognitiva no se centra en emular cómo funcionan las capacidades cognitivas —para más tarde implementarlas— si no que se basan en estudios provenientes de distantes áreas (e.g. ciencias cognitivas, biología) para lograr desarrollar máquinas capaces de interactuar en el mundo de forma congruente. Una razón de utilizar robots es en parte porqué algunas teorías —tomadas por la robótica cognitiva— que describen cómo emerge la cognición consideran fundamental la interacción del cuerpo en el ambiente —algunas de las teorías que sustentan esta afirmación son la de Cognición Cimentada (Barsalou, 2008) además de Cognición Corporizada (Wilson, 2002).

Cabe subrayar que, si las máquinas interactúan en el mundo a través de sensores (e.g. sonares, cámaras) que obtengan información en un ambiente dinámico —como es el mundo— (Brooks, 1990) además de actuadores para modificar su ambiente, es probable que el problema de cimentación de símbolos (Harnad, 1990) se pueda resolver. Es relevante mencionar que los cambios sensoriales percibidos no son únicamente por consecuencias exógenas (i.e. que provienen del ambiente), son también secuela de los cambios motrices.

A pesar de dotar de cuerpo —que pueda sentir e interactuar en el mundo— a una máquina no es condición suficiente para resolver el problema de cimentación de símbolos pero nos permite explorar el desarrollo de procesos cognitivos a través de interacción cuerpo-mundo. Un análisis sobre cognición corporizada en robótica cognitiva se encuentra en (Ziemke, 2016), en este trabajo encontramos un análisis sobre investigaciones en esta área además describe algunos aspectos que no se han tomado en cuenta —no se entrará en detalle a esta discusión— por ejemplo la base biológica.

Las RNA han sido una herramienta de gran utilidad en robótica cognitiva para poder procesar información tanto interoceptiva (proviene del cuerpo) como exteroceptiva (información del ambiente) en agentes artificiales. Las redes neuronales de autoorganización, por ejemplo los Self-organizing Map (SOM) tienen una arquitectura basada en la respuesta cortical hacia estímulos, tal que estímulos semejantes tienden a excitar zonas cercanas entre sí —consecuencia de una etapa de entrenamiento que organiza (codifica) las señales con mayor similitud espacialmente más cercanas entre ellas— además de ser capaz de organizar en forma de mapa las entradas a esta red (Kohonen, 1982).

En pocas palabras la robótica cognitiva se encarga de explicar fenómenos —como aprendizaje y movimiento— sin tener una descripción concreta (programa) de cómo actuar en relación a la información de entrada (exteroceptiva).

1.4. Neuronas

El sistema nervioso es sin duda uno de los principales elementos situado en la base del funcionamiento de las funciones cognitivas (e.g. control motriz, procesamiento de información exteroceptiva). Este sistema está constituido principalmente de neuronas, las cuales son células encargadas de distintas funciones, por ejemplo de procesar información así como de transmitirla hacia otras estructuras (Wilson y Keil, 2001).

Las neuronas son células (ver figura 1.1) que están constituidas por varios componentes. Los tres principales componentes del cuerpo celular (soma) son, dendritas, axón y núcleo. Las dendritas son un canal de comunicación de entrada mientras que el axón funciona para mandar información. La forma de comunicación de estas células deriva del enlace entre neuronas, que se origina en parte por una señal eléctrica que viaja por el axón de una neurona presináptica hacia una neurona postsináptica (Gross, 2015). Las señales eléctricas generadas por las neuronas —que son constituidas por impulsos eléctricos— son llamadas potenciales de acción. Esta señal tiene como consecuencia el liberado de una sustancia llamado neurotransmisor (e.g. serotonina, dopamina). Estos neurotransmisores tiene dos principales funciones, excitación e inhibición (Friedenberg y Silverman, 2011; Cacioppo et al., 2007). El núcleo tiene el objetivo de captar todas las señales de entrada para después transmitir otra señal al axón. Las neuronas —son células excitables— que hacen uso de los canales iónicos para mantener en balance el potencial de membrana (Diferencia entre el potencial

eléctrico entre el interior y el exterior). En el potencial de membrana interactúan sodio (Na), potasio (K), cloro (Cl), entre otros componentes. La variable concentración de estos componentes provoca cambios en las propiedades eléctricas (e.g. corriente, impedancia, conductancia) de estas células (Hodgkin y Horowicz, 1959; Hodgkin y Huxley, 1952b).

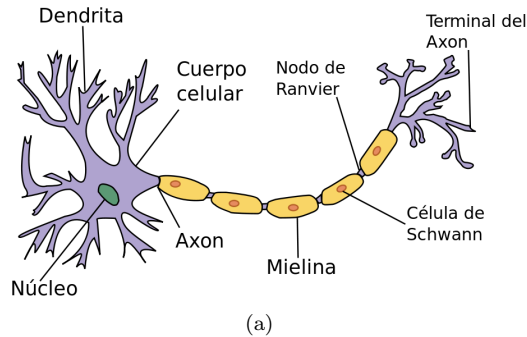


Figura 1.1: Esquema de los principales componentes de una neurona. Imágen obtenida de www.wikimedia.org, el archivo se encuentra bajo la licencia *Creative Commons Attribution-Share Alike 3.0*

Es importante mencionar que en las funciones cognitivas —tanto de bajo como de alto nivel— se ven involucradas un conjunto de neuronas que se encuentran conectadas en redes complejas Wilson y Cowan (1972), razón que nos llevaría a pensar que la función de las neuronas como unidad no sea de gran importancia. Sin embargo su estudio como individuos ha aportado conocimiento importante en cuanto a su comportamiento.

Estas células son bastante importantes en el estudio de RNA dado que son el componente principal del sistema nervioso —situado en la base de las funciones cognitivas— y ciertamente es el encargado de procesar información así de controlar movimiento. En consecuencia, uno de los objetivos en esta área es simular las capacidades de las neuronas biológicas. Considero importante asimilar que la robótica cognitiva hace uso de RNA para modelar el procesamiento de información en los agentes artificiales.

A lo largo de la historia han surgido diversos modelos tanto de neuronas como de conjuntos de estas. Dentro de los modelos existentes podemos encontrar redes que modelan la dinámica de activación entre poblaciones, aquellas enfocadas en el modelado de actividad eléctrica —por ejemplo, los modelos de conductancia (Hodgkin y Horowicz, 1959; Hodgkin y Huxley, 1952b)— además de modelos clasificadores.

1.5. Redes Neuronales Artificiales

Las RNA son modelos computacionales que funcionan como una herramienta de cómputo (Basheer y Hajmeer, 2000) con propiedades tales como la habilidad de aprender, representar información, adaptarse y organizar o agrupar datos. Además de intentar modelar la capacidad de procesamiento de información del sistema nervioso (Rojas, 1996).

Estas redes consisten de un conjunto de unidades de procesamiento también llamadas nodos o neuronas artificiales. Estas unidades se comunican al mandar señales entre ellas a través de conexiones con “peso” (Kröse y van der Smagt, 1996).

Las redes pueden estar organizadas en distintas disposiciones dependiendo de su dimensión, aquí mencionaremos algunas de una y dos dimensiones. Las conexiones únicamente incluyen conexiones a sus nodos alelos —entre las más comunes— pudiendo ser cíclicas o no; las no cíclicas contienen un conjunto de nodos —que llamaremos “frontera”— con menos conexiones acorde a su vecindad (ver figuras 1.2a, 1.3), baste lo anterior para entender que los nodos frontera no están conectados entre ellos, no obstante, en el caso de las cíclicas los nodos frontera sí incluyen conexiones entre ellos (ver figuras 1.2b, 4.5), en particular se puede esbozar el carácter cíclico de dos dimensiones con la superficie llamada toro.

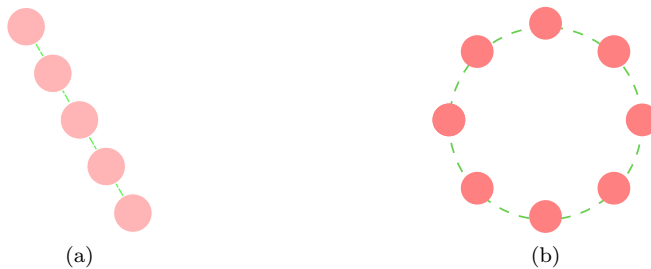


Figura 1.2: Red de una dimensión. (a) Red no cíclica. (b) Red cíclica.

Las redes neuronales pueden organizarse en forma de mapas —usualmente de dos dimensiones— y brinda de características de conexión particulares en relación a la forma de sus vecindades. El conjunto de unidades de procesamiento esta organizado correspondientemente al tipo de vecindad que tiene con sus nodos alelos. Algunas de las vecindades más habituales son la vecindad cuadrada y hexagonal (figura 1.3).

De acuerdo al tipo de que pertenezca la RNA puede tener distintos

componente, pero a grandes rasgos los componentes principales de una RNA son los siguientes:

Dominio de entrada son los datos que se introducirán al sistema (e.g. imágenes, sonido, información del estado del tiempo)

Unidades de procesamiento son las encargados de procesar la información

Valor de la unidad es el valor de salida o respuesta de la unidad

Pesos de conexión entre unidades estos pesos determinan el efecto que tiene x sobre y entre un par de “neuronas” (x, y)

Regla de propagación determina el efecto de la entrada sobre las neuronas

Función de activación determina el nivel de excitación entre la entrada y la activación de la neurona

Regla de aprendizaje se encarga de coleccionar la información

Sí, es verdad que las neuronas son “lentas” al ser comparadas con las capacidades de las compuertas lógicas —dispositivos electrónicos que son la base de aparatos de computo eléctricos— encontrados en las computadoras de hoy en día; estos dispositivos pueden cambiar de estado en nanosegundos. En cambio la velocidad para cambiar de estados de estas compuertas pierde peso en tareas (e.g. reconocimiento facial, intencionalidad, atención

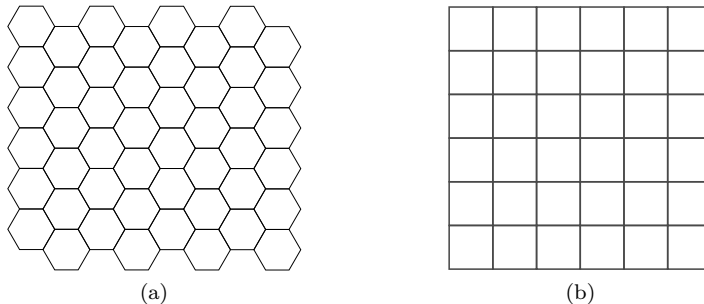


Figura 1.3: Vecindades de redes. (a) Vecindad hexagonal. (b) Vecindad cuadrada.

visual) que el cerebro —con sus redes neuronales— puede resolver y las computadoras todavía no pueden (Rojas, 1996).

La relación de *entrada/salida* del sistema de conexión analógica transmitida por pulsos, es descrita por una función sigmoide además tiene características particulares, por ejemplo, ser no decreciente —se podrá encontrar más detalle sobre las características de la función sigmoide en el capítulo 3.

Las redes usadas para modelar el tejido nervioso pueden dividirse en tres categorías:

Las redes *Feed-Forward* consisten de unidades organizadas en capas —también llamados niveles—, tal que la primera capa es encargada de recibir entrada mientras la última entrega una respuesta (salida). Las capas solamente pueden transferir información hacia la capa siguiente inmediata. Las redes *Feed-Back* a diferencia de las *Feed-Forward* tienen un tipo de conexión especial que permite a las capas superiores transmitir información hacia niveles previos. Las redes Auto-organizadas las unidades compiten en sus actividades mediante interacción lateral para adaptar, los pesos de estas unidades, con el objetivo de ser *codificadoras* de la distribución de datos de *entrada* (Rojas, 1996).

Dos clasificaciones a considerar referente al aprendizaje son las siguientes: El *Aprendizaje supervisado* es una forma de entrenar a una RNA con el objetivo de que entregue una respuesta deseada a partir de una entrada particular; esta meta se puede lograr al etiquetar las entradas provistas a la red. En otras palabras se clasifican los datos de entrada a conveniencia de lo que fuere necesario como salida. El *Aprendizaje no supervisado* utiliza algún procedimiento de clasificación para encontrar similitudes entre las entradas, por lo que a diferencia del aprendizaje supervisado, este no requiere de etiquetas. Los mapas auto-organizados son un nítido ejemplo para comprender el mecanismo de clasificación y sera detallado a continuación.

Mapas Auto-Organizados

Los Mapas Auto-Organizados tienen la propiedad de crear representaciones organizadas espacialmente de distribuciones de datos además son representación generalmente en forma de malla de distribuciones de datos. Este tipo de red neuronal tiene la cualidad de agrupar con mayor cercanía espacial a las unidades más parecidas de las no similares. Estos mapas son particularmente exitosos en tareas de reconocimiento de patrones. Para lograr la organización se calcula la similitud de los datos —pesos de las unidades— contra en torno a los datos de entrada. Este proceso tiene como

finalidad ajustar los pesos de las unidades para hacer más parecidos los datos de las unidades y los de entrada. Ciertamente este proceso se tiene que repetir para lograr una distribución homogénea de características entre unidades y datos (este proceso iterativo se llama entrenamiento) (Kohonen, 1990).

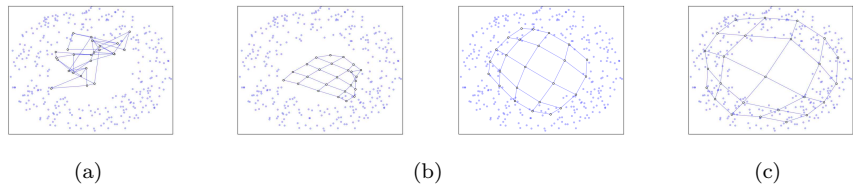


Figura 1.4: Esquema entrenamiento de una distribución de datos en forma de elipse en un SOM con vecindad cuadrada. Los puntos conectados representan los pesos en el dominio de distribución de entrada de los nodos del SOM mientras que los otros puntos son la distribución de entrada. (a) Estado inicial. (b) Estados intermedios. (c) Estado final.

La fase de entrenamiento es la encargada de lograr una semejanza entre las unidades y el conjunto de datos de entrada; al finalizar este proceso se tiene una red “entrenada”. El objetivo del entrenamiento es excitar un conjunto de neuronas de acuerdo a la similitud con el dato de entrada. Al finalizar el entrenamiento, las diferentes entradas parecen estar ordenadas además a cada unidad dentro de la red le corresponde un patrón de la distribución de los datos; en pocas palabras, cada unidad —o grupo de células— funciona equivalentemente a un decodificador de la entrada, con la función de activarse al presentar por entrada un dato semejante (Kohonen, 1982, 1998). Quisiera destacar que la finalidad de esta red es clasificar datos, asunto importante en trabajos de reconocimiento de patrones (e.g. predicción del estado del tiempo).

Existen múltiples variaciones de éstos mapas, en ésta sección se abordarán de manera superficial 2 tipos de ellos: El SOM es una representación (generalmente bidimensional) de distribuciones de datos. Todo elemento de la distribución de datos está en \mathbb{R}^n . El mapa se constituye de unidades llamadas nodos, los nodos tienen posición fija además están organizados en una malla. Las unidades de la malla pueden tener distintas formas de vecindad, por ejemplo, hexágona y cuadrada (ver figura 1.3) (Kohonen, 1989). Es importante resaltar que el dominio de la distribución de datos es distinto del dominio de la ubicación del mapa (ver figura 1.5);

una representación de los pesos del mapa, en el dominio de la distribución de datos, se puede observar en la figura 1.4.

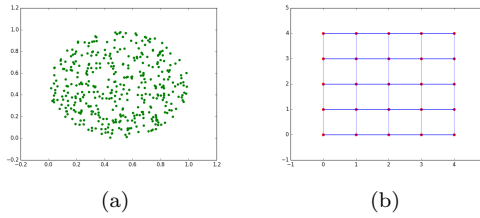


Figura 1.5: Representación de los dominios. (a) Dominio de la distribución de datos de entrada. (b) Dominio del SOM.

Los Dynamic Self-organizing Map (DSOM) tienen la independencia del tiempo para el proceso de aprendizaje, este punto se puede destacar observando su regla de aprendizaje y vecindad. Por lo anterior es posible un aprendizaje continuo (on-line learning) de distribuciones estáticas y dinámicas. Desde un punto de vista biológico éste algoritmo alumbra a la plasticidad cortical vista cómo un acoplamiento dinámico entre el medio ambiente y el modelo (Rougier y Boniface, 2011). Si el ambiente es estático las representaciones se mantienen estables, por el contrario, si el ambiente es dinámico entonces las representaciones deben cambiar modificarse hasta estabilizarse al nuevo ambiente.

CAPÍTULO 2

Atención visual

La atención visual es un fenómeno que involucra un conjunto de procesos que han sido estudiados por distintas disciplinas (e.g. robótica cognitiva y psicología). Este conjunto de procesos complejos resulta difícil abordarlos en un trabajo integral y se ve reflejado en que una gran fracción de los estudios se concentra en algún subconjunto de estos procesos de atención (e.g. saliencia (Hou y Zhang, 2007), *covert attention* (Ling y Carrasco, 2006)).

El fenómeno de atención ha sido estudiado desde distintos enfoques utilizando la metáfora conceptual. La metáfora conceptual es una herramienta que permite expresar una idea en términos de otra, teniendo por objetivo facilitar su comprensión, sin necesidad de tener el significado literal de la palabra usada como metáfora (McGlone, 2007). En otras palabras, la metáfora conceptual consiste de tener un mapeo conceptual entre un dominio (fuente) hacia un codominio (objetivo), tal que el mapeo define una relación entre ideas, entidades, propiedades, etc. hacia otros de diferente tipo.

Algunas de las metáforas utilizadas para abordar el fenómeno de atención pueden agruparse en tres teorías (Fernandez-Duque y Johnson, 2002), dos principales (*causa, efecto*) y una secundaria (teoría híbrida entre las dos principales). La teoría de *causa* se refiere a que el control del procesamiento de información es consecuencia de la atención, esto es que la atención modula el procesamiento de la información. La teoría de *efecto* se refiere a cuando la atención es producto del procesamiento de la información. Por último, la teoría híbrida toma elementos de las dos teorías principales (Fernandez-Duque y Johnson, 2002).

A grandes rasgos podríamos decir que la atención es causa de un mecanismo complejo de los sistemas autónomos que da ponderación a la información que entra al sistema así como a la experiencia previa (memoria). El proceso de ponderación permite definir qué información será procesada —de forma similar a un filtro— en un momento dado. En otras palabras, la información en proceso y la nueva información compiten por obtener una mayor parte de la capacidad de procesamiento. Para dilucidar esto, imagine que se ubica en un cuarto cerca de una ventana, tiene un libro en sus manos y se encuentra leyendo. En algún momento mientras continua la lectura se

produce un relámpago. A causa de ésto se distrae y quita su mirada del texto intentando encontrar la dirección del relámpago. Esta distracción ha permutado su atención de la lectura en torno a la información del relámpago.

Al procesamiento de información resultado de un proceso indirecto que involucra hacer inferencias sobre el mundo en base a la experiencia previa se le conoce por *top-down* mientras que al procesamiento directo de la información que es captada por los sensores es conocida por *bottom-up* (Fernandez-Duque y Johnson, 1999). En pocas palabras el mecanismo *top-down* se encarga de seleccionar a qué debe atender mientras que el mecanismo *bottom-up* es resultado de la información entrante.

En particular, el fenómeno de atención visual puede abordarse desde dos mecanismos diferentes. El mecanismo llamado *covert attention* permite detectar la ubicación de un estímulo sin utilizar movimiento ocular, en algunos trabajos es estudiado desde dos perspectivas: la primera es llamada *sustained* o endógena (*top-down*), la segunda *transient* o exógena (*bottom-up*) (Ling y Carrasco, 2006) El mecanismo conocido por el nombre de *overt attention* permite orientar los ojos hacia la ubicación del estímulo (Findlay y Gilchrist, 2003; Posner, 1980).

Por un lado, el proceso de atención *bottom-up* toma el control del procesamiento para dar prioridad a la información exteroceptiva, tal que por sus características es predominante (e.g. una alarma de incendio). Por otro lado, el proceso de atención *top-down* se puede describir como la capacidad de enfocarse en un conjunto de información, en otras palabras, se refiere a tener control sobre qué información se quiere atender, esta atención es dada por inferencias sobre el mundo. Estos dos procesos también son descritos desde la teoría sobre el marco del cerebro predictivo Clark (2013). En pocas palabras nuestros cerebros hacen un procesamiento predictivo sobre las regularidades del mundo —via *top-down*— las cuales se comparan con la percepción sensorial —via *bottom-up*— que de ser diferentes generan un error de la predicción para ser procesado. Recordemos el ejemplo sobre el relampago cuando nos encontrábamos leyendo, las inferencias (*top-down*) sobre las regularidades del mundo no predijeron que habría un relampago y es por esto que el error predictivo tuvo gran impacto y la atención fue permutada hacia este.

Por todo lo anterior es evidente que la memoria tiene un papel importante en este fenómeno, tanto para generar inferencias *top-down* basadas en las regularidades del mundo aprendidas así como de una memoria de trabajo para —de alguna forma— mantener ubicado el estímulo relevante —proveniente del entorno (*bottom-up*)— que está siendo atendido (Styles, 2006).

2.1. Atención artificial

Con el objetivo de crear modelos se han abordado diversos procesos pertenecientes al fenómeno de atención visual. Estos modelos han demostrado tener aplicaciones en el área de visión por computadora además de otras áreas tales como la robótica cognitiva. En este capítulo se abordará este fenómeno desde tres tipos de procesamientos de la información. En primer lugar se describirá el procesamiento *bottom-up*, esto es, modelos basados en las propiedades de los estímulos visuales —generalmente para buscar elementos sobresaliente es la escena— en el dominio visual (Itti y Koch, 2000; Itti et al., 2005; Hou y Zhang, 2007). En segundo lugar, se examinará al procesamiento *top-down*, por ejemplo los modelos basados en inferencias. Finalmente, se describirán modelos híbridos (Itti y Koch, 2001).

En varios de estos modelos (Itti y Koch, 2000, 2001; Itti et al., 2005) se utilizan *mapas de relevancia* (*saliency map*) bidimensionales que son constituidos por múltiples mapas de características. A través de competencia entre unidades se obtiene la ubicación —que además es única— del objeto más sobresaliente en la escena (Itti y Koch, 2000). En otras palabras los mapas codifican la relevancia de los objetos en el entorno mediante posiciones de los objetos sobresalientes.

Itti y Koch (2000) describen un mecanismo de atención que utiliza mapas de características (e.g. color, intensidad) que son combinados en un *mapa de relevancia*. La entrada al modelo son imágenes previamente procesadas. En resumen, las operaciones que se llevan a cabo son las siguientes:

1. Se aplican filtros ¹ para extraer características. Una peculiaridad de algunos de los filtros utilizados es el escalamiento recurrente para reducir el tamaño de la imagen.
2. Al conjunto de los mapas resultantes, se les aplica procesamientos de normalización, filtro de Diferencia de Gaussianas, etc.
3. Los mapas derivados —de los procedimientos anteriores— son sumados en tres distintos mapas (brillo, color y orientación), por consecuencia de la suposición que características similares compiten fuertemente entre ellas.
4. Se emplea una RNA *integrate-and-fire* (Burkitt, 2006a,b) con el objetivo de calcular el máximo de cada uno de los tres mapas. La

¹En el trabajo de Itti y Koch (2000) se utilizan pirámides Gaussianas y Gabor

neurona que dispara primero propaga una secuencia de potenciales de acción, que permutan el foco de atención hacia la posición contenida por la unidad ganadora.

Itti y Koch (2001) señalan que la selección de las características sobresalientes mediante el incremento del peso del estímulo es una forma plausible de describir cómo la atención *top-down* puede modular el procesamiento visual. La relevancia se obtiene al calcular la probabilidad de que en alguna ubicación esté presente un estímulo. La selección del foco de atención es guiada por el mapa de relevancia hacia la ubicación sobresaliente en la escena. Es a través de la inhibición transitoria de las unidades que determinan el foco de atención actual —esta transición a inhibición es un fenómeno llamado inhibición de retorno— además de las propiedades del algoritmo de la RNA que converge a la ubicación de la siguiente zona más sobresaliente, en consecuencia se producen trayectorias de escaneo de la escena visual.

Como se ha mencionado los modelos basados en saliencia han sido eficientes en tareas de detección de zonas —que por las características de las escenas son— sobresalientes. A estos se añaden los modelos de detección de objetos basados en relevancia (Hou y Zhang, 2007).

Hou y Zhang (2007) utilizan un método para adquirir el mapa de relevancia, este procedimiento se utiliza la representación espectral logarítmica (i.e. $\log(A(f))|A(f)$ es el promedio del espectro de Fourier) de imágenes además de una matriz —de dimensiones $n \times n$ dividida por n^2 — cuadrada. El mapa es calculado al hacer la diferencia entre el espectro logarítmico de la imagen —escalada a 64 píxeles en su largo o ancho— y la convolución de la matriz con el espectro logarítmico. El procedimiento resulta en un mapa de relevancia que contiene las partes significativas de la escena. Al usar un método inverso se puede regresar al dominio de la imagen original para obtener los objetos sobresalientes en la escena.

Ciertamente todos los modelos del fenómeno de atención vistos hasta ahora son precisos al describir procesos fundamentales de este fenómeno además de que obtienen resultados sorprendentes. Sin embargo casi todos utilizan técnicas computacionales para detección de características (e.g. mapas de saliencia (Itti y Koch, 2000), *Scale Invariant Feature Transform* (Lowe, 2004), *Speeded-up Robust Features* (Bay et al., 2008)). La detección de puntos de interés, la construcción de descripciones espaciales de estos puntos así como el emparejamiento de estas descripciones con una base de datos del objetivo son algunas etapas involucradas en el problema de reconocimiento de objetos. Desde luego estas etapas no son las únicas pero en algunos problemas se toman en cuenta para resolver el problema

de detección de objetos. Si bien varios de estos modelos pueden tener un fundamento biológico, dependiendo del modelo, algunos —en cierta medida— han dejado de lado la interacción de poblaciones de neuronas para describir los procesos que pertenecen a este fenómeno. Se considera importante tomar esta vertiente ya que en los procesos cognitivos se ven involucradas un conjunto de neuronas.

Sin duda los modelos descritos anteriormente ofrecen un mecanismo de atención visual, pero se considera que el preprocesamiento de las imágenes tiene una carga exagerada y dejan con menor ponderación el interesante proceso neuronal, inclusive si ocupan RNA para determinar la zona de relevancia. Para realizar un modelo artificial de atención visual que tome en gran consideración a la dinámica neuronal se cree necesario empezar por el mecanismo *covert attention*. Por un lado, podemos observar que *overt attention* requiere de localizar la zona relevante además es necesario un procedimiento para orientar el aparato visual en torno al objeto sobresaliente. Por otro lado, se espera que la simulación —mediante una implementación computacional— neuronal en el marco teórico de CND sea de utilidad para localizar regiones sobresalientes en la escena visual, que podría ser, resultado de la dinámica de activación de estos campos.

Campos Neuronales Dinámicos

Las matemáticas a través de las ecuaciones diferenciales han sido útiles para explicar en forma compacta y elegante fenómenos que se desarrollan en el tiempo. Existen distintos tipos de ecuaciones diferenciales, ecuaciones lineales, las ecuaciones no lineales, pueden describir fenómenos que no pueden caracterizarse linealmente —esto es, una ecuación $a_0 * x^0 + a_1 * x^1 + \dots + a_n * x^n$ para algún $n \in \mathbb{Z}$. Otro ejemplo son las ecuaciones parciales que describen cambios entre distintos parámetros (Zill et al., 2009). Los CND son ecuaciones diferenciales parciales que describen los cambios de potencial de membrana en relación a distintos parámetros (e.g. tiempo, posición de la unidad (neurona), conectividad de neuronas, entrada al sistema, entre otros) (Wilson y Cowan, 1972).

Este trabajo se realizó con el objetivo de simular el proceso de atención visual, para este propósito se utilizó el marco de CND. Los campos neuronales fueron estudiados por primera vez por Wilson y Cowan (1972), en su trabajo proponen estudiar poblaciones de células nerviosas (regiones corticales) para estudiar funciones cognitivas (e.g. procesamiento de información sensorial, aprendizaje) evitando el estudio por unidad. La relevancia de realizar un estudio por población es por una parte que en las funciones cognitivas interactúan un cuantioso número de neuronas (redes), por otra parte las neuronas estudiadas por unidad no tienen gran poder explicativo en funciones cognitivas, además que el requerimiento computacional es, hoy en día, no competente para simulaciones por unidad en poblaciones masivas.

Tomemos por ejemplo la descripción de cambio de temperatura en un espacio, si se realiza un análisis por unidad se requiere tener diversos parámetros sobre las partículas (e.g. espín, velocidad, dirección, tipo) dependiendo de cuánta resolución fuere necesaria. Sin embargo el conocimiento de los parámetros de cada una de las partículas —inclusive con pocos de estos parámetros— sería muy costoso computacionalmente incluso en espacios pequeños. Además, el conocimiento de las propiedades y parámetros de cada partícula no nos habla de cómo se comportarían las demás en el tiempo, al interactuar entre ellas. En cambio, al hacer un análisis por población, es posible describir el cambio de temperatura con

un modelo de ecuaciones diferenciales, que sí pudiere esbozar propiedades de la dinámica de interacción entre sus entidades. Este tipo de ecuaciones permiten describir cambios dinámicos y en el tiempo.

Debemos tomar en cuenta que no siempre es posible expresar la solución de una ecuación diferencial, ya sea en forma explícita o implícita, aún cuando esta solución exista. En estos casos se calcula una aproximación de la solución por un conjunto de puntos en R^n .

Wilson y Cowan (1972) plantean una ecuación diferencial no lineal que resulta en una aproximación descriptiva de la activación de población de neuronas, también muestran que la discretización además de otras modificaciones no afectan a la precisión para describir la activación neuronal. Una descripción sencilla y detallada de un campo neuronal constituido por una capa fue analizada por Amari (1977), quién analizó las dinámicas y patrones de activación.

Los CND ofrecen una descripción de la dinámica de activación neuronal, el enfoque de estos campos está en la proporción de neuronas que se vuelven activas por unidad de tiempo en una población de neuronas. El potencial de membrana es aproximado con ecuaciones diferenciales temporales no lineales para describir las poblaciones excitadora e inhibidora. Estos campos son un modelo matemático descrito por una ecuación diferencial parcial no lineal. Es importante mencionar que encontrar una solución para este tipo de ecuaciones resulta complejo, en consecuencia para resolver este problema se calcula una aproximación de la solución usando métodos numéricos. Uno de los métodos mas simples es el método de Euler, que permite aproximar la solución al hacer uso recursivo de una función, resultando en n coordenadas y dados n valores x , cada y_i, y_{i+1} se encuentran a una distancia constante h (Zill et al., 2009; Kincaid y Cheney, 2008).

Si bien, es cierto que existen métodos numéricos que permiten una mejor aproximación en comparación con el método de Euler, algunos lo logran a un costo computacional mayor y no son mencionados por considerarse no relevantes para los fines del trabajo aquí presentado.

En los trabajos de Wilson y Cowan (1972, 1973) se da énfasis a las propiedades de población, además se analiza la proporción de células que se vuelven activas por unidad de tiempo en la población de células que se encuentran espacialmente cercanas. Es relevante mencionar que incluso en un volumen pequeño de tejido neuronal existen células con respuestas idénticas a estímulos idénticos, lo cual se puede observar cómo redundancia local. La redundancia local se puede observar a lo largo del cerebro y permite la corrección de errores dentro del paso de señales entre neuronas de acuerdo a la alta cantidad de ruido que hay entre éstas. De forma que las

neuronas tienen mayor probabilidad que el mensaje sea procesado y llegue a su destino sin alteraciones (Hopfield, 1988). La respuesta de neuronas se asume cómo una función sigmoide por su similitud con esta función. La forma sigmoidea de la función de respuesta tiene una importancia clara, y ésta es que elementos con un nivel de excitación bajo fallaran en la excitación de otros elementos, mientras que los elementos con una excitación fuerte excitaran otros elementos de la población.

Para determinar que una función $f(x)$ es sigmoidea, es suficiente que cumpla alguna de las siguientes premisas (Amari, 1977):

- Es creciente monótona de x tal que $x \in (-\text{ínf}, \text{ínf})$.
- Los valores asintóticos de 0 y 1 son aproximados así cómo x se aproxima a $-\text{ínf}$ e ínf respectivamente.
- Tiene un solo punto de inflexión.

Expresión de excitación promedio

Si asumimos que las células suman sus entradas y el efecto de la estimulación decae con el tiempo $\alpha(t)$ entonces el nivel de excitación generado en la célula excitadora al tiempo t será:

$$\int_{-\text{ínf}}^{\text{ínf}} \alpha(t-t')(c_1 E(t') + P(t')) dt' \quad (3.1)$$

3.1. Categorías de Amari

La descripción de cómo el tejido neural puede ser tratado matemáticamente cómo un campo neuronal que consiste de una capa, se describe en los trabajos de Amari (1977, 1983). Los campos neuronales forman y propagan patrones de excitación, los patrones de activación, interacción y su respuesta al estímulo, son clasificados en 5 categorías por (Amari, 1977), éstas son:

- Campo mono-estable en donde toda la excitación eventualmente decae.
- Campo mono-estable en donde todo el campo se vuelve excitado.
- Campo biestable donde la excitación se propaga en un rango amplio pero decae si el estímulo se dio en un rango corto.

- Campo biestable donde la excitación inicial causa una excitación localizada de tamaño definido o su decaimiento.
La excitación se mueve hasta encontrar la posición del $max\{ \text{estímulo} \} \ni$ entrada.
- Campo con patrón de excitación periódico.

3.2. Modelo CND

Un campo neuronal dinámico es una descripción matemática de la dinámica de activación de una población neuronal (Rougier y Detorakis, 2013). Las ecuaciones que rigen al campo neuronal dinámico son:

$$\tau \frac{\partial v(x, t)}{\partial t} = -v(x, t) + \int f(v(y, t))w(x - y)dy + s(x, t) \quad (3.2)$$

$$w(x - y) = Ae^{\frac{(x-y)^2}{2\sigma_A^2}} - Be^{\frac{(x-y)^2}{2\sigma_B^2}} \quad (3.3)$$

$$f(v(y, t)) = \begin{cases} v(y, t) & \text{si } v(y, t) > 0 \\ 0 & \text{si } v(y, t) \leq 0 \end{cases} \quad (3.4)$$

(Sandamirskaya, 2015)

- $v(x, t)$ designa el potencial de membrana en posición x al tiempo t
- $w(x - y)$ es la función de peso de conexión lateral entre la posición x y y
- A y σ_A son el ancho y amplitud de la parte excitadora
 B y σ_B son el ancho y amplitud de la parte inhibitora
- f es la función de tasa de disparo
- $s(x, t)$ es la entrada en la posición x que controla al CND
- τ es la constante temporal (resolución temporal)

La integral en la eq. 3.2 es una función de convolución entre la función de disparo —del campo— con el kernel de activación. Una representación de cómo se ve una convolución se puede ver en la figura 3.1, en este esquema se utilizaron las funciones *sin* y *cos* para dar una ilustración asequible de esta función.

Dependiendo de la función de tasa de disparo f , la función de conectividad lateral w y la entrada s , los campos han exhibido un rango de

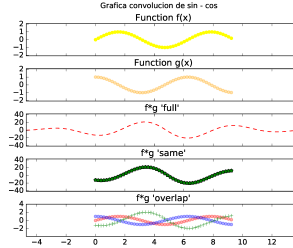


Figura 3.1: Grafica de la convolución de las funciones seno y coseno.

comportamientos dinámicos que van de patrones espaciales y/o temporales periódicos a regiones de actividad localizados. Éste último caso ha sido utilizado en otros trabajos para modelar atención visual, control motriz y memoria de trabajo. Para efectos de rastreo de objetos del fenómeno de atención, los CND han mostrado que las regiones de actividad “fuerte” tienen un vínculo de mapeo espacial hacia la entrada (Detorakis y Rougier, 2012, 2014; Buss y Spencer, 2008; Spencer et al., 2009; Johnson et al., 2008; Kopecz y Schöner, 1995; Johnson et al., 2009; Zibner et al., 2011; Oubbati y Palm, 2007; Sandamirskaya, 2015).

La simulación numérica de dichos campos requiere la discretización espacial del dominio M en n elementos espaciales mientras la integración temporal puede hacerse usando esquemas de integración clásicos (e.g. método de Romberg (Kincaid y Cheney, 2008)).

$$f[n] * g[n] = \sum_{k=-\text{inf}}^{\text{inf}} f(k) \cdot g(n - k) \quad (3.5)$$

$$\tau \frac{\Delta v(x_i, t)}{\Delta t} = -v(x_i, t) + \sum f(v(y_j, t))w(x_i - y_j) \quad (3.6)$$

$$x_i = \frac{i}{n - 1}$$

Bajo éstas suposiciones se ha estudiado un conjunto de parámetros para la conexión de pesos lateral que siguen la propiedad de converger a un paquete de actividad localizado cuyo máximo es aproximadamente igual a s para cualquier entrada uniforme positiva s (Rougier y Detorakis, 2013).

Al utilizar el marco teórico de CND es posible crear un mecanismo de aprendizaje similar al visto por los SOM —por su forma de mapa— además se argumenta una posible implementación con base biológica usando cómputo local, distribuido y numérico (Rougier y Detorakis, 2013). Este tipo de mapa es llamado *Dynamic Neural Field Self-organizing Map*. El aprendizaje se realiza a través de la modificación de conexiones feed-forward usando una regla de aprendizaje invariante en el tiempo que permita aprendizaje dinámico. El marco teórico CND refleja la hipótesis de que las fuertes interacciones recurrentes —en poblaciones locales— de neuronas, forman un mecanismo de procesamiento de la información cortical. Estas interacciones recurrentes causan comportamiento dinámico en poblaciones de neuronas. La actividad de la población que se inicia por señales externas tiempo-dependientes puede llegar a ser auto-sostenida incluso en la ausencia de cualquier entrada externa (Erlhagen y Bicho, 2006; Alecu y Frezza-Buet, 2009).

Experimentos y resultados

4.1. Instrumentos

Para realizar los experimentos presentados en esta tesis, se utilizó:

Un **robot** modelo Pioneer 3-DX con dos ruedas independientes y una rueda *loca* (figura 4.1).



Figura 4.1: Robot Pioneer 3-DX. (b) Vista frontal. (a) Vista lateral.

Como sensor visual se usó una **Cámara** *uEye* modelo UI-1410-C, con una resolución de 640×480 (Horizontal \times Vertical) píxeles y sensor de imagen de $1/3''$ CMOS. Para los experimentos, esta cámara se acopló a un espejo con forma semi-elipsoide (también llamado espejo cónico). El acoplamiento entre cámara y espejo cónico también es conocido por cámara omnidireccional (figura 4.2), el acoplamiento se obtuvo al montar horizontalmente la cámara apuntando hacia el espejo.

Para modelar un **CND** homogéneo de dos dimensiones y una capa, las ecuaciones utilizadas fueron las siguientes:

$$\tau \frac{\partial v(x, t)}{\partial t} = -v(x, t) + \int f(v(y, t))w(x - y)dy + s(x, t) \quad (4.1)$$

$$(4.2)$$

$$w(x - y) = Ae^{\frac{(x-y)^2}{\sigma^2}} \quad (4.3)$$

$$f(x) = \frac{\tanh(B * x) + 1}{2} \quad (4.4)$$

Las posiciones x y y son coordenadas espaciales del campo. Para el kernel de activación se utilizó una gaussiana (eq. 4.3), para la función sigmoide se utilizó una tangente hiperbólica (eq. 4.4).

4.2. Metodología

Se implementó un método de adquisición de imágenes para ser usadas como entrada a un CND. Las imágenes se capturaron en formato png (modelo de color RGB) con resolución 640×480 píxeles.

Directamente de la cámara, se obtienen imágenes panorámicas de 360° del entorno gracias a la forma elipsoide del espejo (figura 4.2b). La forma de estas imágenes se puede observar en la figura 4.4a. A estas imágenes se les aplica un mapeo que da como resultado una imagen *desdoblada* de 960×240 (figura 4.4b) —para facilitar la lectura en este documento la imagen se representó como una imagen desdoblada plana.

Uno de los efectos relacionados a la posición de la cámara (además de la forma del espejo) es que parte de la imágenes contengan información improductiva, por ejemplo información del techo o cielo de la escena no correspondiente al reflejo del espejo. Por consiguiente, al realizar el mapeo, solo se toma en cuenta la región que contiene al espejo, además no tomar en

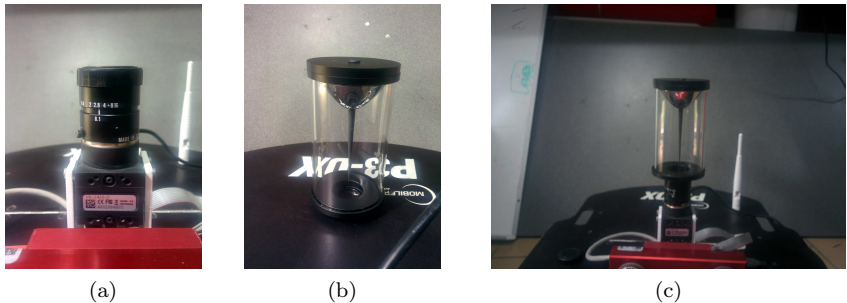


Figura 4.2: Componentes del sensor visual. (a) Cámara *ueye* UI-1410-C. (b) Espejo semi-elipsoidal (espejo cónico). (c) Cámara *ueye* y espejo acoplados (cámara omnidireccional).

cuenta parte de esa información es consecuencia de que parte de la captura no se encontraba a una distancia focal adecuada.

Posteriormente se aplicó un umbral a la imagen, para extraer el canal del color rojo (figura 4.4c). Para lograrlo se cambió al modelo de color *HSV* y se utilizaron los rangos:

Rango	límite	Matiz	Saturación	Valor
Primero	inferior	0	40	10
	superior	10	255	155
Segundo	inferior	164	40	10
	superior	179	255	155

El color rojo se encuentra en dos secciones del matiz, aproximadamente entre $[0^\circ, 15^\circ]$ y $[340^\circ, 360^\circ]$, ver figura 4.3) por lo tanto fueron necesarios dos rangos. En la representación de 24bits del modelo HSV, el matiz tiene sólo 8 bits en consecuencia el valor de 360° es representado en el intervalo $[0, 180]$.

Seguido de este proceso, las imágenes fueron redimensionadas a un tamaño de 100×100 píxeles (figura 4.4d) para obtener una proporción 1 : 1, el método utilizado fue promedio del valor de los píxeles dentro del área. Si bien es cierto que se pierde resolución horizontal, se consideró no relevante por ser homólogo a tener objetos delgados en escena y no debería ser obstáculo para el proceso de rastreo. Cabe mencionar que antes de realizar experimentos con imágenes obtenidas de la cámara omnidireccional, se realizaron pruebas con distintos tamaños de entrada para calcular el tiempo promedio necesario para obtener la activación entre dos tiempos (consecutivos) de la activación del campo. A partir de analizar los resultados de estas pruebas se determinó que para obtener un rendimiento adecuado en la plataforma de prueba (Pioneer 3-DX), es adecuado un campo neuronal de tamaño 100×100 unidades con el objetivo de obtener resultados en tiempo real.

Se implementó un CND en el robot con un tamaño de 100×100 unidades. La forma y tamaño de este campo se construyó correspondiente a las imágenes de entrada. Esto es, las unidades de este campo recibieron entrada a partir de un mapeo espacial 1 : 1 de los píxeles de la imagen. Las imágenes de entrada utilizadas son las imágenes redimensionadas —del canal rojo que fueron desdobladas— obtenidas de la cámara (figura 4.4d).

El campo neuronal se implementó con un tamaño de 10,000 neuronas organizadas en malla de 100×100 unidades (neuronas), esta malla se ajustó a la superficie de un toro (figura 4.5). Para facilitar la interpretación de las gráficas en este documento, el CND está representado como una superficie

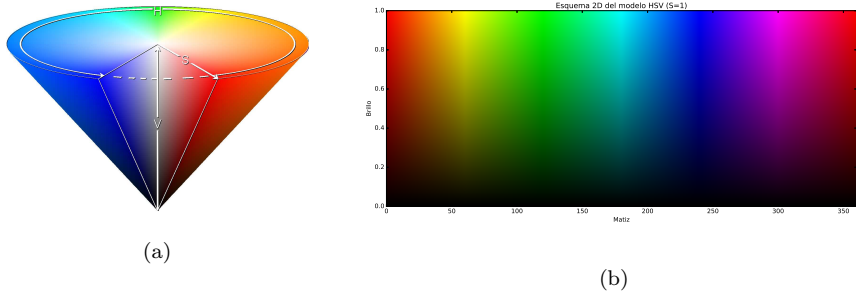


Figura 4.3: Representación HSV. (a) Cono modelo *HSV*, Imágen obtenida de www.wikimedia.org, el archivo se encuentra bajo la licencia *Creative Commons Attribution-Share Alike 3.0*. (b) Esquema 2D *HSV* (H,V). .

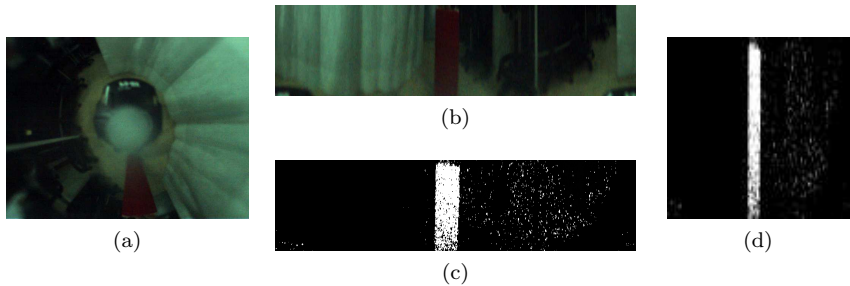


Figura 4.4: Imágenes panorámicas (360°). (a) Imágen original —“doblada”— obtenida de la cámara omnidireccional. (b) Imágen “desdoblada” —cortada a la región de interes— del espejo semi-elipsoide (espejo cónico). (c) Imágen en el umbral del rojo (desdoblada y cortada a la región de interes del espejo). (d) Imágen redimensionada a 100 * 100 píxeles de la imagen en el umbral del rojo.

plana en lugar de la superficie del toro. En otras palabras, un toro es una superficie de revolución —esto es, que puede ser generada a través de rotar una circunferencia, tal que cualquier posición en esta superficie se puede expresar a través de coordenadas $(x, y) | x \in [0, 2\pi], y \in [0, 2\pi]$ (ambas en radianes)— con forma similar a un anillo o dona. La ordenada x determina el ángulo en la circunferencia, la ordenada y determina la posición del círculo en el anillo.

La razón de usar una superficie con forma de Toro es que, algunas de las funciones (e.g. convolución) que describen a los CND tienen problemas para definir —solución— a valores en la frontera con coordenadas rectangulares.

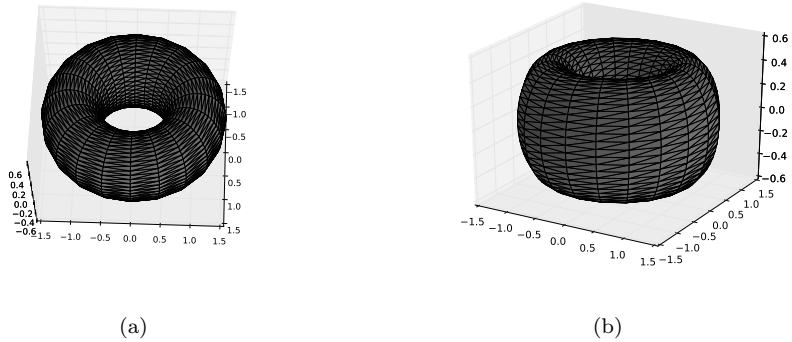


Figura 4.5: Esquema de un toro. (a) Vista de arriba. (b) Vista lateral.

En el ambiente que se encontraba el robot se colocaron como estímulos visuales objetos cilindricos cubiertos por fieltro rojo. Se hicieron distintos experimentos tanto con uno como dos objetos, al igual que con distintos parámetros del campo (e.g. amplitud del kernel de activación σ^2).

Se pudieron observar distintos tipos de patrones de activación —algunos de ellos correspondientes a los categorizados por Amari (1977)— los cuales dependen de los valores utilizados durante las simulaciones del CND. Los valores modificados fueron la varianza σ^2 y la dimensión temporal τ . Los valores de σ^2 utilizadas para el kernel de activación —en estos experimentos— son de rango corto, largo además de una categoría intermedia de medio rango.

Rango del kernel	$\sigma \approx$
Corto	0,08
Medio	0,13
Largo	0,3427

Se esperaba observar algunos de los patrones de activación (descritos en la categorización de Amari (1977)), de los cuales se cree que algunos de los más utiles, para el proceso de rastreo, son:

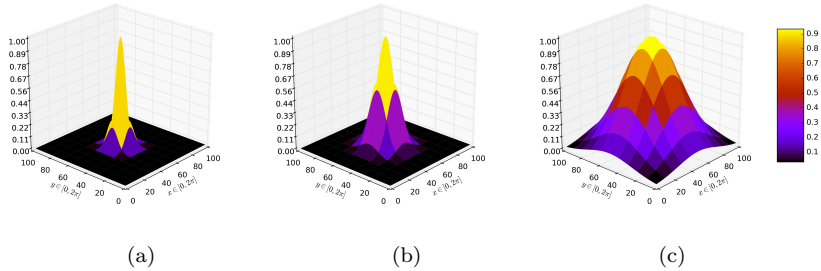


Figura 4.6: (a) Kernel rango corto $\sigma^2 = 0,0064$. (b) Kernel rango medio $\sigma^2 = 0,0169$. (c) Kernel rango largo $\sigma^2 = 0,1174$.

- Campo mono-estable tal que la excitación eventualmente decae
- Campo biestable que causa excitación localizada

Esto dependiendo de la posible saliencia (relevancia) del estímulo en la escena vista por el robot (P3-DX). Para la primera categoría se esperaba que la activación decayera al remover el estímulo de la escena —o al bloquear la visión de éste. Para la segunda —que se pensaba podría ser la más importante para el estudio de atención— se esperaba ver que la activación más fuerte en el campo estuviese en la posición relativa del campo hacía la posición en la imagen —simulara la capacidad de rastreo, tal como fue definida anteriormente. En concordancia con las razones de aparición de los patrones que se esperaba observar, se pensó que podrían ser en gran medida resultado de tanto la forma y dimensiones del kernel de activación —por ejemplo, amplitud y latitud— el cual en esta investigación se utilizó un kernel simétrico con forma gaussiana (ver figura 4.6). Veamos ahora los resultados esperados asimismo de las observaciones del comportamiento del campo.

En el primer experimento se colocó un objeto al frente del robot, este objeto se movió hacia la izquierda, después mantuvo posición del lado izquierdo y finalmente se movió hacia la derecha. Como se pudo observar en la figura 4.7 fue posible localizar el objeto inclusive cuando este objeto se encontraba en movimiento. Para la simulación del CND se utilizó un kernel de rango largo ($\sigma \approx 0,3427$) con una discretización temporal $\tau = 100ms$. Se esperaba observar una alta activación en la región de la posición del estímulo además de muy poca actividad en las regiones no salientes (relevantes). Entre los resultados obtenidos del primer experimento, se observó activación

en la region del campo correspondiente a la posición del objeto observado por el robot además de otras pequeñas regiones que no mantuvieron activación por mucho tiempo —en este caso no fue mayor a 30 avances t en el tiempo— en la simulación. Como se puede percibir en los resultados (ver figura 4.7) los parametros utilizados en el CND tuvieron efecto en el patrón de activación observado (estable con excitación localizada), característica que resulta de gran interés ya que funcionó de forma excepcional para rastreo.

En el segundo experimento se colocaron dos objetos en el campo visual del robot, el objeto de la izquierda se aleja del objeto de la derecha desplazandose a la izquierda del campo visual, despues se acerca al objeto de la derecha hasta que ambos se encuentran adyacentes. Para este experimento se utilizó un kernel de rango medio además de la constante $\tau = 100ms$. En este experimento (ver figura 4.8) se muestra que si dos objetos se encuentran contiguos, pueden provocar un solo pico de activación, tal como si fueran un único objeto. Con el tipo de kernel de rango medio se observó un comportamiento similar que el rango corto del kernel pero con una latitud menor en la activación del campo. Sin embargo esto sólo fue observado cuando el tamaño del objeto sumado a la distancia del robot era percibido de un tamaño menor o igual que el radio de acción del kernel de activación. En otras palabras, si el objeto tiene un tamaño aproximado al área de acción (*atención*) —definido por σ^2 — podía ser rastreado al mantener una activación en el campo.

En el tercer experimento se colocó un objeto que mostraba un movimiento hacia la izquierda (aproximadamente $10cm$), despues se obstruyó la visibilidad del objeto poniendo un obstáculo entre la cámara y el objeto, finalmente se retiró la obstrucción permitiendo la visibilidad del objeto nuevamente (ver figuras 4.9, 4.10). Este experimento se analizó al realizar simulaciones con distintos valores de τ y se pueden observar dos peculiaridades importantes, cada una correspondiente al valor de τ utilizado en el CND, estas se manifestaron con gran resalte en el lapso $t \in [30, 41]$. Se observó que con $\tau = 50ms$ es este caso inclusive cuando el objeto no se encontraba visible el campo mantenía la activación en la región donde se encontraba el objeto antes de perderse de vista además mostró un decaimiento de la activación (ver figura 4.10) al no estar presente el objeto para más tarde recobrar la activación cuando el objeto fue visible nuevamente. En contraste al utilizar un valor de $\tau = 100ms$ se puede observar que en el mismo interval ($t \in [30, 41]$) al ocultar el objeto de la visión, la activación del campo menguó tal como si no hubiera ningun objeto (ver figura 4.10). Es importante resaltar la capacidad de los CND de tener

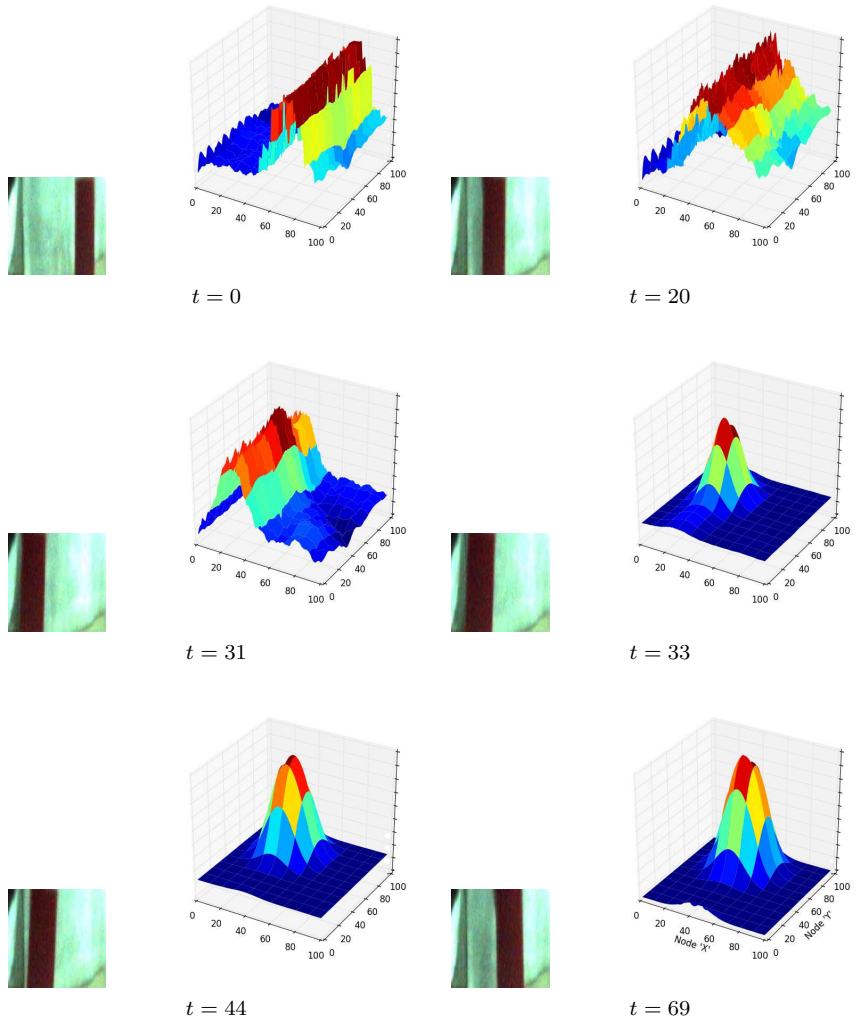


Figura 4.7: Imágenes tomadas de una secuencia en donde el objeto se desplaza en el campo visual del agente, al mismo tiempo se observa la activación del CND.

la capacidad de mantener activación de regiones inclusive cuando no se tiene entrada de los objetos en escena, ya que permite tener una característica semejante a una memoria de trabajo.

Cuando se utilizó un kernel de rango corto, se pensó que podría generar activación en el área donde se posicionaba el estímulo además de una fuerte activación (además de persistencia) en el ruido asimismo de pequeños objetos rojos. Lo cual podría ser útil para obtener más resolución en los detalles de la escena. Por el contrario de lo observado con el kernel de rango largo, con este tipo de rango se observó que no mantenía *estabilidad* el CND. En otras palabras, los patrones de activación observados parecían de alguna forma disparos aleatorios de las *neuronas* del campo aun cuando parte de la activación mantenía una relación a los objetos observados en la escena. Es decir que los disparos eran coherentes con la escena pero sin mantener suavidad del movimiento —relacionado al movimiento de los objetos— en la activación en el campo.

Es necesario agregar que en cualquiera de los casos anteriores la activación se encuentra ligada al tamaño del estímulo (objeto rojo) —debemos recordar que el campo sólo se activa por el espectro del color rojo— lo cual era de esperar por consecuencia de la invariabilidad del kernel de activación (Consecuencia de mantener un tamaño fijo para la capa del campo) en el tiempo. Claro que esto no aclara el hecho de que pequeñas regiones mantuvieran activación, sin embargo se debe agregar que la imagen contenía mucho ruido —parte del cual es generado por píxeles *muertos* en el sensor de la cámara.

En resumen, dependiendo de los parámetros utilizados, en la simulación del CND, el campo generaba una fuerte activación donde se encontraban los cilindros y al paso de iteraciones —que son avances discretos, en el tiempo con magnitud τ — se observaron distintos tipos de activación: patrón de activación en forma de gaussiana, unión de picos de activación (cuando objetos en escena se encontraban adyacentes), decaimiento de la activación al obstruir visión.

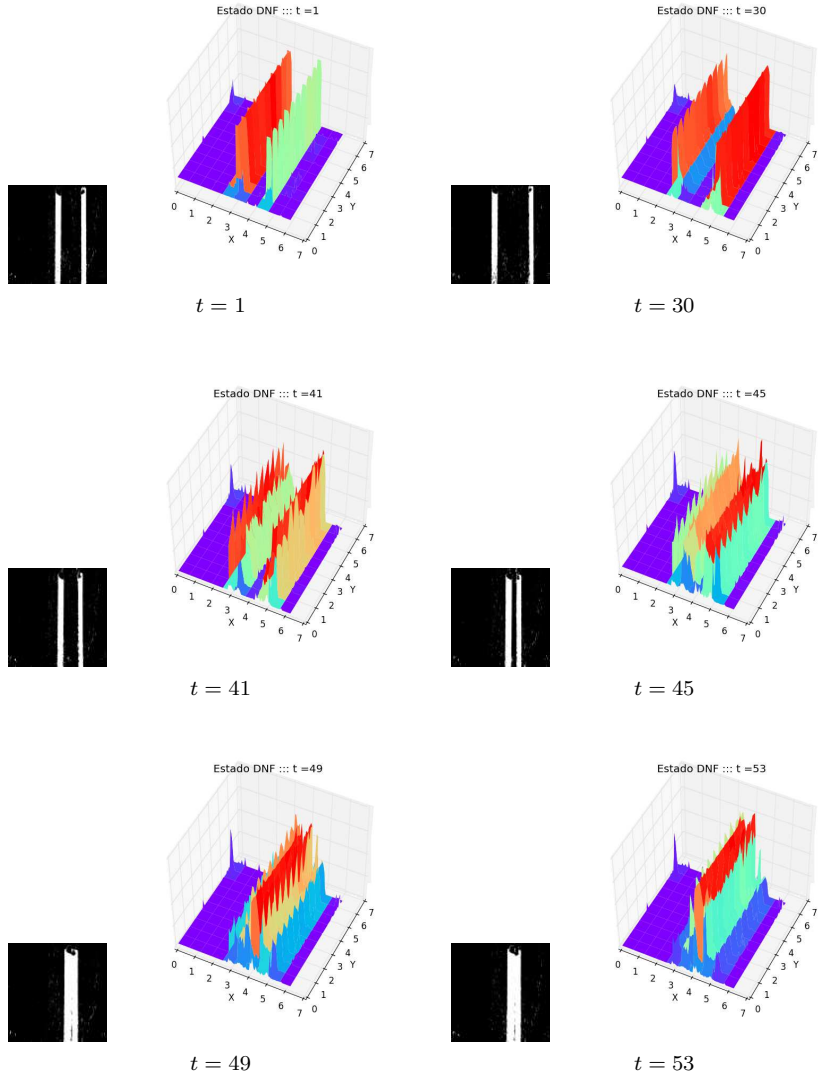


Figura 4.8: Imágenes tomadas de una secuencia en donde se encuentran dos objetos, conforme avanza el tiempo los objetos se alejan y más tarde se acercan hasta aparecer como uno solo. Al mismo tiempo se observa la entrada visual en el canal del rojo asimismo la activación del CND.

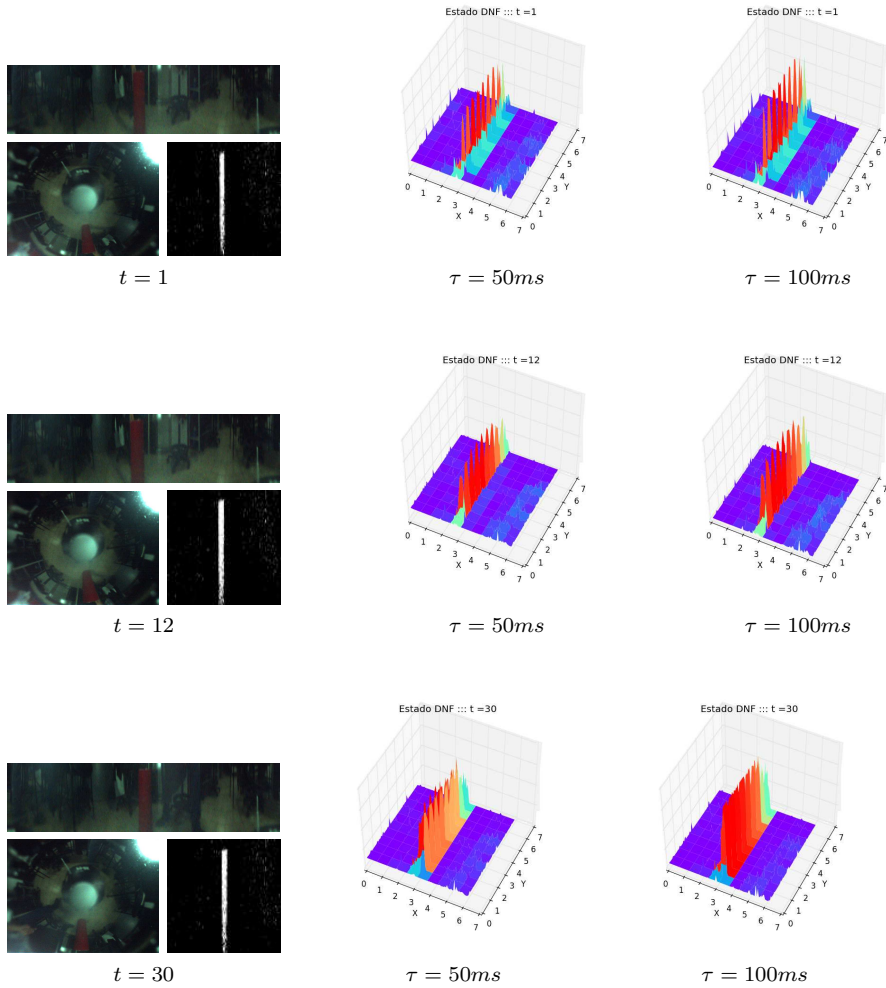


Figura 4.9: Imágenes tomadas de una secuencia en donde se encuentran un objeto en movimiento seguido de una obstrucción entre la cámara y el objeto, finalmente se retira la obstrucción. Al mismo tiempo se observa la entrada visual de la cámara omnidireccional, la imagen desdoblada, la imagen en el canal del rojo asimismo la activación del CND con dos diferentes valores de τ .

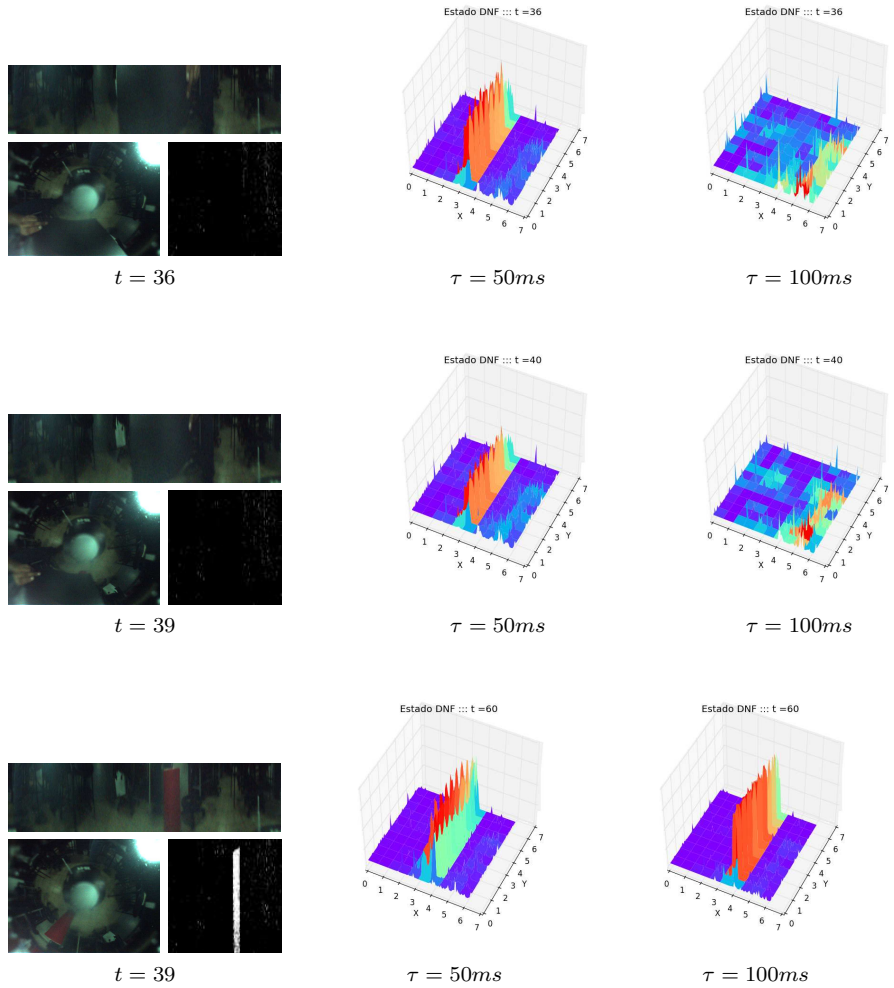


Figura 4.10: Imágenes tomadas de una secuencia en donde se encuentran un objeto en movimiento seguido de una obstrucción entre la cámara y el objeto, finalmente se retira la obstrucción. Al mismo tiempo se observa la entrada visual de la cámara omnidireccional, la imagen desdoblada, la imagen en el canal del rojo asimismo la activación del CND con dos diferentes valores de τ .

CAPÍTULO 5

Conclusiones

El trabajo reportado en esta tesis se realizó con la intención de simular el mecanismo de rastreo (similar a *covert attention*) en un agente artificial. Todo esto está basado en un modelo de dinámica de activación neuronal. En este trabajo se trajeron a discusión problemas importantes en los modelos de atención visual basados en saliencia, relevantes para el área de robótica cognitiva.

Se profundizó en modelos de relevancia (*bottom-up*) para describir al mecanismo de *covert attention* para abordarse desde una perspectiva de dinámica de activación neuronal —provista por el marco teórico de los CND. Estos modelos podrían, en un futuro, ser la base de mecanismos más complejos (e.g. *overt attention*).

Como fue mencionado en el capítulo 1, uno de los objetivos para realizar esta investigación fue modelar un mecanismo de rastreo con base en activación neuronal. En los resultados de los experimentos (ver figuras 4.7, 4.8, 4.9 , 4.10) se muestra evidencia de la capacidad de los CND para modelar atención (tipo *bottom-up*). Así mismo se pueden mencionar dos habilidades significativas:

- Mantener el rastreo de objetos aún cuando estos se encuentran en movimiento (ver figura 4.7).
- Mantener la activación en el CND de objetos, aun cuando estos desaparezcan del campo visual (ver figura 4.10).

Con estos resultados se comprueba la hipótesis de esta tesis, ya que el marco teórico de CND fue de utilidad para simular el proceso de rastreo.

Por lo observado en este tercer experimento se cree que es también posible, a través de la dinámica de activación de los CND, obtener una característica similar a la memoria de trabajo. Ésto debido a que la permanencia de la activación de estos estímulos, sería suficiente para la simulación de ciclos sensorimotrices.

Para los experimentos descritos en este trabajo se utilizó únicamente el canal rojo, sin embargo, se considera que podría ser extendido a modelos que incluyan más colores (e.g. RGB).

Cabe resaltar que para obtener resultados semejantes a los observados en este trabajo es necesario mantener un ajuste de los parámetros que definen a los CND además de la estabilidad en la entrada. Por una parte, el kernel de activación del campo debe tener una amplitud (σ^2) en relación al tamaño del objeto en escena que se requiera rastrear. Esto significa que el objeto pueda ser envuelto por el área bajo la función del kernel. Por otra parte es también imprescindible que la entrada visual no sea muy distinta entre imágenes, con eso último entiendase que aún cuando temporalmente las imágenes obtenidas por el sensor visual se encuentren cercanas, si la información percibida al tiempo t y $t + 1$ los objetos se encuentran distantes espacialmente es posible que no puedan ser rastreados por el campo.

Una ventaja más de estos modelos es que no se requiere de mecanismos computacionales para determinar regiones de interés en la escena visual.

Trabajo futuro

Como trabajo a futuro se propone crear un CND —para cada canal de color RGB— y calcular la activación con otros CND. Así mismo valdría la pena investigar los efectos de la forma de los obstáculos en la activación de los CND.

Para dar continuidad al trabajo presentado en esta tesis, se considera que podría ser de utilidad utilizar varias capas de CND por distintos aspectos. El primero es que al utilizar varias capas, se podrían utilizar distintos rangos de kernel, con la restricción de tener orden, para ello se propone utilizar rangos cortos en las capas más cercanas a la entrada sensorial (exteroceptiva) y rangos largos conforme se alejan de la entrada. En otras palabras, se sugiere utilizar kernel con rangos crecientes en el orden que las capas se alejan de la entrada sensorial. Las capas bajas (cercanas a la entrada sensorial) podrían obtener información de áreas pequeñas —que estarían determinadas por la longitud del kernel— que podrían extraer información específica mientras que las capas más elevadas se encargarían de procesar información de mayor tamaño.

En un segundo aspecto se sugiere utilizar también kernel de tipo *Diferencia de gaussianas* —también conocido por el nombre de *sombrero mexicano*— ya que este tipo de kernel permite describir excitación de corto alcance además de inhibición de largo alcance, que se piensa podría llegar a ser de utilidad para localizar objetos con mayor precisión en el caso de que ruido (e.g. píxeles muertos) esté presente.

Es posible que para observar inhibición por repetición —el cual no fue observado en los resultados descritos en esta tesis— sea necesario considerar un tercer aspecto, que es modelar no sólo población excitadora si no también

población inhibidora tanto en una misma capa como en campos separados.

Un cuarto aspecto a tomar en cuenta en futuro trabajo, es añadir un mecanismo que permita al agente almacenar regularidades observadas sobre el mundo. Hasta ahora sólo se logra obtener un tipo de memoria parecido a la memoria de trabajo. Además se considera de gran importancia facultar al agente de movimiento para que al trabajar en conjunto con la memoria, pueda aprender también de los efectos que producen sus movimientos en su percepción del mundo, con el objetivo de lograr un desempeño más coherente en el ambiente.

Para finalizar se considera importante tomar en cuenta las teorías de predicción como las descritas por Clark (2013) del cerebro predictivo además de la de Friston (2005) en codificación predictiva, todo lo anterior para complementar las teorías sobre cognición que toman al cuerpo (Clark, 2015; Wilson, 2002; Pfeifer, 2002) como un elemento fundamental para la emergencia de fenómenos cognitivos (e.g. atención).

Glosario de términos y abreviaciones

CND	Campos Neuronales Dinámicos. 1–4, 21, 23, 24, 26–31, 33–42, 45
DNF	Dynamic Neural Fields. 45
DNF-SOM	Dynamic Neural Field Self-organizing Map. 27, 45
DSOM	Dynamic Self-organizing Map. 15, 45
Ecuación diferencial	Ecuación que contiene derivadas de una o más variables respecto a una o más variables independientes . 45
Ecuación integrodiferencial	Ecuación que contiene tanto integrales como derivadas . 45
Elipsoide	Cuerpo de forma esferoide cuyas secciones planas son circunferencias o elipses . 45
Esferoide	Cuerpo de forma aproximadamente esférica . 45
GOFAI	Good Old-Fashioned Artificial Intelligence. 6, 45
GOFAI	Inteligencia artificial tradicional (simbólica). . 45
HSV	Hue Saturation Value. 31, 32, 45
Hue Saturation Value	Modelo de color Matiz Saturación Valor (brillo) . 45
IA	Inteligencia Artificial. 1, 2, 4–7, 9, 45
Intencionalidad	Propiedad de los estados mentales para referirse o dar representación a cosas, propiedades y estados en el mundo. . 7, 8, 45

Pioneer 3-DX	Robot utilizado principalmente en investigación. A este robot se le pueden añadir componentes tales como sensores de proximidad laser, camaras, brazos de agarre, etc . 4, 45
Potencial de membrana	Diferencia entre el voltaje interno y externo de una neurona . 45
RGB	Red Green Blue. 41, 42, 45
RGB	Modelo de color Rojo Verde Azul . 30, 45
RNA	Redes Neuronales Artificiales. 1, 9, 11, 12, 14, 19-21, 45
Solución explícita	Una función $\phi(x)$ tal que al sustituirla en vez de y en la ecuación diferencial satisface la ecuación para toda x en el intervalo I . 24, 45
Solución implícita	Se dice que una relación $G(x, y) = 0$ es una solución implícita de una ecuación diferencial en el intervalo I si define una o más soluciones explícitas en I . 24, 45
SOM	Self-organizing Map. 9, 15, 16, 27, 45
SOM	Modelo matemático con organización topografica similar a la encontrada en corteza cerebral . 45
Toro	En geometría, superficie de dos dimensiones con forma de “dona”. Un solido de revolución generado por una circunferencia (ver figura 4.5) . 31-33, 45
VGA	Video Graphics Adapter. 45
VGA	Adaptador gráfico de video que es comunmente referido por su resolución de 640×480 (horizontal \times vertical) píxeles. . 45

Bibliografía

- Alecu, L. y Frezza-Buet, H. (2009). Reconciling neural fields to self-organization. En *ESANN*.
- Amari, S.-i. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological cybernetics*, 27(2):77–87.
- Amari, S.-i. (1983). Field theory of self-organizing neural nets. *Systems, Man and Cybernetics, IEEE Transactions on*, (5):741–748.
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, 59:617–645.
- Basheer, I. y Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods*, 43(1):3–31.
- Bay, H., Ess, A., Tuytelaars, T., y Van Gool, L. (2008). Speeded-up robust features (surf). *Computer vision and image understanding*, 110(3):346–359.
- Brette, R. y Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of neurophysiology*, 94(5):3637–3642.
- Brooks, R. A. (1990). Elephants don't play chess. *Robotics and autonomous systems*, 6(1):3–15.
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208.
- Burkitt, A. N. (2006a). A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological cybernetics*, 95(1):1–19.
- Burkitt, A. N. (2006b). A review of the integrate-and-fire neuron model: II. inhomogeneous synaptic input and network properties. *Biological cybernetics*, 95(2):97–112.
- Buss, A. y Spencer, J. P. (2008). The emergence of rule-use: A dynamic neural field model of the dccc. En *Paper to appear in the Twenty-ninth Annual Conference of the Cognitive Science Society*.

- Cacioppo, J. T., Tassinary, L. G., y Berntson, G. (2007). *Handbook of psychophysiology*. Cambridge University Press.
- Campbell, M., Hoane, A. J., y Hsu, F.-h. (2002). Deep blue. *Artificial intelligence*, 134(1):57–83.
- Church, A. (1936). A note on the entscheidungsproblem. *The journal of symbolic logic*, 1(01):40–41.
- Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03):181–204.
- Clark, A. (2015). Embodied prediction. En *Open MIND*. Open MIND. Frankfurt am Main: MIND Group.
- Detorakis, G. I. y Rougier, N. P. (2012). A neural field model of the somatosensory cortex: formation, maintenance and reorganization of ordered topographic maps. *PloS one*, 7(7):e40257.
- Detorakis, G. I. y Rougier, N. P. (2014). Structure of receptive fields in a computational model of area 3b of primary sensory cortex. *Frontiers in computational neuroscience*, 8:26.
- Erlhagen, W. y Bicho, E. (2006). The dynamic neural field approach to cognitive robotics. *Journal of neural engineering*, 3(3):R36.
- Fernandez-Duque, D. y Johnson, M. L. (1999). Attention metaphors: How metaphors guide the cognitive psychology of attention. *Cognitive Science*, 23(1):83–116.
- Fernandez-Duque, D. y Johnson, M. L. (2002). Cause and effect theories of attention: The role of conceptual metaphors. *Review of general psychology*, 6(2):153.
- Findlay, J. M. y Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Number 37. Oxford University Press.
- Friedenberg, J. y Silverman, G. (2011). *Cognitive science: An introduction to the study of mind*. Sage.
- Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences*, 360(1456):815–836.

- Gross, R. (2015). *Psychology: The Science of Mind and Behaviour 7th Edition*. Hodder Education.
- Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3):335–346.
- Haugeland, J. (1989). *Artificial intelligence: The very idea*. MIT press.
- Hodgkin, A. y Horowicz, P. (1959). The influence of potassium and chloride ions on the membrane potential of single muscle fibres. *The Journal of Physiology*, 148(1):127.
- Hodgkin, A. L. y Huxley, A. F. (1952a). Propagation of electrical signals along giant nerve fibres. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, p'aginas 177–183.
- Hodgkin, A. L. y Huxley, A. F. (1952b). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500.
- Hopfield, J. J. (1988). Artificial neural networks. *Circuits and Devices Magazine, IEEE*, 4(5):3–10.
- Hou, X. y Zhang, L. (2007). Saliency detection: A spectral residual approach. En *2007 IEEE Conference on Computer Vision and Pattern Recognition*, p'aginas 1–8. IEEE.
- Houdé, O., Kayser, D., Koenig, O., Proust, J., y Rastier, F. (2003). *Diccionario de Ciencias Cognitivas: Neurociencia, Psicología, Inteligencia Artificial, Lingüística y Filosofía*. Amorrortu Editores.
- Itti, L. y Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research*, 40(10):1489–1506.
- Itti, L. y Koch, C. (2001). Computational modelling of visual attention. *Nature reviews neuroscience*, 2(3):194–203.
- Itti, L., Rees, G., y Tsotsos, J. (2005). Models of bottom-up attention and saliency. *Neurobiology of attention*, 582.
- Izhikevich, E. M. et al. (2003). Simple model of spiking neurons. *IEEE Transactions on neural networks*, 14(6):1569–1572.
- Jacob, P. (2014). Intentionality. En Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Winter 2014 edition.

- Johnson, J. S., Spencer, J. P., Luck, S. J., y Schöner, G. (2009). A dynamic neural field model of visual working memory and change detection. *Psychological science*, 20(5):568–577.
- Johnson, J. S., Spencer, J. P., y Schöner, G. (2008). Moving to higher ground: The dynamic field theory and the dynamics of visual cognition. *New Ideas in Psychology*, 26(2):227–251.
- Kincaid, D. R. y Cheney, E. W. (2008). *Numerical Mathematics and computing*, volumen 6. Brooks Cole Pub Co.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1):59–69.
- Kohonen, T. (1989). Self-organizing feature maps. En *Self-organization and associative memory*, p'áginas 119–157. Springer.
- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480.
- Kohonen, T. (1998). The self-organizing map. *Neurocomputing*, 21(1):1–6.
- Kopecz, K. y Schöner, G. (1995). Saccadic motor planning by integrating visual information and pre-information on neural dynamic fields. *Biological cybernetics*, 73(1):49–60.
- Kröse, B. y van der Smagt, P. (1996). *An introduction to neural networks*, volumen 8. University of Amsterdam.
- Ling, S. y Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision research*, 46(8):1210–1220.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110.
- McGlone, M. S. (2007). What is the explanatory value of a conceptual metaphor? *Language & Communication*, 27(2):109–126.
- Newell, A. y Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19(3):113–126.
- Oubbati, M. y Palm, G. (2007). Neural fields for real-time navigation of an omnidirectional robot. En *EMCR*.

- Peng, J., Bo, L., y Xu, J. (2009). Conditional neural fields. En *Advances in neural information processing systems*, p'aginas 1419–1427.
- Pfeifer, R. (2002). Embodied artificial intelligence—on the role of morphology and materials in the emergence of cognition. *Informatik bewegt: Informatik*, 32.
- Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, 32(1):3–25.
- Rojas, R. (1996). Neural networks. a systematic approach.
- Rougier, N. y Boniface, Y. (2011). Dynamic self-organising map. *Neurocomputing*, 74(11):1840–1847.
- Rougier, N. P. y Detorakis, G. I. (2013). Self-organizing dynamic neural fields. En *Advances in Cognitive Neurodynamics (III)*, p'aginas 281–288. Springer.
- Sandamirskaya, Y. (2015). Dynamic neural fields as a step toward cognitive neuromorphic architectures. *Neuromorphic Engineering Systems and Applications*, p'agina 145.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, 3(03):417–424.
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.
- Sipser, M. (1997). *Introduction to the Theory of Computation*, volumen 1. Thomson Publishing Inc.
- Spencer, J. P., Perone, S., y Johnson, J. S. (2009). The dynamic field theory and embodied cognitive dynamics. *Toward a unified theory of development: Connectionism and dynamic systems theory re-considered*, p'aginas 86–118.
- Styles, E. (2006). *The psychology of attention*. Psychology Press.
- Thagard, P. (2014). Cognitive science. En Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2014 edition.
- Turing, A. M. (1936). On computable numbers, with an application to the entscheidungsproblem. *J. of Math*, 58(345-363):5.

-
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236):433–460.
- Wilson, H. R. y Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1.
- Wilson, H. R. y Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13(2):55–80.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review*, 9(4):625–636.
- Wilson, R. A. y Keil, F. C. (2001). *The MIT encyclopedia of the cognitive sciences*. MIT press.
- Zhao, X., Kim, J., Robinson, P. A., y Rennie, C. J. (2014). Low dimensional model of bursting neurons. *Journal of computational neuroscience*, 36(1):81–95.
- Zibner, S. K., Faubel, C., Iossifidis, I., y Schoner, G. (2011). Dynamic neural fields as building blocks of a cortex-inspired architecture for robotic scene representation. *IEEE Transactions on Autonomous Mental Development*, 3(1):74–91.
- Ziemke, T. (2016). The body of knowledge: On the role of the living body in grounding embodied cognition. *Biosystems*, 148:4–11.
- Zill, D. G., Cullen, M. R., et al. (2009). *Ecuaciones diferenciales con problemas con valores en la frontera*, volumen 7. Cengage Learning.