

Capítulo 5

CONCLUSIONES

Se ha desarrollado un modelo de representación que, tomando referencias de los trabajos basados en representaciones activas, realiza la caracterización de objetos y el análisis de su movimiento.

Como el objetivo que subyace es el desarrollo de arquitecturas especializadas en visión, en movimiento, de tiempo real, se han seguido planteamientos conexionistas por cuestiones de paralelizabilidad y posibilidad de refinamiento progresivo de la respuesta. Se entiende que este enfoque establece una diferencia sustancial respecto de la práctica habitual en la investigación de visión de nivel medio: en general, los problemas se abordan con planteamientos puramente computacionales, con ignorancia de los aspectos de implantación y de las características de los computadores. El procesamiento se realiza, finalmente, sobre sistemas potentes; casi sin consideraciones de especificidad que permitan explotar las peculiaridades de la naturaleza de los problemas de visión.

Este trabajo, junto con (Mora, 2001), establece los primeros pasos para el desarrollo de arquitecturas específicas de tiempo real, no sólo dirigidas al análisis del movimiento, sino que deben servir para tratar muchos otros aspectos de la visión de nivel medio y bajo, algunos de los cuales ya han sido presentados en esta memoria.

Tomando como base el comportamiento de los gases neuronales, se adapta la topología de su red de interconexión a un espacio de entrada definido por el objeto. El efecto que se consigue es doble: por un lado, el Grafo Preservador de la Topología sirve como estructura de datos reducida, pero rica en información; por otro, se dota al modelo de capacidad de operación bajo restricciones de tiempo real, previa modificación de sus algoritmos de aprendizaje.

Adicionalmente, debido al carácter dinámico de estas redes que no necesitan reiniciar el aprendizaje al introducir nuevos patrones, se modeliza el movimiento realizado por los objetos como las deformaciones que sufre la red a lo largo de toda la secuencia de imágenes.

Para poner de manifiesto las prestaciones de dicha metodología, se ha propuesto resolver una instancia del problema del reconocimiento de gestos, uno de los problemas considerado complicado en el análisis del movimiento, debido al gran número de grados de libertad que posee la mano humana, que conlleva una gran variedad de estados morfológicos de la misma.

5.1 APORTACIONES

A continuación se presentan, de forma más ordenada, las diversas aportaciones presentadas en esta trabajo:

- Se ha realizado un estudio de la capacidad de preservación de la topología de objetos bidimensionales por parte de diversos modelos neuronales auto-organizativos, comparando su comportamiento. Se ha medido de forma empírica dicha capacidad, empleando el producto topográfico. Se ha obtenido que los gases neuronales, los modelos Neural Gas y Growing Neural Gas, son los que mejor preservan la topología de un espacio de entrada bidimensional.

- Se dota a los gases neuronales de capacidad de funcionamiento bajo restricciones de tiempo real. Se han introducido variantes a los procesos de aprendizaje para que la adaptación pueda supeditarse a cotas de tiempo y tener una medida de la calidad de respuesta.
- Se han empleado los resultados de las aportaciones anteriores para la representación de objetos bidimensionales: la red de interconexión entre las neuronas determina la morfología de los objetos, obteniéndose un Grafo Preservador de la Topología. Se obtiene una nueva aplicación de las redes neuronales auto-organizativas, ya que usualmente han sido empleadas como clasificadores de características, pero no como las propias características de los objetos.
- Sin necesidad de modificar los algoritmos de aprendizaje de los gases neuronales, sino modificando únicamente la propiedad visual para la que es sensible la red; se puede extraer tanto la representación externa (contorno) como su representación interna.
- Se han extraído características del Grafo Preservador de la Topología tanto al adaptarse a contornos, como a superficies; de modo que operaciones de clasificación y reconocimiento posteriores se vean simplificadas, evitando la alta complejidad de la comparación entre grafos.
- Dadas las capacidades de compresión de la información y de preservación de la topología del *GPT*, se realiza la síntesis de los objetos a partir del grafo. Se han presentado diversas posibilidades, ya sea obteniendo dicha síntesis contando únicamente con la representación del *GPT* o disponiendo de información adicional, obtenida al calcularlo.
- Se ha extendido la aplicación de los *GPT* al tratamiento de secuencias de imágenes y, más concretamente, al seguimiento de objetos y análisis del movimiento. Debido al carácter dinámico de estas redes, se ha modelizado la evolución de los objetos como las modificaciones que sufre la red a lo largo de toda la secuencia, entendiendo

dichas modificaciones como variaciones en la topología de la red de interconexión así como zona del espacio de entrada para la que es sensible cada uno de los elementos de procesamiento.

- La modelización del movimiento como la dinámica de las neuronas, evita el problema de correspondencia, uno de los más costosos en la mayoría de técnicas de seguimiento de objetos.
- Aunque no ha sido destacado especialmente en la memoria, este modelo permite la caracterización y seguimiento de múltiples objetos presentes en la escena. Asimismo, se puede realizar el seguimiento de un objeto que se disgregue en varias partes o que se fusione.

Resumidamente, la investigación que se contiene en esta memoria ha consistido en la aportación de estructuras de datos y de métodos, desde el enfoque de proporcionar conocimiento para cubrir los objetivos iniciales: establecer un marco de modelización de arquitecturas especializadas para visión, con capacidad de procesamiento en tiempo real, y suficientemente generalista para dar cobertura a técnicas de tratamiento de una amplia diversidad de problemas de visión.

Por cuestiones de extensión del trabajo, la investigación se ha concretado en establecer la base conceptual sobre la que después se materializarán las arquitecturas. La modelización se ha particularizado para un par de problemas representativos: uno de caracterización y otro de seguimiento.

5.2 LÍNEAS DE CONTINUACIÓN

Dada la aplicación de las redes neuronales auto-organizativas como caracterizadoras de los objetos, se espera que este trabajo abra una nueva línea en el desarrollo de arquitecturas especializadas en visión de nivel medio y bajo. Si bien en este trabajo, se ha aplicado el modelo al problema del análisis del movimiento, se debe realizar una generalización del mismo para el tratamiento de otras operaciones de visión a todos los niveles. Asimismo, se debe dotar al modelo de robustez para su aplicación en entornos realistas, con presencia de ruido, oclusiones,...

Como ya se ha expresado en diversas ocasiones, este modelo debe materializarse en el desarrollo de arquitecturas de tiempo real especializadas, con características de miniaturización y empotrabilidad; que muestren la potencia del prototipo presentado en esta memoria.

Se deben estudiar algunas cuestiones que quedan al margen de este trabajo, pero que se encuentran estrechamente ligadas:

- Desarrollo de nuevos modelos de representación, tomando las ideas que se desprenden de esta tesis. En particular, se está desarrollando una red en anillo incremental, la Growing Ring Structure, que tomando el algoritmo de aprendizaje de la Growing Cell Structure, permite una adaptación a los contornos de los objetos sin que se produzcan aristas indeseadas en zonas de gran curvatura.
- Si existen varios objetos presentes en la escena, se puede realizar el seguimiento de todos ellos adaptando varios *GPT* simultáneamente. Esto puede producir interacciones entre las diversas redes, lo que lleva a la posible consideración de miscibilidad de los gases neuronales.
- Desarrollo de sistemas robustos que permitan su empleo en el tratamiento de escenas realistas.
- Diseño de arquitecturas de alto rendimiento, basadas en el modelo presentado, para su integración en el dispositivo de visión, con enfoque de miniaturización y empotrabilidad.

Por último, en el tratamiento del reconocimiento de gestos, el desarrollo de un modelo robusto ante entornos heterogéneos y el desarrollo de nuevos filtros que permitan una buena segmentación de las manos, permitirá su empleo en numerosas aplicaciones, como son el reconocimiento de lenguaje de signos, el desarrollo de video-fonos para sordomudos, el diseño de nuevos interfaces hombre-máquina,...

