

Arquitectura de microservicios para reconocimiento multi-dimensional de características utilizando un modelo de representación de conocimiento, basados en la percepción humana

Microservices architecture for multi-dimensional feature recognition using a knowledge representation model, both based on human perception

Lisardo Prieto González, Beatriz Puerta Hoyas, Antonio de Amescua Seco
Universidad Carlos III de Madrid (España)

DOI: <http://dx.doi.org/10.6036/8254>

El crecimiento de Internet y la aparición de nuevas tecnologías y tendencias, como el IoT (*Internet of Things*, Internet de las cosas) junto con los hábitos recientes, tales como el extenso uso de los teléfonos inteligentes y las redes sociales, desencadenaron un fenómeno denominado *Information Overload* (o sobrecarga de información) [1]. El volumen estimado de datos generados por los usuarios es cercano a los 1,25 millones de imágenes por minuto [2]. Sin embargo, la infraestructura de la Web está construida en base al texto. Incluso muchos de los principales motores de búsqueda se basan en texto. Para

el caso de las imágenes, utilizan (en caso de estar disponibles) las meta etiquetas de información asociadas, que proporcionan una representación textual de las imágenes, pero éste enfoque está cambiando hacia un mejor y más sofisticado esquema.

La motivación tras este trabajo es el emular la manera en la que los seres humanos asimilan la información a través de los sentidos, a fin de desarrollar un nuevo y amplio rango de aplicaciones (tales como biometría mejorada, análisis y seguimiento de vehículos, detección automática de contenido ilegal en sitios web, recomendadores y motores de búsqueda avanzados, entre otros) basadas en un sistema inteligente capaz de aprender autónomamente del entorno, refinar y extender el conocimiento existente y de combinar la información aprendida dentro de una base global de conocimiento, incluyendo un conjunto de propiedades

al menos tan grande como la variedad de sensores involucrados en la captura (ej. visual, acústico, táctil, químico, etcétera), para desarrollar conceptos nuevos, más complejos y completos, ideas y teoremas de forma eficiente.

El presente trabajo extiende uno anterior, centrado en la característica perceptual de la visión [3]. Los principales objetivos de dicho estudio son el encontrar un modelo adecuado para cubrir la falta de funcionalidades en las soluciones existentes, tales como el análisis de imágenes multidimensional, o el análisis semántico complejo, así como la extensión y refinado del marco de referencia propuesto en [4]. El estudio anterior devolvió resultados prometedores en términos de factibilidad de la solución y mejora de rendimiento comparado con el uso de arquitecturas tradicionales no distribuidas. Sin embargo, las pruebas fueron llevadas a cabo

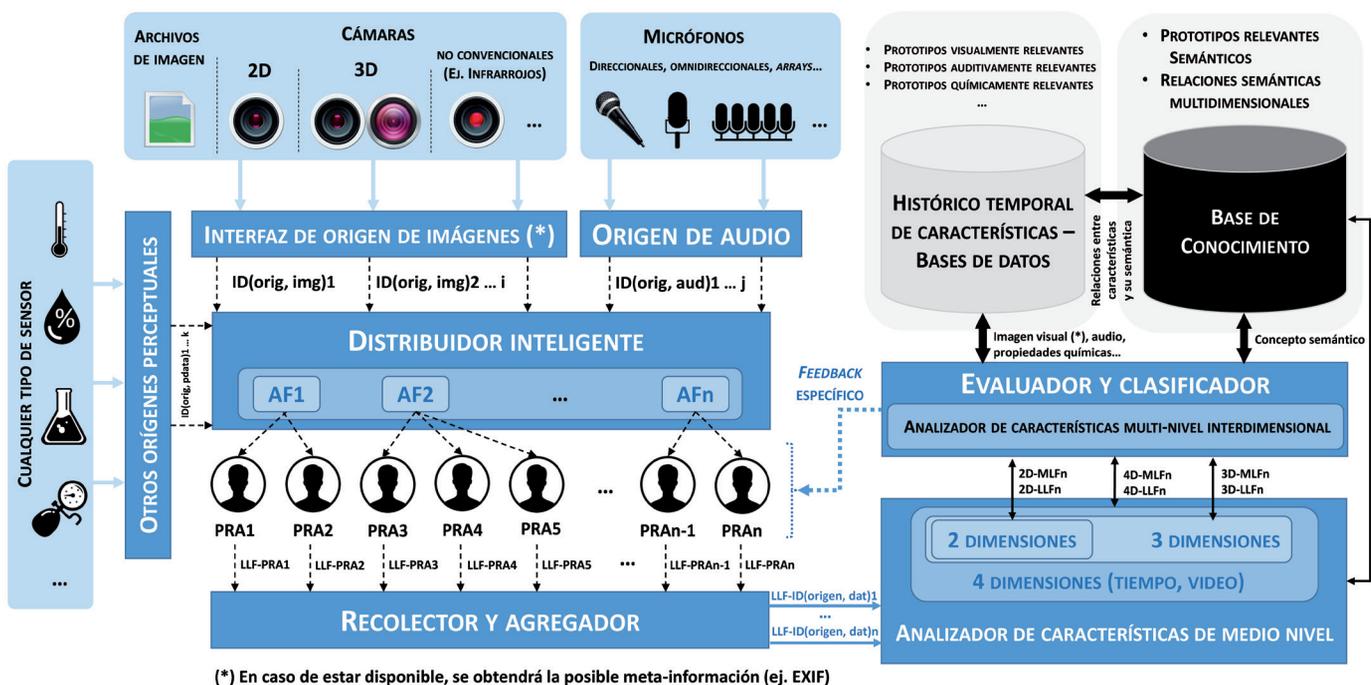


Fig. 1: Arquitectura y modelo reconocedor de características multi-dimensional

utilizando únicamente agentes inteligentes centrados en el reconocimiento simple de patrones, los cuales utilizaban algoritmos voraces simples. Tras actualizar dichos agentes para probar el rendimiento en un escenario más complejo de reconocimiento, aparecieron algunos problemas con la precisión. El nuevo escenario de pruebas consistió en el reconocimiento de frutas cítricas. Específicamente, un conjunto de naranjas, pomelos, limones y limas de diversos tamaños y puntos de maduración. Mediante el uso de cuatro redes neurales específicamente entrenadas para cada tipo de fruta, los resultados fueron bastante buenos en promedio, pero en varios casos, las probabilidades de que una fruta fuese erróneamente reconocida como una naranja en lugar de como un pomelo, o de ser erróneamente reconocida como un limón en lugar de como una lima (o vice-versa), fueron ligeramente elevadas. La arquitectura mejorada y basada en microservicios añadió, vía visión estéreo, el reconocimiento de medidas simples (anchura, altura y profundidad) de la caja circunscrita conteniendo a la fruta, así como el nivel de acidez de la misma (medido e introducido de forma manual en el sistema). Mediante el uso de esta nueva información perceptual, en conjunto con los resultados estadísticos devueltos por las redes neurales, el reconocimiento de cada ejemplar fue casi perfecto, pudiendo identificar inequívocamente cada fruta con su correspondiente clase.

Como conclusión, la adición de más fuentes de información, además de la visión, correspondientes a la información que puede ser percibida de manera similar a como lo hace el ser humano, permite tener mejor conceptualización semántica de los elementos reconocidos, pudiendo ser almacenada en una base general de conocimiento. Esos elementos conceptualizados semánticamente de una forma mejorada y natural también pueden ser utilizados para aumentar las capacidades de reconocimiento de los agentes inteligentes al considerar características adicionales relevantes, las cuales en algunos casos, podrían ser decisivas para el reconocimiento de objetos en el entorno, como ha sido observado en el experimento, o para determinar nuevas subclases en la taxonomía correspondiente al elemento general utilizado como prototipo de clase. El modelo propuesto ha demostrado un mayor rendimiento en términos de tiempo y precisión en comparación con los reconocedores individuales en el campo de pruebas, por lo que parece un enfoque novedoso e interesante hacia un sistema de reconocimiento natural multi-propósito automatizado, que todavía debe ser refinado y extendido.

REFERENCIAS

- [1] Eppler Martin J and Mengis Jeanne. "The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines". The Information Society. Vol.20-5 p.325-344. Nov. 2004 DOI: 10.1080/01972240490507974.
- [2] Griffith Erin. "Algorithms that can 'see' images on the Web". Fortune (online ed.). Jan. 2015.
- [3] Prieto González Lisardo, Puerta Hoyas Beatriz, and de Amescua Seco Antonio. "Distributed multi-dimensional feature recognition and multi-level semantic classification model for computer vision". DYNA New Technologies. Vol.3-1 Dec. 2016 DOI: <http://dx.doi.org/10.6036/NT7791>.
- [4] Prieto-González Lisardo, Stantchev Vladimir, and Colomo-Palacios Ricardo. "Applications of ontologies in knowledge representation of human perception". International Journal of Metadata, Semantics and Ontologies. Vol.9-1 p.74-80. 2014 DOI: 10.1504/IJMSO.2014.059128.

El poder en sus manos

HP ha reinventado la productividad en impresión de gran formato con la impresora HP PageWide XL 5000 para que usted tenga la posibilidad de potenciar su ahorro. Use el mismo dispositivo para trabajos en blanco y negro o en color. Y con una velocidad de hasta 14 páginas A1 por minuto, podrá imprimir con los mismos costes o costes más bajos que con las impresoras LED comparables¹.

Más información:

+34 930 031 133

www.hp.com/go/pagewidexl/5000series



¹ Basado en los costes de suministros y de mantenimiento. Costes de impresión en blanco y negro de documentos técnicos para impresoras LED comparables de volumen medio basados en impresoras para impresión en gran formato de documentos técnicos a velocidades de entre 8 y 13 páginas D/A1 por minuto a fecha de abril de 2015, y que representan más del 80 % de la cuota de impresoras LED de volumen medio en EE. UU. y Europa en 2015, según IDC. Costes de impresión en color de documentos técnicos para impresoras en color comparables de volumen medio basados en impresoras capaces de imprimir 4 páginas D/A1 por minuto o más y de menos de 50 000 USD para la impresión en gran formato de documentos técnicos a fecha de abril de 2015, y que representan más del 80 % del mercado en EE. UU. y Europa en 2014, según IDC. Para consultar los criterios de prueba, véase hp.com/go/pagewidexlclaims.