



## Oferta de contrato predoctoral

El **Grupo de Procesado Multimedia (GPM)** de la **Universidad Carlos III de Madrid** busca candidatos para solicitar un contrato predoctoral para la formación de doctores en la convocatoria 2015. El objetivo sería **realizar una tesis doctoral** en el marco del proyecto de investigación **“Saliency and Attention: MUltimodality, context-awaReness, self-Adaptation and bio-Inspiration”** (TEC2014-53390-P) financiado en la convocatoria 2014 de Proyectos de I+D del Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia.

El contrato tendrá una **duración máxima de 4 años**.

### REQUISITOS:

- Grado en Ingeniería de Telecomunicaciones, Informática o titulaciones afines.
- Se valorará muy positivamente el interés por temas relacionados con interfaces con humanos y procesos cognitivos.
- Buen expediente académico.
- Estar admitido o en disposición de ser admitido en un programa de doctorado o en un máster de investigación de la Universidad Carlos III de Madrid en el curso 2015-16.
- Cumplir los requisitos concretados en la convocatoria que se puede consultar en <http://www.boe.es/boe/dias/2015/06/12/pdfs/BOE-A-2015-6508.pdf>
- Realizar la solicitud entre los días **15 y 29 de junio a las 15:00**.

Los interesados pueden contactar por correo electrónico en las direcciones [gallardo@tsc.uc3m.es](mailto:gallardo@tsc.uc3m.es) o [carmen@tsc.uc3m.es](mailto:carmen@tsc.uc3m.es) enviando una carta de motivación, su CV y expediente académico.

### SOBRE EL GRUPO DE INVESTIGACIÓN:

El Grupo de Procesado Multimedia de la Universidad Carlos III de Madrid es un grupo de investigación joven y dinámico cuyas áreas principales de interés involucran el procesado de **voz, audio, imagen y video**, con especial énfasis en **extracción de contenido multimedia, tecnologías del habla y codificación de video** de última generación.

El GPM combina sus experiencia en el **análisis y la representación multimedia** (codificación, segmentación, detección de bordes, seguimiento de objetos, morfología, etc.) con un conocimiento profundo de cada uno de los modos individuales (regiones de interés, jerarquía y saliencia perceptuales, etc.) sólidamente basado en técnicas modernas de **aprendizaje máquina y descubrimiento de información** para construir **interpretaciones** útiles de los **datos multimedia** para gran variedad de aplicaciones desde anotación automática de video e imagen, interacción hombre-máquina, transcripción de habla, codificación de vídeo o predicción de insolación a corto plazo, etc.

### SOBRE EL PROYECTO:

Nuestro conocimiento del mundo está modelado por la percepción humana. Nuestras capacidades sensoriales y motoras nos permiten comprender e interaccionar con la realidad. La cognición es el resultado de estas interacciones. Imitar estas funciones del cerebro es uno de los retos científicos más estimulantes que los tecnólogos han abrazado con el nombre de computación cognitiva y cuyo objetivo es construir máquinas inteligentes inspiradas biológicamente.



La saliencia es un mecanismo cognitivo por el cual priorizamos ciertos estímulos: nuestro cerebro toma decisiones sobre qué es relevante y qué no en cada situación en su proceso de exploración del mundo.

Desde una perspectiva investigadora identificamos las siguientes direcciones clave para hacer avanzar esta tecnología:

1. **Multimodalidad:** los humanos no podemos concebir el mundo usando una sola modalidad. Sin embargo la mayoría de los resultados de investigación se especializan en una en particular. Basándonos en nuestra experiencia proponemos la integración de dos de las modalidades humanas principales: auditiva y visual. Su integración pivota sobre dos conceptos: primero, la adopción de un marco de evaluación basado en la teoría de la información que mejora la interpretabilidad de los resultados y favorece la fusión y segundo, la comprensión del papel del tiempo.

2. **Bio-inspiración:** los algoritmos de aprendizaje profundo han tenido un gran impacto en una gran cantidad de tareas y se han utilizado también para construir modelos de saliencia, principalmente, para la fusión de mapas provenientes de diferentes características. Sin embargo, su aplicación a la saliencia auditiva no se ha explorado aún. La morfología matemática también se ha mostrado ventajosa para imitar las propiedades psico-acústicas del sistema auditivo.

3. **Consciencia del contexto y auto-adaptación:** en contraste con la abundante literatura existente sobre la saliencia de abajo-arriba (la basada en características de bajo nivel o estímulos), el modelado de arriba-abajo de la atención visual sigue siendo un problema abierto ya que sus soluciones son, en general, dependientes de la tarea. Nos proponemos realizar su integración asumiendo un marco general que permita aprender y adaptar a cada contexto particular los objetivos del usuario y sus relaciones con los estímulos. La capacidad de adquirir conocimiento a través del descubrimiento de clases latentes, tópicos, tareas o eventos así como la adopción del análisis de datos exploratorio guiado por expertos es nuestra propuesta en este área.

Desde un punto de vista metodológico, adoptamos una perspectiva de usuario final ya que el conocimiento de la relevancia perceptual de objetos audiovisuales se puede aplicar a multitud de problemas: p.ej. reconocimiento de objetos, clasificación de acciones o detección de eventos, etc. Esto no sólo conlleva el desarrollo de algoritmos que incorporen la saliencia sino también de modificaciones en el protocolo de evaluación, pasando de un enfoque tradicional de alineamiento de los mapas de saliencia con las fijaciones humanas a otro con mayor significado.

Así, el propósito de este proyecto tiene dos vertientes: primero, contribuir al avance de la tecnología en cada uno de los tres ejes anteriores y segundo, desarrollar un conjunto de herramientas multipropósito que puedan ser ensambladas en diferentes aplicaciones tales como detección de eventos, anotación e indexación de vídeo, sistemas de recomendación, diagnóstico por imagen médica, etc.