ASIGNATURA DE MÁSTER:

VISIÓN ARTIFICIAL

Curso 2013/2014

(Código:31101235)

1.PRESENTACIÓN

La Visión Artificial es una disciplina de creciente interés en el ámbito científico-técnico, en áreas tan diversas como el análisis de imágenes médicas, robótica, teledetección o control de calidad. Su objetivo fundamental, planteado un problema, es la extracción automatizada de información significativa a partir de imágenes digitales. Por esto, forma parte de una de las líneas de especialización dentro de este programa.

Para extraer esta información, los modelos dominantes en visión artificial suponen la transformación de los datos sensoriales en descripciones significativas de la escena mediante la utilización de etapas lógicas, que emplean progresivamente representaciones más y más abstractas de la imagen original. Para cada una de estas etapas se señala el conocimiento específico que es preciso inyectar, así como los modelos matemáticos y algoritmos adecuados para su representación y uso. Se hará especial énfasis en las etapas de alto nivel, donde la IA aplicada puede realizar sus mayores aportaciones.

2.CONTEXTUALIZACIÓN

Esta asignatura se encuadra en el módulo "Aplicaciones de la IA" dentro de la especialidad "IA.1: Sistemas Inteligentes de diagnóstico, planificación y control" de la titulación de posgrado "Master de IA Avanzada. Fundamentos Métodos y Aplicaciones".

El contenido del curso se estructura entorno a la idea de construir sistemas de visión completos. Además de definir la terminología utilizada habitualmente en un sistema de visión artificial, se mostrará al alumno la envergadura del problema, recalcando:

1. La necesidad descomponer la tarea en diferentes subtareas y en distintos niveles de descripción o representación con grado creciente de semántica;
2. La necesidad de inyectar conocimiento en cada una de las etapas de procesado para poder llegar a una solución sin incertidumbre del problema planteado.

Finalmente, para cada una de las subtareas descritas, se presentarán métodos de IA que se utilizan en su resolución.

3.REQUISITOS PREVIOS RECOMENDABLES

Ninguno diferente de los generales de acceso a este programa de postgrado orientado a la investigación. Sin embargo, dado el carácter práctico de la asignatura, es conveniente que el alumno esté familiarizado con la terminología usada en procesado digital de imágenes y tenga conocimientos de aspectos de la IA relacionados con la representación y uso del conocimiento y con el aprendizaje.

4.RESULTADOS DE APRENDIZAJE

- Familiarizarse con la terminología básica utilizada en visión.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
9B0F5D99A0DDA458AB58D5A780F955A3

- Conocer las dificultades que entraña la visión artificial.
- Conocer las etapas de procesado en que se suele descomponer un sistema de visión artificial.
- Distinguir los distintos niveles de descripción con creciente grado de semántica que nos encontramos a lo largo del proceso de interpretación de una imagen o escena.
- Familiarizarse con las estructuras de datos utilizadas y con las librerías software existentes.

5.CONTENIDOS DE LA ASIGNATURA

Parte 1. Introducción a la visión artificial

Parte 2. Visión de bajo nivel

Parte 3. Visión de medio nivel

Parte 4. Visión de alto nivel

Objetivos por tema y orientaciones breves

Parte 1: Introducción a la visión artificial.

1. Introducción a la visión artificial

    1.1. Introducción

    1.2. Terminología

    1.3. Etapas del procesado

    1.4. Niveles de descripción

    1.5. Introducción histórica

Objetivo:
Este primer tema es de carácter introductorio, sin embargo, resulta fundamental para entender el conjunto de la asignatura. Además de definir la terminología utilizada habitualmente en un sistema de visión artificial, se mostrará al alumno la envergadura del problema, recalcando: 1) la necesidad descomponer la tarea en diferentes subtareas y en distintos niveles de descripción o representación con grado creciente de semántica; 2) la necesidad de inyectar conocimiento en cada una de las etapas de procesado para poder llegar a una solución sin incertidumbre del problema planteado.
Este objetivo global puede descomponerse en los siguientes objetivos más concretos:

- Conocer las dificultades que entraña la visión artificial.
- Familiarizarse con la terminología básica utilizada en visión.
- Conocer las etapas de procesado en que se suele descomponer un sistema de visión artificial.
- Distinguir los distintos niveles de descripción con creciente grado de semántica que nos encontramos a lo largo del proceso de interpretación de una imagen o escena.
- Familiarizarse con las estructuras de datos utilizadas y con las librerías software existentes.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
9B0F5D99A0DDA458AB58D5A780F955A3

Orientaciones:

El material docente desarrolla los conceptos fundamentales del capítulo. Posteriormente, el alumno realizará distintas actividades de consulta bibliográfica y de práctica sobre diversos ejemplos para la consolidación de los conceptos objetivo.

Parte 2: Visión de bajo nivel.

2. Visión de bajo nivel: Preprocesado

    2.1. Introducción

    2.2. Filtrado de imágenes digitales: Eliminación del ruido y Realce de características

    2.3. Detectores de bordes

    2.4. Transformaciones basadas en las intensidades del nivel de gris

    2.5. Flujo óptico y campo de flujo óptico

3. Visión de bajo nivel: Segmentación

    3.1. Introducción

    3.2. Segmentación basada en la detección de fronteras

    3.3. Segmentación basada en la umbralización

    3.4. Segmentación basada en la agrupación de píxeles

    3.5. Segmentación basada en el movimiento

Objetivo:

La visión de bajo nivel comprende las transformaciones realizadas sobre la imagen captada con el objetivo de reducir el ruido y aumentar la información contenida en los datos. Los procesos de bajo nivel son guiados por los propios datos, es decir, sin conocimiento del dominio de aplicación.

El contenido se organiza en dos capítulos, uno dedicado al preprocesado de la imagen y otro dedicado a la segmentación. El preprocesado engloba el conjunto de tareas dedicadas a eliminar el ruido presente en la imagen y a realzar aquellas características de interés con el fin de mejorar la imagen captada por el sensor y así facilitar el procesamiento posterior. La segmentación tiene por objetivo descomponer la imagen en distintas zonas con alguna característica visual común.

Este objetivo global puede descomponerse en los siguientes objetivos más concretos:

- Conocer las técnicas básicas de visión de bajo nivel y sus características: descripción del algoritmo, coste computacional, suposiciones e influencia en etapas posteriores.

- Relacionar estas técnicas con el resto de etapas de un sistema de visión.

- Capacitar al alumno para implementar estas técnicas e integrarlas en aplicaciones de visión.

Orientaciones:

El material docente desarrolla los conceptos fundamentales del capítulo. Posteriormente, el alumno realizará distintas actividades de consulta bibliográfica y de práctica sobre diversos ejemplos para la consolidación de los conceptos objetivo.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
9B0F5D99A0DDA458AB58D5A780F955A3

Parte 3: Visión de medio nivel.

4. Visión de medio nivel: Segmentación con conocimiento

- 4.1. Transformada de Hough
- 4.2. Contornos activos

5. Visión de medio nivel: Seguimiento

- 5.1. Introducción
- 5.2. Estimadores demovimiento
- 5.3. Seguimiento de objetosmediante contornos activos

Objetivo:

Normalmente, la informacion obtenida en la etapa de preprocesado contiene mucho ruido, de modo que existen gran cantidad de bordes y/o regiones en la imagen que no corresponden a superficies de los objetos de interés, mientras que otras superficies, que sí son significativas, no quedan bien definidas. Por este motivo, se han desarrollado métodos que tratan de segmentar la imagen en blobs con mayor significado semántico, esto es, que estén relacionados con las superficies de los objetos existentes en la imagen. La entrada a estos métodos serán los elementos primitivos obtenidos en la etapa anterior y la salida serán los blobs asociados a las superficies o los contornos de los objetos.

Por tanto, podríamos decir que la visión de medio nivel se caracteriza por la inyección de conocimiento geométrico genérico. En el capítulo 4, se describirán distintas técnicas de segmentación que utilizan cierto conocimiento de las estructuras esperadas para controlar el proceso de selección de los elementos primitivos de bajo nivel que intervienen en la definición de contornos y superficies de los objetos. Otros temas que encajan dentro de esta categoría, como recuperación de la forma a partir de la textura, del movimiento, de la sombra, de par estéreo, se dejan fuera de esta introducción a la visión artificial por motivos de espacio.

Este objetivo global puede descomponerse en los siguientes objetivos más concretos:

- Conocer las técnicas básicas de visión de medio nivel y sus características: descripción del algoritmo, coste computacional, suposiciones e influencia en etapas posteriores.

- Identificar el conocimiento utilizado en el desarrollo de la tarea.

Relacionar las técnicas de segmentación con el resto de etapas de un sistema de visión.

- Capacitar al alumno para implementar estas técnicas e integrarlas en aplicaciones de visión.

Orientaciones:

El material docente desarrolla los conceptos fundamentales del capítulo. Posteriormente, el alumno realizará distintas actividades de consulta bibliográfica y de práctica sobre diversos ejemplos para la consolidación de los conceptos objetivo.

Parte 4: Visión de alto nivel.

6. Visión de alto nivel: Modelado de objetos

- 6.1. Introducción

- 6.2. Representación de estructuras geométricas bidimensionales

- 6.3. Representación de estructuras geométricas tridimensionales

- 6.4. Representación en un espacio de características discriminantes

7. Visión de alto nivel: Reconocimiento

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
9B0F5D99A0DDA458AB58D5A780F955A3

7.1 Reconocimiento de patrones

7.2 Métodos estadísticos

7.3 Métodos estructurales y sintácticos

7.4 Métodos basados en la apariencia

7.5 Reconocimiento de actividades

7.6 Aplicaciones

8. Visión de alto nivel: Interpretación de imágenes

8.1 Arquitecturas básicas

8.3 Ontología de la tarea

8.4 Modelado del dominio

8.6 Aplicaciones

Objetivo:

El objetivo final del sistema de visión es la descripción de la escena orientada a la tarea. Esta descripción consiste, desde la simple identificación o el reconocimiento de los objetos de interés, hasta la determinación de las actividades que están llevando a cabo un objeto o un conjunto de objetos.

La descripción orientada a la tarea es una descripción abstracta ("el hombre, llamado Pepe, está sentado y comiendo palomitas"). Para llegar a esta descripción a partir de características de la imagen, es necesario reconocer los objetos de interés ("que es un hombre", "que es Pepe", "que está sentado", "que hay palomitas", "que está comiendo", etc). Todos estos objetos pertenecen al dominio de aplicación y, para reconocerlos, es necesario primero modelarlos y, posteriormente, establecer la correspondencia entre el modelo y los datos procedentes de la imagen.

Además, debido a la incertidumbre asociada a las distintas etapas del sistema de visión, salvo en situaciones muy simples, en las que se puede utilizar una estrategia de control guiada por los datos, es necesario incluir conocimiento del dominio en la estrategia de control del sistema de visión para guiar los procesos de bajo y medio nivel, de modo que se facilite el reconocimiento de los objetos de interés y esto haga posible la descripción coherente de la escena.

En resumen, la visión de alto nivel se caracteriza por el uso intenso de conocimiento del dominio e implica procesos de razonamiento abstractos y planificación de operaciones orientadas a cumplir el objetivo final: la descripción de la escena orientada a la tarea.

La organización del tema es la siguiente. En el capítulo 6, se estudiará el modelado de los objetos, en concreto, la representación de la forma. En el capítulo 7, se estudiarán las técnicas de reconocimiento de patrones como método de reconocimiento de objetos y de actividades y se distinguirán las fases de diseño (aprendizaje) y de operación (clasificación). Finalmente, en el capítulo 8, se estudiarán las arquitecturas de control de los sistemas de interpretación de imágenes y se analizarán nuevos enfoques desde la perspectiva de la ingeniería del conocimiento.

Este objetivo global puede descomponerse en los siguientes objetivos más concretos:

- Conocer la problemática de la visión de alto nivel.

- Conocer y manejar las arquitecturas utilizadas en interpretación de imágenes.

- Analizar un problema de visión artificial, representar el conocimiento del dominio necesario y definir una arquitectura de visión que resuelva el problema.

Orientaciones:

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
9B0F5D99A0DDA458AB58D5A780F955A3

El material docente desarrolla los conceptos fundamentales del capítulo. Posteriormente, el alumno realizará distintas actividades de consulta bibliográfica y de práctica sobre diversos ejemplos para la consolidación de los conceptos objetivo.

6.EQUIPO DOCENTE

- MARGARITA BACHILLER MAYORAL
- MARIANO RINCON ZAMORANO

7.METODOLOGÍA

Adaptada a las directrices del EEES, de acuerdo con el documento del IUED. La metodología docente será la general del programa de postgrado, junto a actividades y enlaces con fuentes de información externas. Existe material didáctico propio preparado por el equipo docente.

La asignatura no tiene clases presenciales. Los contenidos teóricos se impartirán a distancia, de acuerdo con las normas y estructuras soporte telemático de la enseñanza en la UNED. El material docente incluye un resumen de los contenidos de cada tema y distintos tipos de actividades relacionadas con la consulta bibliográfica y la implementación de los métodos descritos en la teoría.

Tratándose de un master orientado a la investigación, las actividades de aprendizaje se estructuran en torno al estado del arte en cada una de las materias del curso y a los problemas en los que se va a focalizar en el proyecto final, sobre el que se realizará la evaluación.

8.BIBLIOGRAFÍA BÁSICA

Comentarios y anexos:
Dentro del curso virtual se proporcionará material didáctico introductorio de cada tema, que se complementará con lecturas recomendadas.

9.BIBLIOGRAFÍA COMPLEMENTARIA

Comentarios y anexos:

- J. González: "Visión por computador". Paraninfo, 1999.

- A. De la Escalera: "Visión por computador. Fundamentos y métodos". Prentice Hall, 2001.

- M. Sonka, V. Hlavac y R. Boyle: "Imagen Processing, Analysis and Machine Vision". Chapman & Hall Computing, 1993.

- G. Pajares y J. M. de la cruz: " Visión por computador. Imágenes digitales y aplicaciones". Ra-Ma, 2001.

- Blake, A. and Isard, M.: "Active Contours". Springer, 2000.

- T. Zhao and R. Nevatia. "Tracking Multiple Humans in Complex Situations", IEEE trans. on Pattern Analysis and Machine Intelligence, 26(9), 1208-1221, 2004

- Richard O. Duda, Peter E. Hart y David G. Stork: " Pattern Classification". John Wiley & Sons. 2001.

- H.H.Nagel. "Steps toward a cognitive vision system". AI Magazine 25 (2), pp. 31-50. 2004

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
9B0F5D99A0DDA458AB58D5A780F955A3

- R. Nevatia, J. Hobbs and B. Bolles, "An Ontology for Video Event Representation", IEEE Workshop on Event Detection and Recognition, June 2004

- Hongeng, S. and Nevatia, R. and Bremond, F. " Video-based event recognition: activity representation and probabilistic recognition methods". Computer Vision and Image Understanding, pags.129-162, 2004.

- I. Haritaoglu, D. Harwood, and L.S. Davis, "W4: Real-time Surveillance of People and Their Activities," PAMI, 22(8), pp. 809-830, Aug. 2000.

- Green, R.D. Ling Guan. Quantifying and recognizing human movement patterns from monocular video Images-part I: a new framework for modeling human motion. IEEE Transactions on Circuits and Systems for Video Technology. pgs 179- 190. 14(2). 2004.

- Green, R.D. Ling Guan. Quantifying and recognizing human movement patterns from monocular video Images-part II: applications to biometrics. IEEE Transactions on Circuits and Systems for Video Technology. pgs 191- 198. 14(2). 2004.

- Wang, L., Hu, W., Tan, T. Recent developments in human motion analysis. Pattern Recognition. 36(3), pp. 585-601, March 2003.

10.RECURSOS DE APOYO AL ESTUDIO

La plataforma de e-Learning Alf, proporcionará el adecuado interfaz de interacción entre el alumno y sus profesores. aLF es una plataforma de e-Learning y colaboración que permite impartir y recibir formación, gestionar y compartir documentos, crear y participar en comunidades temáticas, así como realizar proyectos online.

Se ofrecerán las herramientas necesarias para que, tanto el equipo docente como el alumnado, encuentren la manera de compaginar tanto el trabajo individual como el aprendizaje cooperativo.

11.TUTORIZACIÓN Y SEGUIMIENTO

La tutorización y el seguimiento de los alumnos se llevará a cabo a través de la plataforma de e-Learning Alf o por cualquier otro medio de contacto (e-mail, teléfono, etc)

Dª Margarita Bachiller Mayoral

Guardia: Lunes y Martes de 14:30 a 16:30 horas

Asistencia al Estudiante: Martes y Miércoles de 10:00 a 14:00 horas

12.EVALUACIÓN DE LOS APRENDIZAJES

La evaluación global de la asignatura se hará mediante el trabajo individual y personal del alumno a través de varios trabajos que ha de entregar en fechas concretas a lo largo del curso. No obstante, si el equipo docente lo considera oportuno, se puede sugerir la formación de grupos de trabajo que ayuden al alumno en su desarrollo individual. Dichos grupos serán constituidos por el equipo docente dependiendo del grado de interés mostrado en la memoria M2 especificada a continuiación.

Dado que el desarrollo temporal del curso cubre un semestre, la planificación de actividades se repartirá en intervalos temporales por temas hasta mediados de Junio, dejando un plazo hasta el 10 de septiembre para mejorar o ampliar aquellos trabajos que el equipo docente considere insuficientes. La secuencia temporal es la siguiente:

- Desde el inicio del curso hasta el 15 de Enero:

(1) Lectura del material proporcionado por el equipo docente y realización de actividades propuestas al

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
9B0F5D99A0DDA458AB58D5A780F955A3

final de cada tema (actividades de carácter voluntario);

(2) Realización de una pequeña práctica cuyo objetivo es la toma de contacto con las librerías de visión seleccionadas. El alumno debe entregar una memoria (M1) donde se describa el trabajo realizado.

- Desde el 16 de Enero hasta el 28 de Febrero:

Análisis del problema planteado para el trabajo final, revisión bibliográfica y planteamiento de la solución al problema elegido.

Al inicio de este periodo estarán accesibles las propuestas de trabajos para el curso. Al final del periodo, el alumno debe entregar una memoria (M2) con el estado del arte en el tema elegido y la propuesta de solución que pretende desarrollar.

- Desde el 1 de Marzo hasta el 15 de Junio: desarrollo de la propuesta descrita en la memoria M2. Al final de este periodo, el alumno deberá entregar una memoria (M3) con:

- Revisión de la memoria M2 a partir del trabajo desarrollado.

- Validación del sistema desarrollado y evaluación.

- Código fuente, librerías utilizadas, imágenes utilizadas durante el desarrollo.

- Manual de usuario, ejemplos de uso.

NOTA: La memoria M1, M2 y M3 son de carácter obligatorio y tienen que ser entregadas en las fechas indicadas. El trabajo final del curso consiste en estas tres memorias.

13.COLABORADORES DOCENTES

Véase equipo docente.

14.Software recomendado

Una parte importante de las actividades del curso está asociada al desarrollo de prácticas de análisis y/o de implementación de operadores de visión artificial.

El equipo docente recomendará al principio del curso las librerías más adecuadas. Se utilizarán librerías de uso libre, tales como OpenCV o ITK (en C++), y/o el entorno de desarrollo MATLAB, que dispone de una librería específica para el tratamiento de imágenes.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
9B0F5D99A0DDA458AB58D5A780F955A3