

Visió 3D

Codi: 43090
Crèdits: 6

Titulació	Tipus	Curs	Semestre
4314099 Visió per Computador / Computer Vision	OB	0	1

Professor/a de contacte

Nom: Gloria Haro Ortega

Correu electrònic: Desconegut

Equip docent

Josep Ramon Casas Pla

Javier Ruiz Hidalgo

Utilització d'idiomes a l'assignatura

Llengua vehicular majoritària: anglès (eng)

Equip docent extern a la UAB

Antonio Agudo

Federico Sukno

Pedro Cavestany

Prerequisits

Grau en Enginyeria, Matemàtiques, Física o similar.

Objectius

L'objectiu d'aquest mòdul és aprendre els principis de la reconstrucció en 3D d'un objecte o d'una escena a partir de múltiples imatges o vídeos estereoscòpics. Per això, primer s'introdueixen els conceptes bàsics de la geometria projectiva i l'espai 3D. La resta d'aspectes i aplicacions teòriques es basen en aquestes eines bàsiques. S'estudiarà el mapeig del món 3D al pla de la imatge, tot introduint diferents models de càmera, els seus paràmetres i la forma d'estimar-los (calibració de la càmera i auto-calibració). S'estudiarà la geometria que relaciona un parell de vistes. Tots aquests conceptes s'aplicaran per obtenir una reconstrucció 3D en les dues principals situacions possibles: càmeres calibrades o sense calibrar. En particular, aprendrem a: estimar la profunditat dels punts d'una imatge, extreure els punts 3D subjacents donats un conjunt de correspondències puntuals a les imatges, generar vistes noves, estimar l'objecte 3D donat un conjunt d'imatges calibrades en color o imatges binàries, i estimar un conjunt de punts 3D donat un conjunt d'imatges no calibrades. S'estudiarà la representació 3D en voxels i malles. Explicarem la reconstrucció i modelatge a partir de les dades de Kinect, com un model particular de sensors que proporcionen una imatge de l'escena junt amb les seves profunditats. Finalment, veurem algunes tècniques per processar núvols de punts 3D. Els conceptes i tècniques apreses en aquest mòdul s'utilitzen en aplicacions reals que van des de la realitat augmentada, la digitalització d'objectes, la captura demoviment, la síntesi de noves vistes, la generació d'efectes especials, la robòtica, etc.

Competències

- Assumir tasques de responsabilitat en la gestió de la informació i el coneixement.
- Comprendre, analitzar i sintetitzar els coneixements avançats que hi ha en l'àrea, així com proposar idees innovadores.
- Conceptualitzar alternatives de solucions complexes per a problemes de visió i crear prototips que demostrin la validesa del sistema proposat.
- Identificar els conceptes i aplicar les tècniques fonamentals més adequades per resoldre els problemes bàsics de la visió per computador.
- Planificar, desenvolupar, avaluar i gestionar solucions per a projectes en els diferents àmbits de la visió per computador.
- Que els estudiants sàpiguin aplicar els coneixements adquirits i la seva capacitat de resolució de problemes en entorns nous o poc coneguts dins de contextos més amplis (o multidisciplinaris) relacionats amb la seva àrea d'estudi.
- Que els estudiants tinguin les habilitats d'aprenentatge que els permetin continuar estudiant, en gran manera, amb treball autònom a autodirigit.
- Seleccionar les eines de programari i els conjunts d'entrenament més adequats per desenvolupar les solucions per als problemes de visió per computador.
- Tenir coneixements que aportin la base o l'oportunitat de ser originals en el desenvolupament o l'aplicació d'idees, sovint en un context de recerca.
- Treballar en equips multidisciplinaris.

Resultats d'aprenentatge

1. Assumir tasques de responsabilitat en la gestió de la informació i el coneixement.
2. Comprendre, analitzar i sintetitzar els coneixements avançats que hi ha en l'àrea, així com proposar idees innovadores.
3. Identificar els problemes bàsics que s'han de resoldre en la recuperació de la informació 3D de les escenes, així com els algorismes específics.
4. Identificar les millors representacions que es puguin definir per resoldre problemes de recuperació d'informació 3D.
5. Planificar, desenvolupar, avaluar i gestionar una solució a un problema concret de reconstrucció 3D d'escenes.
6. Que els estudiants sàpiguin aplicar els coneixements adquirits i la seva capacitat de resolució de problemes en entorns nous o poc coneguts dins de contextos més amplis (o multidisciplinaris) relacionats amb la seva àrea d'estudi.
7. Que els estudiants tinguin les habilitats d'aprenentatge que els permetin continuar estudiant, en gran manera, amb treball autònom a autodirigit.
8. Seleccionar les tècniques apreses i entrenar-les per a solucionar un projecte concret de reconstrucció 3D d'escenes.
9. Tenir coneixements que aportin la base o l'oportunitat de ser originals en el desenvolupament o l'aplicació d'idees, sovint en un context de recerca.
10. Treballar en equips multidisciplinaris.

Continguts

1. Introducció i aplicacions.
2. Geometria projectiva 2D. Transformacions planars.
3. Estimació de la homografia. Rectificació afí i mètrica.
4. Geometria projectiva 3D i transformacions. Models de càmeres.
5. Calibració de càmera. Estimació de posa.
6. Geometria epipolar. Matriu fonamental Matriu essencial. Extracció de matrius de càmera.
7. Càlcul de la matriu fonamental. Rectificació de la imatge.
8. Mètodes de triangulació. Càlcul de profunditat. Síntesi de noves vistes.
9. Estèreo multi-vista. Structure from motion.
10. Autocalibració. Bundle adjustment.
11. Sensors 3D (kinect).

12. Processament de núvols de punts.

Metodologia

Sessions supervisades:

- Sessions de teoria, on els professors expliquen continguts generals sobre els diferents temes. Alguns d'ells s'empraran per resoldre els problemes.

Sessions dirigides:

- Sessions de projecte, on els problemes i els objectius dels projectes seran presentats i discutits, els estudiants interactuaran amb el coordinador del projecte sobre problemes i idees per resoldre el projecte (aprox. 1 hora/setmana).
- Sessió de presentació, on els estudiants fan una presentació oral sobre com han resolt el projecte i una demostració dels resultats.
- Sessió d'examen, on els estudiants són avaluats de forma individual sobre el coneixement obtingut i les habilitats per resoldre problemes.

Treball autònom:

- L'estudiant estudiarà de forma autònoma i treballarà els materials derivats de les sessions teòriques.
- L'estudiant treballarà en grups per a resoldre els problemes del projecte amb entregues de:
 - Codi
 - Informes
 - Presentació oral

Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Sessions teòriques	20	0,8	3, 4, 9
Tipus: Supervisades			
Sessions de seguiment de projectes	8	0,32	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Tipus: Autònomes			
Treball autònom	113	4,52	1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Avaluació

La nota final d'aquest mòdul es calcularà amb la següent fórmula:

$$\text{Nota final} = 0,4 \times \text{Examen} + 0,55 \times \text{Projecte} + 0,05 \times \text{Assistència}$$

on,

Examen: és la nota obtinguda a l'examen del mòdul (ha de ser ≥ 3)

Assistència: és la nota derivada del control d'assistència a les sessions (mínim 70%)

Projecte: és la nota proporcionada pel coordinador del projecte a partir del seguiment setmanal del projecte i dels lliuraments. Tot plegat amb criteris específics com ara:

- Desenvolupament de codis (estil, comentaris, etc.).
- Participació en sessions de discussió i en el treball en equip (avaluacions entre membres).
- Lliurament de les parts obligatòries i opcionals del projecte.
- Informe (justificació de les decisions en el desenvolupament del projecte).
- Presentació (presentació i demostració del projecte desenvolupat).

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Assistència a sessions	0.05	0,5	0,02	1, 2, 3, 4, 7
Examen	0.4	2,5	0,1	1, 2, 4, 5, 6, 8, 9
Projecte	0.55	6	0,24	1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Bibliografia

Llibres:

1. O. Faugeras, *Three-dimensional computer vision: a geometric viewpoint*, MIT Press, cop. 1993.
2. O. Faugeras, Q.T. Loung, *The geometry of multiple images*, MIT Press, 2001.
3. D. A. Forsyth, J. Ponce, *Computer vision: a modern approach*, Prentice Hall, 2003.
4. R. I. Hartley, A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, 2000.
5. R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2011.

Tutorials:

1. Y. Furukawa and C. Hernández, *Multi-View Stereo: A Tutorial*, Foundations and Trends® in Computer Graphics and Vision, vol. 9, no. 1-2, pp.1-148, 2013.
2. T. Moons, L. Van Gool, M. Vergauwen, *3D Reconstruction from Multiple Images Part 1*, Principles, Foundations and Trends® in Computer Graphics and Vision, vol. 4: no. 4, pp 287-404, 2010.