

Kinect for Windows V2



Què és?

El sensor Microsoft Kinect for Windows V2 (K4W2) és probablement el model més popular i representatiu de les càmeres RGB de profunditat i baix cost (anomenades càmeres RGB-D o càmeres de profunditat –depth cameras en anglès) desenvolupades recentment i, per aquest motiu, s'utilitza àmpliament en una àmplia gamma de tecnologies i aplicacions científiques i de desenvolupament (R + D). Les càmeres RGB-D, apart de proporcionar imatges en color de l'objecte mesurat, també proporcionen informació de les distàncies dels punts de la seva superfície a la càmera, es a dir, estimen la profunditat de cada píxel de l'escena gravada. El sensor Kinect for Windows V2 es la segona generació del dispositiu Kinect desenvolupat per l'empresa Microsoft Corporation, EUA. Dissenyat per ser utilitzat com a complement de videojocs domèstics, el sensor K4W2 es va desenvolupar per al seu ús en entorns interiors i, particularment, amb nivells d'il·luminació baixos. És per això que la majoria de desenvolupaments tecnològics i treballs de recerca se centren en l'ús de Kinect en aplicacions d'interior, si be, en condicions de baixos nivells d'il·luminació o utilitzant llum artificial durant la nit, pot utilitzar-se en aplicacions a l'exterior, com en el cas de les aplicacions a l'agricultura. Les mesures del sensor K4W2 en mode estàtic, és a dir, sense moviment, són similars a les que realitza una càmera de vídeo estacionària, però amb la diferència que K4W2 també proporciona dades de distància, cosa que permet obtenir núvols de punts 3D. Així, en aquest mode operatiu, el sensor Kinect té un rendiment similar a un sistema LiDAR terrestre (TLS) estacionari, tot i que amb un abast més curt i un camp de visió (FOV) més reduït. Els sensors RGB-D de baix cost disponibles comercialment, com ara el sensor K4W2, poden ser una alternativa interessant als sistemes mòbils

Activitat finançada a través de l'Operació 01.02.01 de Transferència Tecnològica del Programa de desenvolupament rural de Catalunya 2014-2022





terrestres (MTLS) basats en LiDAR, ja que l'alta precisió pot no ser crítica en determinades aplicacions, per exemple en agricultura. Per això, cal incorporar un sistema per aconseguir el desplaçament d'aquests sensors al llarg d'una trajectòria determinada. A més, aquests sensors poden proporcionar informació addicional, com ara color i infrarojos, per a cada punt del núvol.

Característiques generals

El sensor Kinect for Windows V2 està format per una càmera RGB passiva i una càmera del tipus Time of Flight, ToF, activa. Mitjançant la càmera ToF, el sensor emet llum làser en múltiples direccions dins del seu camp de visió i mesura la distància als objectes detectats pels diferents feixos de llum mitjançant el principi del temps de vol –time of flight-, de manera similar a un LiDAR del tipus temps de vol: mesurant el temps transcorregut entre l'emissió i el retorn de la llum entre el sensor i l'objecte. El resultat de la mesura és un núvol de punts amb fins a set variables associades a cadascun dels punts o píxels: 1) informació d'ubicació, normalment coordenades cartesianes locals x, y i z relacionades amb l'origen de les coordenades del sensor, 2) informació del color, generalment en format RGB i 3) intensitat infraroja (IR) de la llum retrodispersada. Les dades de color i IR proporcionen informació addicional per permetre l'extracció de les característiques de l'objecte mitjançant algorismes de post-processament. El sensor proporciona les dades del color, reflectància a l'infraroig (IR) i distància (profunditat) de punts situats a la superfície de l'objecte mesurat amb una velocitat màxima de 30 fotogrames per segon (30 fps). Per mesurar la reflectància corresponent a l'infraroig i estimar la profunditat, utilitza una càmera IR amb una resolució de 512 × 424 píxels i un camp de visió (FOV) de 70,6 × 60 graus. D'altra banda, la càmera de color (RGB) té una resolució de 1920 × 1080 píxels i un camp de visió de 84,1 × 53,8 graus. Els camps de visió d'ambdues càmeres no són, per tant, coincidents. Les mesures infraroges es basen en la detecció activa, mentre que la profunditat s'estima segons el principi del temps de vol. Bàsicament, consisteix a mesurar la quantitat de llum rebuda pels detectors sincronitzats que funcionen de manera complementària. A més, el sensor té un filtre de llum ambiental integrat que detecta si un píxel està saturat i restableix el valor de reflectància mesurat per aquest píxel. Aquestes característiques fan que el dispositiu de detecció sigui prou versàtil i robust en diverses aplicacions a l'aire lliure.

A la **Taula 1** es recullen les principals característiques del sensor Kinect for Windows V2.



Taula 1. Principals característiques del sensor Kinect for Windows V2.

Resolució de la càmera de color (píxels)	1920 x 1080 a 30 fps
Resolució de la càmera de profunditat (píxels)	512 x 424 a 30 fps
Rang de mesura de profunditat	0,5 ~ 4,5 m (fins a 8 m en condicions favorables)
Principi de mesura de profunditat	Temps de vol (ToF)
Camp de visió horitzontal de la càmera de profunditat	70°
Camp de visió vertical de la càmera de profunditat	60°
Resolució angular	0,14°/pixel
Camp de visió horitzontal de la càmera de color	84,1°
Camp de visió vertical de la càmera de color	53,8°
Interfase	USB 3.0
Sistema operatiu	Windows 8-8.1 (WSA)
Dimensions (cm)	249 x 66 x 67
Pes sense cables (g)	970
Tensió d'alimentació	12 VDC
Consum (W)	~ 15

Aplicacions en AP

L'aplicació dels sensors RGB-D o sensors de profunditat en agricultura de precisió prové, principalment, de la seva capacitat per generar models tridimensionals dels arbres i cultius, permetent extreure'n paràmetres geomètrics (alçada, amplada, profunditat ...) i estructurals (densitat foliar, capacitat de penetració de la llum, estructura i geometria de les plantes ...) que poden ser molt valuosos a l'hora de prendre decisions relatives al maneig de les plantacions. A més, també es pot obtenir el model digital del terreny (MDT) de les parcel·les amb un alta resolució.

El sensor RGB-D Kinect for Windows V2 es del tipus 3D. Això vol dir que permet obtenir directament models 3D de la part dels objectes mesurats que estan dintre del seu camp de visió (field of view, FOV) i dins del seu rang de distàncies de mesura. Ara bé, per obtenir models tridimensionals d'objectes grans que tenen parts fora del camp de visió i/o del rang espacial de mesura (en aquest cas, d'arbres o plantacions) cal prendre múltiples mesures cobrint tot l'objecte d'interès. Per això, cal moure el sensor per situar-lo en diferents posicions, ja sigui de manera discontinua o contínua. Finalment, juntant les mesures obtingudes s'obté un model 3D de l'objecte mesurat. La integració o unió dels resultats de les diferents mesures realitzades per obtenir un únic





model tridimensional de l'objecte es pot realitzar mitjançant diverses tècniques, com ara: i) la integració i funcionament sincronitzat amb un sistema de posicionament global (GNSS, de l'anglès Global Navigation Satellite System) com ara el popular GPS o GALILEU, de manera que cada punt del núvol de punts obtingut te unes coordenades espaials conegudes; ii) una altra tècnica freqüentment utilitzada es coneix com SLAM (de l'anglès Simultaneous Localization and Mapping), basats en l'ús de la informació de diferents tipus de sensors i l'aplicació d'algoritmes matemàtics per determinar les coordenades espaials.

Els resultats finals de les mesures consisteixen en un conjunt de punts, anomenat "núvol de punts (point cloud)". Es a dir, el model 3D de l'objecte mesurat (en el nostre cas un arbre o una plantació completa) consisteix en un núvol de punts pertanyents a la superfície de l'objecte, de manera que cada punt del núvol de punts obtingut te unes coordenades espaials conegudes, i es pot situar amb gran exactitud en mapes, facilitant la seva anàlisi juntament amb altres mapes de la mateixa plantació, en el context de l'Agricultura de Precisió. A més de les coordenades espaials, per cada punt del núvol obtingut es disposa d'informació del seu color i de la intensitat d'IR reflectida.

Limitacions

Les principals limitacions de l'ús dels sensors RGB-D Kinect for Windows V2 consisteixen en: i) el grau d'especialització i coneixements del sistema que ha de tenir l'operador d'aquest sensor; ii) en el cas de voler tenir les coordenades geogràfiques de cada punt mesurat, per poder obtenir models 3D de les plantacions, cal que aquest sensor treballi alhora i sincronitzadament amb un sistema GNSS, per tal de poder georeferenciar els núvols de punts resultants; iii) a l'igual que succeeix amb d'altres tecnologies 3D, el magatzematge dels resultats de les mesures obtingudes pot donar lloc a arxius molt grans, de l'ordre d'alguns GB; iv) en aplicacions agrícoles (i, en general en ambients exteriors) cal realitzar les mesures al començament o final del dia per a que el nivell d'il.luminació ambient natural es mantingui en nivells baixos (alternativament, es poden realitzar les mesures durant la nit, i, en cas de voler enregistrar la informació del color, il.luminar l'objecte amb llum artificial); v) per algunes aplicacions, el camp de visió del sensor K4W2 pot resultar massa reduït per enregistrar tot el camp d'interès, éssent necessari, en aquest cas, utilitzar diversos sensors treballant simultàniament o be realitzar diverses passades canviant la seva posició (normalment l'alçada de muntatge).



Referències

- Yandun, F.J., Gregorio, E., Zúñiga, M., Escolá, A., Rosell-Polo, J.R., Auat Cheein, F.A. (2016). Classifying Agricultural Terrain for Machinery Traversability Purposes. IFAC-PapersOnLine, 49(16), 457-462. DOI: 10.1016/j.ifacol.2016.10.083
- Kinect v2 Sensor-based Mobile Terrestrial Laser Scanner for Agricultural Outdoor Applications. Joan R. Rosell-Polo, Eduard Gregorio, Jordi Gené, Jordi Llorens, Xavier Torrent, Jaume Arnó, Alexandre Escolà. 2017. IEEE/ASME Transactions on Mechatronics, 22(6), 2420-2427. DOI: 10.1109/TMECH.2017.2663436
- Multi-modal deep learning for Fuji apple detection using RGB-D cameras and their radiometric capabilities. Gené-Mola J, Vilaplana V, Rosell-Polo JR, Morros JR, Ruiz-Hidalgo J, Gregorio E. 2019. Computers and Electronics in Agriculture, 162, 689-698. DOI: 10.1016/j.compag.2019.05.016

L'autoria d'aquest document és del **Grup de Recerca en AgròTICa i Agricultura de Precisió** i ha estat elaborat per Joan Ramón Rosell Polo amb la coordinació d'Àlex Escolà i el suport de Carla Román.



Aquesta obra està sota una llicència de Creative Commons Reconeixement-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0)

Activitat finançada a través de l'Operació 01.02.01 de Transferència Tecnològica del Programa de desenvolupament rural de Catalunya 2014-2022

