



(51) International Patent Classification:
G06K 9/00 (2006.01) *G06K 9/62* (2006.01)
G06F 17/00 (2006.01)

(21) International Application Number:
PCT/GB2017/051679

(22) International Filing Date:
09 June 2017 (09.06.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1611033.0 24 June 2016 (24.06.2016) GB

(71) Applicant: IMPERIAL COLLEGE OF SCIENCE,
TECHNOLOGY AND MEDICINE [GB/GB]; Exhibition
Road, London Greater London SW7 2AZ (GB).

(72) Inventors: MCCORMAC, John; 5 Killaney Lane, Ban-
bridge Down BT32 3TA (GB). HANDA, Ankur; 103 Pe-
ters Court, Porchester Road, London Greater London W2

5DS (GB). DAVISON, Andrew; Department of Comput-
ing, Huxley Building, 180 Queens Gate, London Greater
London SW7 2RH (GB). LEUTENEGGER, Stefan; 54
Shorrolds Road, London Greater London SW6 7TP (GB).

(74) Agent: EIP; Fairfax House, 15 Fulwood Place, London
Greater London WC1V 6HU (GB).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,
KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,

(54) Title: DETECTING OBJECTS IN VIDEO DATA

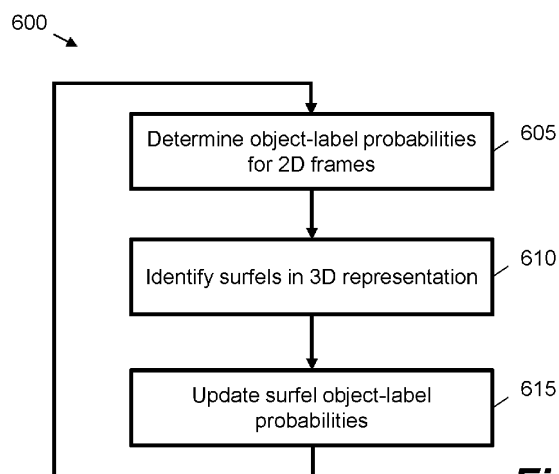


Fig. 6

(57) Abstract: Certain examples described herein enable semantically-labelled representations of a three-dimensional (3D) space to be generated from video data. In described examples, a 3D representation is a surface element or 'surfel' representation, where the geometry of the space is modelled using a plurality of surfaces that are defined within a 3D co-ordinate system. Object-label probability values for spatial elements of frames of video data may be determined (605) using a two-dimensional image classifier. Surface elements that correspond to the spatial elements are identified (610) based on a projection of the surface element representation using an estimated pose for a frame. Object-label probability values for the surface elements are then updated (615) based on the object-label probability values for corresponding spatial elements. This results in a semantically-labelled 3D surface element representation of objects present in the video data. This data enables computer vision and/or robotic applications to make better use of the 3D representation.

GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))