



- (51) **International Patent Classification:**
G06K 9/46 (2006.01) *G06T 7/20* (2017.01)
- (21) **International Application Number:**
PCT/SG2018/050420
- (22) **International Filing Date:**
17 August 2018 (17.08.2018)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
62/546,641 17 August 2017 (17.08.2017) US
- (71) **Applicant:** NATIONAL UNIVERSITY OF SINGAPORE [SG/SG]; 21 Lower Kent Ridge Road, Singapore 119077 (SG).
- (72) **Inventors:** REN, Tongwei; Room 915, FeiYimin Building, Nanjing University, Nanjing, Jiangsu 210093 (CN). GUO, Jingfan; Room 915, FeiYimin Building, Nanjing University, Nanjing, Jiangsu 210093 (CN). CHUA, Tat-Seng; c/o National University of Singapore, School of Computing, Department of Computer Science, 21 Lower Kent Ridge Road, Singapore 119077 (SG). SHANG, Xindi; c/o National University of Singapore, School of Computing, Depart-

ment of Computer Science, 21 Lower Kent Ridge Road, Singapore 119077 (SG).

- (74) **Agent:** LINDSAY, Jonas Daniel; Marks & Clerk Singapore LLP, Tanjong Pagar, P O Box 636, Singapore 910816 (SG).

(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, 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,

(54) **Title:** VIDEO VISUAL RELATION DETECTION METHODS AND SYSTEMS

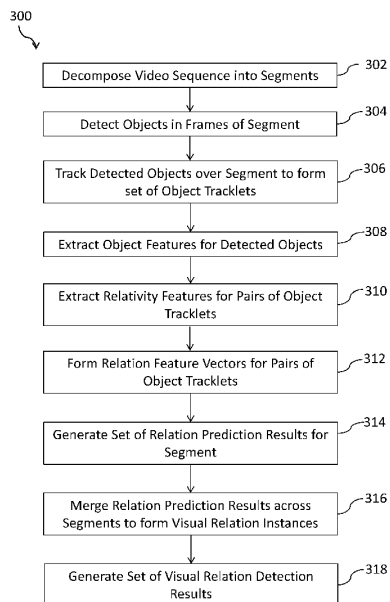


Figure 3

(57) **Abstract:** Methods and systems for detecting visual relations in a video are disclosed. A method comprises: decomposing the video sequence into a plurality of segments; for each segment, detecting objects in frames of the segment; tracking the detected objects over the segment to form a set of object tracklets for the segment; for the detected objects, extracting object features; for pairs of object tracklets of the set of object tracklets, extracting relativity features indicative of a relation between the objects corresponding to the pair of object tracklets; forming relation feature vectors for pairs of object tracklets using the object features of objects corresponding to respective pairs of object tracklets and the relativity features of the respective pairs of object tracklets; and generating a set of segment relation prediction results from the relation features vectors; generating a set of visual relation instances for the video sequence by merging the segment prediction results from different segments; and generating a set of visual relation detection results from the set of visual relation instances.

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

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