



- (51) International Patent Classification:
G06K 9/62 (2006.01) *G06K 9/46* (2006.01)
- (21) International Application Number:
PCT/NL2013/050780
- (22) International Filing Date:
31 October 2013 (31.10.2013)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
12190869.3 31 October 2012 (31.10.2012) EP
- (71) Applicant: **NEDERLANDSE ORGANISATIE VOOR TOEGEPAST-NATUURWETENSCHAPPELIJK ONDERZOEK TNO** [NL/NL]; Schoemakerstraat 97, NL-2628 VK Delft (NL).
- (72) Inventor: **BURGHOUTS, Gerardus Johannes**; c/o TNO-PLT, Schoemakerstraat 97, NL-2628 VK Delft (NL).
- (74) Agent: **JANSEN, C.M.**; V.O., Johan de Wittlaan 7, NL-2517 JR Den Haag (NL).
- (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,

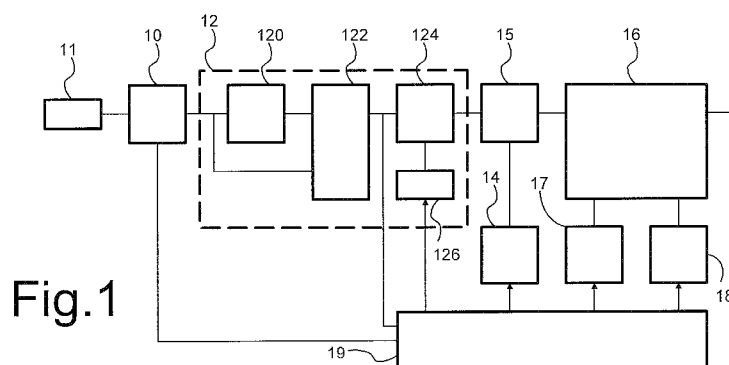
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, 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))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: DETECTION OF HUMAN ACTIONS FROM VIDEO DATA



(57) Abstract: A video action detection system uses feature a data extractor to extract feature data from video data at detected spatiotemporal interest points. A feature data quantizer assigns the extracted feature values to bins of a feature vector. Bin values are computed from a sum of contributions of spatiotemporal points of interest that have been assigned to the bin, with a bin dependent adjustment of a size of the sum and/or the contributions. The video action detection system computes a sum of match scores between the feature vector and reference vectors for the predetermined type of action. The bin dependent adjustment of the size of the sum and/or the contributions is adapted in a training step, based on partial match scores for individual bins obtained using an initial action detector.