AAL-Oriented TOF Sensor Network for Indoor Monitoring

Giovanni Diraco, Alessandro Leone and Pietro Siciliano
Institute for Microelectronics and Microsystems, Italian National Research Council (CNR)
Via Monteroni, c/o Campus Università del Salento - Palazzina A3, Lecce, Italy

Keywords: Time-of-Flight Sensor, Sensor Network, Indoor Monitoring, Event Detection, Ambient Assisted Living.

Abstract: One distinctive feature of ambient assisted living-oriented systems is the ability to provide assistive services in smart environments as elderly people need in their daily life. Since Time-Of-Flight vision technologies are increasingly investigated as monitoring solution able to outperform traditional approaches, in this work a monitoring framework based on a Time-Of-Flight sensor network has been investigated with the aim to provide a wide-range solution suitable in several assisted living scenarios. Detector nodes are managed by a low-power embedded PC to process Time-Of-Flight streams and extract features related with person's activities. The feature level of detail is tuned in an application-driven manner in order to optimize both bandwidth and computational resources. The event detection capabilities were validated by using data collected in real-home environments.

1 INTRODUCTION

During the last years, the interest of scientific community for smart environments has grown very fast especially gained from the European Ambient Assisted Living (AAL) program aiming to increase the quality of life of older people by helping them to live more independently and longer at their homes. Moreover, smart environments present new criticalities: 1) devices and applications are often isolated or proprietary, preventing the effective customization and reuse; 2) traditional monitoring systems are strongly dependent from ambient conditions and/or invasive. The answare to the first criticality comes from the use of AAL-oriented middleware acting as a flexible intermediate layer able to adapt to evolving requirements and scenarios. AAL middleware architectures have been described, among others, by (Wolf et al., 2010) and by (Schäfer, 2010). Concerning the second criticality, recently TOF (Time-Of-Flight) sensors have been increasingly investigated as solutions being able to overcome drawbacks affecting traditional cameras (included the privacy issue). (Wientapper et al., 2009) described a model-free approach for classifying human postures in TOF images. The usage of cheaper (non-TOF) 3D sensors have been recently reported especially for fall detection application (Mastorakis, 2012). However non-TOF 3D sensors estimate distances from the distortion of an infrared light pattern projected on the scene, and so their accuracy and distance range result seriously limited (distance up to 3-4 m with 4 cm accuracy) (Khoshelham, 2011). This paper presents a novel event detection solution for elderly monitoring in smart environments composed by a TOF SN (Sensor Network) and an application-driven processing architecture integrated within an AAL-oriented middleware infrastructures.

2 SYSTEM ARCHITECTURE

The SN topology includes *M* detector nodes managing *N* TOF sensor nodes for each, and one coordinator node that receives high-level reports from the detectors. TOF sensors (Figure 1.a) and embedded PCs (Figure 1.b), managing coordinator and detector nodes, are low-power, compact and noiseless devices, in order to meet typical requirements of AAL applications. The computational framework is conceived as a modular, distributed and open architecture implemented by coordinator and detector nodes, and integrated into a larger AAL system through open middleware. Figure 1.c shows sixteen sample TOF frames of the collected dataset in real home environments. The typical apartment is shown in Figure 1.d with the

Diraco G., Leone A. and Siciliano P..
AAL-Oriented TOF Sensor Network for Indoor Monitoring.
DOI: 10.5220/0004253202360239
In *Proceedings of the 2nd International Conference on Sensor Networks* (SENSORNETS-2013), pages 236-239
ISBN: 978-989-8565-45-7

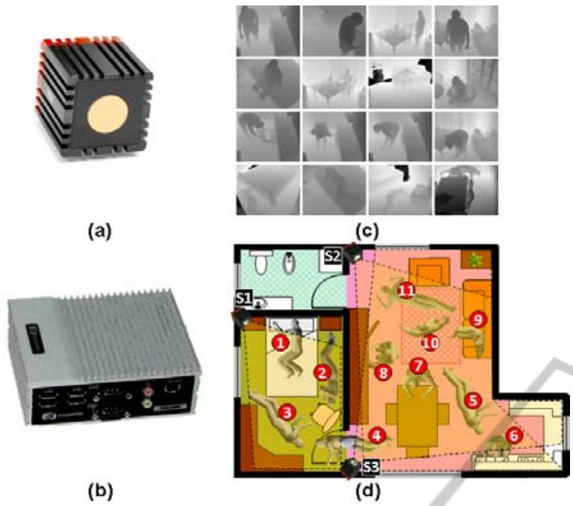

Figure 1: a) MESA SR-4000 TOF sensor; b) Intel Atom Processor-based Embedded PC managing the detector nodes and the coordinator node; c) Some TOF frames collected in real homes; d) Plan of the apartment used for data collection with the indication of performed actions (from 1 to 11) and sensor positions (S1, S2, S3).

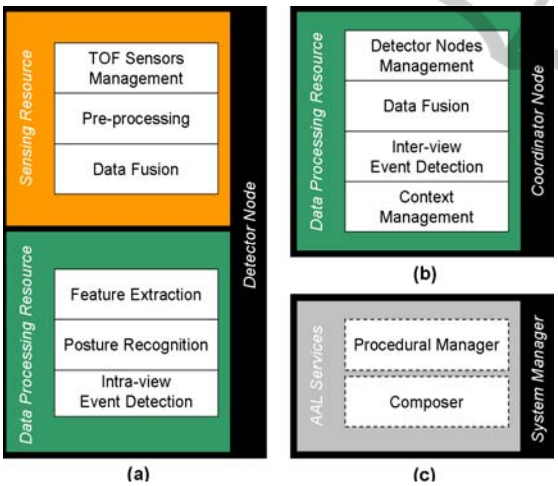

Figure 2: Conceptual representation of the computational framework: a) Detector node, b) Coordinator node, c) System Manager.

locations (from 1 to 11) in which actions have been performed. The sensor network used during experiments included three sensor nodes, S1 in the bedroom, and S2 and S3 in the living room with overlapping views. The Figure 2 shows the overall architecture of the computational framework. There are three main logical layers: Data Processing Resource, Sensing Resource and System Manager. Data Processing Resource layers are implemented by both the detector nodes (Figure 2.a) and the coordinator node (Figure 2.b). The detector nodes implement, in addition, the Sensing Resource layer. The coordinator includes architectural modules for detector nodes management (control and data gathering), high-level data fusion, inter-view event detection and context management. The System Manager (Figure 2.c) manages the whole AAL system. It is inspired to OpenAAL (Wolf, 2010), that is an open distribution of the UniversAAL (2012) middleware, achieving global AAL services' goals.

The sensing node is represented by the MESA SwissRanger SR-4000 (MESA Imaging AG, 2011), a state-of-the-art TOF sensor with compact dimensions (65×65×68 mm), noiseless functioning (0 dB noise), QCIF resolution (176×144 pixels), long distance range (up to 10 m) and wide (69°×56°) FOV (Field-Of-View). The pre-processing module includes functionalities to perform the extrinsic calibration (wall-mounting self-calibration), fuse together range data gathered from overlapping views and identify a person into range data. A cascade of well-known vision processing steps, namely background modelling, segmentation and people tracking, is implemented according to a previous authors' study (Leone, 2011). Finally, a middleware module plugs in the TOF sensors into the system providing also a semantic description of acquired range data.

The detector data processing resource includes the following modules: feature extraction, posture recognition, and intra-view event detection. Features are extracted from TOF range data by using two body descriptors having different level of detail and computational complexity. Coarse grained features are extracted by using a volumetric descriptor that exploits the spatial distribution of 3D points represented in cylindrical coordinates (h, r, θ) corresponding to height, radius and angular locations respectively. The 3D points are grouped into rings orthogonal to and centred at the h-axis while sharing the same height and radius, as depicted in Figure 3.a. The volumetric features are represented by the cylindrical histogram shown in Figure 3.b obtained by taking the sum of the bin values for each ring. Fine grained features are achieved, instead, by using a topological representation of body information embedded into the 3D point cloud. The intrinsic topology of the body shape is graphically encoded by using the Discrete Reeb Graph (DRG) proposed by (Xiao et al., 2004). The DRG is extracted by subdividing the geodesic map (Figure 3.c) in regular level-sets and connecting them on the basis of an adjacency criterion as described by (Diraco et al., 2011) that suggest also a methodology to handle self-occlusions (due to overlapping body parts). The

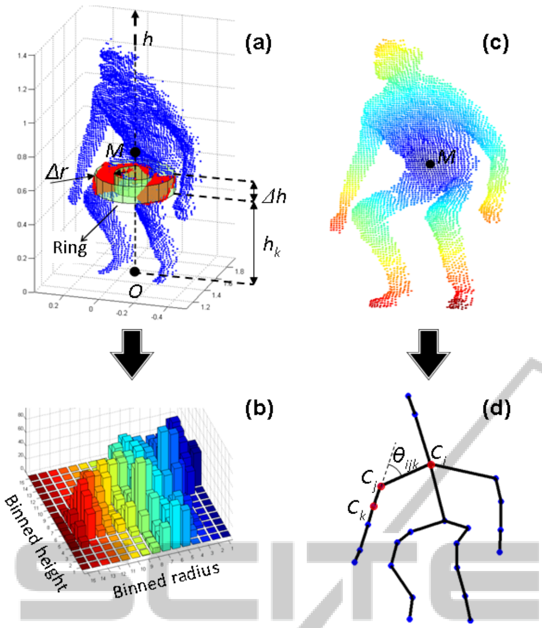

Figure 3: a) Volumetric representation of body 3D point cloud. b) Cylindrical histogram features extracted from the volumetric representation. c) Geodesic map (near points are blue, far red). d) Extracted Reeb graph features.

DRG-based features are shown in Figure 3.d. The topological descriptor includes DRG nodes and related angles. Given the coarse-to-fine features extracted as previously discussed, a multi-class formulation of the SVM (Support Vector Machine) classifier (Debnath, 2004) based on the one-against-one strategy is adopted to classify the person's postures. The coordinator data processing resource includes the following functional modules: detector nodes management, data fusion, inter-view event detection and context management. Human actions are recognized by considering successive postures over a time period. Instead, global events are recognized by using Dynamic Bayesian Networks (DBNs) specifically designed for each application scenario, following an approach similar to (Park and Kautz, 2008).

3 EXPERIMENTAL RESULTS

The event detection performance was evaluated in real home environments by involving ten healthy subjects, 5 males and 5 females, having different physical characteristics: age 31±6 years, height 173±20 cm, weight 75±22 kg. Four relevant AAL application scenarios have been considered, namely fall detection, wandering detection, ADLs (Activities of Daily Living) recognition and training exercises recognition. One dataset for each scenario was collected and characterized by different combinations of occlusions, angles and distances as reported in Table 1 by the first four columns.

Table 1: Detection performance of four event classes.

Detected Event	Semi-occlusions	Angles	Distances	Detection Performance (%)	
				Specificity	Sensitivity
Falls	Yes	All	All	97.5	83.0
Wandering	Yes	All	All	92.7	81.6
Activities of Daily Living	Few	All	Med	98.3	96.4
Training Exercises	No	Front	Low	99.2	95.6

The last two columns report the achieved detection performance in terms of specificity and sensitivity measures defined as follows:

$$Specificity = \frac{TN}{TN + FP}, \quad Sensitivity = \frac{TP}{TP + FN},$$

where TP, TN, FP, FN stand for True Positive, True Negative, False Positive and False Negative, respectively. Concerning the fall detection scenario, falls have been correctly discriminated from non-falls even in presence of semi-occlusions, achieving 97.5% and 83% of specificity and sensitivity, respectively. However, since fall events were detected by single detector nodes, ambiguous situations such as those in which a fall was located between non-overlapping views (i.e. location 4 in Figure 1.d) given rise to false negatives. The wandering activity was detected by the coordinator since it normally involves several non-overlapping views (i.e. sensors in different rooms). Wandering was discriminated from ADLs with 92.7% and 81.6% of specificity and sensitivity, respectively. A low level of feature detail was involved for fall detection with prevalent adoption of the volumetric representation, instead in the case of wandering detection a slightly higher detail level was needed with a moderate use of topological representation. In order to evaluate the ADLs recognition capability, the following seven kind of activities were performed: sleeping, waking up, eating, cooking, housekeeping, watch TV and physical training. ADLs were recognized with 98.3% and 96.4% of specificity and sensitivity, respectively. A moderate misclassification was observed for housekeeping activities since occasionally were erroneously recognized as wandering state. For the physical training scenario, a virtual trainer was developed instructing participants to follow a physical activity program and perform the recommended exercises correctly. The recommended physical exercises were of the following kind: biceps curl, squatting, torso bending, etc. The involved feature detail level was high with prevalent use of the topological representation. Performed exercises were correctly recognized achieving 99.2% and 95.6% of specificity and sensitivity, respectively. The most computationally expensive steps were pre-processing, feature extraction, and posture classification. They were evaluated in terms of processing time that was constant for pre-processing and classification resulting respectively in 20 ms and 15 ms per frame. The volumetric descriptor taken an average processing time of about 20 ms, corresponding to about 18 fps (frame-per-second). The topological approach, on the other hand, required a slightly increasing processing time among hierarchical levels from an average value of 40 ms to 44 ms due to the incremental occurrence of self-occlusions, achieving up to 13 fps.

4 CONCLUSIONS

The main contribution of this work is to design and evaluate a unified solution for TOF SN-based in-home monitoring suitable for different AAL application scenarios. An open (OpenAAL inspired) computational framework has been suggested able to classify a large class of postures and detect events of interest accommodating easily (i.e. with self-calibration), at the same time, wall-mounting sensor installations more convenient to cover home environments avoiding large occluding objects. Moreover, the suggested computational framework was optimized and validated for embedded processing to meet typical in-home application requirements, such as low-power consumption, noiselessness and compactness. The system was able to adapt effectively to four different AAL scenarios exploiting an application-driven multilevel feature extraction to reliably detect several relevant events and overcoming, at the same time, well-known problems affecting traditional monitoring systems in a privacy preserving way. The ongoing work concerns the on-field validation of the system that will be deployed in elderly dwellings at support of two different ambient assisted living scenarios concerning the detection of dangerous events and abnormal behaviours.

ACKNOWLEDGEMENTS

The presented work has been carried out within the BAITAH project funded by the Italian Ministry of Education, University and Research (MIUR).

REFERENCES

Debnath R., Takahide N., Takahashi H., 2004. A decision based one-against-one method for multi-class support vector machine. Pattern Analysis and Applications;7(2):164-175.

Diraco, G., Leone, A., Siciliano, P., 2011. Geodesic-based human posture analysis by using a single 3D TOF camera. In: Proc of ISIE;1329-1334.

Fuxreiter, T., Mayer, C., Hanke, S., Gira, M., Sili, M., Kropf, J., 2010. A modular platform for event recognition in smart homes. In: 12th IEEE Healthcom;1-6.

Khoshelham, K., 2011. Accuracy analysis of Kinect depth data. In: Proc of ISPRS Workshop on Laser Scanning.

Leone, A., Diraco, G., Siciliano P., 2011. Detecting falls with 3D range camera in ambient assisted living applications: a preliminary study. Med Eng Phys;33(6):770-81.

Li, W., Zhang, Z., Liu, Z., 2010. Action recognition based on a bag of 3D points. In: Proc of CVPRW;9-14.

Mastorakis, G., Makris, D., 2012. Fall detection system using Kinect's infrared sensor. J Real-Time Image Proc;1-12.

MESA Imaging AG, 2011. SR4000 Data Sheet Rev.5.1. <http://www.mesa-imaging.ch>.

Park, S., Kautz, H., 2008. Privacy-preserving recognition of activities in daily living from multi-view silhouettes and RFID-based training. In: AAAI symposium on AI in eldercare, Arlington, Virginia USA.

Schäfer, J., 2010. A Middleware for Self-Organising Distributed Ambient Assisted Living Applications. In: Workshop Selbstorganisierende, Adaptive, Kontextsensitive verteilte Systeme (SAKS).

UniversAAL project 2012. http://www.universaal.org/.

Wientapper, F., Ahrens, K., Wuest, H., Bockholt, U., 2009. Linear-projection-based classification of human postures in time-of-flight data. In: Proc of ICSMC;559-564.

Wolf, P., Schmidt, A., Parada Otte, J., Klein, M., Rollwage, S., Koenig-Ries, B., Dettborn, T., Gabdulkhakova, A., 2010. openAAL - the open source middleware for ambient-assisted living (AAL). In: AALIANCE conference, Malaga, Spain.

Won-Seok, C., Yang-Shin, K., Se-Young, O., Jeihun, L., 2012. Fast iterative closest point framework for 3D LIDAR data in intelligent vehicle. In: IEEE Intelligent Vehicles Symposium (IV);1029–1034.

Xiao, Y., Siebert, P., Werghi, N., 2004. Topological segmentation of discrete human body shapes in various postures based on geodesic distance. In: Proc of ICPR 2004;3:131-135.