WiCoSens - a Wearable, Intelligent Color Sensing Platform for non-invasive Storage Shelf Identification

Orkhan Amiraslanov
DFKI GmbH

Hymalai Bello
DFKI GmbH

Gerald Pirkl
DFKI GmbH

Peter Hevesi
DFKI GmbH

Paul Lukowicz
DFKI GmbH, TU Kaiserslautern
[firstname.lastname]@dfki.de

ABSTRACT
Reaching out onto a shelf or into a cabinet and taking something out is a fundamental part of many activities. In our daily life we carry out several picking and placing actions from/to shelves, cupboards, racks, containers, *etc.*. With enough instrumentation identifying the shelf from which something has been taken is not a fundamental problem. However, reliably detecting specific shelf from which something has been taken with an unobtrusive and cheap setup remains, in general, an open problem. As a solution we developed *WiCoSens* - a wrist worn color sensor (CS) array to detect color coded surfaces. We describe the hardware design, the identification method and a simple color coded shelf setup to evaluate the system. Initial results show 100% accuracy with user independent training.

Author Keywords
Color Sensor; Wearables; Order Picking; Warehouse Management; Activity Recognition.

ACM Classification Keywords
I.2.9 Computing Methodologies: Sensors

INTRODUCTION
Reaching out onto a shelf or into a cabinet and taking something out is a fundamental part of many activities. In everyday life we reach into cupboards to take dishes, food, clothing, medicine, *etc.*. In maintenance and production settings tools and raw materials have to be taken out and put back into storage shelves. In logistics manual product placement and order-picking are still a major factor. As a consequence, tracking *"take out"* and *"put in"* actions are important considerations in activity recognition. In this paper we focus on the question of identifying the specific shelf from which an item has been taken or placed. In general, items are not placed randomly on shelves and this provides an indication about the type of the object that may have been taken out. Thus, in a household scenario depending on the shelf, which the user has reached we may know that he/she has taken out a cup, a plate, spice, or a particular type of food. In structured environment such as warehouses or shops the type of items connected to a particular shelf is even more strictly defined.

Clearly, with enough instrumentation effort identifying the shelf from which something has been taken is not a fundamental problem. The most obvious solution is to place an RFID reader with a sufficiently large antenna on each shelf and have the user wear a RFID-tag on a wrist or *vice versa*. Unfortunately, in most scenarios (*e.g.* activity tracking in a typical home, order picking at a large warehouse) the required instrumentation effort is not justifiable. Placing a RFID reader on the wrist and tags on every shelf is problematic because of range limitations of small wrist worn coil antennas. Another possible approach is exact location tracking of the user's hand. Here the problem is that cheap, common indoor localization systems (*e.g.* WiFi location or Bluetooth beacons) do not have the required accuracy. Even more advanced systems such as UWB cannot always reliably distinguish between neighboring shelves.

AR based optical tracking systems with markers such as[4] are sufficiently accurate, but very expensive and prone to occlusion effects. Occlusions and reflection are also an issue with ultra sound systems (which are more cost effective). While in principle a viable approach is using vision, it has number of practical issues such as privacy, occlusions, sensitivity to light conditions and computational complexity.

In summary, **reliably** detecting specific shelf from which something has been taken **with an unobtrusive and cheap setup** still remains as an open problem.

Paper Contributions
We have developed a novel wrist worn color sensor array – the *WiCoSens*, which we believe addresses the above mentioned problems and could be used to identify color coded shelves or in activity tracking (see Figure 1.**a**). The main contributions of the paper are:

- Present a detailed design and implementation of the wristband hardware.
- Present a novel concept of using color coding and a wrist band based color sensor array for shelf identification: upper and lower walls of each shelf is color coded with dis-

ISWC '17, September 11–15, 2017, Maui, HI, USA.
ACM ISBN 978-1-4503-5188-1/17/09...$15.00
https://doi.org/10.1145/3123021.3123062

tinct colors (see Figure 2). Essentially this gives us a two digit code. By using four four colors (which are very easy to separate) allows up to 16 distinct shelves to be uniquely identified. User wears the wristband and when user's arm enters the cabinet top and the bottom facing color sensors scan the surface and report detected colors.

- Description of the respective sensor signal processing and shelf recognition chains.
- The evaluation of the system in a university kitchen with 12 (color coded with four basic colors) shelves and six subjects each randomly picking various items (*e.g.* glasses, cups, plates, *etc.*). Given the good separability of the selected colors and small scale setup, the system achieves perfect recognition result of (100%).

Our paper first describes the sensor hardware implementation steps in great detail to ease reproducibility. We then discuss software architecture to read and process sensor data accordingly. Later we present the test setup and evaluate the system in a simple kitchen scenario. Results from evaluation show that the wristband can be used in real-world applications.

RELATED WORK

Here we shortly illustrate the currently available industrial approaches for rack identification and asset tracking systems. These systems mainly utilize vision based (QR-code, barcode) tracking and RFID tags.

Wearable barcode and QR-code scanner (*e.g.* Honeywell 8670) are widely used in current asset identification systems[7]. This labor intensive approach requires human input (a button press), which introduce scan time delays every time wearer wants to identify and register the rack.

RFID is another popular option[3, 1] and it has been widely studied. However, this approach is better suited for the object recognition. Key issues are the need to equip many shelves with RFID tags, interference from metal structures, high installation costs and the limited range[10] of $(5 - 10\ cm)$.

A combination of RFID and computer vision based for the detection of interaction with objects was described in [2]. The use of computer vision by itself for the recognition of complex activities is a broad research field. Examples of work that involves access to specific cupboards and shelves (in a kitchen scenario similar to our evaluation) are[6, 8].

Today various high end, expensive commercial camera based systems use printed markers for *sub-cm* tracking. Similar accuracy can be achieved using short range, high end magnetic trackers, which are, however, not suitable for big scale, real life activity recognition. In[5] an oscillating magnetic field positioning system is used. With sufficient number of coils the system can provide reasonable accuracy with a added deployment effort and cost for the purpose of shelf identification.

Unlike above mentioned methods, our system could also be used to track user activity and automate color coded shelf identification seamlessly without introducing identification delays, extra installation costs or human input .

WICOSENS: A WEARABLE COLOR SENSING DEVICE

WiCoSens (see Figure 1.**a**) is a wrist worn color sensing platform that combines circular array of RGBC (Clear) color sensors, onboard IMU, barometer sensors and illumination LEDs. The wrist worn design could potentially be integrated into wide range of consumer devices such as smart watches or fitness arm bands. Depending on a lighting level in the environment the brightness of the illumination LEDs is dynamically adjusted and with IMU data the system also orients its angular position and can illuminate only the part of the wrist where it faces pre-colored parts of the surfaces (*i.e.* bottom and top walls of a shelf).

Hardware Design

The hardware combines three board types: the microcontroller unit (MCU), main sensor board and multiples of identical mini color sensor boards. The overall architecture of the system is shown in Figure 1.**b**.

There are 16 mini-boards ($1\times$onboard, $7\times$on the left and $8\times$on the right side of the sensor board) that are connected with a common communication and power distribution bus to the main sensor board. Each mini-board contains a 16 bit RGBC color sensor (TCS34725) with IR filter, two neutral white illumination LEDs (XQ-A), an LED driver (TLC5973), and an I^2C address translation chip (LTC4316). By default all color sensors have the same bus address, this IC appends unique, predefined translation bits to each sensor board on the fly, enabling access to all colors sensors, which share the same bus. To protect the sensors from stray light $4mm$ high 3D printed light tube is also attached to the mini-boards. To improve comfort and wearability flexible textile string and magnetic clips were added to connect mini boards.

The main sensor board (see Figure 1.**c**) has one color sensor and accompanying LEDs, an absolute orientation sensor (BNO055), a combined barometer, humidity and temperature sensor (BME280) to determine height and environment, $3\times I^2C$ buffer ICs (TCS9617B) for ESD-safe communication with off-board sensors, and an expansion port for extra wireless connectivity option.

The MCU board (Teensy 3.6) is a 32bit ARM Cortex-M4 MCU and it is placed on top of the main sensor board. This board controls data acquisition from sensors, provides wired connectivity to PC and has on board microSD slot for local data logging.

Power consumption

Main power hungry parts are illumination LEDs. At maximum brightness (current limited to $25mA$) resulting current consumption is $800\ mA$ for total of 32 LEDs. However, only 4 pairs of LEDs are turned on at one time and brightness is dynamically adjusted based on ambient light levels to save more power. Color sensors and main sensor board consumes only $11mA$ and the MCU board itself $100\ mA$. Moreover, the system can be put in low-power sleep mode by disabling all the color sensors, LEDs and MCU by using the IMU in standalone motion-triggered interrupt mode to detect a sudden movement and wake the system up and do readings.

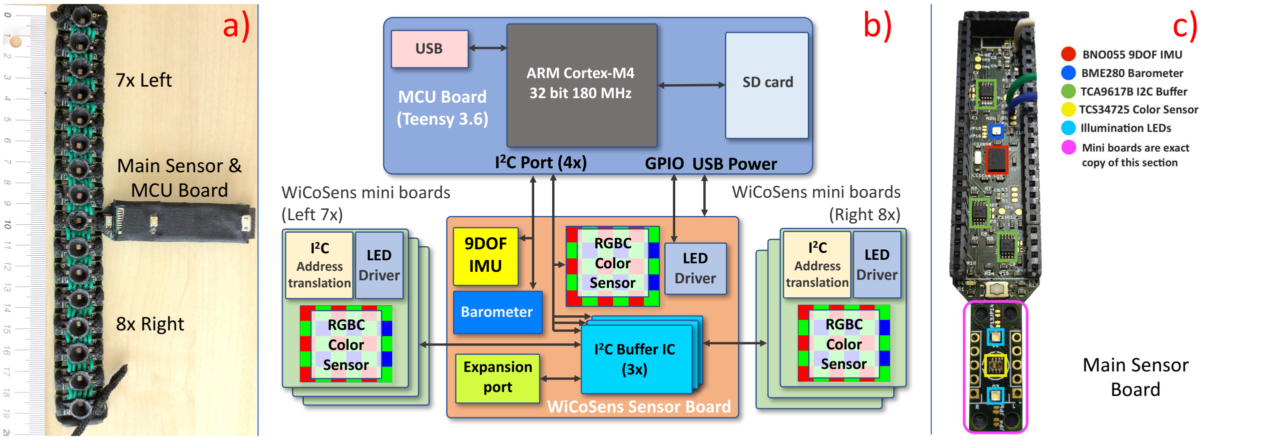

Figure 1. a) WiCoSens - A wrist worn color sensing device with magnetic locking. b) WiCoSens system architecture. c) Main sensor board (top view).

Software Design

At power up MCU initializes all the sensors and calibrates them based on the coefficients stored on its non-volatile memory (EEPROM). If needed user can press the onboard button and recalibrate the IMU sensor. Color sensors provide RGB and clear light levels at 16 bits resolution.

To achieve higher sampling rate and reduced bus conflicts we allocate two I^{2}C ports only for color sensors (CS) (*i.e.* port 1 for 7×left side and onboard CS, port 2 for 8× right side sensors). Each individual sensor has a typical $2.4ms$ color integration time and thus for eight sensors on each bus we achieve $20ms$ sampling time, which is suitable for real-time processing. A top facing sensor readings are also used to dynamically adjust integration time and gain coefficients of the active CSs to achive uniform readings over wide range of ambient light conditions. Luminous flux (lux) value is calculated and used to adjust the illumination LEDs and sensor coefficients accordingly to achieve steady readings when the ambient luminance parameters change.

We also extract gravity vector from accelerometer data and selectively read two adjacent CSs from each bus (2× facing upwards, 2× facing downwards) at one time (also illuminate respective LEDs) and transmits the sensor values over USB for processing. As a result we reduced the number of CSs to be read from 16 to four, thus increased frame sampling time from $50Hz$ to around $130Hz$.

Colors and Color Model

In order to uniquely color code surfaces we first need to define basic colors. In practical applications there are 11 basic colors[9]: *white, gray, black, red, yellow, green, blue, orange, purple, pink and brown*. If we color two walls of a shelf uniquely, then total of $11^2 = 121$ distinct shelves can be coded. Our color identification method relies on HSV *(hue, saturation, value)* color model. RGBC readings are converted to HSV on the MCU before transmission to PC. We mainly use hue values, which are associated with the dominant wavelength of a color. Main advantage of using hue is that the color readings are not affected by the changing luminance in environment and by the surface distance (safe readings up to $20cm$).

EVALUATION

Sensor Evaluation

We have extensively characterized CS parameters, in terms of maximum readable surface distance $(0 - 32cm)$, best separable colors, optimal sensor gain $(1,\ 4,\ 16,\ 60\times)$, sensor integration times $(2.4,\ 24\ ms)$ and effects of varying ambient illumination on sensor readings. Eight basic colors have been tested against above mentioned parameters and usable surface distance of max *20cm*, with gain of 60× and integration time of $(2.4ms)$ have been selected. As mentioned earlier the system uses one of the CS readings to calculate ambient light level and if its too low/high (typical 100 Lux) calibration coefficients and illumination LED intensity will be dynamically adjusted.

Setup and Experiment

A simple university kitchen with four upper cabinets (each with three shelves) was selected to evaluate the system. Each shelf has dimensions of $27cm \times 60cm \times 40cm$ *(H, W, D)*. The cabinets are used to store ordinary dishes and food items. We added colored A3 sized paper markers at the top and bottom of each shelf as seen in Figure 3.**a**, while retaining their content unchanged.

Figure 2. Cabinet color coding.

Figure 2 presents the color coding of the kitchen cabinets. Four opposite colors from the color circle has been picked (*i.e. red, green, blue, yellow*) to color code all 12 shelves. The cabinets are treated as matrix rows and columns, thus top walls of each shelf in every row and bottom walls in every column have common colors. This dual color coding scheme allows us to address up to 16 shelves. Colored papers can also be replaced by paint or vinyl adhesive paper. When selecting a colored material care must be taken to not have glossy surface, since this type of surfaces will add reflections from ambient light fixtures.

Identification Methods

Identification by Hue

Simple, yet effective method is to use a look-up table based approach. This self-sustained method does not require high-level machine learning (ML) algorithm or a processing PC, and it runs in real-time on the MCU. Every time the MCU detects sudden change in gravity vector (picking started) it starts to sample up and down facing color sensors. If the measured hue (color) fits in one of the predefined color combination table then it reports detected shelf number. We have tested it with our current setup and it works reasonably reliable ($\leq 5\%$ error rate), however this method is only reliable up to (≤ 16) shelf count. To increase addressable shelf count we must also increase number of colors used and add more color stripes for each surface.

Identification by Machine Learning

In our second approach we used ML based approach. Six test subjects were asked to pick any object from each of the 12 shelves in random order. Each action was labeled and stored alongside the experiment by a bystander for supervised learning.

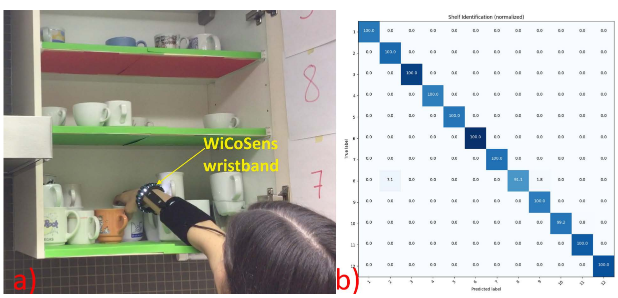

Figure 3. ***a)*** **Participant picks objects from color coded shelf.** ***b).*** **Shelf identification confusion matrix.**

We take one picking action as one instance and divide it into $100ms$ long frames. Each frame contains 10 averaged consecutive measurements. On average each picking instance has around 15 frames. As the event consists of reaching into the shelf, grabbing an object and pulling it out number of frames provides significant diversity in orientation and distance from the top/bottom to create a reliable model. The color combinations are distinct enough so that 15 points per class suffice for a good model. We segmented the data using the video stream into individual events, each related to a single shelf. We consider the recognition on the basis of individual frames, containing 10 measurements. One of the test subjects data was used to train a decision tree classifier and used it to test on remaining five subjects. Shelf recognition is then done on a frame level by applying the trained classifier to the data. We have also used recurrent neural network (LSTM) for shelf identification. Hue, saturation and brightness levels of each sensor are used as main features. Again one of the test subjects data was used to train the system.

In decision tree classifier we achieve 100% accuracy, classifying over 5000 samples correctly. In LSTM we achieved 99.1% accuracy (Figure 3.**b**).

RESULTS

We use the data from *just one subject* to create a model and test model on the remaining *five* subjects. However, we believe due to the well separation of the colors, a small number of identifiable shelves and limited number of participants (limited size of the dataset) both of our classification models are overfitted. We plan to carry out more extensive experiments with larger number of participants and picking actions.

We also plan to increase number of identifiable shelves (address space) to be $\geqslant 100$ to reflect more real world usage. To achieve this we will add colored paper stripes to increase color variation on every surface.

ACKNOWLEDGMENTS

This work has been funded by the SmartWerk project (FKZ 16SV6272). The authors would like to thank the subjects who volunteered to participate in the experiment.

REFERENCES

1. Flipse, M. RFID techniques for indoor warehouse location sensing indoor location sensing.
2. Park, S., and Kautz, H. Hierarchical recognition of activities of daily living using multi-scale, multi-perspective vision and rfid. In *Intelligent Environments, 2008 IET 4th International Conference on*, IET (2008), 1–4.
3. Patterson, D. J., Fox, D., Kautz, H., and Philipose, M. Fine-grained activity recognition by aggregating abstract object usage. In *ISWC 2005*, IEEE (2005), 44–51.
4. Pickl, S. Augmented Reality for Order Picking Using Wearable Computers with Head-Mounted Displays. 154.
5. Pirkl, G., and Lukowicz, P. Robust, low cost indoor positioning using magnetic resonant coupling. In *Proceedings of the UBICOMP 2012,ACM*, ACM (2012), 431–440.
6. Rohrbach, M., Amin, S., Andriluka, M., and Schiele, B. A database for fine grained activity detection of cooking activities. In *CVPR 2012 IEEE Conference*, IEEE (2012), 1194–1201.
7. Sivakami, N. Comparative Study of Barcode , QR-code and RFID System in Library Environment.
8. Tenorth, M., Bandouch, J., and Beetz, M. The tum kitchen data set of everyday manipulation activities for motion tracking and action recognition. In *ICCV Workshops, 2009 IEEE 12th International Conference*, IEEE (2009), 1089–1096.
9. van de Weijer, J., Schmid, C., Verbeek, J., and Larlus, D. Learning color names for real-world applications. *IEEE Transactions on Image Processing 18*, 7 (2009), 1512–1523.
10. Wölfle, M., and Günthner, W. A. Wearable rfid in order picking systems. In *Smart Objects: Systems, Technologies and Applications, Proceedings of RFID SysTech 2011 7th European Workshop on*, VDE (2011), 1–6.