Monitoring Activity of Taking Medicine by Incorporating RFID and Video Analysis

Faiz M. Hasanuzzaman, Xiaodong Yang, and YingLi Tian*
Department of Electrical Engineering
The City College of New York
New York, NY, USA
{fhasanu00, xyang02, ytian}@ccny.cuny.edu

Qingshan Liu
School of Information & Control
Engineering, Nanjing University of
Information Science and Technology, China
qsliu@nuist.edu.cn

Elizabeth Capezuti
College of Nursing, New York University,
New York, NY USA
ec65@nyu.edu

* Corresponding author. E-mail: ytian@ccny.cuny.edu.

Abstract

In this paper, we present a new framework to monitor medication intake for elderly individuals by incorporating a video camera and Radio Frequency Identification (RFID) sensors. The proposed framework can provide a key function for monitoring activities of daily living (ADLs) of elderly people at their own home. In an assistive environment, RFID tags are applied on medicine bottles located in a medicine cabinet so that each medicine bottle will have a unique ID. The description of the medicine data for each tag is manually input to a database. RFID readers will detect if any of these bottles are taken away from the medicine cabinet and identify the tag attached on the medicine bottle. A video camera is installed to continue monitoring the activity of taking medicine by integrating face detection and tracking, mouth detection, background subtraction, and activity detection. The preliminary results demonstrate that 100% detection accuracy for identifying medicine bottles and promising results for monitoring activity of taking medicine.

1. Introduction

As the fraction of the elderly in the total population continues to rise, specifically in developed countries, providing appropriate technology to assist them is essential. According to the UN [1], "Population ageing is unprecedented, without parallel in the history of humanity." The number of elderly people in the world who are 60 years or older will increase from 10 percent currently to around 20 percent in 2050 [1]. In some countries, 16 to 18 percent of the population is already 65 or older [2]. The proportion requiring personal assistance with everyday activities increases with age, ranging from 9 percent, for those who are 65 to 69 years old, to 50 percent, for those who are 85 or older. Furthermore, the likelihood of dementia or Alzheimer's disease increases with age for those over 65 [3]. Alzheimer's is predicted to affect 1 in 85 people globally by 2050 [4, 5]. This data indicates that the demand for caregivers will reach far beyond the number of individuals able to provide care. One solution to this growing problem is to find ways to enable elders to live independently and safely, at least in their own homes, for as long as possible [6].

Recent technology developments in computer vision, digital cameras, RFID and computers make it possible to assist the independent living of older adults by developing safety awareness technologies to analyze the elder's activities of daily living (ADLs) at home. For example, older adults take more medications than any other age group, with elders taking on average of about 5.7 prescription medicines and 2 to 4 nonprescription drugs each day [6]. Taking medications is one of the most important activities in an elder's daily life. Furthermore, older people with multiple diseases are often required to take more than one medicine. In addition, those with dementia or Alzheimer's disease require prompts to take their medication safely. Therefore, it is very common for the elderly to experience adverse drug effects when they do not take their prescribed dose (taking too many, too few or forgetting to take their medications) [2]. New methods are required to take care of these aged persons in their home, effectively and automatically, and with the lowest possible cost. Knowing that technology can significantly improve the quality of life for elders and those experiencing mental illnesses, we are interested in the problem of monitoring medication intake through the detection of medicine bottles by RFIDs and tracking their movement by cameras. This can help elderly people live in technology-assisted environments instead of human-assisted environments.

RFID is a type of automatic identification technology, relying on storing and remotely retrieving data using RFID readers and tags, which can be attached to or embedded within objects to be identified. Tags contain silicon chips and antennas to enable them to receive and respond to radio-frequency queries from an RFID transceiver. Passive tags require no internal power source, whereas active tags require a power source [7]. An RFID system may consist of several components: tags, tag readers, edge servers, middleware, and application software. The purpose of an RFID system is to enable data to be transmitted by a mobile device, or the tag, which is read by an RFID reader and processed according to the needs of a particular application. The data transmitted by the tag can provide identification, location information, or specifics about the product tagged, such as price, color, data of purchase, etc. In a typical RFID system, individual objects are equipped with a small, inexpensive tag. The tag contains a digital memory chip that is given a unique electronic product code. The interrogator, an antenna packaged with a transceiver and decoder, emits a signal activating the RFID tag so it can read and write data to it. When an RFID tag passes through the electromagnetic zone or the antenna, it detects the reader's activation signal. The reader decodes the data encoded in the tag's integrated circuit and the data is passed to the host computer.

Computer vision is the science and technology of machines that see and understand the surrounding environments from images or videos. Computer vision has produced vital applications in fields such as industrial automation, video surveillance, object recognition, image retrieval etc. [8-13]. Computer vision is also playing a role in aiding elderly people as well as visually impaired people. This is achieved by vision technologies in order to improve monitoring systems where data is used to determine trends in the quality of life of disabled older people. The idea of monitoring covers a potentially large area, involving a range of conceptual, methodological and instrumental issues [14, 15].

Camera-based object detection and recognition is becoming increasingly common in computer vision community. There are many challenges to detect objects under conditions of occlusion, variations of scales and viewpoints, large lighting changes, and cluttered situations. Common approaches to general object detection is to slide a window across the image and classify whether the window is containing the target or background. This type of approach has been successfully applied to detect faces, cars, and pedestrians [16-21].

Passive monitoring of the elder's activities of daily living allows them to stay at home longer. In this paper, we focus on designing a system for monitoring the actions of taking medications by combining RFID sensors and cameras. Since the RFID-based object identification generally has high accuracy, instead of detecting objects directly through a camera, we detect medicine bottles from RFID sensors and then employ a vision-based method to continue track them in a large range. The RFID sensors can effectively identify medicine bottles in the range of their antennas. Once a medicine bottle with a RFID tag is removed from the range of the antenna, a video-based vision method is triggered to track the movement of the tagged medicine bottle for further action analysis.

The paper is organized as following: Section 2 summarizes the related work. Section 3 describes the RFID-based medicine bottle identification and camera-based detection of medicine intake action. Section 4 demonstrates the preliminary results. Section 5 concludes the paper and lists the future research directions.

2. Related Work

Research in object detection pertaining toward sensing is a relatively new method, and work on scene analysis and world modeling has been tedious work for many researchers. Much of the current research work has been focused on location information since it provides an important source of context. Sensing a person standing before a particular machine allows automatic reconfiguration to suit preferences accordingly; tracking people allows for new security measures, while tracking objects permits for a more controlled and efficient working environment. Certain devices for remote object perception for blind individuals convey to users information about distal objects in the environment. Typically, such systems are either vision-based [23-26], or RFID tag-based [27-30]. Vision-based systems typically use computer vision algorithms to extract object features such as shape, size, and texture, from object images. However, these systems are with limited accuracy in uncontrolled real-world environments, due to lighting changes, scale changes, pose changes, motion blur, video noise, etc. [31].

According to VIDET [23], a vision system with a wearable assistive device for object perception permits users to sense objects from a certain distance. 3D models of objects are extracted through stereo vision. The user can input the 3D models through a wire actuated hepatic user interface wherein the movement of the user's finger is constrained by a wire as the user's finger moves across the 3D model. However, experimentation shows that users have difficulty understanding or recognizing objects through the devices provided by VIDET. Since the difficulty of perceiving physical objects through hepatic user interfaces, many researchers have taken different approaches to solve this problem, specifically the information of perceptual level [23-26]. Object features may be classified into predetermined perceptual categories, and then subsequently conveyed to users to invoke mental concepts [24]. This method is much easier as compared to working with physical representations, and can enable users to perceive remote objects in real-time. A few examples of approaches that work at a perceptual level include a suite of vision-haptic transfer algorithms [24] that extract hepatic features, specifically hepatic shape, size, texture and material, from an object's visual image; a vision-based wearable assistive device for landmark recognition [25]; and a handheld vision-based device for object feature detection including color and size [26].

In Badge3D [27], a recent assistive device for object recognition is designed to avoid obstacles for blind users while each object is attached to a bar code which can be found and extracted by computer vision algorithms. This device can identify objects, as well as estimate the object position. Objects that are untagged lying on the ground are identified through ultrasonic signals to assist users about the obstacles. Alternative visual tagging information is Cyber Code by Rekimoto and Ayatsuka [28], which presents a new visual tagging system, based on visual patterns. However, visual tags have several weaknesses. First, visual tags must be within the camera view otherwise it will not be detected. Secondly, visual tags can modify the appearance of objects where in contrast RFID tags may be embedded or attached inside objects. Finally, information stored in visual tags is more difficult to comprehend because there can be some loss of information in comparison to RFID tags. These limitations cause researchers to choose RFID tags for remote object perception rather than visual tags. A simple example of the use of RFID tags for navigation and remote object perception is RoboCart [29]. This is a robotic shopping assistant for individuals who are visually impaired. RFID tags are attached to grocery items, assisting users in attaining their products. Even though the system is designed for navigation, it can be applied to object detection. Another example of using RFID information is to assist users with navigation and remote object perception [30]. In this work, RFID tags are placed in a grid formation, both indoors and outdoors, to assist with navigation. Tags may be placed at entrances to inform users about objects within rooms, such as tables, chairs, etc., and their relative locations.

The existing work of integrating RFID and computer vision has been limited to the field of robotics with the purpose of simplifying object recognition, object localization, object tracking, and task planning [32-37]. A tag-based system for object recognition was developed by Mat *et al*. [32]. The appearance model of an object is extracted from its tag and is used to recognize the object in the scene. If there is no presence of the model contained in the tag that matches the object, a new appearance model is accumulated and stored in the tag. When enough models with different postures are collected for an object, a robot is able to recognize the object without the use of RFID. One limitation of this approach is that it is difficult to collect new appearance models in cluttered real world conditions. Boukraa and Ando [33] attempted to solve problems of object recognition from vision-based perspective with RFID. The object model is extracted from an object's tag and used for object registration via projective geometry. This approach can solve problems in real-time model-based object recognition as only a subset of object models need to be considered during object recognition. According to the work of Takermura *et al.* [34], the process of creating CAD object models in [23] are time consuming and expensive. Furthermore, not every object will have a CAD model available from the manufacturer.

Another approach to an object tracking system is that it uses a CAD model obtained from an object's tag, which was developed by Hotani *et al.* [35]. In this system, visual tags are used rather than RFID tags, yet some tags may not be identified if they are not in the camera's view. Chong *et al.* [36] have developed a system that combines both RFID and computer vision for task planning in robotics. An experimental setup was developed where a robot's task was to clear dishes from a table after customers have finished eating. The robot obtains CAD models via RFID tags placed on each object. The models are used to find the location and position of each object. This information can guide the robot in learning how to grasp objects so that it may pick up and remove each tagged object from the table. A more recent approach developed by Kim *et al* [37] utilizes a robotic system for object recognition and localization. They proposed the use of smart tags because these tags contain an active landmark (IRED) and a data structure consisting of geometrical, physical and semantic information. Whenever a tagged object is detected or read, its IRED is activated. Then, the robot searches the active landmark, which has a flickering light. When the light is found, stereo vision on a pan-tilt mechanism is used to find the object's depth, size, and pose. This information is then used to grasp the object.

3. Monitoring Activity of Taking Medicine by Incorporating RFID and Video Analysis

In this section, we will first briefly introduce the framework of our proposed medicine intake algorithm. Then describe the methods of RFID-based medicine bottle identification and vision-based method for monitoring activity of taking medicine.

3.1 Overview of the Proposed Framework

The flowchart of our proposed medicine intake monitoring system is displayed in Figure 1. Tagged medicine bottles are stored in a medicine cabinet. RFID readers detect the tagged medicine bottle in cabinets simultaneously as camera-based vision is running in the background. The RFID tags data are read from the database of medicine bottles. Once a bottle is removed, its ID is identified to trigger the camera-based method to continue monitor the activity by integrating face detection and tracking, mouth detection, background subtraction, and activity detection.

The background subtraction based on blob tracking method is applied to detect moving blobs in determining if the medicine is a distance close enough to the mouth.

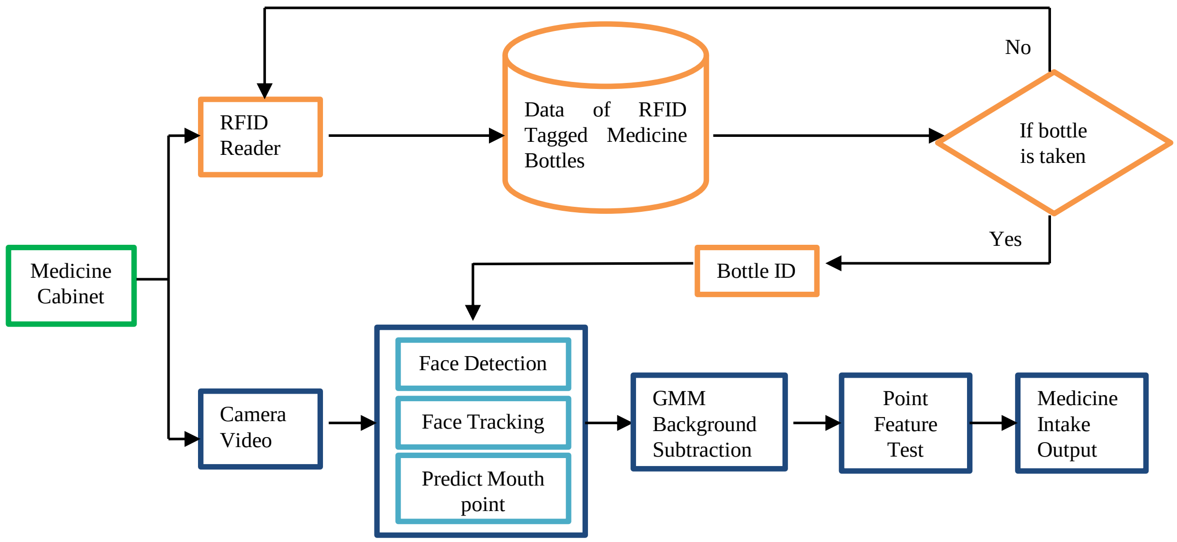


Figure 1. System diagram of the proposed medicine bottle identification by combining video camera and RFID sensors.

3.2. RFID-based Medicine Bottle Identification

We employ RFID readers, antenna, and host interface boards for identifying tagged medicine bottles, where the reader identifies tags in the field of the antenna. As shown in Figure 2(a), we use SkyeTek M2 RFID readers [22], which can read a radio frequency of 13.56 MHz in our experiments. Remote identify tags IS015693 [38] are used to provide faster readability identification of RFID tags. The RFID tags can record a wide range of information including a unique ID, and other supplementary information such as manufacturer, product type, etc. These tags can be used in objects of any shape, including ones like bottles, chairs, or clothes, as long as it is not near metal, steel, or liquid.

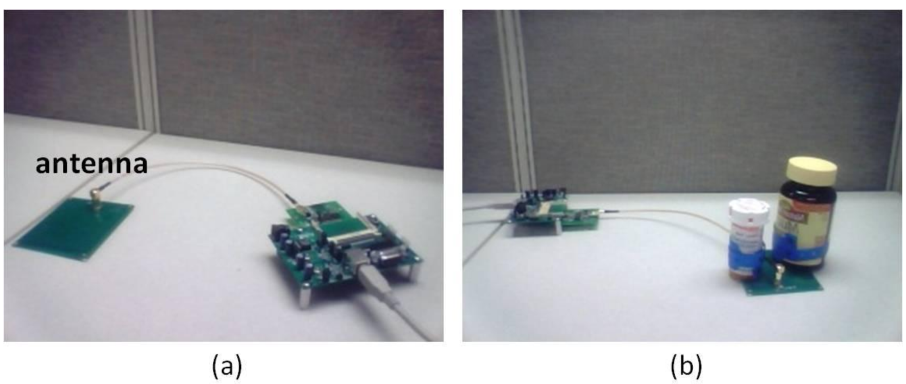


Figure 2. SkyTek M2 RFID system includes a USB connection, antenna, RFID reader, and a host interface board for identifying medicine bottles.

In our experiments, we observed that the SkyeTek M2 reader can read or detect a tag at a maximum height of 2.937 inches, with a radius of 1.375 inches. The detecting range for the M2 takes a volume of 5.8143 cubic inches, which can keep several tagged medicine bottles within our application. It can support multiple communication interfaces such as TTL Serial, SP1, and USB 2.0. Each RFID tag has a special unique ID that can be read by the SkyeTek M2 reader. Therefore, each RFID tagged medicine bottle can be retrieved and used to obtain the information of the medicine saved in the tag database. When a user encounters multiple tagged medicine bottles, RFID reader and antenna typically resolve any collisions, and eventually retrieve each tag ID.

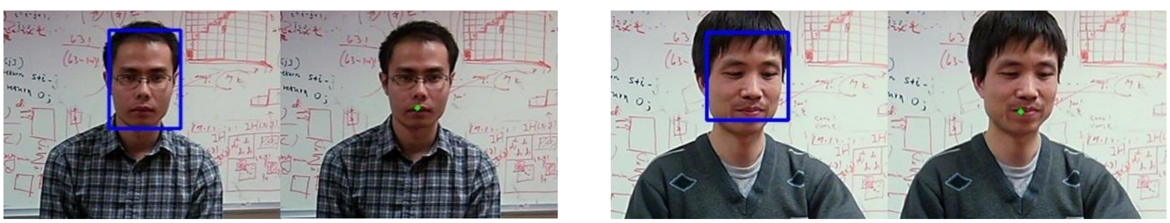
In our development, as tags are retrieved, object information is conveyed to the user. Figure 2(b) demonstrates the setup of identifying medicine bottles with RFID tags (in blue color). All the RFID tagged medicine bottles have to be within a certain range of the antenna in order to be detectable. In order to validate the effectiveness of RFID tags in this system, we evaluate its performance in detecting 2 medicine bottles in range of the antenna with 10 trials. Each trial consists of different ways to align the medicine bottles (e.g. sideways and upside-down) in different angles. Our RFID-based method achieves an accuracy rate of 100% for identifying two medicine bottles. However, due to the limited range of the antenna, the tagged medicine bottle cannot be identified by RFID-base method once it is moved out of the range of the antenna. Therefore, a vision-based method for monitoring activity of taking medicine will be triggered to continue track an RFID medicine bottle and detect if the medicine is taken.

3.3. Vision-based Method for Monitoring Activity of Taking Medicine

Although RFID can achieve a very high accuracy in medicine bottle identification, the medicines must be in range of the RFID antenna. To continue to identify the object to see if the medication intake has occurred once the object is out of range, we propose using a video camera to monitor the person who is taking the medicine. The vision system consists of three main components: (a) detecting face-tracking and estimating mouth position, (b) detecting moving objects by Gaussian Mixture Model (GMM) based background subtraction, and (c) monitoring activity of taking medicine.

3.3.1. Real-Time Face Detection, Tracking, and Mouth Position Estimation

Viola and Jones [40] proposed an image-based face detection system, which can achieve remarkably good performance in terms of detection accuracy as well as speed. In our system, we employ the method of Viola and Jones to detect face [40] and then employ the Camshift algorithm to track the face [41]. Camshift is an effective and efficient tracking algorithm. Since the mouth is located in the lower 1/3 of the face, once the face is detected and tracked, we then simply estimate the mouth position based on the face coordinate information as shown in Figure 3. The blue boxes indicate the detected faces and the green points represent the estimated mouth positions. The location of mouth does not vary much for different users.



(a) (b)

Figure 3. Illustration of face detection, tracking (blue boxes), and mouth detection (green points).

3.3.2. Moving Object Detection based on GMM Background Subtraction

We employ a GMM-based adaptive background subtraction method [39] to detect moving objects. This method is to model each background pixel as a mixture of Gaussian models. The Gaussian models are evaluated using a simple heuristics to hypothesize which are most likely to be a part of the background process. The distribution model of a particular pixel (x_0, y_0) is defined in an image sequence. At any time, t, the set of intensity values to this pixel is $\{X_1, .., X_t\} = \{I(x_0, y_0, i) : 1 \leq i \leq t\}$, where I indicates the image sequence. If each historic value of the pixel is modeled by a mixture of K Gaussian distributions, the probability of the current observed pixel value can be written as:

$$P(X_t) = \sum_{k=1}^{K} w_{k,t} * \eta(X_t, \mu_{k,t}, \Sigma_{k,t}) \quad (1)$$

where $w_{k,t}$ is the weight of the K[th] Gaussian model at time t, and $\sum_{k=1}^{K} w_{k,t} = 1$. It reflects the reliability of the pixel value described by current Gaussian model. $\mu_{k,t}$ is the mean and $\Sigma_{k,t}$ is the covariance of the K[th] Gaussian component at time t. K is the number of Gaussian distributions which is depended on the computing power of the system, and it is usually 3 to 5. A Gaussian probability density function can be written as:

$$\eta(X_t, \mu, \Sigma) = \frac{1}{2\pi^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t-\mu_t)^T \Sigma^{-1}(X_t-\mu_t)} \quad (2)$$

where n indicates the dimensions of a Gaussian distribution function. Generally, the Gaussian model with the higher weight $w_{k,t}$ and the lower variance $\sigma_{k,t}$ is considered to match the background model. When a new object appears in the scene and none of the current K distributions matches it, therefore a new model will be generated or variance of known models increased, and its variance will always be greater than that of the background pixel before a moving object stops. Firstly, to find the model of representing the background, the Gaussian models are ordered by the value of $\frac{w}{\sigma}$. If a Gaussian model gains more information about the background, this value will increase while the variance decreases. After approximation of the Gaussian mixture model parameters, there is an effective order from the model, then the model is matched, followed by a distribution of the candidate background, which becomes the most likely background distributions, being on top and the less probable background distributions are at the bottom of the order. The first B distributions are used as the background, where B is estimated as

$$B = argmin(\sum_{k=1}^{b} w_k > T) \quad (3)$$

where the threshold T is a measure of the minimum fraction of the background model. Moving object detection provides a classification of the pixels in the video sequence into either foreground (moving objects) or background. GMM background subtraction method is used to segment the moving blobs in an image sequences taken from a static camera. In this paper, we track the center points of each moving blob obtained by the GMM-based background subtraction method as shown in Figure 4. We further calculate the distance between the center positions of the moving blobs and the mouth position to detect if an activity of taking medicine has occurred. Given a set of blob points which changes from frame to frame in an image sequence, $(x_1, x_2, \ldots, x_n)$ where each point is a 2-dimensional position vector, the n points into k number of frames $(k \leq n)$ $\boldsymbol{S} = \{S_1, S_2, \ldots, S_k\}$ we then apply a distance formula from the mouth to the center of the blob points

$$Y = \sum_{i=1}^{k} \sum_{x_j \in S_i} ||\boldsymbol{x}_j - \boldsymbol{\mu}_i||^2 \quad (4)$$

where $\boldsymbol{\mu}_i$ is the set of mouth position in S_i.

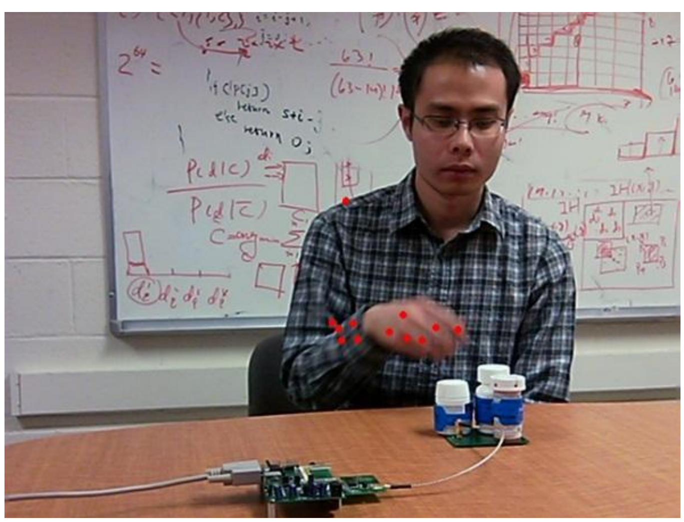


Figure 4. GMM-based background subtraction to calculate the center positions of each moving foreground blob (red points).

The distance of each moving blob and the mouth position in each frame is employed to determine whether the user has taken a medicine or not. If the distance is less than a threshold, this moving blob reaches the users mouth. Therefore, we classify that the user has taken their medication. We obtain this threshold from empirical observation based on a public available dataset of daily activity monitoring containing high resolution video sequences of complex actions such as cooking, eating, drinking, etc. [42]. In our experiment, we selected all videos expect those that involve an action or gesture close to a person's face. Therefore, videos that include eating a banana and answering a phone were all being tested to calculate a threshold point value tested with a total of 15 videos. Figure 5 demonstrates several successful video frames where hands (red points) are very close to the mouth (green points) using the proposed camera-vision method.

Figure 5. Some samples of the activities of daily living dataset where hands (red points) are very close to the mouth (green points). First row represents answering phone and the second row for eating a banana.

4. Experiments and Discussion

To evaluate the performance of the proposed method, we conduct a test with 14 users trying to acquire RFID tagged medicine bottles under different conditions such as a user actually taking the medicine (see Figure 6), a user picking up the medicine bottle but refusing to take it (see Figure 7), and a user has been interrupted after the user picked up the medicine bottle but before he/she takes it (See Figure 8). In total we capture 20 videos including 12 positive videos (a user actually taking the medicine) and 8 negative videos (a user picked up a medicine bottle but does not actually taking the medicine). The medicine bottles were selected randomly, varying in shape, size, texture and material that can fit in the given antenna's size. A component was developed in order to enable the portable RFID reader to communicate with our vision system. When RFID tags are read by the reader (the reader is set up to continuously read), unique tag IDs are stored in a resizable array. This array is periodically checked by the system for new entries. When a tag ID is detected, the system looks it up in the database and retrieves the medicine information from the database. When all the tagged medicine bottles are returned back to the RFID antenna's range, the vision-based method for monitoring activity of taking medicines will be terminated.

In our experiments, the proposed RFID-based method for medicine identification is developed by using the SDK on Windows in C++ for RFID sensing provided by SkyeTek. The medicine identification achieves 100% accuracy. The vision-based method for monitoring activity of taking medicines was implemented by using Intel's Open Source Computer Vision library. Once a tagged medicine bottle leaves the field of detection, the vision-based method is activated to continue tracking human actions, determining if the medication has been taken. The resolution of our camera is 320x240 pixels. If the user does not take the medication as pre-designed schedule, the system will prompt the user in every 2-minute intervals to take the medication. In our design, if a user does not pick a tagged medicine bottle within 2-minute intervals, the system will provide a reminder to the user to pick up the medicine bottle. In addition, if the user selects a wrong medicine bottle, the system will prompt the user to pick the correct medicine bottle with more detailed description.

Our preliminary experimental results are promising for detecting if an activity of taking medications occurred by combining RFID sensors and cameras. Figure 6 demonstrates some example results of videos with the user actually taking the medicine by using our proposed vision-based system. The simple method of mouth position detection is not always accurate to find the exact location of the mouth since it is only based on face detection and tracking. Our system can handle the negative videos where the user picked up the medicine bottle but does not take the medicine. Figure 7 and Figure 8 illustrate two examples of video sequences where the user does not take medicine respectively.

One main limitation of the proposed method is that the range of the RFID antenna is too small. Therefore, if the medicine bottles are placed at a position out of the range of the RFID antenna, the medicine bottle will not be able to be identified. Furthermore, our method cannot handle the complex situations if there are more than one people appeared in the camera view.

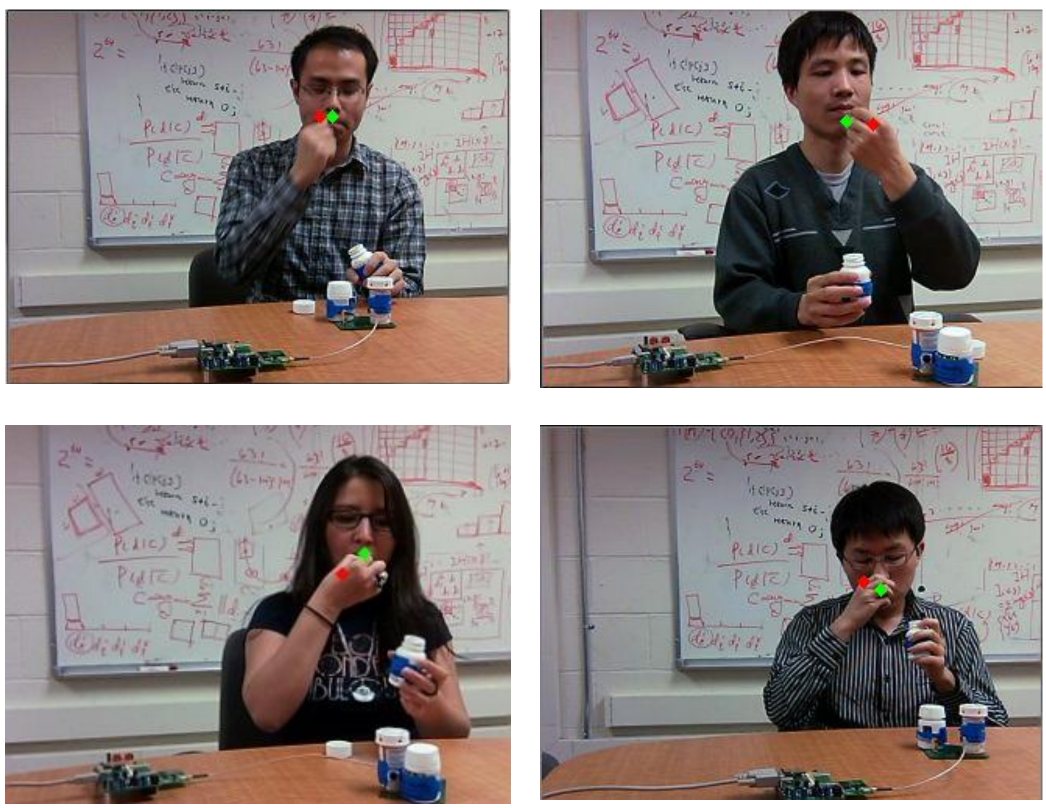


Figure 6. Some example results of monitoring the activity of taking medications by using computer vision-based method.

Figure 7: The proposed method can handle video sequences with negative examples of a user refuses to take the medicine after he picked up the medicine bottle.

5. Conclusion and Future Work

In this paper, we have presented a real-time system to identify medicine bottles and monitor the activity of taking medicines by integrating RFID sensors and cameras which has the potential to be used in applications for analyzing elderly people's activities. RFID enables the system to detect the medicine bottles through RFID tags and also detect if the user taking the bottle away from the range of the RFID antenna. When a medicine bottle is taken away from the range of the RFID antenna, the vision-based method will be activated to monitor if the user takes the medicine by combining face detection, mouth detection, and moving blob tracking based on GMM background subtraction. The preliminary results are promising. Our future work will focus on extending the system to identify and monitor the whole medicine cabinet by using more antennas and handle more complex situations.

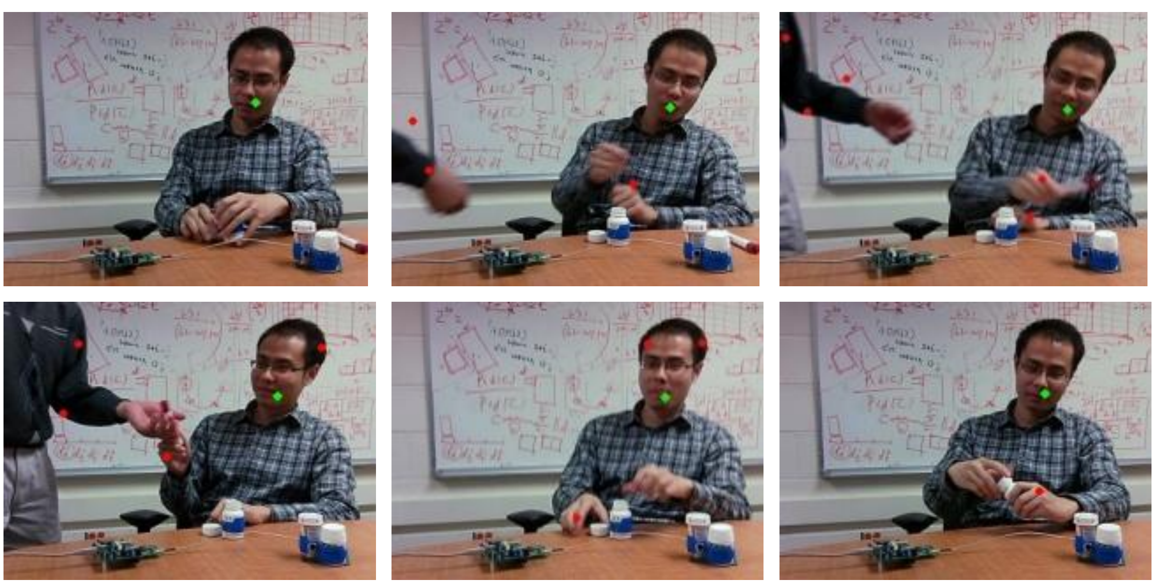


Figure 8: The proposed method can handle video sequences with negative examples of a user was distracted to take medicine after he picked up the medicine bottle.

Acknowledgement

This work was supported in part by NIH 1R21EY020990, NSF grant IIS-0957016, the REU program at DHS CCICADA of Rutgers University, and DHS Summer Research Team Program for Minority Serving Institutions.

References

[1] U. N. Population Division, DESA, Executive summary," January 2010. "World population ageing 1950-2050: [online] Available: http://www.un.org/esa/population/ publications/.

[2] Helping older people to take prescribed medication in their own home [online]. Available: http://www.scie.org.uk/publications/briefings/files/briefing15.pdf.

[3] R. Brookmeyer, S. Gray, C. Kawas, 1998. "Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset". Am J Public Health 88 (9): 1337–42. PMID 9736873.

[4] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, H. Arrighi, 2007. "Forecasting the global burden of Alzheimer's disease". Alzheimer's and Dementia 3 (3): 186–91. PMID 19595937.

[5] World population prospects: the 2006 revision, highlights. Working Paper No. ESA/P/WP.202. Population Division, Department of Economic and Social Affairs, United Nations. 2007. http://un.org/esa/population/publications/wpp2006/WPP2006_Highlights_rev.pdf.

[6] C. Johnston, UCSF Division of Geriatrics Primary Care Lecture Series May 2001

[7] K. Romer, T. Schoch, F. Mattern, and T. Dubendorfer, "Smart identification frameworks for ubiquitous computing applications," in the First IEEE International Conference on Pervasive Computing and Communications, Fort Worth, TX, 2003, pp. 253-262.

[8] Y. Tsin, R. Collins, V. Ramesh, and T. Kanade, "Bayesian color constancy for outdoor object recognition," in the IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, 2001.

[9] T. Gevers and A. Smeulders, "Color-based object recognition " Pattern Recognition, vol. 32, pp. 453-464, 1999.

[10] D. Forsyth and J. Ponce, Computer vision: a modern approach Upper Saddle River, NJ: Prentice Hall, 2002.

[11] A. Torralba, K. Murphy, W. Freeman, and M. Rubin, "Context-based vision system for place and object recognition," in the Ninth International Conference on Computer Vision, Nice, France, 2003.

[12] P. Chang and J. Krumm, "Object recognition with color cooccurrence histograms," in the International Conference in Computer Vision and Pattern Recognition, Ft. Collins, CO, 1999.

[13] C. Papageorgiou and T. Poggio, "A trainable system for object detection," International Journal of Computer Vision, vol. 38, pp. 15–33, 2000.

[14] S. Giroux, A. Ayers, and H. Pigot, "An Infrastructure for Assisted Cognition and Telemonitoring," in the International Conference on Aging, Disability and Independence, Washington, DC, 2003.

[15] A. Sixsmith, "Pervasive computing and the well-being of older persons," in the International Conference on Aging, Disability and Independence, Washington, DC, 2003.

[16] H. Rowley, S. Baluja, and T. Kanade. Human face detection in visualscenes. In Advances in Neural Info. Proc. Systems, volume 8, 1995.

[17] H. Schneiderman and T. Kanade. A statistical model for 3D object detection applied to faces and cars. In CVPR, 2000.

[18] P. Viola and M. Jones. Robust real-time object detection. Intl. J. Computer Vision, 57(2):137–154, 2004.

[19] P. Viola, M. Jones, and D. Snow. Detecting pedestrians using patterns of motion and appearance. In IEEE Conf. on Computer Vision and Pattern Recognition, 2003.

[20] G. Bouchard and B. Triggs. A hierarchical part-based model for visual object categorization. In CVPR, 2005.

[21] R. Fergus, P. Perona, and A. Zisserman. A sparse object category model for client learning and exhaustive recognition. In CVPR, 2005.

[22] Readers and technology that optimize inventory asset usage and enhance customer service, "Products & Services Overview" 2004 – 2006. [Online], Available: http://www.skyetek.com/.

[23] K. Romer, T. Schoch, F. Mattern, and T. Dubendorfer, "Smart identification frameworks for ubiquitous computing applications," in the First IEEE International Conference on Pervasive Computing and Communications, Fort Worth, TX, 2003, pp. 253-262.

[24] L. Stefano,, and S. Mattoccia, 2002. Real-Time stereo within the VIDET project. In Proceedings of Real-Time Imaging, 8, 439-453.

[25] K. Kahol, 2006. Distal object perception through haptic user interfaces. Doctoral Thesis. Arizona State University.

[26] A. Luo,X. Zhang,W. Tao,, and H. Burkhardt, 1999. Recognition of artificial 3-D landmarks from depths and color: a first prototype of electronic glasses for blind people. In Proceedings of SCIA99.

[27] A. Hub, J. Diepstraten, and T. Ertl, 2004. Design and development of an indoor navigation and object identification system for the blind. In Proceedings of ACM Sigaccess Accessibility and Computing, 77-78, 147-152.

[28] G. Iannizzotto, C. Costanzo, P. Lanzafame, and F. La Rosa, 2005. Badge3D for visually impaired. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3, 29-36.

[29] J. Rekimoto and Y. Ayatsuka, 2000. CyberCode: designing augmented reality environments with visual tags. In Proceedings of Designing Augmented Reality Environments (DARE 2000).

[30] V. Kulyukin, C. Gharpure, and J. Nicholson, 2005. RoboCart: toward robot-assisted navigation of grocery stores by the visually impaired. In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, 2845-2850.

[31] S. Willis, and S. Helal, 2005. RFID information grid for blind navigation and wayfinding. In Proceedings of IEEE International Symposium on Wearable Computers, 34-37.

[32] K. Dana, B., van Ginneken, S. Nayar, and J. Koenderink, 1997. Reflectance and texture of real-world surfaces. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 151.

[33] Y. Mae, T. Umetani, T. Arai, and K. Inoue, 2000. Object recognition using appearance models accumulated into environment. In Proceedings of International Conference on Pattern Recognition, 4, 845-848.

[34] M. Boukraa and S. Ando, 2002. A computer vision system for knowledge-based 3D scene analysis using radiofrequency tags. In Proceedings of International Conference on Multimedia and Expo, 2, 245-248.

[35] K. Takemura, K. Ohara, K. Ohba, N. Chong, S. Hirai, and K. and Tanie, 2004. Knowledge distributed tag-based vision system. In Proceedings of the 1stInternational Workshop on Networked Sensing Systems.

[36] H. Hontani, K. Baba, T. Kugimiya, K. Sato, and M. Nakagawa, 2003. Visual tracking system using an ID-tag and the network. In Proceedings of SICE 2003 Annual Conference, 3, 2375-2380.

[37] N. Chong, H. Hongu, M. Miyazaki, K. Takemura, K.Ohara, K.Ohba, S. Hirai, and K. Tanie, 2004. Robots on selforganizin knowledge networks. In Proceedings of International Conference on Robotics and Automation, 4, 3494-3499.

[38] Y. Kim, C. Im, S. Lee, and H. Lee, 2005. Object recognition using smart tags and stereo vision system on pantilt mechanism. In Proceedings of ICCAS2005.

[39] Stauffer C, Grimson W. Adaptive background mixture models for real-time tracking[C]. Proceedings of 1999 IEEE Conference on Computer Vision and Pattern Recognition, 1999, vol. (2):246-252.

[40] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", in IEEE Conference on Computer Vision and Pattern Recognition, pp. 511-518, Kauai, USA, December 2001.

[41] Bradski Gary R. Computer vision face tracking as a component of a perceptual user interface. Proceedings of IEEE Workshop Applications of Computer Vision, Princeton, NJ: IEEE, 1998, p.214-219.

[42] Messing R, Pal C, Kautz H. “Activity recognition using the velocity histories of tracked keypoints” ICCV '09: Proceedings of the Twelfth IEEE International Conference on Computer Vision, 2009, Washington, DC, USA.