Visual System for Student Attendance Monitoring with Non-standard Situation Detection

O. Kainz*, D. Cymbalák*, J. Lámer* and F. Jakab*
*Department of Computers and Informatics, Košice, Slovakia
Ondrej.Kainz@tuke.sk, David.Cymbalak@tuke.sk, Jaroslav.Lamer@tuke.sk, Frantisek.Jakab@tuke.sk

***Abstract*— In this paper we propose a visual system for monitoring of student attendance in seminars and lectures. Basic idea is to estimate the number of people in the room using face detection algorithms and subsequently utilize face recognition algorithms to determine the actual identification of persons (students). Presented approach may be used for multiple purposes. Principal and primary purpose is to monitor attendance, which is possible thanks to university database. When implemented, system is expected to evaluate the attendance automatically or if necessary using collaborative authentication. Non-standard or anomaly detection is another feature that is to be provided by system, subject to tracking are hands, eyes and movement.**

Proposed solution is expected to improve and facilitate attendance monitoring of students at seminars and lectures. Further it may be used for anomaly prevention (e.g. cheating) and in specific cases for security or legal matters.

I. Introduction

The growing number of students at the educational institutions may increase the requirements for lecturers and increase the complexity of attendance monitoring. Solution presented in this paper may serve for multiple purposes: monitor the number of students in the specific area (for our case lecture hall), recognize each student within this area and thus provide information about the attendance, provide information about room usage (if implemented in the real-time provide instant information about room status).

Except for room attendance monitoring, another feature to be considered is security and cheating prevention. Information as acquired in the image format may be used for legal purposes for the case of some offence or rules violation, i.e. kind of abnormal situation when compared to standard behavior. By abnormal or non-standard situation we mean any situation that deviates from normal, subsequently lecturer or other authority may be notified. Analogous approach is taken for cheating prevention - specific student is being actively tracked and based on predefined model is estimated level of deviation.

Monitoring of number of students present during lecture or seminar may be beneficial regarding many factors. First of all, it eliminates the fraud, when some students sign for some friends or leaves during the lecture - this may be utilized by face recognition algorithms which allow to distinguish persons. Face recognition abolishes the need for attendance list and allows permanent checking of students. Another advantage is active monitoring of total count during lectures - this way attendance of specific lectures may be estimated for specific period of time.

Often might happen that some room is not being utilized, e.g. when specific room is assigned to some study group and lecturer is absent. This may result in vacant room while other lecturers are unaware of its status. Using proposed system we eliminate this disadvantage and propose solution that will inform other lecturers that room is still vacant. Overall statics of room utilization may be eventually stored in database and used in future for statistical purposes.

II. State of Art

As an approach for detection we chose discriminative method [1] that utilize training and features extraction techniques. In order to detect human being in the image there is a need to implement some descriptor for feature extraction. Descriptor that outperforms existing ones proposed by Dalal and Triggs[2] is Histograms of Oriented Gradient (HOG). Ha, Kang and Lee [3] utilized HOG in learning-based human segmentation algorithm, another study by Neagoe, Tudoran and Neghina[4] utilized HOG together with Concurrent Self-Organizing Maps (CSOM) with success rate of 99.7 percent in the detection of pedestrians. Approach taken in this study was following proposal first authors of HOG descriptor [1] - image is divided into overlapping small pieces (windows) containing extracted HOG vectors, these vectors are processed by linear SVM.

Another options to enable learning and classification is to utilize artificial neural networks (ANN), e.g. [5]Kohail used Active Appearance Model (AAM) for feature extraction and subsequently using this inputs trained Multi-layer perceptron neural networks (MPL). However we decided not to utilize ANN bearing in mind specific disadvantages, as presented byW. Ma and M. Lu[6]:

- ANN might miss the "big picture" since they converge on local minima not global maxima,
- ANN might end up considering noise as a part of pattern.

978-1-4799-7740-6/14/$31.00 ©2014 IEEE

There has been many studies regarding attendance itself and many of them were successfully implemented in the real world environment. Gatsheni,Kuriakose, and Aghdasi[7] utilized mostly hardware solution for this kind of tracking - RFID automatic registration, where radio-frequency identification (RFID) tags were embedded in student cards. Study however doesn't point out the case when student carries multiple student cards.

Analogous research using RFID technology was perfomed by Kassim et al. [8]. Benyó et al. [9] on the other hand used near field communication (NFC) technology together with biometric identification (fingerprint), for checking was utilized contact-less terminal connected to server entity using wireless connection. Terminal included NFC and fingerprint reader. System was put into operation in real-world environment and was used for over 200 000 identifications. What authors do not address is the case of identity switch after the identification itself. Another research by Moksin and Yasin [10] utilized barcode scanner, again study does not address the possibility of multiple cards per person or identity theft. Avireddy et al. [11] proposed attendance tracking system MITSAT (Madras Institute of Technology Student

Attendance Tracking system) using Bluetooth technology, where each student is provided with Bluetooth transmitter with unique ID that is being paired with access point. Authors do not address above stated drawbacks. Agulla, Rúa and Castro [12] developed multibiometric-based student access control and attendance modules, study presents several option of student attendance tracking - face tracking, voice verification and fingerprint authentication. Study reports that 3 out of 16 videos required collaborative authentication which is 18.75%. Module was designed for single user only, i.e. presumption that student is sitting in front of the computer machine.

III. System Design Proposal using static images

The main role of proposed system is attendance monitoring and this is to be the main aim considered during the design stage. Several of above presented studies did not apply deeper identification solution for student, most of them were hardware-based, i.e. some external device such as RFID or NFC chip had to be used, often together with biometric identification. Our solution is rather similar to Agulla, Rúa and Castro [12], however we do not utilize voice nor biometric (fingerprint) verification. Reason why is that our solution is focused to cover larger area, that is to track multiple persons (students) at the time - this eliminates the possibility of voice authentication. Biometric authentication was omitted due fairly good face recognition results. Bhati et al. [13] carried out a comparative analysis of neural networks for face recognition with recognition rate reaching mostly from 80 to 100 percent. Wang et al. [14] used for face recognition three methods using SVM with accuracy of 80 to 100 percent. Due to high recognition rate and availability of students' images from university system we will be able to sufficiently well estimate student attendance. Our system thus does not include any hardware devices that is to be carried by student. If some error occurs, images are to be available subsequent collaborative authentication.

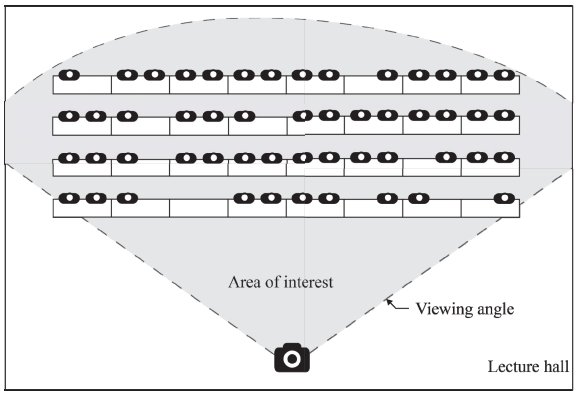

Figure 1. Camera setup: Placement example

During system design we take into the consideration several factors:

- students vary based on specific lecture/seminar,
- lecture halls have various capacity,
- setup complexity,
- hardware resources,
- cost.

Fluctuation of number of students that attend specific lecture hall is dependent on size of study group, size of lecture hall and actual number of students that attend lecture. Bearing this in mind proposed solution has to be adaptable to fit these constraints. Camera placement is therefore very important and must provide view on all the places within the room. See Fig. 1 for an illustrative example. Depending on size of the room different cameras may be used, for larger halls are to be required cameras with wide or ultra wide lenses. Room size has direct correlation to room capacity, based on this predefined cameras may be utilized. System setup and type of camera is highly dependent on above mentioned. Resources that are being utilized in this project are mostly open source,

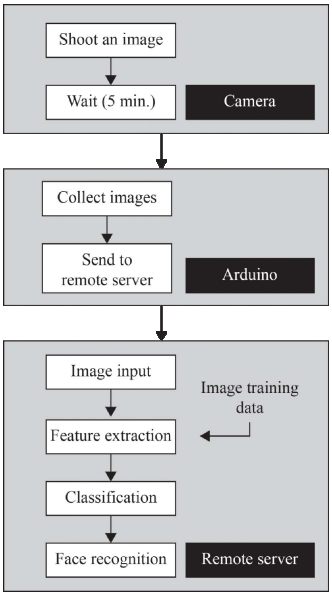

Figure 2. System design proposal

except for hardware digital cameras. As for the hardware part of this system is utilized Arduino micro-controller that provides collection of images that are send to remote server for further processing.

Fig. 2 depicts actual functionality of the system itself. Three components are essential for system functionality, these are: camera, Arduino micro-controller and remote server. Intermediary devices such as router or switch are not being considered and are not described in the solution. Topology is determined by specific needs. Camera is hardware device directly connected to micro-controller and based on set interval it shoots the image, once image is shot it is transferred to micro-controller which subsequently sends data to remote server. Server is the core part which provides processing of image data as described in the picture. Server then evaluates the image or image sequence based on the image data as provided from the outer database. Once data is collected and evaluated they are saved in database and accessible for learner. Same holds for images that are stored on remote server. As mentioned earlier - image data may be used for various purposes i.e. collaborative authentication, legal matters or security issues.

IV. System Design Proposal using Video Input

Using video camera it is possible to improve proposed system by deployment of advanced computer vision techniques. The whole problem of anomaly or threat detection in video is based on few steps. We assume the most important:

- Object detection and recognition: the main feature, which provides object and data mining from input images.
- Object tracking: helps to re-recognize previously detected object and provides the background for dynamic motion anomalies.
- Contextual analysis: it adds possibility to apply the designed model in many environments.
- Anomaly detection: detect any anomalous behavior in static or dynamic scenes.
- Anomaly prediction: helps to predict recognized objects behavior.
- Knowledge sharing (in a case of dynamic systems): provides knowledge data exchange between systems.

A. Object detection methods

There are four main principles for object detection in image or video presented in [15]:

- Based on object color: threshold, image moments
- Based on object contours: contour detector, edge detector
- Based on object pattern: pattern finders
- Based on similarity: based on similarity features extraction

As authors in [15] describes, the main method, to detect object based on color difference is thresholding. This means, that object is in color, which is set as threshold (+- offset) and this color is set on output to white. All other colors out of threshold range are converted to black. Result is a black/white image, where white represents the detected object. But this output is not final because it actually contains just information, that the object was detected. Other information is still encapsulated in result. Export of object position is therefore needed. For this purpose is possible to use object positioning based on image moments. For example image moments of order 1 can be used to calculate the object center position.

Second principle usage of geometrical description of object is described by same authors in [15]. For example it is possible to use count of edges, points, size of bordered area etc. The keystone for this kind of detection is one of edge detection algorithm, which detects edges.

Last two methods are based on object pattern or similarity. These methods have better detection results and resistivity from noises, geometrical and affine transformations as previous techniques [15].

Pattern based techniques work with all image pixels. The template pattern is continuously sliding through image in each direction. Through this action, the metric matrix calculated and stored in memory. This metric represents the quality of match between template and image. On the next image is the visual representation of input image, finding pattern and output metric matrix. The most white area in output aims the most probably match region [16].

One of the strong similarity based techniques is SURF

Figure 3. Face detection and face recognition in room

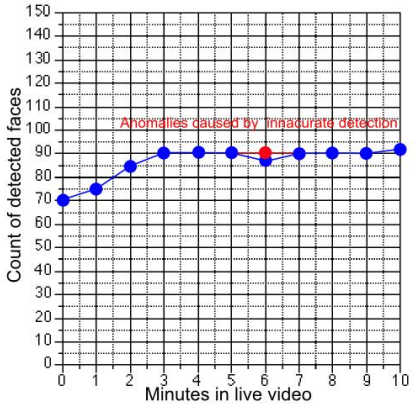

Figure 4. Trendline of count of detected faces in room

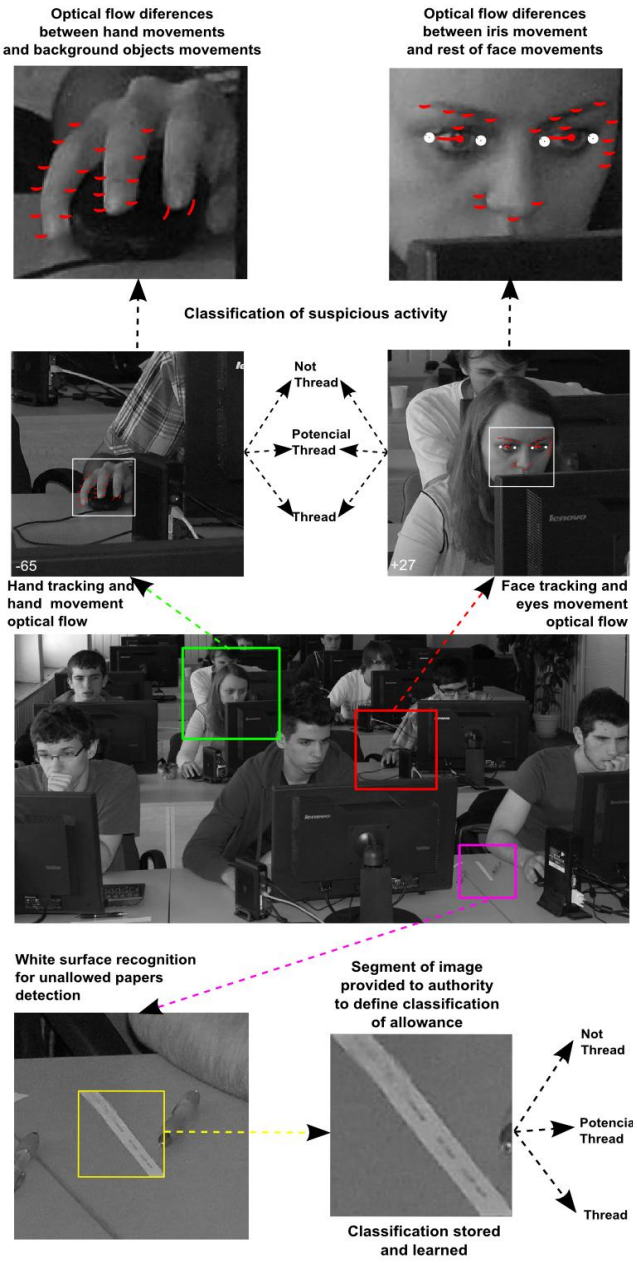

Figure 5. Threat alert and threat classification by eyes and hands tracking and suspicious objects recognition.

algorithm, which is based on similarity points and neighborhood vector representation. This feature makes algorithm resistant from geometric or affine deformations. It consists from two main parts – detector and matcher. Detector plays the main role in data mining from image. Matcher then compare obtained features with defined pattern [17].

In our solution the methods of object detection and object recognition will be use in cooperation of face detection, face identification and for counting attendance in the room which is monitored by camera (Fig. 3).

Related to the reliability of detection method and the limited view angle from capturing camera the count of detected people in the room must be approximated and local drops caused by unclearly view on head will be flatten (Fig. 4).

B. Object tracking methods

The most simply, but strong method for tracking a known object is Optical flow. It is based on pixel position difference between two or more frames in sequence (video). Method can track one single pixel, or more pixels in group. The computing requirements are low comparing to other tracking methods. In case of object tracking, the method has no dependency to object shapes. It only tracks the group of pixels. This can be considering as object shape self-learning mechanism.

The combination of object tracking methods with object detection will be used as tool for anomalies

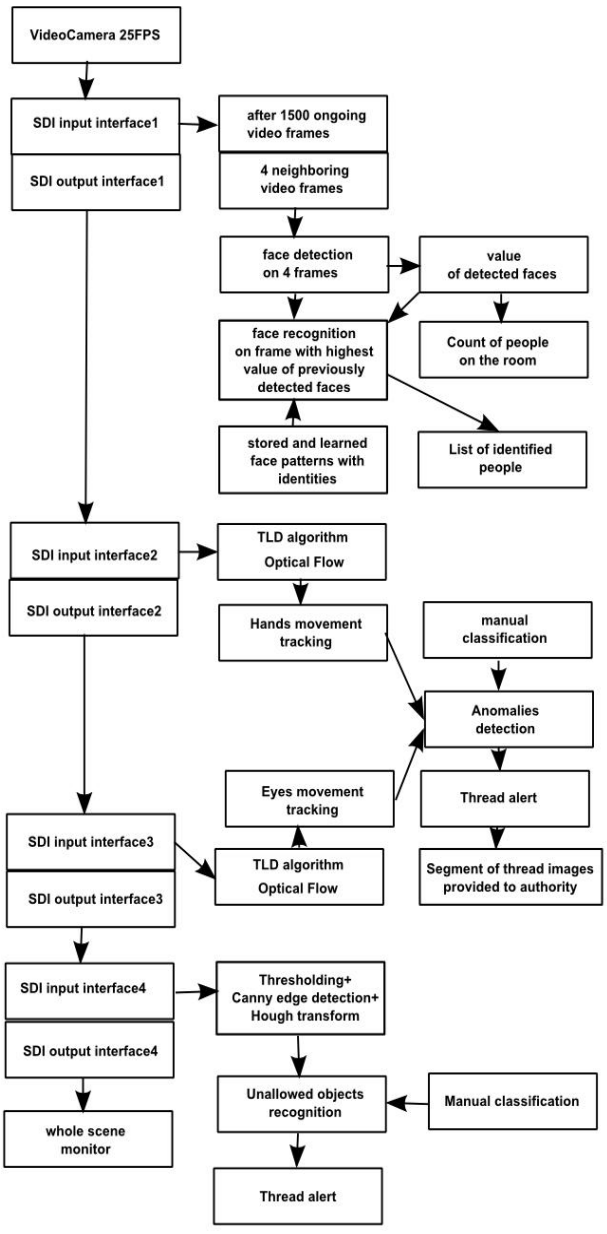

Figure 6. Interconnection between proposed system components

detection and threat alerting.

We are considering 3 functionalities of threat alerting (Fig. 5):

- Hand tracking and movement anomalies
- Eyes tracking and movement anomalies
- Suspicious objects recognition

The proposition of interconnection between components (Fig. 6) is based on using four fast SDI interface input cards which clones live image from one video camera to each others. This live image is processed separately with object recognition tools and object tracking algorithms. Also there is modules for anomalies detection and classification of threat with learning ability to next phase of new generation full automatic threat detection system.

V. Conclusion and Future Works

Student attendance used to be usually achieved by classical (rather traditional) way – this means papers or more novel approaches by hardware tools such as radio-frequency identification (RFID), near field communication (NFC), biometric identification or combination of just presented. What we proposed and partially implemented in our solution does not require student to carry any hardware device nor to perform some kind of direct biometric identification.

Our approach utilizes more indirect approach where student is identified by camera, i.e. face is matched to the one stored in the database. In this way students are automatically and indirectly monitored during classes and lectures. Another output information that is provided to instructor is actual number of students in the room or lecture hall. Object tracking methods are utilized to track the anomalies in picture and if threat (non-standard situation) is detected it is being manually classified.

This study was tested on rather small number of classes and specific sample number of students. As for the future work, more classes and larger databases are to be used in the process of testing. Utilized databased corresponded to university database however it included smaller sample. Solution is to be implemented also in large lecture halls, e.g. lecture hall with the capacity of 500 students, and possibilities of utilizing multiple devices for monitoring are being analyzed. Utilization of multiple devices, i.e. some form of multi-camera system with multiple view (or views that will expand the view angle) is expected to enhance effectivity and accuracy of the evaluation. Another goal is to utilize neural networks for classification of threats and face recognition and eventually provide comparison of both methods.

Acknowledgment

We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

We would like to thank instructors of courses the at Computer networks laboratory at Technical university of Kosice. Especially Michal Kovalcik for his cooperation during Computer networks courses.

References

[1] G. Bouchard, and B. Triggs, “Hierarchical Part-Based Visual Object Categorization”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) vol. 1*, pp. 710-715, June 2005.

[2] N. Dalal, and B. Triggs, “Histograms of Oriented Gradients for Human Detection”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1*, pp. 886-893, June 2005.

[3] J-E. Ha, D-J. Kang, and W-H. Lee, “Learning-based Human Detection Using Difference Image in Visual Surveillance”, *International Conference on Control Automation and Systems (ICCAS)*, pp. 2318-2319, October 2010.

[4] V-E. Neagoe, C-T. Tudoran, and M. Neghina, “A Neural Network Approach to Pedestrian Detection”, *Proceeding ICCOMP'09 Proceedings of the WSEAES 13th international conference on Computers*, pp. 374-379, July 2009.

[5] S.N. Kohail, “Using Artificial Neural Network for Human Age Estimation Based on Facial Images”, *International Conference on Innovations in Information Technology (IIT)*, pp. 215-219, March 2012.

[6] W. Ma, and M. Lu, Modeling wine preference by Machine Learning. [online] [Cited: September 30, 2014] https://sites.google.com/a/u.northwestern.edu/wensen/analysis

[7] B.N. Gatsheni, R.B. Kuriakose, and F. Aghdasi “Automating a student class attendance register using radio frequency identification in South Africa”, *4th IEEE International Conference on Mechatronics (ICM2007)*, pp. 1-5, May 2007.

[8] M. Kassim, H. Mazlan, N. Zaini, and M.K. Salleh, “Web-based Student Attendance System using RFIDTechnology“, *IEEE Control and System Graduate Research Colloquium (ICSGRC)*, pp. 213-218, July 2012.

[9] B. Benyó, B. Sódor, T. Doktor, and G. Fördős, “Student attendance monitoring at the university using NFC”, *Wireless Telecommunications Symposium (WTS)*, pp. 1-5, April 2012.

[10] M.I. Moskin, and N.M. Yasin, “The Implementation of Wireless Student Attendance System in an Examination Procedure”, *International Association of Computer Science and Information Technology - Spring Conference (IACSITSC '09)*, pp. 174-177, April 2009.

[11] S. Avireddyl, P. Veerapandianl, S. Ganapatil, M. Venkat, P. Ranganathant, and V. Perumal, “MITSAT-An Automated Student AttendanceTracking System using Bluetooth and EyeOS”, *International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s)*, pp. 547-552, March 2013.

[12] E.G. Agulla, E.A.Rúa, and J.L.A. Castro,“Multimodal Biometrics-based Student Attendance Measurement in Learning Management Systems”, 11th IEEE International Symposium on Multimedia (ISM '09), pp. 699-704, December 2009.

[13] R. Bhati, S. Jain, and D.K. Mishra, "A Comparative Analysis of Different Neural Networks for Face Recognition Using Principal Component Analysis and Efficient Variable Learning Rate", *International Conference on Computer and Communication Technology (ICCCT)*, pp. 526-531, September 2010

[14] Ch. Wang, L. Lan, Y. Zhang, and M. Gu, "Face Recognition Based on Principle Component Analysis and Support Vector Machine", 3rd International Workshop on Intelligent Systems and Applications (ISA), pp. 1-4, May 2011

[15] J. Lamer, D. Cymbalak, F. Jakab: Computer vision based object recognition principles in education, *ICETA 2013 : 11th IEEE International Conference on Emerging eLearning Technologies and Applications : proceedings* : October 24-25, 2013, Stary Smokovec. - Danvers : IEEE, 2013 S. 253-257. - ISBN 978-1-4779-2161-4

[16] Bradski, Gary; Kaehler, Adrian. *Learning OpenCV: Computer vision with the OpenCV library.* " O'Reilly Media, Inc.", 2008.

[17] Gossow, David; Decker, Peter; Paulus, Dietrich. *An evaluation of open source SURF implementations*. In: RoboCup 2010: Robot Soccer World Cup XIV. Springer Berlin Heidelberg, 2011. p. 169-179.

[18] OpenCV, Optical Flow [online]. [cit. 2014-10-10] Available online: http://docs.opencv.org/trunk/doc/py_tutorials/py_video/py_lucas_kanade/py_lucas_kanade.html