www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

Detection and classification of the behavior of people in an intelligent building by camera

Henni Sid Ahmed[1], Belbachir Mohamed Faouzi[2], Jean Caelen[3]

[1] Universite of sciences and technology USTO in Oran Algeria, laboratory LSSD, Faculty genie electrique, department electronique, BP 1505 el menouar Oran 31000 Algeria

[2] Universite of sciences and technology USTO in Oran Algeria, laboratory LSSD, Faculty genie electrique, department electronique, BP 1505 el menouar Oran 31000 Algeria

[3]Universite Joseph Fourier, Grenoble, F , LIG Grenoble computer laboratory ,domaine universitaire BP 53, 220 rue de la chimie 38041 Grenoble cedex 9 France

Emails: [1] sidhenni@yahoo.fr

Submitted: Apr. 10, 2013 ***Accepted: July 30, 2013*** ***Published: Sep. 3, 2013***

Abstract- an intelligent building is an environment that contains a number of sensor and camera, which aims to provide information that give the various actions taken by individuals, and their status to be processed by a system of detection and classification of behaviors . This system of detection and classification uses this information as input to provide maximum comfort to people who are in this building with optimal energy consumption, for example if I workout in the room then the system will lower the heating . My goal is to develop a robust and reliable system which is composed of two fixed cameras in every room of intelligent building which are connected to a computer for acquisition of video sequences, with a program using these video sequences as inputs, we use RGB color histograms

bench. Real-time synchronization between Energy Plus and Energy Manager was performed using a technology developed by [9] and [10] platform. Validation of energy management systems can be achieved through studies sociological via surveys or questionnaires. A classification of activities in the habitat was proposed in [11]. It relies on the automation level and the number of activations of a device. The recognition of people's behavior provide information that will enable the system to better manage and optimize energy consumption in the intelligent building. The detection of events is defined by the detection of attracting human attention [12] situations. The recognition of human activities is widely studied in recent years by the scientific community in computer vision. Reference articles [13, 14, 15] describe carefully the methods used in the state of the art. There are several studies on the analysis of human behavior in scenes. Recent studies by Forsyth et al. [16] and Pope [17] interested in the recognition of actions from video streams. In recent years, many approaches to the recognition of actions been proposed. They are described in the literature studies [18, 19]. These techniques have been classified according to the method of image representation and algorithm for classification of event as follows: Image representation: the calculation of features from video frames takes into account the temporal dimension. It is generally optical flow vectors [20], spatio-temporal characteristics such as cuboids [21] or Hessian characteristics [22]. A descriptor is then was designed to represent the video sequence. [23] calculate the descriptors by learning classifiers Ada-boost from low-level features. Some descriptors such as HOG / HOF [24], HOG3D [25] and ESURF (SURF extended) [22] are based on the analysis of local spatio-temporal points movement. The best methods of image representation are those who discern effective actions in different classes and running in real time. Classification of the action: the mechanism to classify an action. It can be performed using a classifier as SVM [26], SOM (Self-Organizing Map) [27], a Gaussian process [28] a distance function [29], or as a discriminant model HCRF (Hidden Conditional Random Field, Hidden Conditional Random Field) [30]. In order to test or compare different approaches, such as video databases KTH [31] is used. Activity Recognition is, from low-level information such as the numerical value of a set of pixels, to obtain a semantic representation is a natural language scene. The process of recognition of activity can be considered as a classification problem, where the various Representations of activities and recognition techniques, are involved, it is a very complex problem. There are many difficulties; we can mention a few here: There is a large variable intra class; same action performed by the

comfort. We selected 11 behaviors to know people's behaviors and each behavior is represented by a set of images, the fixed cameras are with great qualities which gave us images with high precision, we used the RGB color histograms and texture LBP represented for the different images obtained from video sequences, and SVM Light (Support Vector Machine) for implementation because it is a very powerful programming tool which is with a very low calculation and learning times with a very such as learning times our obtained results show the advantages of the use of histogram RGB and LBP in the field of detection and classification of the behavior of people in the building smart camera

II. Local Binary Pattern 2 (LBP)

LBP method, introduced by Ojala [45] is defined as a measure of invariant texture derived from the general definition of texture in a local neighborhood. The concept of LBP is simple binary code describing the local texture of a region is determined by calculate neighborhood thresholding with the gray level of the center pixel. All neighbors will then have a value of 1 if the value is greater than or equal to the current pixel and 0 otherwise, we will then multiply this matrix of 0 and 1 by weight and LBP and calculate the totals of these elements together to get the value of the current pixel LBP. This will thus of pixels whose intensity is between 0 and 255 as in a normal 8-bit image. Rather than, describe the image by the sequence motifs LBP, you can choose to texture descriptor histogram of size 255. Two variants of the LBP method were presented in [46] the defines first LBP for neighborhoods of different sizes, which can handle the texture at different scales, the second defines what is called the uniform LBP. Patterning uniform as a pattern having exactly 0 or 2 transitions (01or10) in a circular path. The notion of consistency in the LBP method is important to represent the information structural primitives such as corners and edges. Ojala found that only 58 of the 256 LBP patterns are uniform but experimentally we find in [47] only 90% of the motifs encountered in images are uniform. In this case, the size of the LBP histogram can be reduced significantly with histogram dimension 59. Each of the 58 first categories will contain the number of occurrences of the uniforms motifs. The last one contain the number of occurrences of all non-uniform patterns, this grouping to reduce the size without losing too much information. LBP method has proved very effective for the

representation scalable in terms of number of cells and the representation of cells. In [57], regions are described by two parameters representing the color density and the probability of co-occurrence of a color. These two parameters are derived from a multidimensional histogram RGB 3D quantified in 216 cells (6 £ 6 £ 6). In Blobworld the system [58], regions are described by histograms.

Histogram processing is a technique with numerous applications. The goal of exact histogram specification (HS) is to transform an input image into an output image having a prescribed histogram. Histogram equalization (HE) is a particular case of HS. Among the applications of HS let us mention invisible watermarking, image normalization and enhancement, object recognition, [59], [60]. Exact HS is straightforward for images whose pixels values are all different from each other. However exact HS (and also exact HE) is an ill-posed problem for digital (quantized) images since the number of pixels3 n is much larger than number the possible intensity levels L [61] . The clue to achieving *exact* HS is to obtain a *meaningful* total strict ordering of all pixels in the input digital image. Research on this problem has been conducted for four decades already [62]. The Local Mean (LM) method of Coltuc, Bolon and Chassery [61], the wavelet-based approach (WA) of Wan and Shi in [63] and the specialized variational approach (SVA) of Nikolova, Wen and Chan [64] are the state-of-the-art methods.

IV. SUPPORT VECTOR MACHINES

a. Introduction

Support Vector Machines (SVM) are a classification methods show good performance in solving various problems such as pattern recognition or classification of texts, particularly well suited to deal with high dimension data such as text, and images. SVM is a learning algorithm, to learn a separator [65]. This reduces the problem to know what a splitter. Give us a finite set of vectors in Rn, separated into two classes. Belonging to a group or another is defined by a label associated with each vector, on which is inscribed "Class 1" or "Class 2". Find a separator returns to build a function that takes a vector of our set, and can tell which group it is. SVMs are a solution to this problem, as would be simple learning by heart the classes associated with vectors of our set. Theoretically we encounter an infinite number of separators to distinguish between the two classes. The objective of the SVM method is to decide the best separator that maximizes the

Figure 1. Postures detected by the SVMs .Lights; **(a)** I am sitting in the office listening to music; **(b)** I am sitting in the office and I work in the computer; **(c)** I'm sitting on the bed and I read a document ; **(d)** I'm sliping in bed ; **(e)** I'm sliping in bed ; **(f)** I sat on the bed and I watch television; **(g)** I eat in the kitchen table; **(k)** I do the gym in the room; **(l)** I walked from one room to another in DOMUS (intelligent building); **(m)** I walked from one room to another in DOMUS (intelligent building); **(n)** I prepared coffee and I drink it up; **(o)** I prepared coffee and I drink it up; **(p)** I prepared coffee and I drink it up; **(q)** I entered DOMUS (intelligent building) and I sit; **(w)** I do the washing up

Fourier of Grenoble under the direction of Mr Jean Caelen. This lab contains an intelligent building called DOMUS Figure 3, Figure4, which is an apartment type F2 and fully equipped space, consisting of a kitchen and a bedroom containing a bed and a TV and window shutters, and a shower, toilet and a bedroom office that contains a desk and a computer and a stereo and a hallway, two fixed cameras in each room and two fixed cameras in the kitchen. There are a some difficulties in recognizing scenarios, among these difficulties that can be cited the confusion between the scenarios and the lighting problem precisely people’s shadow where objects reduce the rate of recognition. to solve these problems we have to play each scenario several times by the same person and other persons in different age, sex, and size, and the length of the body. we do vary the lighting at each scenario, to have a model for those who represent all the people we would like to acknowledge their behavior in order to obtain a very good learning, in addition we set a processing time scenario to avoid the confusion between scenarios; eg scenario to go in and to settle into the domus can be confused with the scenario of walking from one room to another in domus, in a interval time, it is for this reason that we must define a processing time scenarios to avoid confusion between the scenarios, although sure this processing time is low to preserve the notion of real time. After you complete this step that we spent a lot of time and from which the data will be treated, so be sure that have real data to get real results, we started the second stage and which is the conversion of videos sequences into images we look for the best software that could make this conversion with very good results. Then there's another step which is the construction of the input corpus behaviors recognition system in an intelligent building. In this step, we classified all images of each scenario in a directory, 70% of the input corpus is devoted for learning, and 30% for the classification or test, which gives us 2129 images for learning and 913 images for testing. After completing this very important step, we began the next step is preprocessing and parameterization of the images, different images that we obtained have a very good resolution and high quality this is due to the quality of camera used for filming scenarios, although we have it used the pretreatment. For the parameterization of the images as we opted for normalized RGB histograms and color histograms for the LBP texture, we used two histograms for each image one for color (RGB standard) and the other for texture (LBP) as it has been shown that the use of these two histograms gave a very good results. After applying the RGB and LBP histograms of images, each image is represented by a vector of the same size which is 319 elements. For the phase of implementation and after a long search we opted for

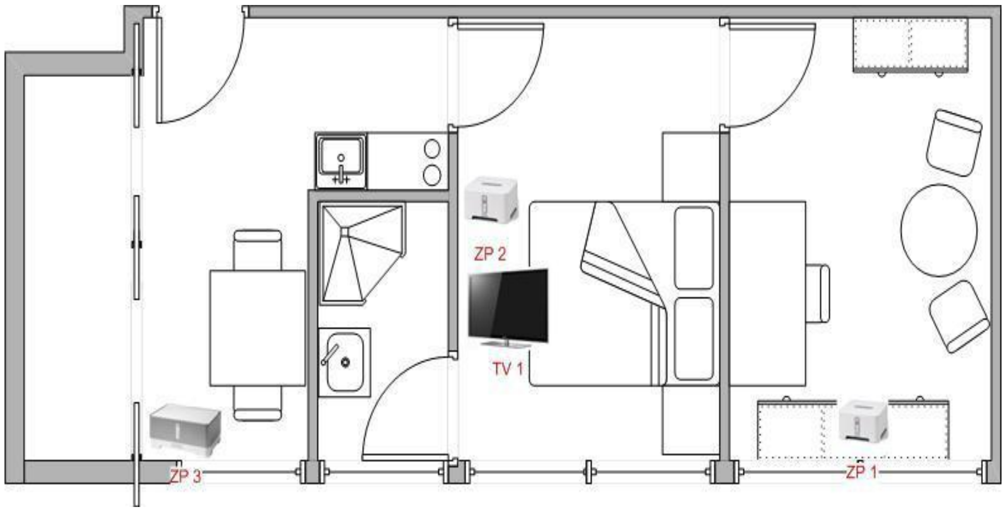

Figure 4. Archetecture intelligent building "domus Laboratory CTL team Multicom Université Joseph Fourier Grenoble France

b. Results

11 As we recognize behaviors, SVM will generate 11 models in the learning phase or training phase, so that each model corresponds to a behavior Figure 5 for the classification of behavior, the SVM use the 11modeles obtained in the learning phase, for the classification of a behavior, we use the models and the images that correspond to each behavior. The classification of the behaviors is according to the classification rate it means the behavior that we want is the one that have the highest rate of classification Figure 6.

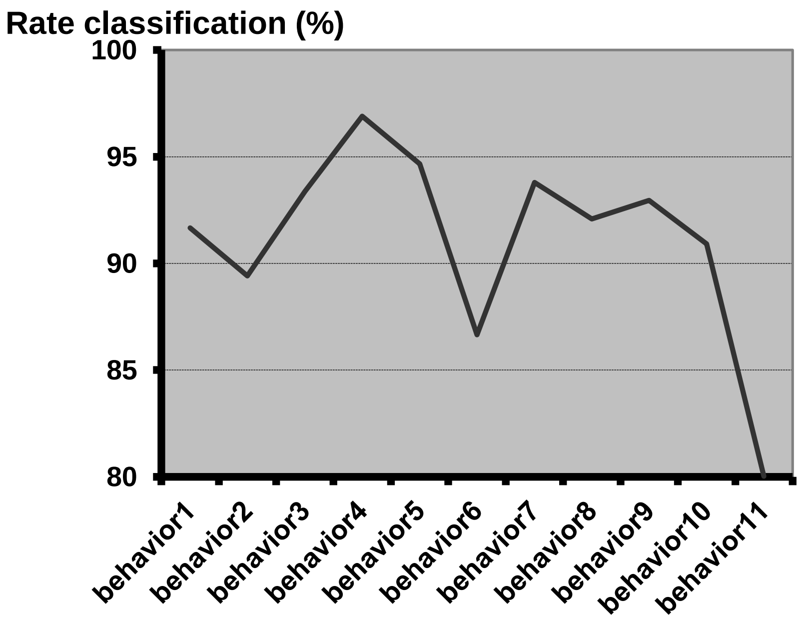

Figure 7. Showing the percentage of correctly Detected behavior for the linear kernel

Table 1: Painting of the recall and precision for the Detection of linear behavior for the core

behavior classified	**precision linear kernel**	**reminder for linear kernel**
behavior 1	46.15%	41.10%
behavior 2	32%	46%
behavior 3	100.00%	27.91%
behavior 4	89.47%	68.92%
behavior 5	56.98%	67.12%
behavior 6	33.62%	44.83%
behavior 7	59.32%	87.50%
behavior 8	70.59%	37.89%
behavior 9	59.71%	89.25%
behavior 10	64.71%	23.16%
behavior 11	28.51%	77.91%

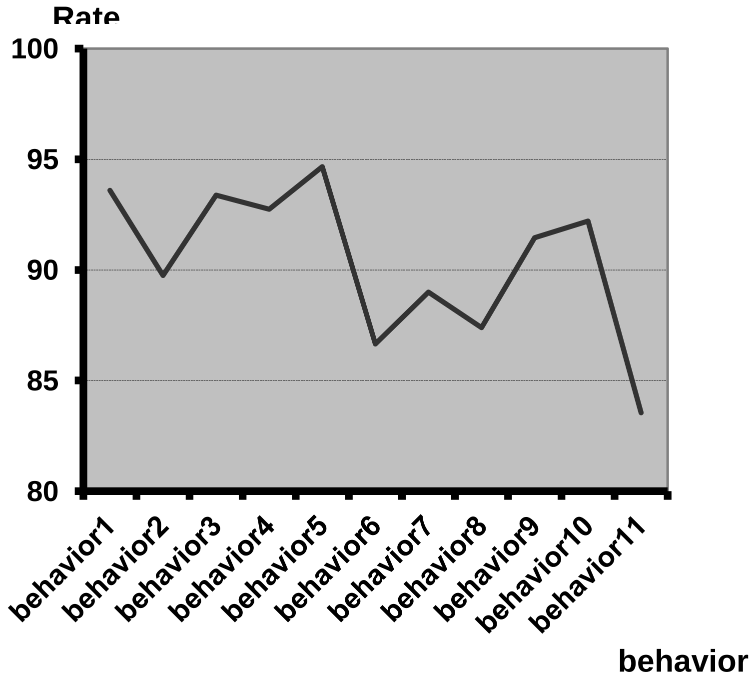

Figure 9. Showing the percentage of correctly Detected behavior Polynomial kernel

Table 3: Painting of the recall and precision for the Detection of the core behaviors for Polynomial

behavior classified	**precision Polynomial kernel**	**reminder for polynomial kernel**
behavior 1	66.67%	35.62%
behavior 2	33.5%	44.96%
behavior 3	100%	27.91%
behavior 4	100%	30.11%
behavior 5	65.75%	65.75%
behavior 6	33.62%	44.83%
behavior 7	42.94%	87.50%
behavior 8	34.51%	44.95%
behavior 9	74.07%	21.51%
behavior 10	71.15%	38.95%
behavior 11	31.04%	44.95%

According to the results in the tables (1,2,3,4) and in the figures (7,8,9,10) above, we can make the following deductions: 1) it is more difficult to detect certain behaviors than others, and this is due to the nature of the behavior detected, the behavior of going into the domus and settle in, is more difficult to be detected than the behavior of doing the washing up in the kitchen, 2) the classification rate of color images is higher than the black and white images, 3) The results presented above show a very satisfactory performance especially for the RBF kernel because in addition to the highest recognition rate the response time is the rapid compared to other nuclei especially the time of learning, 4) We noticed that there is behaviors that has a good classification rate but with reduced precision and recall in all nuclei used, we can see it clearly in the 2nd , 6th, 8th, and 11th behaviour; for the 2nd behaviour the cause id due to the rapid movements of the person, but for the 6th and 11th behaviour is due to the number of activities that this two behaviours contained. Finally, for the 8th behaviour the cause is the confusion between the behaviors which is due to the great similarity of the images.

We selected the best learning models that give us the best results are to say the highest rate of classification. According to the results obtained in the various tables the best model is the RBF.

c. Execution Time

Our application must generate a response in real time when an event is detected. The system is tested on a computer with a Pentium 3, 4 GHz and 1 GB RAM. The application scans 6-10 frames per second for resolutions respectively 640x480. Extracting images from videos and calculated their RGB and LBP histograms are a phase that consume the most time of calculation.

VI. CONCLUSIONS

We present in this paper a system for detecting and classifying behaviors of people in an intelligent building that class 11 behavior in real time. This work that has not been done in this field of research and this is what presents the originality of our work. This system allows us to characterize the activity of people in a room. This information will be useful to the management system of the building regulate the consumption of electrical energy in order to optimize (lighting, heating, etc..), We chose 11 behaviors for classification in order to increase the capacity of the system that manages and that optimizes the electrical energy consumption of intelligent building, because the increase of the classified behavior give additional information on the system behavior which will help people for better management of electricity consumption. The main difficulty to detect people is the great variability Intra class, By their clothes, their size, their weight, their haircuts, the appearance of two people can be very different. In addition, the human body is highly articulated, the number of possible poses is very large and the silhouette of a person changes over time. Our classification system was evaluated on a broad base of videos to get a good learning. We used several nuclei in the learning phase to select the best model, the experimental results are very satisfactory one the used method which allows us to say that the classifications of 11 behaviors in real time gave very good results using the RBF nucleu SVMLights and for learning and RGB histograms and LBP for the parameterization of images that is to say the images conversion into vectors of the same size. Future work will improve these results with the acquisition of new data such as speech and adding a class describing transitions between two activities. Our classifier is, for the moment, time on the windows of 1 minute regardless of previous and subsequent windows. To implement this extraordinary event data indexed (and therefore cut correctly), we end up with windows that are part of the end of one activity and the beginning of another. The classifier in this case react in an unpredictable manner. It would be interesting to add this new class. We could also incorporate this knouledge in priority. Indeed, the location will restrict the possible activities and the time of day will give us an indication of the activity wich can be performed. We worked on the automatic classification without taking any data, but we can improve the results by adding its knowledge. A final interesting finding is that the benefit that can be gained from the use of complex methods are not always significant. There are many situations where a simple method is as efficient (in terms of quality of detection) than complex methods; simple methods in this case are as effective as the complex methods. If one adds to this the stresses induced by complex methods in terms of speed and memory usage, the interest of simple methods becomes even more important.

REFERENCES

[1] T. Emmanuel, S. Intille and K. Larson, “Activity Recognition in the Home Using Simple and Ubiquitous Sensors”, In Proceedings of 2nd International Conference on Pervasive Computing in LNCS, Springer, Vol. 3001, 2004, pp. 158-175.

[2] A. Mahajan, C. Oesch, H. Padmanaban, L. Utterback, S. Chitikeshi and F. Figueroa, “Physical and Virtual Intelligent Sensors for Integrated Health Management Systems”, International Journal on Smart Sensing and Intelligent Systems, Vol. 5, No. 3, September 2012, pp. 559 – 575.

[3] T.Jayakumar, C.Babu Rao, John Philip, C.K.Mukhopadhyay, J.Jayapandian, C.Pandian, “Sensors for Monitoring Components, Systems and Processes”, International Journal on Smart Sensing and Intelligent Systems, Vol. 3, No. 1, March 2010, pp. 61-74.

[4] P.Wide, “Human-Based Sensing – Sensor Systems to Complement Human Perception”, International Journal on Smart Sensing and Intelligent Systems, vol. 1, no.1, 2008, pp. 57 – 69.

[5] S. Boukhenous, “A Low Cost Three-Directional Force Sensor”, International Journal on Smart Sensing and Intelligent Systems, vol. 4, no. 1, 2011, pp. 21-34.

[6] M.F. Rahmat, N.H. Sunar, Sy Najib Sy Salim, Mastura Shafinaz Zainal Abidin, A.A Mohd Fauzi and Z.H. Ismail, “Review on Modeling and Controller Design in Pneumatic Actuator Control System”, International Journal on Smart Sensing and Intelligent Systems, vol. 4, no. 4, 2011, pp. 630-661.

[7] T. K. Dakhlallah, M. A. Zohdy, “Type-2 Fuzzy Kalman Hybrid Application for Dynamic Security Monitoring Systems based on Multiple Sensor Fusion”, International Journal on Smart Sensing and Intelligent Systems, Vol.4, No.4, 2011, pp. 607-629.

[8] X.Pang, P.Bhattacharya, Z.O’Neill, P.Haves, M.Wetter, and T.Bailey; “ Real time building energy simulation using Energy Plus and the building controls virtual test bed”. Proceedings of Building Simulation, 12th Conference of International Building Performance Simulation Association, Sydney, November 2011. Proceedings of Building Simulation 2011, pp. 2890-2896.

[9] M.Wetter, “Co-simulation of Building Energy and Control Systems with the Building Controls Virtual Test Bed”, Journal of Building Performance Simulation, Vol.4, no.3, 2011 pp. 185-203.

[10] T.S. Nouidui, M. Wetter, Z. Li, X. Pang, P. Bhattacharya et P. Haves, “BACnet and analog/digital interfaces of the Building Controls Virtual Test Bed”, Proceedings of 12th International IBPSA Conference Building Simulation, , Sydney,Australia, November 2011, pp. 294-301.

[11] D.L. Ha, H. Joumaa, S. Ploix, M. Jacomino. “An optimal approach for electrical management problem in dwellings”. Energy and Buildings, Vol 45, , February 2012, pp 1-14.

[12] Mei-Ling SHYU, Zongxing Xie abd MIN CHEN and Shu-Ching CHEN, ‘‘Video semantic event/concept detection using a subspace-based multimedia data mining framework’’, IEEE transactions on multimedia ISSN 1520-9210, Vol 10, 2008, pp. 252–259.

[13] J. K. Aggarwal and Q. Cai, ‘‘Human motion analysis: a review’’, Computer Vision and Image Understanding, Vol 73, 1999, pp. 90-102.

[14] D. M. Gavrila, ‘‘The visual analysis of human movement: a survey’’, Computer Vision and Image Understanding, Vol 73, 1999, pp. 82-98.

[15] W. Hu, T. Tan, L. Wang, and S. Maybank, ‘‘A survey on visual surveillance of object motion and behaviors’’, Systems, Man, and Cybernetics, Part C:Applications and Reviews, Vol 34, no. 3, 2004, pp. 334-352.

[16] David A. Forsyth, Okan Arikan, Leslie Ikemoto, James O’brien and amanan, ‘‘Computational studies of human motion: part 1, tracking and motion synthesis’’, Found. Trends. Comput. Graph. Vis, Vol 1, 2005, pp. 77–254.

[17] Ronald Poppe, ‘‘A survey on vision-based human action recognition’’, Image and Vision Computing (IVC), Vol 28, no. 6, 2010, pp.976 – 990.

[18] Poppe, R. ‘‘A survey on vision-based human action recognition ’’, Image and Vision Computing (IVC), Vol 28, no. 6, 2010, pp. 976 – 990.

[19] Turaga, P., R. Chellappa, V. S. Subrahmanian, and O. Udrea , ‘‘ Machine recognition of human activities A survey ’’, IEEE Transactions on Circuits and Systems for Video Technology Vol 18, no. 11, 2008, pp.1473–1488.

[20] Ali, S. and Shah, ‘‘ Human action recognition in videos using kinematic features and multipleinstance learning’’, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), Vol 32, no. 2, 2010, pp. 288–303.

[21] Dollar, P., V. Rabaud, G. Cottrell, and Belongie , ‘‘ Behavior recognition via sparse spatiotemporal features’’, In 2nd International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (PETS), 2005, pp. 65–72.

[22] Willems, G., T. Tuytelaars, and V. Gool, ‘‘An efficient dense and scale-invariant spatiotemporalinterest point detector’’, In European Conference on Computer Vision (ECCV), Vol 102, 2008, pp. 650-663.

[23] Fathi, A. and G. Mori, ‘‘Action recognition by learning mid-level motion features ’’, In International Conference on Computer Vision and Pattern Recognition (CVPR), Vol 2, 2008, pp. 726-733.

[24] Laptev, I., M. Marszałek, C. Schmid, and Rozenfeld , ‘‘ Learning realistic human actions from movies’’, In International Conference on Computer Vision and Pattern Recognition (CVPR), Vol 64, 2008, pp. 107-123.

[25] Kläser, A., M. Marszałek, and C. Schmid,‘‘ A spatio-temporal descriptor based on 3dgradients ’’, In British Machine Vision Conference (BMVC), 2008, pp. 995-1004

[26] Mauthner, T., P. M. Roth, and H. Bischof, ‘‘Instant action recognition ’’, In 16th Scandinavian Conference on Image Analysis (SCIA), 2009, pp1-10.

[27] Huang, W. and J. Wu,‘‘Human action recognition using recursive self organizing map and longest common subsequence matching ’’, In International Workshop on Applications of Computer Vision (WACV), 2009, pp. 1 –6.

[28] Wang, L., H. Zhou, S.-C. Low, and Leckie,‘‘Action recognition via multi-feature fusion and gaussian process classification ’’, In International Workshop on Applications of Computer Vision (WACV), 2009, pp. 1-6.

[29] Yang, W., Y. Wang, and G. Mori,‘‘ Efficient human action detection using a transferable distance function’’, In Asian Conference on Computer Vision (ACCV), Vol 5995, 2009, pp. 417- 426.

[30] Zhang, J. and S.Gong ,‘‘ Action categorization with modified hidden conditional random field’’, Pattern Recognition (PR), Vol 43, no.1, 2010, pp. 197- 203.

[31] Laptev, I. and T Lindeberg,‘‘Velocity adaptation of space-time interest points’’, International Conference on Pattern Recognition (ICPR), 2004, pp. 52–56.

[32] R. Kehl, M. Bray, and L.Van Gool, ‘‘Full body tracking from multiple views using stochastic sampling’’, interantional conference on Computer Vision and Pattern Recognition, Vol 2, 2005 pp. 129-136.

[33] D. Weinland, R. Ronfard, and E. Boyer, ‘‘Free viewpoint action recognition using motion history volumes’’, Computer Vision and Image Understanding, Vol 104, no. 2, 2006, pp. 249-257.

[34] F. Lv and R. Nevatia, ‘‘Single view human action recognition using key pose matching and viterbi path searching’’, international conference on Computer Vision and Pattern Recognition, 2007, pp. 1-8.

[35] C. Rao, A. Yilmaz, and M. Shah, ‘‘View-invariant representation and recognition of actions’’, International Journal of Computer Vision, Vol 50, no. 2, 2002, pp. 203-226.

[36] V.Parameswaran and R. Chellappa, ‘‘View invariance for human action recognition’’, International Journal of Computer Vision, Vol 66, no. 1, 2006, pp. 83-101.

[37] A. Gritai, Y. Sheikh, and M. Shah,‘‘On the use of anthropometry in the invariant analysis of human actions’’, International Conference on Pattern Recognition, Vol 2, 2004, pp. 923-926.

[38] A. Yilmaza and M. Shah, ‘‘Matching actions in presence of camera motion’’, Computer Vision and Image Understanding, Vol 104, no. 2, 2006, pp. 221-231.

[39] C. Rao, A.Gritai, M.Shah, and T. Syeda-Mahmood, ‘‘View-invariant alignment and matching of video sequences’’, International Conference on Computer Vision, Vol 2, 2003, pp. 939-945.

[40] T. Syeda-Mahmood, A. Vasilescu, and S. Sethi, ‘‘Recognizing action events from multiple viewpoints’’, Detection and Recognition of Events in Video Workshop, 2001, pp. 64-72.

[41] Qiang He and C. Debrunner, ‘‘Individual recognition from periodic activity using hidden markov models’’, Human Motion Workshop, 2000, pp. 47-52.

[42] A.A. Efros, A.C. Berg, G. Mori, and J. Malik, ‘‘Recognizing action at a distance’’, International Conference on Computer Vision, Vol 2, 2003, pp. 726-733.

[43] R. Cutler and M. Turk, ‘‘View-based interpretation of real-time optical _ow for gesture recognition’’, International Conference on Automatic Face and Gesture Recognition, 1998, pp. 416-421.

[44] J.W. Davis and A.F. Bobick, ‘‘The representation and recognition of action using temporal templates’’, International conference on Computer Vision and Pattern Recognition, 1997, pp. 928-934.

[45] Ojala, T., Pietikainen, M., and Harwood, D, ‘‘A comparative study of texture measures with classification based on feature distributions’’, *In Pattern Recognition*, Vol 29, 1996, pp. 51–59

[46] Ojala, T., Pietikainen, M., and Maenpaa, T, ‘‘Multiresolution gray-scale and rotation invariant texture classification with local binary patterns’’, Vol 24, no. 7, 2002, pp. 971–987.

[47] Ahonen, T., Hadid, A., and Pietikäinen, M.,‘‘Face description with local binary patterns : application to face recognition’’, IEEE Trans Pattern Anal Mach Intel*l*, Vol 28, no. 12, 2006, pp. 2037–2041.

[48] Tan, X. and Triggs, B, ‘‘Enhanced local texture feature sets for face recognition under difficult lighting conditions’’, In IEEE Conf. on AMFG, 2007, pp. 168 –182.

[49] Kolesnik, M. and Fexa, A, ‘‘Multi-dimensional color histograms for segmentation of wounds in images’’, Lecture Notes in Computer Science, Vol 3656, 2005, pp. 1014–1022.

[50] Swain, M. and Ballard, D, ‘‘Color indexin’’, International Journal of Computer Vision (IJCV*)*, Vol 7, no. 1, 1991, pp. 11–32.

[51] . Huang, J., Kumar, S., Mitra, M., Zhu, W.-J., and Zabih, R, ‘‘Image indexing using color correlograms’’, In *Proc* IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, pp. 762–768.

[52] Messing, D., van Beek P., and Errico, J, ‘‘The mpeg-7 colour structure descriptor : image description using colour and local spatial information’’, In Proc. International Conference on Image Processing, Vol 1, 2001, pp. 670–673.

[53] Boujemaa, N. and Vertan, C, ‘‘Upgrading color distributions for image retrieval : can we do better ? In *Proc’’,* of International Conference on Visual Information System (VIS00), *2000,* pp. 178–188.

[54] Vertan, C. and Boujemaa, N, ‘‘Embedding fuzzy logic in content based image retrieval’’, In Proc. NAFIPS Fuzzy Information Processing Society 19th International Conference of the North American, 2000, pp. 85–89.

[55] Zhao, R. and Grosky, W, ‘‘From features to semantics : some preliminary results’’, In *Proc.* IEEE International Conference on Multimedia and Expo ICME 200*0*, Vol 2, 2000, pp. 679–682

[56] Smith, J. R. and Chang, S. F, ‘‘Tools and techniques for color image retrieval’’, In IST/SPIE Proceedings, 1996, pp. 426–437.

[57] Carson, C., Belongie, S., Greenspan, H., and Malik, J, ‘‘Blobworld : image segmentation using expectation-maximization and its application to image querying’’, IEEE Trans on Pattern Anal and Machine Intill. (PAMI), Vol 24, no. 8, 2002, pp.1026–1038

[58] C. Cortes and V. Vapnik ,“Support-vector network,” Mach. Lear*n.,* Vol 20,1995, pp. 273–297.

[59] V. Caselles, J. L. Lisani, J. M. Morel, and G. Sapiro, "Shape preserving local histogram modification", IEEE Trans. on Image Processing, Vol 8, 1999, pp. 220–229.

[60] D. Sen and P. Sankar," Automatic exact histogram specification for contrast enhancement and visual system based quantitative evaluation ", IEEE Trans. on Image Processing, Vol 20, 2011, pp. 1211–1220.

[61] D. Coltuc, P. Bolon, and J.-M. Chassery, "Exact histogram specification ", IEEE Trans. on Image Processing, Vol 15, 2006, pp. 1143–1152.

[62] E. L. Hall, "Almost uniform distributions for computer image enhancement ", IEEE Transactions on Computers, Vol 23, 1974, pp. 207–208.

[63] Y. Wan and D. Shi, "Joint exact histogram specification and image enhancement through the wavelet transform", IEEE Trans. on Image Processing, Vol 16, 2007, pp. 2245–2250.

[64] M. Nikolova, Y. Wen, and R. Chan," Exact histogram specification for digital images using a variational approach ", J. of Mathematical Imaging and Vision, 2012, pp. 1-17

[65] B. Boser, I. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," in Proc. 5th Annu. Workshop on Computational Learning Theory, 1992, pp.144-152.

[66] T. Joachims, ‘‘Making large-scale support vector machine learning practical, In A. Smola B. Scholkopf, C. Burges, editor, Advances in Kernel Methods : Support Vector Machines", editors IEEE transactions on information theory, Vol 44, no.2, MIT Press, Cambridge, MA, 1998, pp. 525-536;

[67] E.Osuna, R. Freund, and F. Girosi, ‘‘Training Support Vector Machines: an Application to Face Detection’’, roceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97), New York, 1997, pp.130-136.

[68] T.Joachims, ‘‘Making large-scale support vector machine learning practical, In A. Smola B. Scholkopf, C. Burges, editor, Advances in Kernel Methods : Support Vector Machines", Cambridge, MIT Press, MA, USA, 1999, pp. 169-184