Data Fusion for User Presence Identification

Vincenzo Di Lecce*, Alberto Amato
D.I.A.S.S.
Politecnico di Bari
Taranto, Italy
{v.dilecce, a.amato}@aeflab.net

Vincenzo Piuri
D.T.I.
Università degli Studi di Milano
Crema, Italy
vincenzo.piuri@unimi.it

***Abstract*—Aim of this work is to present a new approach to the problem of user presence monitoring in working environments. Particularly, this work is focused on the evaluation of the presence or absence of a user in front of a terminal. This question is of paramount importance in applications requiring the user's presence e.g. video surveillance systems, control centrals, etc. The authors propose a technique of data fusion using signals from various low cost sensors.**

Keywords-component; data fusion, signal analysis, fuzzy inference system

I. INTRODUCTION

Nowadays, ICT applications are becoming ever more intrusive in our everyday life. Distributed measurement systems are used in many applications such as: air quality monitoring, surveillance of critical infrastructure, measurement of traffic flows, climate sensing and control in office buildings and homes, etc. Sensor networks are generally deployed to measure some characteristics about a particular environment of interest. The data, they gather, can then be analyzed to extract important information related to the occurrence of events in that environment. The evaluation of these events can be automatic or semi automatic. In the latter case, the system is used as an aid in the decision making process. A good example of such a system can be a video surveillance system showing to a human operator scenes from an area where the presence of some moving objects has been revealed. It will fall to the human operator to evaluate if there is a dangerous situation or not. In this kind of system, the presence of the human operator is of paramount importance since this one is the final decision maker of the whole system. If for any reason he/she is far from his/her workstation, the whole system becomes unreliable. Another example of application in which the user presence evaluation is a parameter of primary importance is the e-learning. Indeed, if a user starts a lesson and then, for any reason, he goes away, the e-learning system does not suspend the lesson thus wasting resources.

In literature there are also other examples of applications that can derive benefits from a system monitoring the user presence. In [1] the authors propose an interesting system applying techniques of user presence and activity monitoring to the last trend of Internet technology, namely the social networks like MySpace® or FaceBook®. In [2] the authors propose a system revealing the user presence around a node of a sensors network using a humidity sensor.

Some authors propose systems based on computer vision to assess the user presence. [3] presents the prototype of a FPGA implementing basic image processing techniques (filtering, color-based segmentation, thresholding) producing a signal when a human-skin colour segment is detected in the image frame. In [4] the authors discuss a computer vision based approach to assess the user presence near a device. This approach uses an IR sensor to reveal the presence of a badge that a user should dress. As the same authors state, probably this is the biggest limit of their approach.

In this work, the authors propose an approach based on multi sensor data fusion to assess the presence of a user. Multi-sensor data fusion is one of the effective approaches for solving different sets of problems having common characteristics. It uses the data from multiple sensors to perform inferences that may not be possible from a single sensor alone. Data from different sensors are combined using signal processing, pattern and image recognition, artificial intelligence, and information theory.

The proposed system is based on a set of low-cost sensors (based on tin oxide semiconductor sensing elements) measuring some atmospheric parameters and a technique based on computer vision applied to the images recorded by a simple webcam. The atmospheric parameters analyzed are CO, CO_2, NO_x, ethanol, humidity, temperature, etc. In an indoor environment, concentrations of these parameters have a strong variability related to human presence and activity. The computer vision based technique analyzes the edge of the objects in the recorded scene (user's edges when this one is present) using an algorithm tested in various works such as [5, 6, 7]. A Fuzzy Inference System (FIS) is used to perform data fusion at features level. The combined use of these sensors, characterized by different dynamics (a low dynamic for the gas sensors and a high dynamic for the camera), allows the system to identify the human presence both in presence of fake images and in situation with high variability of air conditions (unstable air flow in an indoor environment). Local operation is directed to support privacy protection for the users' biometries.

The remaining part of this abstract is so organized: section II describes the proposed system, in section III some preliminary tests are shown and the conclusions and final remarks are reported in section IV.

II. PROPOSED SYSTEM

Figure 1 shows a schematic overview of the proposed system. In the left side there are the atmospheric sensors (S_1,

*contacting author

S_2, …, S_n) and the webcam. The output of each sensor is sent to a module called "feature extraction" that performs all the required processes of signal analysis to obtain a feature feasible for the decision system. After this processing, the extracted features are sent to the decision system that, by applying a fuzzy inference method, assesses the presence or absence of the user. In the following sub-paragraphs the various modules of the proposed system are described.

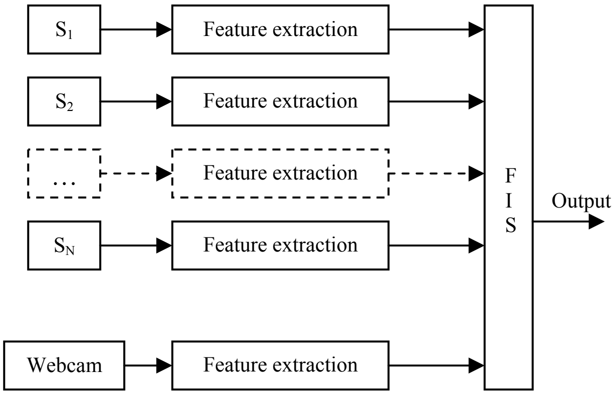


Figure 1 - a schematic overview of the proposed system

A. Environmental sensors

Over the last years, tin-oxide based resistive sensors have been extensively used to analyze gases. These sensors are inexpensive and highly sensitive to a broad spectrum of gases, including atmospheric pollutants such as CO, CO_2, NO_x and H_2S [8]. However, they are well known for such disadvantages as lack of selectivity (see Figure 2) and response drift [9], which is why they are used in low cost alarm-level gas monitors for domestic and industrial applications. The sensing element is composed of a metal oxide semiconductor layer formed on a sensible substrate of a sensing chip together with an integrated heater. In the presence of a detectable gas, the sensor's conductivity changes depending on the gas concentration in the air. A simple electrical circuit can convert the change in conductivity to an output signal which corresponds to the gas concentration. Figure 2 shows an example of sensitivity characteristic relative to one of the used sensors. In this work three different sensors are used and they are produced by Figaro Inc. Each sensor is sensible to different gas-mixture and has a different response to human presence.

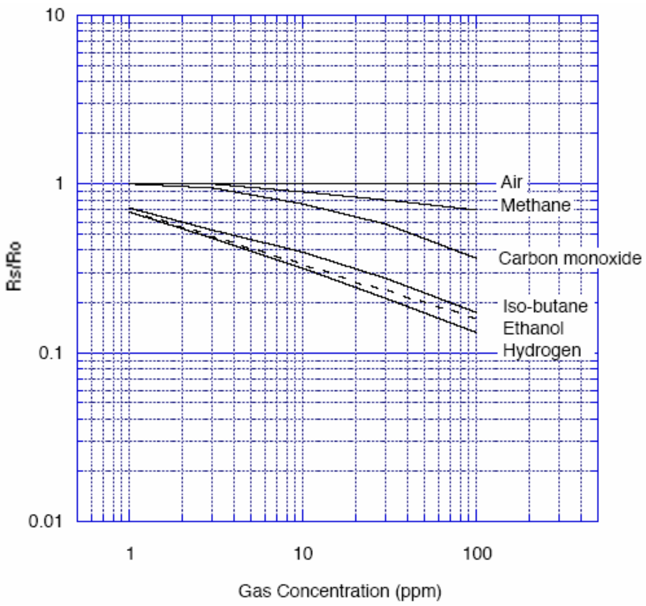


Figure 2 - An example of sensitivity characteristic relative to one of the used environmental sensors

B. Image analysis

The stream of a common USB webcam is used to generate a series of frames that are analyzed by the feature extraction module. The adopted algorithms are proper of the Content-based Image Retrieval research area. Content-based image retrieval (CBIR) emerged as an important area in computer vision, multimedia computing, and data base management as an effective strategy to deal with these needs.

Some content-based image retrieval systems were recently proposed, e.g., QBIC [10], Photobook [11], VisualSeek [12], MARS [13], El Niño [14], and CIRES [15]. These systems characterize an image as a set of syntactic components of the image itself. Visual features are extracted automatically from the image (syntactic or low-level data representation) [16] to characterize the syntactic components; the desired images are retrieved from the database by searching the ones having similar these visual features. Many researches are still trying to identify which are the most significant visual features and how to extract them in the most efficient and effective way. Color, texture, shape, and 2D spatial relationships [17, 18, 19, 20] are typical (although sometimes too rough) image features that are used in CBIR systems to describe the syntactic components and to retrieve automatically (i.e., without user intervention) the desired images.

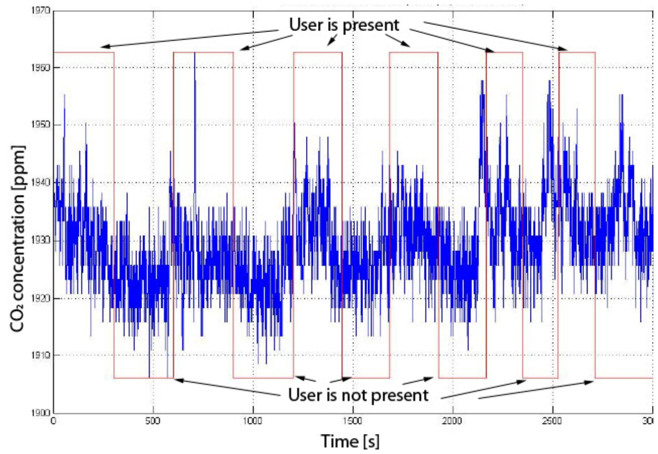
In this work, the authors use a signature presented in [5] called "string signature" that is focused on the analysis of the edges relative to the main objects in an image. This signature derives from the following hypotheses :

- all the image objects are characterized by edges (the best case is a shape or a boundary);
- the fundamental objects (related to the semantic content of the image) are in the foreground;
- edges can be synthesized using the 4+1 main direction (-,|,/,\,null) and represented as characters string.

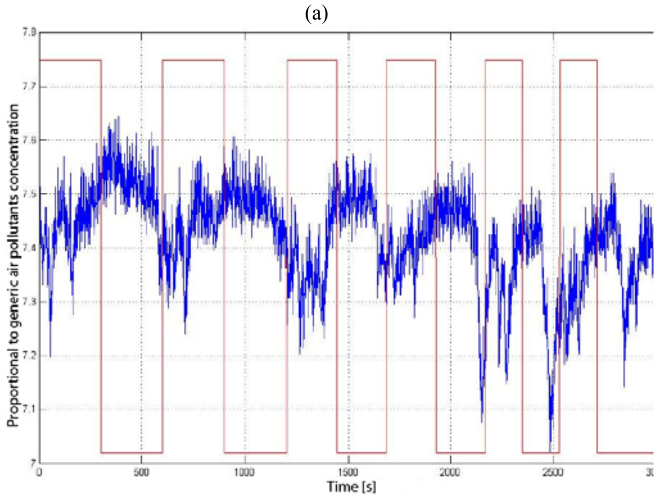
The algorithm to obtain this signature from a given image could be summarized as follows:

1) Noise reduction
2) Find the edges using the Canny algorithm
3) The obtained edge image is divided in windows using a 9x9 regular grid

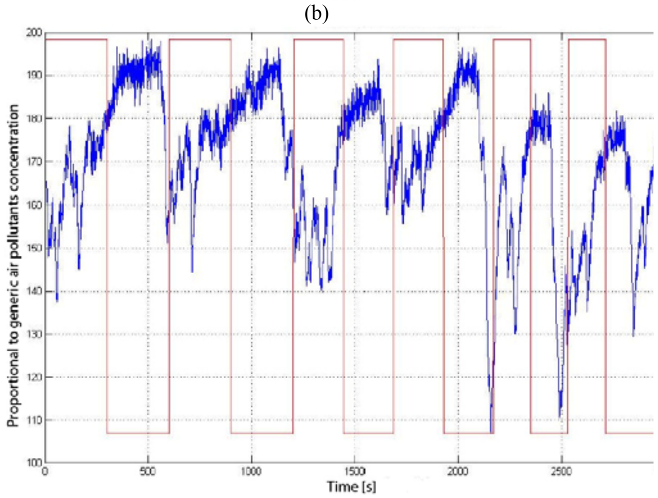
4) For each window the prevalent direction of the lines is computed according to Tamura [21] and the string symbol is associated to it.



(a)



(b)



(c)

Figure 3 - An example of sensor response to human presence. This diagram is referred to the sensor Figaro TGS-4161 (a) that is sensible to CO_2, TGS-2442 (b) and TGS-2600 (c) that are sensible to various air pollutants.

C. *Data fusion*

With the development of the technology, the approach of obtaining information is continuously increasing. The need of using and disposing information is ever more high, and the multi information fusion technology makes the information exactly and more timely arranged. Multi-sensor integration can be defined as the synergistic use of the information provided by multiple sensory devices to assist the accomplishment of a task by a system [22].

Depending on sensors and the goals of the multi sensors data fusion (MSDF), the sensor fusion can take place at different levels. In [23] the authors divide the MSDF into four levels: signal, pixel, feature and symbol. Signal level fusion decreases the covariance of the sensory data. Pixel-level fusion is intended to increase the information content associated to each pixel of an image. Feature-level fusion combines features derived from signals into meaningful representations or more reliable features. Symbol-level fusion allows the information to be fused at the highest level of abstraction and it is usually used in decision-based systems. In practice, using similar sensors, it is sufficient to apply only one level of fusion. In more complex cases, when different groups of sensors are involved, each group is fused at a lower level, and the sensory data coming from different groups is fused at a higher level into one representational format.

In this work the feature level data fusion approach was used. Indeed, in this application the signal and pixel level are not suitable for a series of reasons such as:

- The raw signal has some instantaneous variations due to various events such as noise, sensor precision, and so on
- Analyzing only the instant values sampled by each sensor, previous states of the system are lost. This fact is of paramount relevance because the system should be able to operate in various environmental conditions. In other words, if in a room there is a certain concentration of CO_2 (or any other parameter), the system should be able to distinguish the percentage of CO_2 produced by the user and the ones present in the air before his arrival.

III. EXPERIMENTS

In this work various experiments were carried out to evaluate the proposed approach. A first set of experiments were carried out to test separately the performance of the single sensors in assessing the user presence. In these experiments a workstation was equipped with a USB webcam attached to its LCD monitor. The environmental sensors are mounted in a short pipe (15cm) with a square section of $100cm^2$. A computer fan is mounted at one side of the tube to aid a correct air flow near the sensors. This system is placed on the user desk under the monitor. The distance between user and sensors is about 60 cm.

Figure 3 shows the results obtained by using some environmental sensors: Figaro TGS-4161 (a) that is sensible to CO_2, TGS-2442 (b) and TGS-2600 (c) that are sensible to various air pollutants (CO, H_2, methane, ethanol, etc.). The red

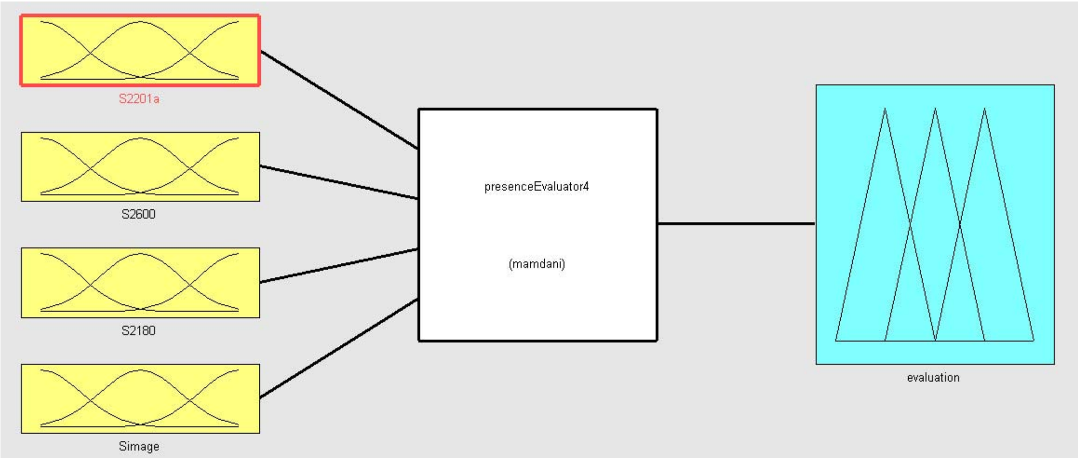


Figure 4 - block diagram of the proposed FIS

line (square wave) represents the actual user's state; the high values stand for "user is present" and the low values stand for "user is not present". The blue line shows the values measured by the various sensors. These figures show some interesting aspects related to the use of these sensors in human presence assessment:

- High variability of the signal
- The dynamic response of the system at human presence or absence has a variable delay due to the air flow propagation.
- The phases, in which the user is present, are characterized by a stronger variability than those where user is absent. This fact is due to user behavior. When user is not present, the air conditions are quite stable (there is a progressive reduction of CO_2 due the natural effect of air cleaning) but when user is present, the emission varies in response to his/her state (e.g. he/she can speak, yawn, modify his/her position relative to the sensors and so on).

This first set of experiments was used to demonstrate the suitability of the selected sensors for a data-fusion system aiming at user presence evaluation.

After this stage, the data-fusion system was implemented using a FIS.

The proposed FIS was developed using the visual tools present in the Matlab® environment. Figure 4 shows the block diagram of the proposed FIS as it appears in the Matlab® FIS editor environment. As this figure shows, a Mamdani-type FIS was used [24]. The features extracted from each sensor and the signature extracted from the image sampled by the webcam are used as input for the FIS. The sigmoidal membership functions were used for all the input. The FIS uses seven rules to evaluate the user presence. These rules try to equilibrate the contribute that each sensor gives to the final decision. For example, if the used visual feature assesses the presence of a user and all the other sensors do not, then the user is considered absent. But if the visual feature and one of the gas sensor reveal the presence of a user while the other gas sensors do not, than the user is considered present. These kinds of rules allow the system to consider the differences in the dynamic of the various sensors.

The proposed system was tested in various conditions using different users in order to stress the system with a high variance both in terms of images and in terms of produced gas concentrations. For example, Figure 5 proposes an example of the results obtained using the proposed system. The red line indicates the user state (present when 1 and absent when 0) while the blue line shows the system evaluations. In this example the user was present for 5 minutes and then he/she was absent for the same time. In this example the system had a good classification performance.

In the various experiments carried out, the worst results gave the 72% of successful classification rate while the better one gave the 97%.

This strong variability in the performance is due to the high variability introduced into the experimental set-up. In particular, the environmental conditions were intentionally changed in various experiments. For example, a parameter that has a strong influence on the system performance is the air flow condition. Indeed, it has a strong influence on the fluid-dynamic processes at the base of the diffusion of the air

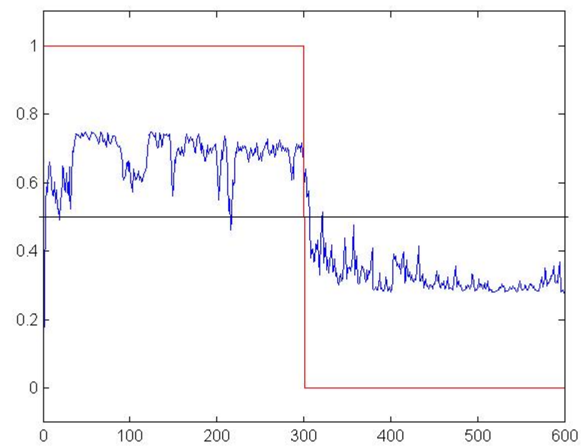


Figure 5 - an example of the results obtained using the proposed system

breathed out by user. For this reason, augmenting the air flow in the experiment set-up, there was a strong reduction of the gas sensors' contribute to the user presence evaluation process.

Another parameter that was in depth analyzed is the system reaction speed to user presence variations. This analysis was conducted carrying out a series of experiments where the user was present and then absent for a given interval of time. This interval of user presence/absence was progressively reduced.

IV. CONCLUSIONS

In this work, a data fusion based method to solve the problem of user's presence or absence in front of a terminal was proposed. The first experimental results are showing good performance in terms of effectiveness for user presence/absence identification.

The main reasons for the selected approach have practical roots. In our opinion the set of sensors (environmental and visual) is sufficient to cover the main common situations in an indoor environment. If one tries to remove one portion of this system, a series of possible shortcomings becomes quite evident. Supposing to use only the environmental sensor set, as the shown results highlight, these sensors are quite able to identify, with sufficient precision, the presence of a human operator. But:

- The human presence identification is not instantaneous, but it has a delay due to the phenomenon of air flow propagation.
- If inside the indoor environment there is an unstable air flow (e.g. an opened window) the results obtained by this set of sensors are quite unpredictable.

Vice versa, supposing to use only the computer vision portion of the proposed system, it is still possible to obtain good results in user presence identification, but the system could be mistaken in various modes such as posing a simple picture of the users in front of the webcam.

The combined use of the two main modules allow for solving many of the posed problems.

Further experiments will be carried out in order to explore the possible biometric extensions of the proposed method. In other words the authors will try to define a biometric signature identifying a user univocally.

ACKNOWLEDGMENT

This work was supported in part by the EU, within the 7FP project, under grant agreement 216483 "PrimeLife".

REFERENCE

[1] Emiliano Miluzzo, Nicholas D. Lane, Shane B. Eisenman, and Andrew T. Campbell, CenceMe – Injecting Sensing Presence into Social Networking Applications, in G. Kortuem et al. (Eds.): EuroSSC 2007, LNCS 4793, pp. 1–28, 2007 © Springer-Verlag Berlin Heidelberg 2007

[2] Jun Han, Abhishek Jain, Mark Luk, Adrian Perrig. "Don't Sweat Your Privacy: Using Humidity to Detect Human Presence". Proceedings of 5th International Workshop on Privacy in UbiComp (UbiPriv'07), September, 2007.

[3] Moshnyaga, Vasily G. Koji Hasimoto, Tadashi Suetsugu, FPGA design for user's presence detection, in 15th IEEE International Conference on Electronics, Circuits and Systems, 2008. ICECS 2008. pp. 1316-1319.

[4] B. Raducanu, S. Subramanian, P. Markopoulos. "Human Presence Detection by Smart Devices". Proc. of Engineering of Intelligent Systems (EIS2004), pp. N/A, Funchal, Island of Madeira, Portugal, 29 Feb-2 March 2004

[5] A. Amato, T. Delvecchio, V. Di Lecce, Image Retrieval Using Simbol String, The IASTED International Conference on Databases and Applications. Lugano, pp.220-225, February 17-19, 2004.

[6] A. Amato, V. Di Lecce, Edge detection techniques in image retrieval: The semantic meaning of edge, EC-VIP-MC 2003, 4th EURASIP Conference, Zagreb - Croatia, pp. 143-148, July 2-4, 2003.

[7] A. Amato, V. Di Lecce, A knowledge based approach for a fast image retrieval system, Image and Vision Computing (2008), Volume 26 , Issue 11 (November 2008), pp. 1466-1480, ISSN:0262-8856, doi:10.1016/j.imavis.2008.01.005

[8] P. Tasai, T. Chen, and M. Tzeng, "Tin oxide carbon monoxide sensor fabricated by thick-film methods," Sens. Actuators B, Vol. 25, pp. 537–539, 1995.

[9] P. Moseley and B. Tofield, Solid-State Gas Sensors. Bristol, U.K.: Adam Hilger, 1987.

[10] M. Flickner et al., "Query by Image and Video Content: The QBIC System", in IEEE Computer, Vol. 28, No. 9, September 1995.

[11] A. Pentland, R. W. Picard and S. Sclaroff, "Photobook: Content-based manipulation of image databases", in Proc. SPIE Storage Retrieval Image Video Databases II, pp. 34-37, Feb. 1994.

[12] J. R. Smith and S. F. Chang, "Visualseek: A fully automated content based image query system", in Proc. ACM Multimedia, Boston, MA, Nov. 1996, pp. 87-98.

[13] S. Mehrotra, Y. Rui, M. Ortega and T. S. Huang, "Supporting content-based queries over images in MARS", in Proc. IEEE Int. Conf. Multimedia Computing Systems, Canada, June 3-6, 1997, pp. 632-633.

[14] S. Santini and R. Jain, "The 'El Niño' Image Database System", in IEEE International Conference on Multimedia Computing and Systems, 1999. Florence, Italy.

[15] Q. Iqbal and K. Aggarwal, "Feature Integration, Multi-image Queries and Relevance Feedback in Image Retrieval", Invited Paper, To Appear, 6th International Conference on Visual Information Systems (VISUAL), Miami, Florida, Sep. 24-26. 2003.

[16] J. P Eakins and M. E Graham, "Content-based Image Retrieval", A report to the JISC Technology Applications Programme, Institute for Image Data Research, University of Northumbria at Newcastle, January 1999.

[17] M. Swain and D. Ballard, "Color Indexing", in International Journal of Computer Vision, Vol. 7, N°1, 1991.

[18] K. Hirata and T. Kato, "Query by Visual Example – Content-Based Image Retrieval", in Proc. EDBT'92, Springer LNCS 1992.

[19] A. Del Bimbo and P. Pala, "Visual Image Retrieval by Elastic Matching of User Sketches", in IEEE Transactions on Pattern Analysis and Machine Intelligence 19(2), 1997.

[20] S.K. Chang and E. Jungert, "Pictorial Data Management based upon the Theory of Symbolic Projections", in Journal of Visual Language and Computing 2, 1991.

[21] V. Di Lecce and A. Guerriero, A Comparative Evaluation of retrieval methods for Duplicate search in Image database, Journal if Visual Languages and Computing, 2001, 12, 150-120

[22] R. C. Luo, and M. G. Kay, "A tutorial on multisensor integration and fusion," in 16" Annual Conference of IEEE Industrial Electronics Society (IECON), vol. 1, Nov 1990, pp. 707-722.

[23] R. C. Luo, Y. Cbih-Chen, and K. L. Su, "Multisensor fusion and integration: approaches, applications, and future research directions," IEEE Sensors Journal,vol. 2, no. 2, pp. 104119, April 2002.

[24] Mamdani, E.H. and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," International Journal of Man-Machine Studies, Vol. 7, No. 1, pp. 1-13, 1975.