INDUSTRIAL SUPERVISION SYSTEM BASED ON VISUAL DATA MINING AND MOTION TRAJECTORY ANALYSIS

J.J. Fuertes[*], P. Reguera[*], M. Domínguez[*], I. Díaz[+] and A. A. Cuadrado[+]

[*] *Instituto de Automática y Fabricación. Área: Automática y Control. Escuela de Ingenierías Industrial e Informática, Campus Universitario de Vegazana, Universidad de León. León, (SPAIN), 24071. Tfno: +3498 729 1694. E-mail: {diejfu, diepra, diemdg}@unileon.es; http://aaciaf.unileon.es*

[+]*Dept. Ingeniería Eléctrica, Universidad de Oviedo, Campus de Viesques, Edificio Departamental Nº2, Gijón, Asturias,(SPAIN), 33204, Tfno: +3498 518 2532 E-mail:{idiaz, cuadrado} @isa.uniovi.es*

Abstract: The current trends in video surveillance systems aim to incorporate mechanisms that understand and remember the activity in a scene to make decisions. These decision modules have as input the object trajectories in a scene resulting from the treatment of the images captured by a video camera. In this paper we propose a novel industrial supervision system for complex multivariable processes that incorporates this decision module. The scenario in this case does not come from the treatment of a video image sequences but from the projection of process variables on a 2D plane using dimension reduction techniques.

Keywords: Supervision, scene analysis, Man-Machine Interfaces, neural networks, Self-Organizing system, monitoring, residues, visual surveillance, complex systems.

1. INTRODUCTION

The growing computation power and storage capacity of current computers makes a vast amount of data coming from the acquisition of industrial process variables to be available. Every day, more and more companies incorporate into their classic supervision methods -based on MMI (Man-Machine interfaces*)* with historical and management alarms-, other techniques that take advantage of the huge amount of hidden information existing in this data warehouse. One of the novel approaches that is being incorporated to these supervisory systems is Data Mining. Data Mining is one step in the Knowledge Discovery in Databases process that consists in applying data analysis and discovery algorithms that produce a particular enumeration or patterns over the data (Fayyad, *et al.*, 1996).

Visual Data Mining systems (Keim, 2001, 2002) in industrial supervision make possible to extract the hidden knowledge in the data and, at the same time, to incorporate the intuitive power of the human mind to make decisions (the visual skills of classic methods are keeped). One instance of this approach, used in the supervision of industrial systems, are dimension reduction techniques, that project multidimensional data of the process variables in a low dimensional space, typically a 2D plane, without significant loss of information. In this 2D visual space, the system state appears as a point that changes in a similar way as a video stream of an object moving over a place.

On the other hand, in the last years, visual surveillance is currently one of the most active research topics in computer vision (Wang, *et al.*, 2003; Buxton, 2003). The actual efforts make possible that the current surveillance systems, that typically work offline (after something has happened), turn into real-time video security systems. The surveillance data analysis in real-time can alert of an unusual situation (e.g. a driver parking in a noparking area) or even a criminal one (pedestrian examining cars in a parking).

Research has been centered in five action fields (Collins, *et al*., 2000): detection, tracking and objects classification from images, human motion analysis, and mainly, behavior understanding in scene. The advances have been applied in varied fields: from security applications as, for example parking control, traffic flow measurement, accidents warning in freeways, pedestrian congestion in public spaces, etc; humanitarian applications, as for example analysis of the refugees flow in conflicting areas; and also commercial applications as, for example, distribution of customers in commercial centers. However, in the industrial environment their application is still rather unexplored, although progressively video surveillance systems make it possible to predict danger situations as detection of escapes or temperature critical increases.

Video surveillance systems have a 2D scenario (typically obtained from image sequences of a three-dimensional scene), that is the analysis target, and a decision module, that understands the activity in the scenario and acts accordingly. In Figure 1, the characteristic block diagram of the surveillance visual task is shown. As seen, a 2D space is obtained after executing the first four stages from the analysis of input images. Finally, this 2D space is the input of the decision module.

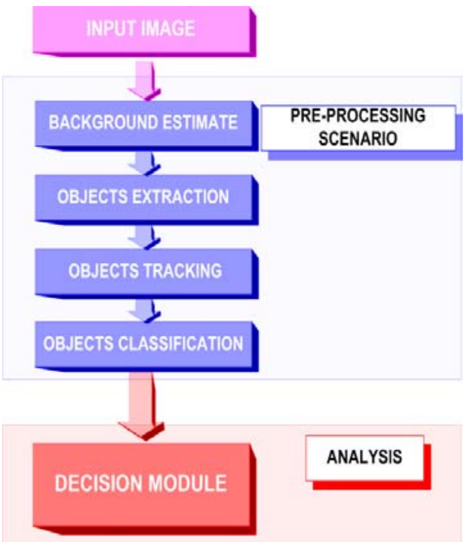

Fig. 1. Visual Surveillance Stages.

If the 2D visualization space in an industrial supervision system based on dimension reduction techniques is treated as the scenario of a visual surveillance system, it can incorporate the advances in this field to supervise complex industrial systems. We will develop this idea along the paper.

2. MOTION TRAJECTORY ANALYSIS

Understanding the activity in a scene requires the recognition of the events that take place in it. Learning the patterns of the trajectory described by an object, can be used for anomalies detection and behavior prediction.

The decision module in a visual surveillance system learns which movements are right and which ones go away of the normality, producing a classification of trajectories as normal or abnormal. Abnormal trajectories can include accesses to areas that are not passed habitually, movements of unusual speed or fluctuating movements. Different approaches to implement this module have been proposed: Finite State Machines (FSM) (Wilson, *et al*., 1997); Hidden Markov Models (HMM) (Oliver, *et al*., 2000); or Non Deterministic Finite Automaton (NFA) (Wada and Matsuyama, 2000). However, the most frequent approach is based on artificial neural networks. These nets are trained to recognize trajectories using video images. During the surveillance, if the net identifies strange behaviors (that is, objects trajectories that are not present during the training phase) it is reported. Trajectories are currently represented by sequences of discrete states (Johnson and Hogg, 1996). These are coded as flow vectors, (x, y, dx, dy) that include position (x, y) and velocity (dx, dy), or extended versions, $(x, y, dx, dy, a_1, ..., a_n)$ that add extra information (parameters a_1, a_2,...). These vectors are inputs to the decision module.

Hu, *et al*. (2004a) propose to use a hierarchical Self-Organized Map (Hierarchical SOM) where they define two neighborhood types, on one hand the neuron neighborhood (instantaneous movement classification) and on the other hand, the internal net neighborhood (trajectory classification). Owens and Hunter (2000) incorporate second order information to the flow vectors that code the trajectory, besides position and instantaneous speed. In particular, they add the acceleration and a function $s(x)$ that incorporates abruptness of change in speed (abrupt changes in, for example, the pedestrian speed in a parking of cars can be cause of robbery alarm). They simplify the learning method with a single layer SOM net. Hu, *et al*. (2004b) get an even bigger simplicity using a fuzzy SOM as classification method: a single layer net whose input is the whole trajectory.

Johnson and Hogg (1996) proposed a statistical model of object trajectories based on a structure with two layer competitive learning networks that are connected by another layer of leaky neurons. This is a memory mechanism to record activation history. The first layer models the probability density function (*pdf*) of possible instantaneous movements by means of a neural net of competitive unsupervised learning (Vector Quantization) while the second layer models the trajectories inside the scene in a similar way. The input space of the first layer is a hyper cubic space in which the flow vectors (x, y, dx, dy) represent position and instantaneous speed. It has four input nodes, one for each component of the dimensional feature, and k output nodes, one for each prototype. After the learning stage, each prototype represents a certain number of training vectors, and the prototype point density approaches the *pdf* of instantaneous movements. Output of this layer is the input to another one that has as many leaky neurons as prototypes, and it is a mechanism to register the activation history, allowing this way, the trajectory classification in the last layer.

3. INDUSTRIAL SURPERVISION SYSTEM

In the industrial supervision system based on Visual Data Mining proposed by Diaz *et al.* (2003) a SOM net is used as dimension reduction technique and projection maps as visualization tool. The current state of an industrial process is projected on a two-dimensional color map representing the inter-neuron distances (distance map) that reveals the cluster structure of the process states. Thus, on one hand, this map represents the static states visualization in the process, and on the other hand, it shows up the process dynamic behavior, represented by the transitions in its frontiers. See Figure 2.

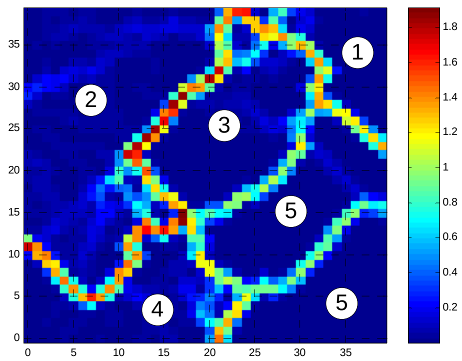

Fig. 2. State Visualization map. SOM Distance map.

Making an analogy with a video surveillance system, the industrial process state moves within a scenario (immobile), similar to one that a video camera makes of a three-dimensional scene.

The data flow, coming from the process, is projected on the SOM neurons in the input space, and therefore in the visualization map resulting in 2D trajectories. According to the classic projection algorithm, an input vector $\mathbf{x}$ is projected in a neuron $\mathbf{g}_c$:

$$S(\mathbf{x}) = \mathbf{g}_c,\ con\ c = \arg\min_i \{d(\mathbf{x},\mathbf{m}_i)\} \qquad (1)$$

where $d(.,.)$ is the distance function and $\mathbf{m}_i$ the codebook vector in input space. It produces a discontinuous projected trajectory. However, if the projection algorithm proposed in Diaz *et al.* (2001) is used (the generalization of SOM to the continuous case using a kernel regression neural network GRNN), the described trajectories are softer and continuous. Therefore, there is a projection in a 2D space for the system dynamic evolution. See Figure 3.

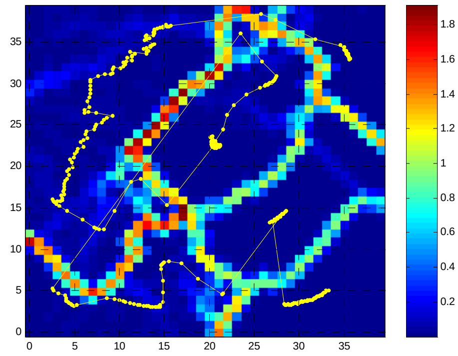

Fig. 3. Trajectory described by an industrial process in the states visualization map.

In this paper, we propose to use the trajectories produced by SOM projection as input to the decision module in a visual surveillance system, so a novel industrial supervision system is made. In Figure 4, the similarity between a visual surveillance system and the industrial supervision system proposed can be observed. The decision module is similar in both cases. However, in the case of a classic surveillance visual system, the preprocessing to obtain 2D scenario needs background estimation, extraction, tracking and objects classification using video camera images; but in the proposed industrial supervision system, only dynamic modeling and later 2D projection (using data-mining techniques to process data) are needed.

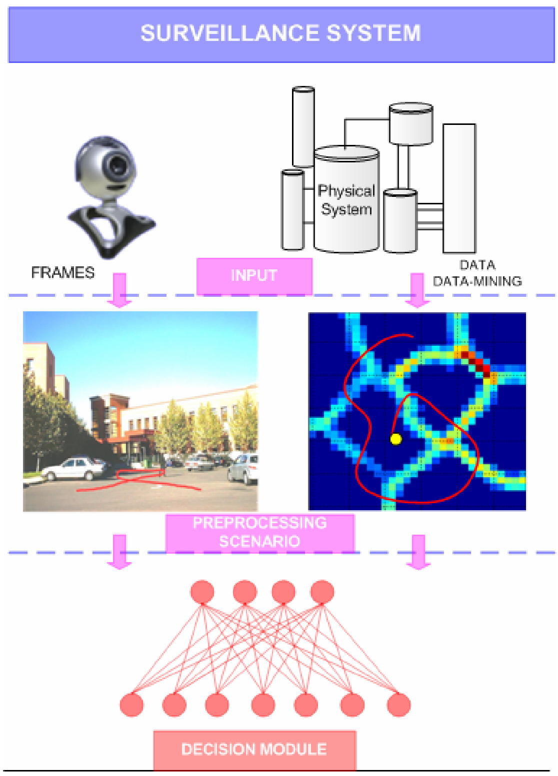

Fig. 4. Similarity between the visual surveillance systems and proposed industrial supervision system based on data mining.

In summary, the visualization map constitutes the 2D scenario in which the object (the state of the industrial system) moves describing two-dimensional trajectories. These are the inputs to the decision module in industrial process visual supervision system.

4. DECISION MODULE

The approach presented in this paper is similar to Johnson and Hogg (1996) and Owens and Hunter (2000), although with differences motivated by the different nature of the systems to keep watch: in our case multivariable industrial processes whose scenario is generated by data projection. The net structure is represented in Figure 5.

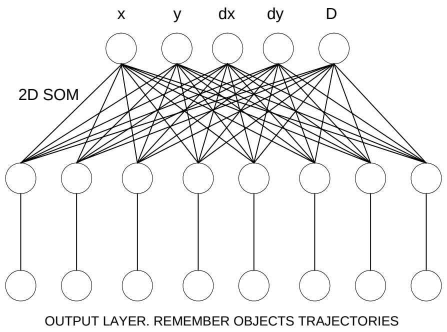

Fig. 5. Decision module architecture.

It is a SOM neural net whose input space is 5-dimensional and it also includes an output layer able to remember object trajectories.

The input layer has five neurons, four that represent system position and system speed, and another one that incorporates information not implicit in projection map: system multidimensional speed in input space. The trajectory flow vector is: (x, y, dx, dy, D).

The fifth dimension is justified by the next reasoning: as it is broadly demonstrated, SOM representation amplifies regions with high data density (the amplification ratio is called *magnification factor*), therefore, a large displacement in the 2D output space can correspond to a small displacement in input multidimensional space, and viceversa, that is to say, the trajectory, described in the 2D projection map, represents faithfully the current state of the process and its direction of movement, but not the speed. Information regarding the state speed is obtained from weights of the prototypes in input space as the distance between the positions of two consecutive neurons of a trajectory. It has been denominated multidimensional speed (D).

$$D = d(\mathbf{x}_i, \mathbf{x}_{i+1}) \qquad (2)$$

As in Owens and Hunter (2000), the SOM is used as learning method for the decision module. This net models possible instantaneous movements in the map. This new net should not be confused with the one used to obtain the analysis scenario starting from the process samples.

The leaky neuron layer contains as many neurons as SOM prototypes. A leaky neuron has a simple input and a simple output. The activation at epoch $t+1$ is calculated from the previous activation $a(t)$ and the current input I :

$$a(t+1) = \begin{cases} I & \text{if } I > \gamma a(t) \\ \gamma a(t) & \text{otherwise} \end{cases} \qquad (3)$$

Where γ is a coefficient in the range (0,1) that governs the rate of decay. The value of γ is 0.5 in our case.

After training, the two-layer neural net learns a model of the normal trajectories present in the training dataset. The D component plane and the residual visual representation of flow vectors can be used to inform the supervisor on deviations from the learned trajectories. The component plane of *j*-th variable associates the value that takes the variable for the neuron $\mathbf{m}_i$ in input space to the corresponding node $\mathbf{g}_i$ in the grid. The leaky neuron layer is used to generate a trajectory trace in the D component plane, and this way, the supervisor can know the system tendency. The residual is the difference between the estimation of the current flow vector, according to the SOM pattern, and the actual flow vector when the process describes novel (not modeled) trajectories. The residual visualization proposed in Díaz and Hollmen (2002) is used.

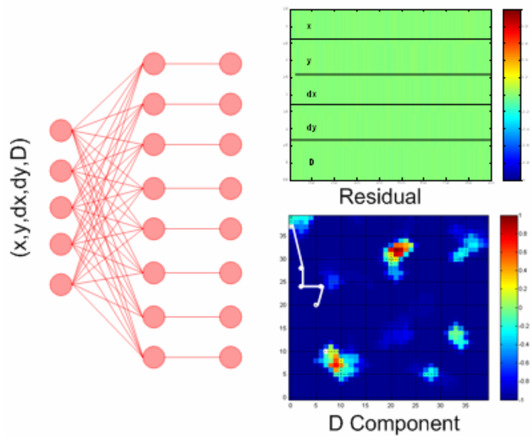

Fig. 6. Decision Visualization. Residual Visual Representation and D component plane with trace generated by leaky neurons.

If the D component plane and the residual map online are showed simultaneously (see Figure 6), supervisor can know if the process is in a significant change state (elevated multidimensional speed, bright colors in the component map), and he can detect erroneous flow trajectory vectors and therefore failure from trajectory deviation (residuals).

In Figure 7, a fault situation due to trajectory deviation when the system is in a significant state change is presented: the system representative neuron has a high D component and the residual map detects residual in speed.

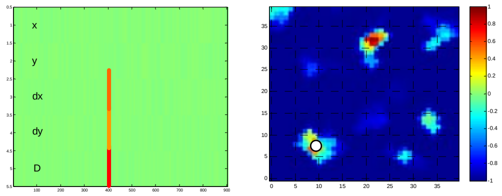

Fig. 7. Fault example on speed variation.

5. APPLICATION IN INDUSTRIAL PROCESS

To analyze the proposed industrial supervision system, it has been used to supervise an industrial process plant (See Figure 8). This plant is formed by a vertical steel reactor of 50 liters of capacity, it has a double cold-heat circuit with heat exchangers, in which both pressurized and not pressurized mixtures and thermal processes can be carried out. Intelligent industrial instruments measure and indicate: temperature, level, pressure, flow, pH and conductivity in the plant. These variables, along with performance and security variables, make a total of 32 process variables (Domínguez, *et al.*,2001).

Fig. 8. Supervised Industrial Pilot Plant.

In the 2D scenario in Figure 9, it is observed how the process describes 13 states: intake of the main circuit with flow control, high pressure recirculation, level control in the reactor, low pressure recirculation, temperature control (heating), temperature control (cooling), reactor drain, and the transition states between them.

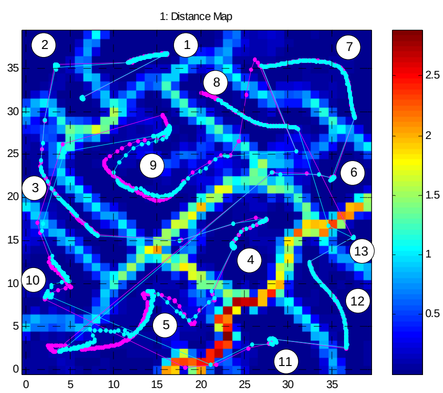

Fig. 9. Industrial plant supervision. 2D Scenario.

The input to the decision module is the flow vector trajectory described by process in a normal way. In Figure 10 the flow vectors that belong to state 9 can be seen.

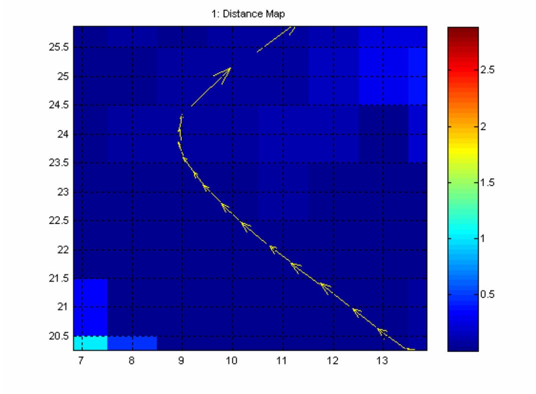

Fig. 10. Flow Vectors in state 9.

Three faults are caused: water supply fault, general pressure fault and heating resistance fault. In Figure 11, it is observed how the decision module, through the D component map and visual residual representation alerts of fault situations.

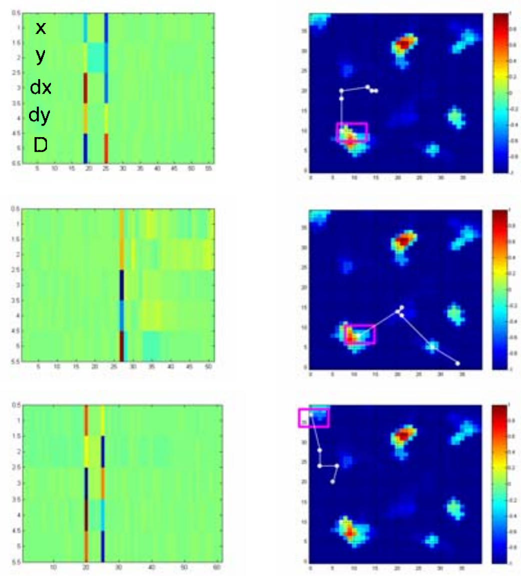

Fig. 11. Decision module alerts before fault situations.

6. CONCLUSIONS

In this paper, we have intended to incorporate the developed advances in the last years in visual surveillance systems to fault detection in complex industrial processes. We have defined a novel industrial supervision system.

This novel industrial supervision system exploits the powerful techniques developed for automatic surveillance systems to 2D scenarios consisting of visualization maps of the industrial process variable projections based on dimension reduction techniques. In particular, we have chosen the SOM distance map as scenario where the process state moves describing 2D trajectories.

To implement the decision module in the supervision system that models the typical trajectories described by the industrial system, we have proposed a SOM neural net with a 5D input space. A leaky neuron layer, on output, remembers the described trajectories. This module informs the supervisor of failure events through the D component map and the visualization of the residuals that appear as consequence of the trajectory deviation.

The proposed supervision system has been implemented in the supervision of an industrial plant of the *Instituto de Automática y Fabricación* at the University of Leon.

REFERENCES

Buxton, H. (2003). Learning and understanding dynamic scene activity: a review. *Image and Vision Computing*, **Vol. 21**, 125-136.

Collins, R.T., A.J. Lipton and T. Kanade (2000). Introduction to de Special Section on Video Surveillance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **Vol. 22**, nº8, 745-746.

Díaz, I., A. Diez and A. Cuadrado (2001). Complex Process Visualization Through Continuous Feature Maps Using Radial Basic Functions. *Lecture Notes in Computer Science (LNCS)*, **Vol. 2415**, 1005-1010.

Díaz, I., and J. Hollmen (2002). Residual generation and visualization for understanding novel process conditions. *Proceedings of IEEE IJCNN Conference*, 2070-2075.

Díaz, I., A. Cuadrado, A. Diez, L. R. Loredo, F. O. Carrera and J. A. Rodríguez (2003). Visual Predictive Maintenance Tool based on SOM Projection Techniques. *Revue de Métallurgie*, nº 3, 307-315.

Domínguez, M., P. Reguera, J.J. González, D. Marcos and L.F. Blázquez (2001). Connection Pilot Plant to Internet. *Workshop on Internet based control education*, IBCE'01, 43-47.

Fayyad, U., D. Haussler and P. Stolorz (1996) Mining scientific Data. *Communications of the ACM*, **Vol. 39**, nº 11, 51-57.

Hu, W., X. Dan and T. Tan (2004a). A Hierarchical Ser-Organizing Approach for Learning the Patterns of Motion Trajectories. *IEEE Transactions on Neural Networks*, **Vol. 15**, nº1, 135-143.

Hu, W., D. Xie, T. Tan and S. Maybank (2004b) Learning activity patterns using Fuzzy Self-Organizing Neural Network. *IEEE Transaction on Systems, Man and Cybernetics*, **Vol. 34**, nº 3, 1618-1626.

Johnson, N. and D. Hogg (1996). Learning the distribution of object trajectories for event Recognition. *Image and Vision Computing*, **Vol. 18**, nº 8, 609-615.

Keim, D. A. (2001). Visual exploration of large Data sets. *Communications of ACM*, **Vol. 44**, nº 8, 39-44.

Keim, D. A. (2002). Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics* **Vol.8,** nº1, 100-107.

Oliver, N. M., B. Rosario and A. P. Pentland (2000). A bayesian computer vision system for modeling human interactions. *IEEE Trans. Pattern Anal. Machine Intell.*, **Vol. 22**, 831–843.

Owens, J. and A. Hunter (2000). Application of the Self-Organising Map to Trajectory Classification. *Third IEEE Visual Surveillance Workshop*.

Wada, T. and T. Matsuyama (2000). Mutli-object behavior recognition by event driven selective attention method. *IEEE Trans. Pattern Anal. Machine Intell.*, **Vol.22**, 873-887.

Wang, L., W. Hu and T. Tan (2003). Recent developments in human motion analysis. *Pattern Recognition*, **Vol. 36**, 585-601.

Wilson, A. D., A. F. Bobick and J. Cassell (1997). Temporal classification of natural gesture and application to video coding. *Proc. IEEE Conf. Computer Vision Pattern Recognition*, 948–954.