Available at www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

Pattern Recognition Letters 24 (2003) 2575–2583

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Stereovision matching through support vector machines

Gonzalo Pajares [*,1], Jesús M. de la Cruz

Departmento Arquitectura de Computadores y Automática, Facultad de Ciencias Físicas, Universidad Complutense de Madrid, Ciudad Universitaria 28040 Madrid, Spain

Received 20 February 2003

Abstract

This paper presents an approach to the local stereovision matching problem using edge segments as features with four attributes. In this paper we design a Support Vector Machine classifier for solving the stereovision matching problem. We obtain a matching decision function to classify a pair of features as a true or false match. The use of such classifier makes up the main finding of the paper. A comparative analysis among other existing approaches is included to show that this finding can be justified theoretically. From these investigations, we conclude that the performance of the proposed method is appropriate for this task.

Keywords: Support Vector Machines; Stereovision; Matching; Classifier

1. Introduction

A significant amount of research in the computer vision community has been aimed at the study of the three-dimensional (3-D) structure of objects using machine analysis of images (Lee and Leou, 1994). Analysis of video images in stereo has become an important passive method for extracting the 3-D structure of a scene.

The key step in stereovision is that of image matching, namely, the process of identifying the corresponding points in two images that are generated by the same physical point in space. This paper is devoted solely to this problem. The stereo correspondence problem can be defined in terms of finding pairs of true matches that satisfy three competing constraints: similarity, smoothness and uniqueness (Marr and Poggio, 1979). The similarity constraint is associated with a local matching process where a minimum difference attribute (properties of features) criterion is applied. The results computed in the local process are later used by a global matching process where other constraints such as smoothness (Marr and Poggio, 1979), minimum differential disparity (Medioni and Nevatia, 1985) and figural continuity (Pollard et al., 1981) are imposed. A good choice of a local matching strategy is the key for good results in the global matching process.

This paper presents an approach to the local stereopsis correspondence problem by designing a

[*] Corresponding author. Tel.: +34-91-394-4477; fax: +34-91-892-2217.
E-mail address: pajares@dacya.ucm.es (G. Pajares).
[1] Tel.: +34-1-394-7600; fax: +34-1-394-4687.

0167-8655/$ - see front matter
doi:10.1016/S0167-8655(03)00102-8

Support Vector Machine classifier (SVC) where a decision function is to be derived. This function is then mapped as a posterior probability so that the performance of this matching strategy can be compared against other existing approaches using probabilities, as we will see below. The use of the SVC, applied to stereovision matching, makes up the main finding of this paper. This classification technique has a number of properties that make it particularly attractive and has recently received much attention in the machine learning community.

SVC uses only the similarity constraint in order to obtain local matching results and then map it in the global matching strategy described in (Pajares et al., 1998b) and based on the Hopfield neural network. This network maps the similarity, smoothness and uniqueness constraints in an energy function, which is then minimized to find the optimal correspondence. This procedure is also carried out in both, Pajares and Cruz (2002) and Pajares et al. (1998a) where the similarity constraint is mapped as a probability, which is computed from a probability density function (PDF). Such PDFs are previously estimated by using the Parzen's Windows classifier (PWC) (Pajares and Cruz, 2002) and a Bayesian classifier (BYC) (Pajares et al., 1998a) for solving the same problem.

The SVC improves the matching results compared to the results obtained with PWC and BYC. Hence, this improvement can be extended to other recent matching methods using only the similarity constraint (Pajares et al., 1998c, 1999) as they are compared in (Pajares and Cruz, 2002). The improvement is also considerable when the global matching strategy maps the similarity constraint by using SVC. This justifies the choice of our SVC as a good strategy to measure the similarity between features in stereovision images.

Two types of techniques have been broadly used for stereovision matching, namely the correlation-based and the feature-based methods. In the correlation-based method, the elements to be matched are image windows of fixed size and the similarity criterion is a measure of the correlation between windows in the two images. The corresponding element is given by the window that maximizes the similarity criterion within a search region. The number of pairs of features to be considered becomes high, because all pixels in the left image must be matched with all pixels in the right one. The feature-based methods use sets of pixels with similar attributes, usually either pixels belonging to edges or the corresponding edges themselves. Instead of image windows they use numerical and symbolic properties of features. The feature-based methods lead only to a sparse depth map, leaving the rest of the object surface to be reconstructed by interpolation. They are faster than area-based methods because there are fewer points (features) to be considered. See Wei et al. (1998) and associated references for both types of techniques.

The matching is made difficult in part by changes in the images of the corresponding points due to different viewpoints and to the different physical cameras. Hence, the corresponding attributes in the two images may display different values. This may lead to incorrect matches. Thus, it is very important to find features in both images that are independent of possible variations in the images. Our experiment has been carried out in an indoor space where edge segments are abundant, making suitable such features (Trucco and Verri, 1998). Moreover, they have been used in previous stereovision matching works (Pajares and Cruz, 2002; Pajares et al., 1998a,b,c, 1999; Medioni and Nevatia, 1985), and in (Wei et al., 1998) we can find additional references where edge segments are used. This fact justifies our choice of features, although such features are produced by intensity changes. Four attribute values (module and direction of the gradient vector, Laplacian and variance) are computed for each edge segment.

This paper is organized as follows. In Section 2 the stereo matching system is considered, which comprises three stages: (1) extraction of features and attributes; (2) SVC design and (3) matching for the current stereo pairs. To show the effectiveness of the proposed method, in Section 3 a test strategy is designed and a comparative analysis among other existing strategies is performed. A generalization of the method is proposed in Section 4. Finally, in Section 5 the conclusion is presented.

2. The support vector classifier in stereovision matching

Our local stereo matching system is designed with a parallel optical axis geometry working in three stages:

1. Extracting information (features and attributes) from the images.
2. Performing a training process, called OFF-LINE, with the samples (true and false matches) which are supplied to the SVC in order to compute an output function derived from the SVC. According to the function value a pair of features is classified as belonging to one of the two classes: true or false.
3. Performing a matching process, called ON-LINE, for the current pairs of features.

The first stage is common for both the OFF-LINE and the ON-LINE processes.

2.1. Feature and attribute extraction

The contour edge pixels in both images are extracted using the Laplacian of the Gaussian filter in accordance with the zero-crossing criterion, (Huertas and Medioni, 1986). At each zero-crossing in a given image we compute the magnitude and the direction of the gradient vector as in (Leu and Yau, 1991), the Laplacian as in (Lew et al., 1994) and the variance as in (Krotkov, 1989). These four attributes are computed from the gray levels of a central pixel and its eight immediate neighbors. The gradient magnitude is obtained by taking the largest difference in gray levels of two opposite pixels in the corresponding eight-neighbourhood of a central pixel. The gradient direction points from the central pixel towards the pixel with the maximum absolute value of the two opposite pixels with the largest difference. It is measured in degrees, quantified by multiples of 45. The normalization of the gradient direction is achieved by assigning a digit from 0 to 7 to each principal direction. The Laplacian is computed by using the corresponding Laplacian operator over the eight neighbors of the central pixel. The variance indicates the dispersion of the nine gray level values in the eight-neighborhood of the same central pixel. In order to avoid noise effects during edge-detection that can lead to later mismatches in realistic images, the following two globally consistent methods are used: (1) the edges are obtained by joining adjacent zero-crossings following the algorithm in (Tanaka and Kak, 1990), in which a margin of deviation of $\pm 20\%$ and $\pm 45^\circ$ is tolerated in magnitude and direction, respectively; (2) then each detected contour is approximated by a series of line segments as in (Nevatia and Babu, 1980); finally, for each segment an average value for the four attributes is obtained from all computed values of its zero-crossings. All average attribute values are scaled, so that they fall within the same range. Each segment is identified by its initial and final pixel coordinates, its length and its label.

Therefore, each stereo pair of edge-segments has two associated 4-D vectors $\boldsymbol{x}_\mathrm{l}$ and $\boldsymbol{x}_\mathrm{r}$, where the components are the attribute values and the sub-indexes l and r which denote features belonging to the left and right images, respectively. A 4-D difference vector of the attributes $\boldsymbol{x} = \{x_\mathrm{m}, x_\mathrm{d}, x_\mathrm{p}, x_\mathrm{v}\}$ is obtained from $\boldsymbol{x}_\mathrm{l}$ and $\boldsymbol{x}_\mathrm{r}$, whose components are the corresponding differences for the module of the gradient vector, the direction of the gradient vector, the Laplacian and the variance, respectively.

2.2. Training process: the support vector machine classifier

The SVC is based on the observation of a set X of n pattern samples to classify them as true or false matches, i.e. the stereovision matching is the well-known two classification problem. The outputs of the system are two symbolic values $y \in \{+1, -1\}$ corresponding each to one of the classes. So, $y = +1$ is with the class of true matches.

The finite sample (training) set is denoted by: $(\boldsymbol{x}_i, y_i)$, $i = 1, \ldots, n$, where each $\boldsymbol{x}_i$ vector denotes a training element and $y_i \in \{+1, -1\}$ the class it belongs to. In our problem $\boldsymbol{x}_i$ is as before the 4-D difference vector.

The goal of SVC is to find, based on the information stored in the training sample set, a decision function capable of separating the data into two groups. The technique is based on the idea of mapping the input vectors into a high-dimensional feature space using nonlinear transformation functions. In the feature space a separating hyperplane (a linear function of the attribute variables) is constructed (Cherkassky and Mulier, 1998; Vapnik, 2000). By using different mapping functions, different types of SVC are implemented. The SVC decision function has the following general form

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} \alpha_i y_i H(\boldsymbol{x}_i, \boldsymbol{x}) - b \tag{1}$$

where b is a constant.

Eq. (1) establishes a representation of the decision function $f(\boldsymbol{x})$ as a linear combination of kernels centred in each data point. Using different kernels $H(\boldsymbol{x}, \boldsymbol{y})$ we get different functions. In this paper, we have used Gaussian Radial Basis functions $H(\boldsymbol{x}, \boldsymbol{y}) = \exp\{-|\boldsymbol{x} - \boldsymbol{y}|^2/\sigma^2\}$ where σ defines the width of the kernel, set to 3.0 after different experiments. The choice of this kernel is motivated by the use of similar kernels in the PWC approach (Pajares and Cruz, 2002).

The parameters α_i, $i = 1, \ldots, n$, in Eq. (1) are the solution for the following quadratic optimisation problem:

Maximise the functional

$$Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j H(\boldsymbol{x}_i, \boldsymbol{x}_j)$$

$$\text{subject to } \sum_{i=1}^{n} y_i \alpha_i = 0, \quad 0 \leqslant \alpha_i \leqslant \frac{c}{n}, \; i = 1, \ldots, n \tag{2}$$

and given the training data $(\boldsymbol{x}_i, y_i)$, $i = 1, \ldots, n$, the inner product kernel H, and the regularization parameter c. As stated in (Cherkassky and Mulier, 1998), at present, there is not a well-developed theory on how to select the best c, although in several applications it is set to a large fixed constant value, such as 2000, which is the used in this paper.

The data points $\boldsymbol{x}_i$ associated with the nonzero α_i are called *support vectors*. Once the support vectors have been determined, the SVC decision function has the form

$$f(\boldsymbol{x}) = \sum_{\text{support vectors}} \alpha_i y_i H(\boldsymbol{x}_i, \boldsymbol{x}) - b \tag{3}$$

we have chosen in our experiments $b = 0$.

The SVC generates a scalar output $f(\boldsymbol{x})$ whose polarity, sign of $f(\boldsymbol{x})$, determines the class membership. The magnitude can usually be interpreted as a measure of belief or certainty in the decision made. As PWC and BYC use posterior probabilities, we use a warping function that maps $f(\boldsymbol{x})$ to a posterior probability. This is carried out assuming that posterior probabilities take the form of a sigmoid and directly estimating the sigmoid (Platt, 2000)

$$p(\boldsymbol{x}) = \frac{1}{1 + \exp\{-(af(\boldsymbol{x}) + v)\}} \tag{4}$$

In order to avoid severe bias in the distances for the training data, the parameters a and v are estimated experimentally and set to 0.2 and 0 in our experiments.

2.3. The current stereo matching process

This is an ON-LINE or decision process in which two new features are to be matched. This is carried out by obtaining the 4-D difference vector of the attributes $\boldsymbol{x}$. With this $\boldsymbol{x}$ we compute the matching probability $p(\boldsymbol{x})$ according to (4). Then, this incoming $\boldsymbol{x}$ is classified as a true or false match.

During the decision process there are unambiguous and ambiguous pairs of features, depending on whether a given left image segment corresponds to one and only one, or several right image segments, respectively. In any case, the decision about the correct match is made by choosing the pair with the greater probability value (in the unambiguous case, there is only one) provided that it surpasses the threshold of 0.50. This value is the intermediate probability value in the interval where the probability ranges. This is the *uniqueness* stereovision matching constraint also applied in PWC and BYC. When the probability value for a pair of features (edge-segments) does not exceed 0.5, it is considered a false match.

3. Comparative analysis and performance evaluation

To assess the validity and performance of our method, we designed a test strategy with two goals:

1. To verify the performance of the SVC against PWC and BYC individually in a local matching strategy and also when it is used for mapping the similarity constraint in a global process (Pajares et al., 1998b), as explained in the introduction.
2. To verify that the matching performance increases as the number of training patterns increases.

3.1. Design of a test strategy

The objective is to test the method by varying indoor environmental conditions in two ways: by using new images with different features (different objects) and by changing the illumination. With this aim in mind a set SP0 of 12 pairs of stereo-images captured with natural illumination was used to extract initial training patterns. Figs. 1–4 show four representative left images of this set.

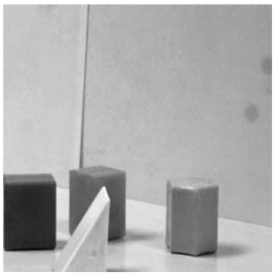

Fig. 1. Left original training image (blocks).

Fig. 2. Left original training image (furnitures).

Fig. 3. Left original training image (computers I).

Fig. 4. Left original training image (computers II).

Five additional sets of stereo-images, SP1, SP2, SP3, SP4 and SP5, all different to each other and to SP0, were used for the test. They were composed of 10, 10, 15, 12 and 11 stereo-images, respectively. The sets SP1 and SP4 were captured with natural illumination, as was the initial set SP0, and the sets SP2, SP3 and SP5 with artificial illumination. Two representative stereo-image pairs are shown for sets SP2 and SP3 in Figs. 5(a) and (b) and 6(a) and (b). The remaining stereo-image pairs belong to the same indoor environment and it is irrelevant to show representative pairs. The total number of pairs of edge-segments extracted from all stereo-images is 4638, which is the number n of training patterns used during the OFF-LINE processes for estimating the function $p(\boldsymbol{x})$ in (4). This number of pairs of edge-segments is completely representative of the environment where our mobile robot, equipped with our stereo-vision system, navigated. From this number, during all the ON-LINE processes, 2925 pairs of edge-segments were classified as true matches. Therefore, this is the total number of training patterns used for estimating the PDFs in PWC and BYC methods.

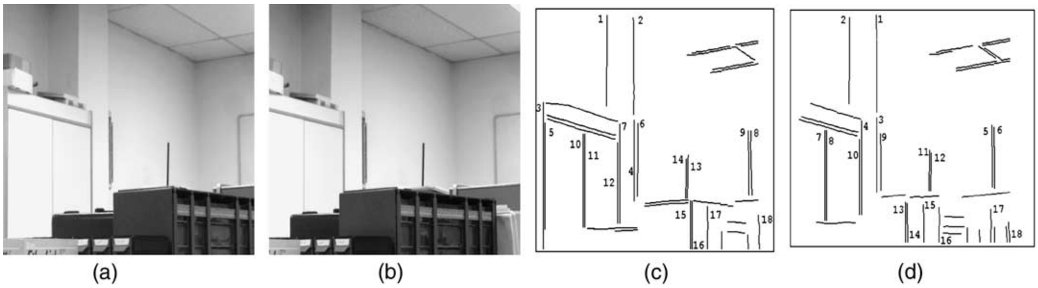

Fig. 5. (a) SP2: original left stereo-image; (b) SP2: original right stereo-image; (c) SP2: labeled segments left image and (d) SP2: labeled segments right image.

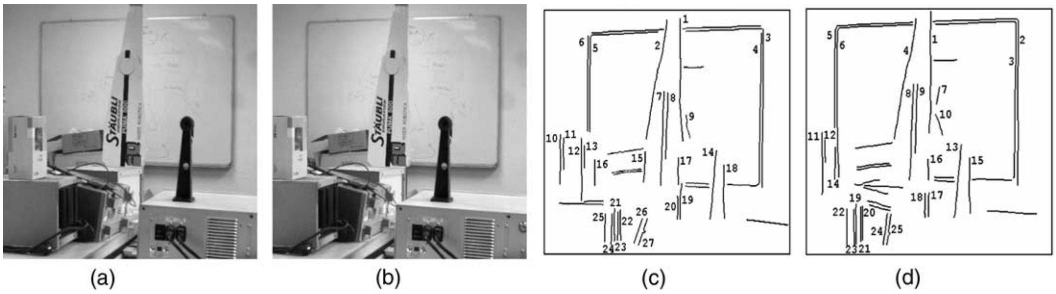

Fig. 6. (a) SP3: original left stereo-image; (b) SP3: original right stereo-image; (c) SP3: labeled segments left image and (d) SP3: labeled segments right image.

The process can be summarized as follows:

STEP 0: ON-LINE → Classify the pairs of features in the set SP0 as true or false matches by using the unsupervised learning strategy described in (Pajares et al., 1999).
OFF-LINE → Estimate $p(\boldsymbol{x})$ through Eq. (4) with the matches classified as true or false during the previous ON-LINE process.
For $\boldsymbol{n} = 1$ to $\boldsymbol{n} = 5$ do

STEP n: ON-LINE → Classify each pair of features, represented by $\boldsymbol{x}$, as a true or false match for the set SP***n*** with the $p(\boldsymbol{x})$ obtained during ***STEP n*** − 1.
OFF-LINE → Estimate $p(\boldsymbol{x})$ through Eq. (4) with the matches classified as true or false during all the ON-LINE processes in ***STEP***s ***0*** to ***n***.

The training patterns and the parameters are stored after each ON-LINE process, so that they can be used for estimating $p(\boldsymbol{x})$ in (4).

The matches required during STEP 0 can be supplied to the system by using any unsupervised stereovision matching method. We have used the method of Pajares et al. (1999), as it has already been tested. When this is not possible, the use of a minimum distance criterion is appropriate and the Euclidean distance is sufficient. Also, a human expert could provide such true matches, although this implies that the system loses its automatic capability at this stage.

4. Comparative analysis

We analyzed the results in more detail to see the performance of our SVC approach. Table 1 displays the percentages of successes for STEPs 1–5. A distinction is made for local and global processes.

Local processes. We call local (L) matching strategies those which use only the similarity constraint, from the known probability function $p(\boldsymbol{x})$. First, based on such values separately, we classify

Table 1
Global (G) and local (L) results for the sets of stereo pairs SP1–SP5

% Successes classifiers	SP1	SP2	SP3	SP4	SP5
LSV	79.3	88.2	94.1	94.3	94.8
LPW	79.3	88.4	94.0	94.3	94.6
LBY	78.2	83.3	86.7	86.8	87.2
GSV	96.3	96.1	98.0	98.1	98.2
GPW	96.1	96.2	97.8	97.9	98.1
GBY	95.1	95.3	95.6	96.0	96.7

a pair of features, represented by $\boldsymbol{x}$, as a true match depending on if its probability value is greater than a threshold, fixed to 0.5 as it is the intermediate value in the range of probability values, otherwise it is a false match. The percentage of successes appear in Table 1 as Parzen's window (LPW), Bayesian (LBY) and Support Vector Machine (LSV).

Global processes. As mentioned before, we select the global matching strategy described in (Pajares et al., 1998b) where the similarity constraint is mapped as a probability function. SVC uses $p(\boldsymbol{x})$ in (4) and PWC and BYC the corresponding estimated PDFs. The results are given in Table 1 as GPW, GBY and GSV, respectively, where G means global and PW, BY, SV refer to the PWC, BYC and SVC strategies, respectively.

From results of Table 1 the following conclusions may be inferred.

1. As expected, global approaches provide better results than local ones. This is the consequence of using a global relaxation approach with more stereovision matching constraints than in the local methods.
2. The results obtained by both GSV and LSV are similar or even better than those obtained by GPW and LPW. Since the PWC compare favorably with other existing learning strategies (Pajares and Cruz, 2002), we can conclude that the SVC appears to be a valid method for local stereovision matching.
3. The matching performance increases as the number of training patterns increases. Indeed, the best results are obtained for SP5, which has been processed with a number of training patterns greater than the remainder SP*i*.

Table 2
Number of training patterns used by PWC and BYC and support vectors in SVC for the stereo pairs SP1–SP5

	SP1	SP2	SP3	SP4	SP5
# Training patterns	528	1157	1897	2378	2925
# Support vectors	56	65	73	73	73

Table 2 displays the number of training patterns used for the PWC and BYC methods for estimating the corresponding PDFs for the different SPs from 1 to 5 and the number of support vectors obtained for the SVC approach for each set of stereo pairs.

From results of Table 2 the following conclusions may be inferred.

1. The number of support vectors is significantly less than the number of training patterns. As the performance of the SVC has proven to be acceptable, this implies that we require less number of patterns to be stored than using PWC or BYC.
2. The number of support vectors does not vary for SP3, SP4 and SP5. Moreover, they are practically the same support vectors, i.e. the SVC requires less number of training images than PWC or BYC for achieving a best performance.

We have also verified the performance of the SVC for solving ambiguities during the local matching process. This is carried out by computing a coefficient μ which provides a decision margin when ambiguities arise. It is computed as follows:

(a) Without loss of generality, assume the following set of pairs of edge segments as an ambiguous case. The left edge segment l matches with q right edge segments, $r = 1, 2, \ldots, q$; with matching probabilities p_{lr}. Let a one of the q right edge segments, so that $p_{la} = \max_{r=1,2,\ldots,q} \{p_{lr}\}$. As mentioned before, the match la is considered a correct match.

Table 3
Decision margin measures for the sets of stereo pairs SP1–SP5

μ	SP1	SP2	SP3	SP4	SP5
LSV	0.14	0.18	0.23	0.28	0.31
LPW	0.08	0.10	0.15	0.18	0.20
LBY	0.04	0.09	0.11	0.12	0.17

(b) Compute $q = \min\{|p_{la} - p_{lr}|\}$, for $r = 1, 2, \ldots, q$; $r \neq a$.
(c) For each ambiguous case j on each set of stereo-image pairs SPh, where $h = 1, \ldots, 5$, compute m_j as in (b).
(d) For each SPh at each step h, compute the coefficient μ as follows:

$$\mu = \frac{1}{k} \sum_{j=1}^{k} m_j \tag{5}$$

where k is the number of ambiguous cases in SPh.

Table 3 displays the decision margin when ambiguities arise. It is clear that the best decision margin is achieved with the LSV approach, i.e. the number of errors is minimized under this strategy.

5. Generalization

In this paper, we have used edge segments as features and the similarity constraint has been proved separately and also mapped in a global strategy. We have proven the performance of the SVC for stereovision matching. The proposed method can be generalized and applied as follows:

1. For any type of features, i.e. curved segments, regions, edge points, corners, etc., according to the salient features in the environment (indoor or outdoor). This is because our approach is based on the estimation of one function where the patterns are the attribute values for the edge segments. Hence, for any given feature the only problem is to compute its attributes, which is beyond the scope of this paper.
2. For a multiresolution scheme where the similarity matching constraint is used. Three types of multiresolution are commonly used:
   (2.1) classical pyramidal structures using Laplacian of the Gaussian filters extract information at different resolution levels based on the standard deviation parameter (Huertas and Medioni, 1986). The matching process at a given resolution level can be carried out through Eq. (4) and is driven by disparity values produced at a coarser level;
   (2.2) multiscale decomposition (Zhang and Blum, 1999) including wavelets Kim et al. (1997). At each level an activity-level measurement is required. In stereovision matching this is performed through a similarity measurement (Kim et al., 1997) which can be achieved by using Eq. (4);
   (2.3) a two level scheme, where the first level matches intensity edges, i.e. a feature-based matching, which drives the disparity at a second level based on area-based matching (Baillard and Dissard, 2000). As before, the similarity measurement can be obtained through Eq. (4).
3. Studying in depth the influence of different kernels in the SVC such as: polynomials of degree $qH(\boldsymbol{x}, \boldsymbol{y}) = [(\boldsymbol{x} \cdot \boldsymbol{y}) + 1]^q$ or two layer neural networks $H(\boldsymbol{x}, \boldsymbol{y}) = \tanh\{b(\boldsymbol{x} \cdot \boldsymbol{y}) + c\}$. Other kernels can be found in (Vapnik, 2000; Cherkassky and Mulier, 1998).

6. Concluding remarks

The SVC improves results when it is used to measure the similarity constraint in both: local and global strategies. An interesting conclusion is that SVC has proven to be an acceptable classifier when compared individually with PWC and BYC and in a indirect fashion against the methods which have been compared with PWC and BYC in (Pajares and Cruz, 2002; Pajares et al., 1998a), respectively.

The number of stereo-images, features and experiments was adequate for our stereo matching approach. This number of pairs of edge-segments was completely representative of the environment where our mobile robot, equipped with our stereovision system, navigated.

The performance of the SVC approach improves as the number of training patterns increases. This behavior is affected neither by the nature of the different objects nor by the illumination conditions.

We have made a generalization of the proposed method and given guidelines for its extension and application to other matching strategies where the similarity constraint is used. This generalization is applicable to other types of environments, such as outdoor scenes or aerial images, where edge segments are probably inappropriate features and therefore different features and attributes would be more suitable.

Acknowledgement

Part of the work has been performed under project CICYT TAP94-0832-C02-01.

References

Baillard, C., Dissard, O., 2000. A stereo matching algorithm for urban digital elevation models. Photogrammet. Eng. Remote Sensing 66 (9), 1119–1128.

Cherkassky, V., Mulier, F., 1998. Learning from Data: Concepts, Theory and Methods. Wiley, New York.

Huertas, A., Medioni, G., 1986. Detection of intensity changes with subpixel accuracy using Laplacian–Gaussian masks. IEEE Trans. Pattern Anal. Machine Intell. 8 (5), 651–664.

Kim, Y.S., Lee, J.J., Ha, Y.H., 1997. Stereo matching algorithm based on modified wavelet decomposition process. Pattern Recognition 30 (6), 929–952.

Krotkov, E.P., 1989. Active Computer Vision by Cooperative Focus and Stereo. Springer-Verlag, Berlin.

Lee, S.H., Leou, J.J., 1994. A dynamic programming approach to line segment matching in stereo vision. Pattern Recognition 27, 961–986.

Leu, J.G., Yau, H.L., 1991. Detecting the dislocations in metal crystals from microscopic images. Pattern Recognition 24 (1), 41–56.

Lew, M.S., Huang, T.S., Wong, K., 1994. Learning and feature selection in stereo matching. IEEE Trans. Pattern Anal. Machine Intell. 16 (9), 869–881.

Marr, D., Poggio, T., 1979. A computational theory of human stereovision. Proc. Roy. Soc. London B 207, 301–328.

Medioni, G., Nevatia, R., 1985. Segment based stereo matching. Comput. Vision Graphics Image Process. 31, 2–18.

Nevatia, R., Babu, K.R., 1980. Linear feature extraction and description. Comput. Vision Graphics Image Process. 13, 26–257.

Pajares, G., Cruz, J.M., 2002. The non-parametric Parzen's window in stereovision matching. IEEE Trans. Systems Man Cybernet. 32 (2), 225–230.

Pajares, G., Cruz, J.M., López, J.A., 1998a. Pattern recognition learning applied to stereovision matching. In: Amin, A., Dori, D., Pudil, P., Freeman, H. (Eds.), Advances in Pattern Recognition. Springer-Verlag, Berlin, pp. 997–1004.

Pajares, G., Cruz, J.M., Aranda, J., 1998b. Relaxation by Hopfield network in stereo image matching. Pattern Recognition 31 (5), 561–574.

Pajares, G., Cruz, J.M., López-Orozco, J.A., 1998c. Improving stereovision matching through supervised learning. Pattern Anal. Applicat. 1, 105–120.

Pajares, G., Cruz, J.M., López-Orozco, J.A., 1999. Stereo matching using Hebbian learning. IEEE Trans. Systems Man Cybernet. 29 (4), 553–559.

Platt, J., 2000. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: Advances in Large Margin Classifiers. MIT Press, Cambridge, MA.

Pollard, S.B., Mayhew, J.E.W., Frisby, J.P., 1981. PMF: a stereo correspondence algorithm using a disparity gradient limit. Perception 14, 449–470.

Tanaka, S., Kak, A.C., 1990. A rule-based approach to binocular stereopsis. In: Jain, R.C., Jain, A.K. (Eds.), Analysis and Interpretation of Range Images. Springer-Verlag, Berlin.

Trucco, E., Verri, A., 1998. Introductory Techniques for 3-D Computer Vision. Prentice-Hall, Upper Saddle River.

Vapnik, V.N., 2000. The Nature of Statistical Learning Theory. Springer-Verlag, New York.

Wei, G.Q., Brauer, W., Hirzinger, G., 1998. Intensity- and gradient-based stereo matching using hierarchical Gaussian basis functions. IEEE Trans. Patt. Anal. Machine Intell. 20 (11), 1143–1160.

Zhang, Z., Blum, R.S., 1999. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. Proc. IEEE 87 (8), 1315–1325.