PATH PLANNING AND NAVIGATION IN A SPARSE WORLD SPACE ENVIRONMENT

Marek B. Zaremba

Département d'informatique et d'ingénierie
Université du Québec (UQO)
Gatineau, Quebec J8X 3X7, Canada
zaremba@uqo.ca

Abstract: This paper investigates the mobile robot navigation problem in a situation where the information about the navigation world can be available in a form of a sparse data set. Such a problem arises when the information about the navigation environment, obtained usually in a raster format from a vision sensor, is subject to noise, continuity and connectivity distortions. In order to deal with the data sparseness, an approach based on spatial filtering, a derivative of a Gabor filter, is proposed. The results of the filtering procedure are used in a different way for path planning and for navigation tasks. The mobile robot path is defined in terms of a skeleton that retains the connectivity information of the shape of the admissible navigation area. The navigation algorithm applies gradient-based neuromorphic processing on the environment map obtained by using a procedure for adaptive thresholding of the filtered sparse data.

Keywords: Robot navigation, image analysis, neural networks, path planning, spatial filtering, reactive navigation control, potential field method.

1. INTRODUCTION

The issue of vision-based navigation of a mobile robot has attracted considerable attention, and resulted in a host of methods, approaches and techniques from different areas, such as artificial intelligence, optimal control, geometrical analysis, etc. In order to derive the navigation path, all the methods use some sort of a model of the obstacles or, alternatively, of the free navigation space.

The navigation problem has traditionally been dealt with by adopting one of the following general approaches: numerical, geometric, or reactive. In the numerical approach, robot path planning and control is regarded as the solution of a non-linear optimization problem with constraints (Glasius *et al*, 1995; Seshadri and Gosh, 1993). In the geometric approach, solution of the navigation task is based on a detailed geometric model (Rowe and Richbourg, 1990). With the reactive type of navigation controller (Payton, 1986), the robot actions are determined by the currently sensed state of the environment.

In situations of incomplete or uncertain information about the robot environment, the sensory data have to be incorporated into the motion control functions. Vision systems in particular have found numerous applications in a variety of image-based control schemes. An efficient use of visual information for navigation purposes in the area of autonomous robot systems was discussed by Talluri and Aggarval (1996). The applications of image-based servoing have been reported by Weiss *et al*. (1987) and Zaremba *et al*. (1987). Visual servoing has been extended to 2.5-D and 3-D setups as presented in (Malis and Chaumette, 2000; Wilson *et al*., 1996). The problem of path planning in image space was dealt with in (Cowan and Koditschek, 2000). An approach consisting in coupling path planning in image space and image-based control, and taking into account constraints in the realized trajectories was discussed in

(Mezouar and Chaumette, 2002). A vast majority of the vision-based navigation algorithms assume a precise acquisition of the environment imagery, usually by a grey scale camera.

Poor lighting conditions, incorrect thresholding in image processing algorithms, and image subsampling are examples of the situations where sparseness in planar shapes and object image regions may occur. In general, shape sparseness is defined as a lack of pixel level connectivity in the raster image of an object. Shape analysis and recognition is one of frequently encountered problems in computer vision and image processing. However, most of the methods and algorithms used for shape analysis are not applicable to sparse objects. The problem becomes even more complicated when the image is obtained using a multispectral camera or another type of multidimensional raster-type sensor. At the same time, recent developments in remote sensing have resulted in the availability of multispectral (tens of spectral frequencies) and hyperspectral (hundreds of spectral frequencies) sensors. For example, the HYperspectral Digital Image Collection Experiment (HYDICE) sensor developed by Naval Research Laboratory has 210 spectral channels and 1–4 m spatial resolution. Because of the high spatial resolution such sensors can be used for surveillance and reconnaissance purposes, often with the involvement of navigation tasks. One of the major challenging problems about multispectral image analysis for navigation is how to deal with the massive data volume in real time while taking full advantage of plentiful spectral information.

This paper presents an integrated approach for extracting an appropriate representation of the navigation environment from multispectral sparse images using statistical learning and spatial filtering. The representation is further employed for path planning and navigation control. For both tasks, the use of an image operator that has local maxima at the centers of location of the objects of interest (Palenichka *et al*., 2000) is investigated. The operator combines neighbouring image components related to selected image properties. The adopted approach can be described as a general, model-based attention mechanism. This attention mechanism is a bottom-up image analysis process based on given low-level intensity and shape properties in order to respond to various unknown stimuli, as well as to provide a reasonable response when no object of interest is present.

Using the results of the spatial filtering, the task of planning a path between a pair of points can be defined as first, finding a piece-wise linear curve connecting successive maxima of the output surface, and second, smoothing the curve as a function of the maximum robot acceleration. By applying gradient-based neuromorphic processing on the thresholded output surface, the presented method can be extended to the task of local navigation. With the use of neuromorphic networks, the sensing process can be integrated with the controller. The navigation control law presented in the paper combines retinal processing of the IRF with subsequent generation of the parameters required by the control scheme based on the potential field method, and enables instantaneous, on-line computation of the robot displacement vectors.

2. EXTRACTION OF NAVIGATION SPACE FROM MULTIDIMENSIONAL DATA

2.1 Data dimensionality

The task of extracting the navigation space from multidimensional data, such as a multispectral sensor image, can be presented in terms of a classification problem, where each pixel is assigned a class label depending on the combination of the spectral intensities associated with the pixel. Typically, two classes are assigned, those associated with the free navigation space and the obstacles. However, the final two-class segmentation of the image can be a result of aggregating areas assigned to several classes. For example, a navigation space can include asphalt, gravel and sand surfaces. The design of a good classifier becomes rapidly more difficult as the dimensionality of the input space increases.

One way of dealing with the problem of the curse of dimensionality is to reduce its dimensionality before applying a classification procedure. Dimensionality reduction can be defined as the process of mapping high dimensional objects to a lower dimensional representation. Formally, dimensionality reduction produces $\hat{x} = j(x)$, where $x \in \Re^n$, $\hat{x} = \Re^m$ and m < n. Mapping j may be linear or nonlinear, and may be obtained using supervised or unsupervised learning procedures. In the navigation type applications, as opposed to many problems in data mining, exemplars of the classes related to the navigation areas are well defined and readily available. Consequently, supervised learning methods can be used.

A standard technique in pattern recognition aiming to achieve optimal linear dimensionality reduction is Fisher's linear discriminant analysis – LDA (Fisher, 1950). The LDA method, stated for a two class separation problem, aims at finding the linear combination $Z = a^T X$ such that the difference between the projected class means, normalized by a measure of the within-class scatter along the direction of vector $\boldsymbol{W}$ of adjustable weight parameters is maximized. It projects the original high dimensional data onto a low dimensional space, where all the classes are well separated by maximizing the ratio of between-class scatter matrix determinant to within-class scatter matrix determinant.

$$\max_a = \frac{a^T B a}{a^T W a} \tag{1}$$

Fisher's problem can be, therefore, defined as a generalized eigenvalue problem, with a given by the largest eigenvalue of $W^{1}\boldsymbol{B}$.

2.2 Multi-class image segmentation

The spatial filtering computations discussed in the next section are performed using the values of a function obtained as a result of the multispectral image segmentation. The function is defined on each pixel (i,j) as:

$$f(i,j)=\sum_{l} t_l \cdot j_l(i,j) \tag{2}$$

where $\{t_l\}$ are constant intensity values of image plane segments corresponding to the objects of interest and the background, and $j_l(i,j)$ is the function describing the lth object. The function $j_l(i,j)$ is equal to zero in the whole image plane except for the points belonging to the lth object.

Multispectral image classification has been recently a topic of considerable research interest. Filter vector algorithm (FVA) was proposed by Bowles *et al*. (1995). Constrained energy minimization (CEM) operator was investigated in (Farrand and Harsanyi, 1997). Algorithms for real-time data processing capable of processing data on-line and providing the information for immediate response were reported in (Stellman *et al*., 2000; Chang *et al*., 2001). The algorithm presented in (Du and Chang, 2001) is a generalization of FVA. The algorithm called constrained linear discriminant analysis (CLDA) projects the original high dimensional data onto a low dimensional space as done by Fisher's LDA. In this low dimensional space different classes are aligned with different directions. This constraint has to be satisfied when the algorithm maximizes the ratio of inter-class distance to intra-class distance, and obtained by using an orthogonal subspace projection (OSP) method (Harsanyi and Chang, 1994) coupled with a data whitening process. As demonstrated in (Du and Chang, 2001), the CLDA algorithm provides more accurate classification results than other popular methods in multispectral image processing. This algorithm has been adopted for the calculation of $j_l(i,j)$.

3. SPATIAL FILTERING OF SPARSE DATA

Sparseness in the visual data sets occurs when the information obtained usually in a raster format from a vision sensor is subject to noise, continuity and connectivity distortions. Pixel-by-pixel classification algorithms, such as the CLDA algorithm, may generate results in the form of sparse images. This is due to the necessity of unmixing sub-pixel data, the classification errors, and the inherent sparsity of the image itself. In order to obtain the environment data useful for navigation planning and control, a spatial filtering procedure has to be performed. The choice of the filter has to take into account the requirements for handling multi-resolution data and the scale and rotation invariance.

Gabor filter banks (Gabor, 1946) are good models of visual processing in primary visual cortex and are one of the most successful approaches for multi-scale image processing. In the space domain, the impulse response of Gabor filters is a Gaussian kernel modulated by a sinusoidal planewave. The spatial frequency can also be expressed in polar coordinates as magnitude and direction. Current Gabor filter bank systems are overly complex, especially for real-time implementations often required in navigation control.

An efficient approach to adopt Gabor filters to multi-resolution sparse data and, at the same time describe a planar shape (as shown in Section 4) is replace Gaussian kernels by structuring elements. In our approach, an initial structuring element S_0 is defined as a symmetrical disk shape in order to insure rotational invariance as well as a reasonable approximation of different possible shapes. A structuring element with diameter d_m at scale m in the uniform scales system is formed, as shown in Fig. 1, by a consecutive binary dilation by S_0, $S_m = S_{m-1} \oplus S_0$, m = 1, 2, ..., M - 1, where M is the total number of scales.

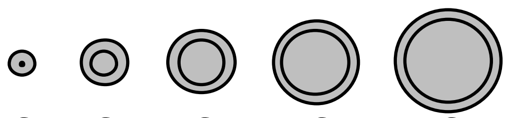


Fig. 1. An example of formation of 5 uniform scales by consecutive scale dilation of S_0.

The analytical expression for the filter can now be defined as a local image operator which takes larger values at points belonging to an object of interest than to the background points. The operator R applied to the function $f(i,j)$ (2) can be defined as in (3), where $|O|$ and $|B|$ denote the number of points in object and background sub-regions $B(i,j)$ and $O(i,j)$, and α is a constant coefficient.

4. PATH PLANNING

Path planning in an environment described by sparse data sets requires representations that retain the connectivity information of the shape. Such a representation is the skeletal representation frequently used in machine vision. Formally, the skeleton of a shape is the locus of the symmetric points of the local symmetries of the shape. Most of the existing skeletonization methods are based

$$R\{f(i,j)\} = \left(\frac{1}{|O|}\sum_{(m,n)\in O(i,j)} f(m,n) - \frac{1}{|B|}\sum_{(m,n)\in B(i,j)} f(m,n)\right)^2 - a\cdot\left(a - \frac{1}{|O|}\sum_{(m,n)\in O(i,j)} f(m,n)\right)^2 \tag{3}$$

on thinning techniques (Chen and Yu, 1996). The basic principle of thinning methods is to repetitively remove the outer layer of pixels of a shape until no more can be removed without altering the topology of the shape. The remaining pattern is the skeleton of the shape. The thinning techniques are sensible to the values of the selected parameters (Singh *et al*., 2000), which is illustrated in Fig. 2. The third and fourth form show the skeletons obtained by thinning the first image, whereas the skeletal form depicted in the second image would be preferred

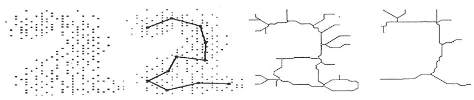
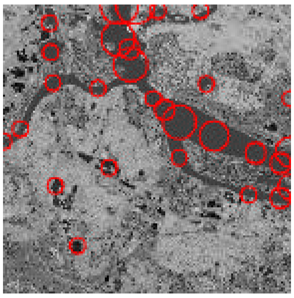


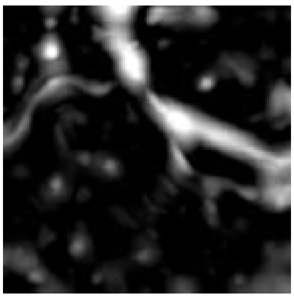
Fig. 2. Skeletons of a sparse image.

Using the spatial filtering method described in Section 3, the task of planning a path between a pair of points in the image space can be defined as finding a piece-wise linear curve connecting successive maxima of the function (4). The idea of using this approach to path planning is shown in Fig. 3.

a)



b)



c)

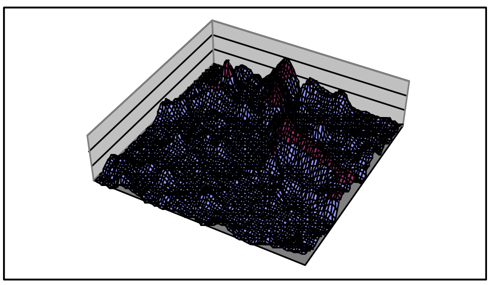


Fig. 3. Path planning.

a) Local structuring elements,
b) Results of the spatial filtering,
c) 3-D plot of $R\{f(I,j)\}$.

The original image, obtained, in this case, as a satellite photo, with superimposed structuring elements is shown in Fig. 3 a). The 2-D image of the resulting filtering using function $R\{f(i,j)\}$ is depicted in Fig. 3 b), and the 3-D plot in Fig. 3 c). The segmentation map $f(i,j)$ used by the spatial filter is a function of the intensity of the image. The robot path can be found by connecting successive maxima of IRF between the initial and the final point. Figure 4 shows an example of a path obtained in this way.

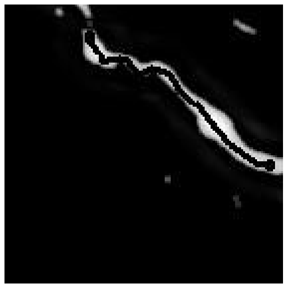


Fig. 4. Navigation path.

The calculated path is superimposed on the intensity image.

5. NAVIGATION

The navigation algorithm that applies the IRF approach is based on the potential field method (Khatib, 1986). It has been adapted to the image-based operation by using a reactive control scheme proposed by Zaremba and Porada (1998) and defined in a discrete space equivalent to a raster image.

The robot is pulled toward the goal point P_{target} by the attractive potential, and at the same time pushed from the obstacles by the repulsive potential. The resulting robot displacement vector, i.e., a planar vector **d** tangent to the robot trajectory at the current robot position P, is a linear combination of an attraction force **U** directed toward a target point P_{target} and a repulsive force **V** associated with the local configuration of the obstacles:

$$\mathbf{d} = a\mathbf{V} + \beta\mathbf{U} \qquad (5)$$

where a and β is are scalar coefficients. The local visual information is acquired in the form of an image $\mathbf{I}_{loc}$ that appears within the perimeter centred at current point P. The local image $\mathbf{I}_{loc}$ and the vector **U** are fed to the controller processor, which then generates the displacement vector **d**, after calculating vector **V**. The repulsive vector **V** originating at P is directed outwards from the obstacle and orthogonal to the obstacle border. It can be defined as a normalized weighted mean of unitary normal vectors within $\mathbf{I}_{loc}$. If no border point appears within the vision perimeter, or if the local image is symmetrical with respect to the centre of the image, then $\mathbf{V} = \mathbf{0}$.

The above reactive navigation control algorithm is applied to the image obtained by thresholding the IRF function at different levels. The obstacle area is defined by the points for which $R\{f(i,j)\} < T$, where T is a variable threshold. The value of T depends on the size of the mobile robot and the precision of the navigation trajectory with respect to the optimal path. The changes of the value of T are also induced dynamically by the dead-end situations, i.e., when the robot is no longer able to move toward the target. In such cases the value of T decreases until the size of the structuring element S reaches a maximum value S_{max}. The effect of the variations of T is illustrated in Fig. 5.

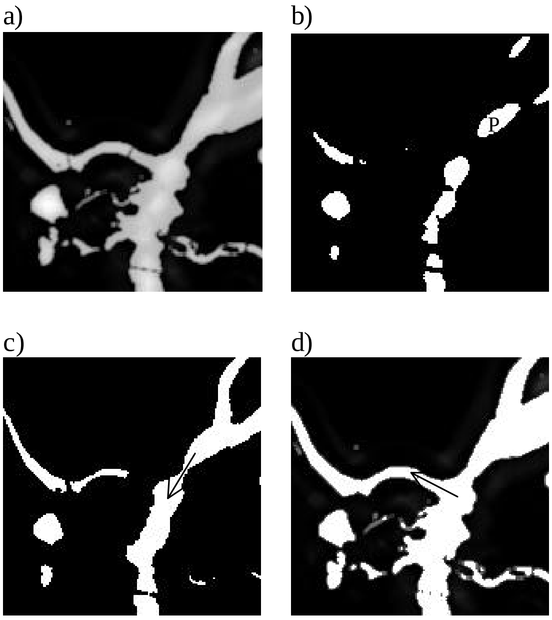


Fig. 5. Obstacle configuration.

a) IRF shown as input grey-scale image,

b) - d) secure navigation zones obtained from IRF at three different thresholds: $T_b > T_c > T_d$.

Initially, the obstacle free area around the initial point P is limited by a small value of the structuring element S_P. Assuming that the navigation target point is located on the left-hand side of the image, the mobile robot would tend to move in the SW direction, and would soon stop at a dead-end. A decrease of the threshold value will reconfigure the obstacle space, and allow the robot to negotiate the passage to a zone closer to the target point, as illustrated in Fig. 5c). Further decrease of T will allow the robot to move toward the left, closer to the target. It should be noted that – in order to keep the actual trajectory close to optimal, i.e., adjacent to the maximal $R\{f(i,j)\}$ values - the variation of T is performed dynamically during the navigation, without being limited to the dead-end situations.

The functional architecture of the entire navigation control system is shown in Figure 6. Also, an optional rule-based reasoning module is shown.

6. CONCLUSIONS

It has become more and more frequent to obtain visual information about the navigation environment from multispectral sensors. The results of processing of the data can be obtained and available in a form of sparse images, which makes the use of most machine vision and navigation control algorithms inadequate. The visual navigation problem in such an environment was discussed in this paper. The presented approach consists in three stages: extraction of a sparse image of the navigation space, spatial filtering of the image, navigation and/or path planning process. The supervised classification algorithm performs dimensionality reduction of the input data and optimal, in terms of discriminatory power, multi-class segmentation. The use of a spatial filtering operator, based on multi-resolution Gabor filters, provides a mechanism for a quick localization of obstacles invariantly to their size and orientation. The path planning operation consists in joining the points of local maximum value of a 3-D function obtained by the spatial filtering. The reactive navigation controller

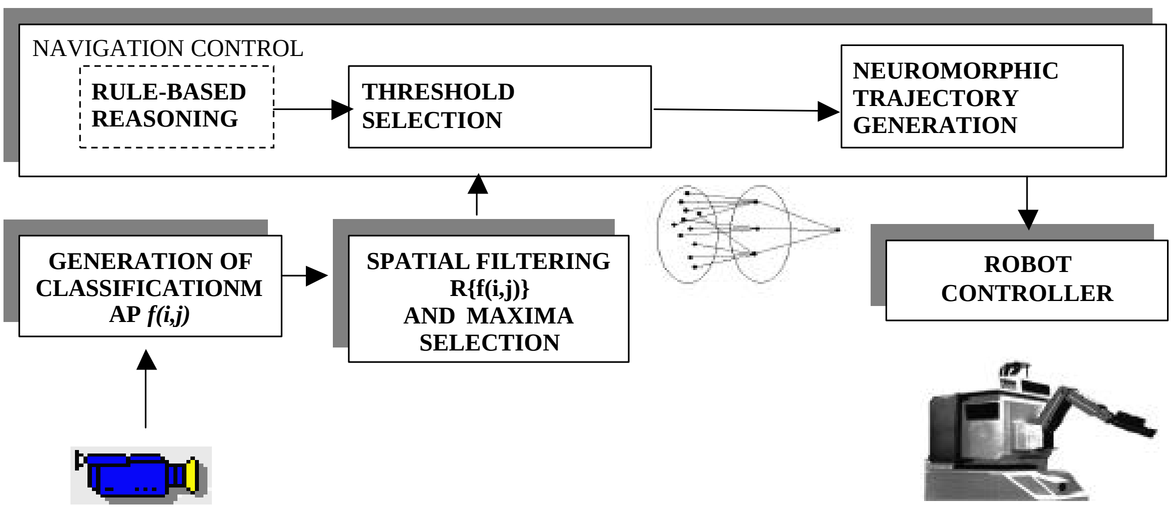


Fig. 6. Navigation control system.

comprises two main elements, the dipole network generator and the controller producing the displacement vector. The modular structure of the controller assures flexibility of the operation and easy extension of its functionality by, for example, interfacing to a supervisory, knowledge-based level of control. The proposed architecture offers the advantage of being oriented to direct processing of discrete input data, instantaneous execution, flexibility, and reactive operation that does not require memory-resident models or sequential procedures. It can be easily implemented in a form of a programmable hardware circuitry, allowing for fast, real-time operation on large size images.

REFERENCES

Bowles, J., P. Palmadesso, J. Antoniades and M. Baumback (1995). Use of filter vectors in hyperspectral data analysis, *Proc. SPIE*, **2553**, 148–152.

Chang, C.-I., H. Ren and S.S. Chiang (2001). Real-time processing algorithms for target detection and classification in hyperspectral imagery, *IEEE Trans. Geoscience and Remote Sensing*, **39**, 760–768.

Chen, Y.S. and Y.T. Yu (1996). Thinning approaches for noisy digital patterns. *Pattern Recognition*, **29**, 1847-1862.

Cowan, N.J. and D.E. Koditschek (2000). Planar image-based visual servoing as a navigation problem. *Proc. IEEE Int. Conf. Robotics and Automation*, San Francisco, CA, 611-617.

Du, Q. and C.-I. Chang (2001). Linear constrained distance-based discriminant analysis for hyperspectral image classification, *Pattern Recognition* **34**, 361–373.

Fisher, R.A. (1950). The use of multiple measurements in taxonomic problems. In: *Contributions to Mathematical Statistics*, John Wiley, New York.

Gabor, D. (1946). Theory of communication. *Journal of the Institute of Electrical Engineers*, **93**, 429–549.

Glasius, R., A. Komoda and S. Gielen (1995). Neural network dynamics for path planning and obstacle avoidance, *Neural Networks*, **8**, 125-133.

Farrand, W.H. and J.C. Harsanyi (1997). Mapping the distribution of mine tailing in the Coeur d'Alene River Valley, Idaho, through the use of constrained energy minimization technique. *Remote Sensing Environment*, **59**, 64–76.

Harsanyi, J.C. and C.-I. Chang (1994). Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection. *IEEE Trans. Geoscience and Remote Sensing*, **32**, 779–785.

Khatib, O. (1986). Real time obstacle avoidance for manipulators and mobile robots. *Int. J. Robotics Research*, **5**, 90-98.

Malis, E. and F. Chaumette (2000). 2.5-D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement. *Int. J. Comput. Vision*, **37**, 79-97.

Mezouar, J. and F. Chaumette (2002). Path planning for robust image-based control. *IEEE Trans. on Robotics and Automation*. **18**, 534-549.

Palenichka, R. M., P. Zinterhof, and M. Volgin (2000). Detection of Local Objects in Images with Textured Background by Using Multi-Scale Relevance Function. *Proc. SPIE Int. Annual Symposium*, San Diego, USA, **4121**, pp. 158-170.

Payton, D.W. (1986). An architecture for reflexive autonomous vehicle control. *Proc. IEEE Int. Conf. Robotics and Automation*, 1838-1845.

Rowe, N.C. and R.F. Richbourg (1990). An efficient Snell's law method for optimal-path planning across multiple two-dimensional, irregular, homogeneous-cost regions, *Int. J. Robotics Res*. **9**, 48-66.

Seshadri, C. and A. Gosh (1993). Optimum path planning for robot manipulators amid static and dynamic obstacles. *IEEE Trans. on Systems, Man, and Cybernetics*, **SMC-23**, 576-584.

Singh, R., V. Cherkassky and N. Papanikolopoulos (2000). Self-organizing maps for thr skeletonization of sparse shapes, *IEEE Trans. Neural Networks*, **11**, 241-248.

Stellman, C.M., G.G. Hazel, F. Bucholtz and J.V. Michalowicz (2000). Real-time hyperspectral detection and cuing, *Optical Engineering*, **12**, 1928–1935.

Talluri, R. and Aggarval, J.K., (1996). Mobile robot self-location using model-image feature correspondence, *IEEE Trans. Robotics and Automation*, **RA-12**, 63-77.

Weiss, L.E., A.C. Sanderson and C.P. Neuman (1987). Dynamic sensor-based control of robots with visual feedback. *IEEE J. Robotics and Automation*, **3**, 404-417.

Wilson, W.J., C.C.W. Hulls and G.S. Bell (1996). Relative end-effector control using Cartesian position-based visual servoing. *IEEE Trans. Robotics and Automation*, **12**, 684-696.

Zaremba, M.B., M.A. Poissant and O.L. Hageniers (1987). Visual robot guidance for materials handling operations. *Proc. 2nd Int. Conf. Robotics and Factories of the Future*, San Diego, CA, 112-116.

Zaremba, M.B. and E. Porada (1998). A Reactive Neuromorphic Controller for Local Robot Navigation. *Journal of Intelligent & Robotic Systems*, **23**, 129-146.