Vergence Control System for Stereo Depth Recovery

Luis M. Jiménez*[a], Oscar Reinoso[a], Rafael Aracil[a], Fernando Torres[b], José M. Sebastian[c]
[a]DCEYT, University "Miguel Hernández", Elche, Alicante, Spain, 03202
[b]DFISTS, University of Alicante, Alicante, Spain, 03080
[c]DISAM, Polytechnical University of Madrid, Madrid, Spain, 28006

ABSTRACT

This paper describes a vergence control algorithm for a 3D stereo recovery system. This work has been developed within framework of the project ROBTET. This project has the purpose of designing a Teleoperated Robotic System for live power lines maintenance. The tasks involved suppose the automatic calculation of path for standard tasks, collision detection to avoid electrical shocks, force feedback and accurate visual data, and the generation of collision free real paths. To accomplish these tasks the system needs an exact model of the environment that is acquired through an active stereoscopic head. A cooperative algorithm using vergence and stereo correlation is shown. The proposed system is carried out through an algorithm based on the phase correlation, trying to keep the vergence on the interest object. The sharp vergence changes produced by the variation of the interest objects are controlled through an estimation of the depth distance generated by a stereo correspondence system. In some elements of the scene, those aligned with the epipolar plane, large errors in the depth estimation as well as in the phase correlation, are produced. To minimize these errors a laser lighting system is used to help fixation, assuring an adequate vergence and depth extraction .

Keywords: Vergence control, stereo vision, depth recovery

1. INTRODUCTION

The work presented in this paper has been developed within framework of the project ROBTET cooperatively with the companies IBERDROLA S.A. AND COBRA S.A. This project has the purpose of designing a Teleoperated Robotic System for live power lines maintenance. The tasks of repair in live lines are, at present, very costly and hazardous due to danger of electric shock. These labors are being replaced by teleoperated systems, especially in those countries where the safety regulations are strict. The efficient utilization of these systems requires the adoption of new control strategies that can be designated as *Intelligent Teleoperation*. These tasks suppose the automatic calculation of path for standard tasks, collision detection to avoid electrical shocks, force feedback and accurate visual data, and the generation of collision free real paths. To accomplish these tasks the system needs an exact model of the environment. Three-dimensional vision techniques permit to approach these problems [1,2,3].

1.1 General System Description

The system has been designed to develop maintenance and repairing tasks on overhead distribution lines of up to 49 kv (figure 1). Figure 2 shows its principal elements: telescopic boom, two hydraulic slave robots with seven degrees of freedom, two masters with force feedback, a hydraulic jib, and a stereo head with a laser illuminator. The cabin evolves all control systems and visual feedback for the operator. The architecture is designed to allow the operator to perform tasks in an optimal way and in the fewest time, and to achieve the highest degree in telepresence and therefore, to increase the performance of the system. The system consists of tree main modules:

- HIC, Human Interactive Computer
- TIC, Task Interactive Computer
- SPC, Sensor Processing Computer

* Correspondence: Email: luis.jimenez@umh.es; Telephone: +34 96 665 8620; Fax: +34 96 665 8798

The HIC is responsible for the interaction of the system with the operator, so that the control loop of the system is closed. The TIC receives commands from the HIC and takes the proper actions on the manipulators and other controlled devices. Finally, the SPC processes the information that the remote sensors provide. Communication between the different modules is performed via ethernet, whereas the exchange established between the TIC and SPC is made on a VME bus. A more detailed information about the architecture of the project may be found in[2,4,5].

Figure 1. ROBTET system repairing a distribution line

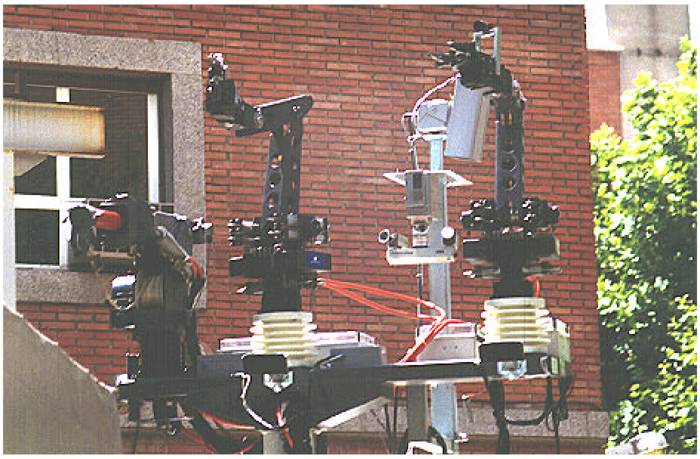

Figure 2. Description of the boom elements: two hydraulic slave robots with seven degrees of freedom, a hydraulic jib, panoramic camera, and a stereo head with a laser illuminator.

1.2 Structure of the Vision System

Robtet has two primary sensorial systems: force feedback and visual feedback. Visual feedback is accomplished not only using video from the scene but including additional information coming from a stereo vision system in order to automatically avoid collisions with high voltage elements, and to introduce augmented reality to the operator, superimposed over video streams. The visual system is made up of a stereo head, a laser scanner and the SPC.

The stereo head (figure 3) is a TRC (HelpMate Robotics Inc., at present) stereo mount equipped with motorized lenses. It is controlled by an 8-axis PMAC controlling card set on a VME bus. The head has four degrees of freedom, that is: pan, tilt and vergence on each camera. Besides this, both cameras include motorized lenses with three degrees of freedom: zoom, focus and iris. The laser scanner is equipped with a 810 nm,/1W diode, projecting a plane that can be recorded even at sunlight.

The SPC is made up of a Datacube Maxvideo 200 and Max860 video processing boards together with a Sun workstation.

Figure 3. TRC Stereo head mount.

1.3 Functionality of the vision System

The Sensor Processing Computer (SPC) updates the database with the model of the environment, needed for the HIC and TIC. The environments that have to be recognized are constituted by different elements: high tension power, their isolating elements and the cables (figure 1). Therefore, it is necessary to segment these objects of the scene, to recognize them and to calculate their location updating the database of the system. The cable is the most delicate element since this object can provoke serious damages to the equipment if a collision occurs. For this reason, the cable has to be localized continuously from the work scene.

The database is update by a feature oriented stereo algorithm based on a set of local and global geometric constrains that solves the correspondence problem between primitives of both images. The stereo algorithm recovers the three-dimensional information of the tower contours and matches this information with a model database in order to generate an accurate location of the tower and its elements such as insulators, fittings and cables. For some types of standard towers an additional recognition stage is introduced to improve his location accuracy allowing automatic maintenance tasks.

In the solution of the 3D-reconstruction problem through computer stereo vision, many authors assume that the optical axes of the cameras are parallel and ortogonal to the image sensor planes. This simplification reduces the complexity in both, the search of the stereo correspondence upon converting the epipolar restriction to a horizontal line, and the 3D reconstruction upon establishing a simple equation between the horizontal disparity and depth. In spite of this, this simplification is difficult to achieve in a real mechanical system. Generally the rectification technique is introduced for projecting the original images on a parallel axes system. This projection outlines the quantification problem of the images and it does not suppose a decrease of the computational cost since this is moved to the rectification process.

1.4 Why Verge?

Using binocular systems to extract the depth of the objects in a scene demands to know the geometry of projection axes according to the nature of the scene. This assumption is imposed by the need of establishing the interest objects in the scene and the evaluation of the overlapping necessary for the stereo correspondence. The vergence of the focal axis constitutes the basic tool to obtain good results in the stereo correspondence algorithms because the disparity of the interest objects is reduced. In this way, objects located from different distances can be used.

On the other hand, vergence gives also a measure of distance on the fixation point (intersection point of the focal axes on an object surface in the scene) through the geometrical triangulation obtained by the intersection from the focal axes. Therefore the vergence process arise as an adjustment mechanism of the sensor system geometry to the analyzed scene, providing coarse depth estimation. Also the control of convergence acts keeping the interest region of the three-dimensional analysis on the fixation point.

This relationship between vergence and depth permits cooperative algorithms with other depth recovery techniques as focus or stereoscopy. In different bibliograpy[3,6,7] this type of interrelationships are shown. Therefore, the vergence system permits the sensor system to be concentrated in interest areas to obtain a greater precision and confidence in the depth measure. The verging process presents in this way several advantages. One them is the math simplification due to the proximity of object points to the optical axis of both cameras, allowing simplifications in the geometric model of the stereo pair (weak perspective). In other way, it also facilitates the stereo correspondence, since the fixation point has zero disparity, and the neighboring areas will have small disparities. This reduces the number of candidates in the correspondence and therefore reduces the possibility of false correspondences, specially when repetitive exist patterns in the scene[8,9,10]. Other advantage is that it makes possible the utilization of an object-linked coordinate system[11] related with the calibration of the sensor system.

2. THE VERGENCE PROBLEM

The vergence problem of a binocular system consists of keeping the two coincident focal axes of an object called fixation point. Figures 3 and 4 shows the TRC stereo head used and the simplified model. The vergence angle is the angle constituted by both focal axes (figure 4). The vergence angle together to the base line, and the gaze angle configure geometrically the fixation point in the space. Therefore the distance from camera to the fixation point can be determined from the vergence angle. In this way, the vergence problem can be considered as the control of the distance from the object to the camera. Two problems can be approached: the first one is to select a fixation point (target) in the scene, and the second one is to estimate the vergence angle for this fixation point.

For the first problem it is necessary to analyze the two-dimensional scene, searching properties that allow to segment the object: edges, contours, textures, regions, etc. These properties will allow selecting regions of interest. In this sense, it can be considered as a general problem. The second problem is more specific to the vergence control, assuming the evaluation of the angle formed by the intersection between both focal axes over the fixation point belonging to the target. In this case, any depth information allows estimating the vergence angle, or, on the contrary, from the value of the vergence angle the depth distance can be estimated.

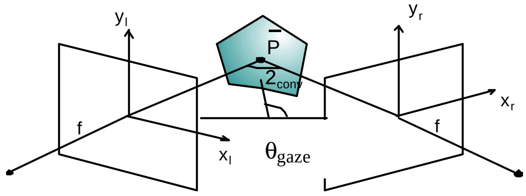

Figure 4 Stereo vergence system. P: fixation point θ_{verg}: vergence angle θ_{gaze}: gaze angle

The vergence of a binocular system can be formed either through independent procedures in each camera, or be coordinated taking into account not only the two-dimensional information, but also the three-dimensional information generated from vergence process. Depth information is usually coarse or it doesn't exist because the vergence process is a previous step to the extraction of depth estimation. Some systems make a feedback between both processes (vergence and depth estimation) allowing in successive iterations to improve the results of each method acting separately. In any case, the vergence process can be considered from two different points of view:

Static: Given the fixation point, the vergence angle keeping the region of interest on this point is evaluated. This procedure produces sharp steps in the vergence angle.

Dynamic: Keeping the vergence angle over a mobile fixation point. The variations in the location of the target generally produce small variations in the vergence angle.

A complete vergence system must consider both aspects: to obtain the vergence and to keep it on the target.

2.1 Strategies for Vergence Control

The vergence process can be split into two stages: the first one consists of selecting the target. This stage can be accomplished through a filtering to detect the contours of the surfaces, or by means of the selection of a camera as dominant. The second stage is based on changing the vergence angle in function of an error estimator until this is null. To calculate the vergence angle that reaches the fixation point, an estimation of the vergence error is necessary. The goal is to calculate this error from visual information and to execute a control action that make it null, that is to say, to achieve that the intersection of the two focal axes stay over the surface of the object. Different techniques can be found in the literature to generate an estimation of this vergence error[6,12,13,14].

A first approximation is to use a depth measure obtained from binocular disparity between the focal centers from the images, but this requires to solve the stereo problem previously, but this process is difficult and costly if the disparity is important. The advantage of this approach is that the relationship between the vergence angle and the disparity is simple and direct, and even it is not necessary to use absolute vergence angle values. The drawback is that the estimators of stereo disparity tend be optimized to obtain a good location and density characteristics but not to be robust and capable of being executed in real time as usually need a global optimization stage.

$$d^2(s,t) = \frac{\left[\sum_{x,y \in W} w_v \, I_L(x,y) \, I_R(x+s, y+t) \, dx \, dy \right]^2}{\left[\sum_{x,y \in W} w_v \, I_L^2(x,y) \, dx \, dy \right] \left[\sum_{x,y \in W} w_v \, I_R^2(x+s, y+t) \, dx \, dy \right]} \qquad (1)$$

Other approaches are based in registration techniques, which operate on the space of the image moving a window over the image until the underneath regions are similar. The cross correlation between both image windows is used as disparity measure. This type of measures are used also in stereo correspondence when the intensity variations are relatively continuous and smooth[6]. These measures have problems when the gradient of the surfaces is large, or when there is occlusion of the fixation point in each one of the images. The phenomenon of occlusion is bound with some limit of the disparity gradient[15,16]. This problem also arises in stereo vision. Cross correlation between selected windows in both images, are used as disparity measure. Equation 1 gives the correlation measure between a point of left image and a shifted point *(s,t)* in the right image using a small pondered neighborhood window using $\mathbf{w}_v$ as weights.

The Fourier transform results useful to accomplish error estimation due to the fact that allows to detect the spatial echo between the two images. Two convergent images will be similar and with a certain displacement that will be translated in an echo in the frequencies spectrum. The main problem of this technique is the computational cost of the Fourier Transform. One of the implementations based in this principle is the use of cepestral filter[17]. The *"cepestrum"* of a signal is the Fourier transform of the logarithm of the power spectrum of the signal (figure 5). The power spectrum is also the Fourier transform of the function of autocorrelation of the signal. This filtering was initially developed by Bogert[18] to analyze signals with echo. Coombs[13] presents an estimator derived from the cepestral filter characterized by a similar behavior but with a smaller computational cost. Considering two windows in left and right images, for a non zero disparity can be see as the original image plus a echo in $(w+d_h, d_v)$, where (d_h, d_v) are the horizontal and vertical disparities. The cepestral filter presents isolated peaks in $(\pm(w+d_h), \pm d_v)$.

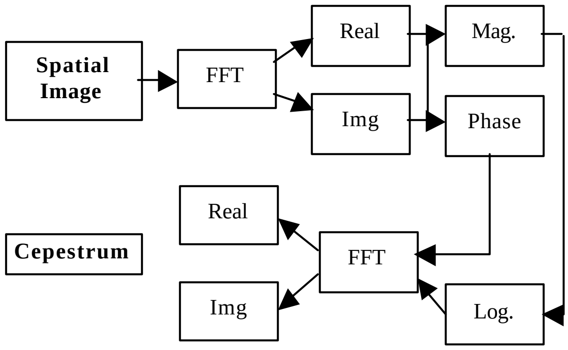

Figure 5 Schematic view of cepestral filter

This vergence estimator is related to the phase correlation. This filter can be expressed as the inverse Fourier transform of logarithm of the power spectrum (eq.2) . What gives the difference with autocorrelation is the non-linear logarithmic filtering that reduces the contribution of narrow band signals produced by periodic patterns and large and smooth objects that have poor correlation characteristics, and lets the broad band signal almost undisturbed (figure 6).

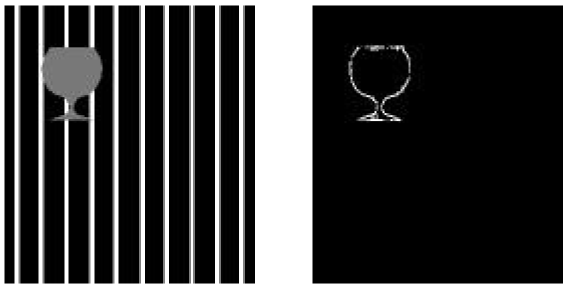
Figure 6 Bad correlation patterns (periodic columns) are filtered showing the properties of the cepestral filter

$$F^{-1}[\log |F|^2] = F^{-1}\left[\left| \frac{\sqrt{\log |F|^2}}{|F|} F \right|^2 \right] \qquad (2)$$

From this filter are derived the phase correlation filters, that have a smaller computational cost through smaller processing windows that the cepestral filter. These filters rely in the fact that a disparity between two signals can be measured as a phase shift in the Fourier Transform. The filter uses the phase information lost in the power spectrum. Kuglin and Hines[19] and Person[20] describe an implementation of such algorithm requiring a smaller processing window (h, w) than the cepestral filter (h, 2w).

2.2 Vergence Control

The proposed verge control is carried out through an algorithm based on a modified cepestral filter[12], trying to keep the vergence on the interest object minimizing the disparity distance estimated through the cepestral filter. The sharp vergence changes produced by the variation of the interest objects are controlled through a coarse depth map generated by a stereo correspondence system (sections 3,4), avoiding local maxima or background objets.

3. STEREO MATCHER

The goal of the stereo vision is to solve two problems [16,21]: the correspondence problem that consists in deciding for a point of the left image, which point in the right image matches (are projections of the same physical point). The second problem is the reconstruction problem: given two corresponding points in left and right images to obtain the 3D coordinates of the point in the space with respect to a world coordinate system. Within this second problem underlies the calibration problem of the visual system that determines with what approximation we know the positions of the focal centers of each optical system, and the position of the image sensors in the world coordinate system. The solution of the problem of the stereo vision must assume from the first moment that the positions of the points in the image and the calibration parameters of the visual system are known in a imperfect way. Therefore it can not be relied on rigid geometric models of the imaging and reconstruction processes.

3.1 The Correspondence Problem

The correspondence problem is indertiminate the way that has been defined: given a point m_1 in the image plane R_1 it can be the corresponding of any point m_2 in the image plane R_2. To solve this difficulty should be identified a set of restrictions that permit to reduce the number of potential correspondences. These restrictions can be classified in four basic types [16,22]:

Geometric restrictions imposed by the imaging system: the most important one is the epipolar constrain that restricts the corresponding point in the right image to lie in the epipolar line (intersection of epipolar plane and the image sensor)[23]. This restriction allows to reduce the problem from a two-dimensional search to one unidimensional. Other restrictions are the vergence of focal axes that impose the zero disparity in the center of the image[13].

Geometric restrictions from the visualized objects: a wide set of conditions of this type have been identified that allowing, in a general way, to solve the ambiguity of the stereo problem. The application of these conditions depends in a large extent on the type of analyzed scenes:

- Uniqueness constrain[21]: if there are not transparent surfaces in the scene the correspondence for a point in the image R_1 must be only one point in the image R_2.
- Surface continuity constrain[21]: the depth of the scene varies smoothly within the surfaces, therefore the disparities should vary also smoothly.
- Disparity gradient constrain[24]: it restricts the maximum disparity gradient allowed between matched primitives.
- Figural continuity constrain[15,25,26]: the disparity varies smoothly along the contours of the surfaces, but it can change abruptly across contours.
- Order constrain: the primitive correspondence must occur in the same left-to-right order along epipolar lines[26]. The order inversion can occur in certain environments (ie. transparent objects, narrow objects, etc.) but they are scarce in natural scenes.
- General position constrain[22]: certain events occur quite infrequently, in a statistical sense, so allowing to dismiss unlikely correspondences.

Physical restrictions: extracted from models of how the objects interact with the illumination. The reflected light by a surface is function of the type of surface, its orientation, the position of the light sources and the point of view[27].

Local restrictions of the primitive: these conditions determine the probability of correspondence for primitives selected of the scene. These restrictions can be the sign and direction of the intensity gradient, or constrains associated with high level primitives as lines or regions that holds a richer set of characteristics.

3.2 Matching techniques

The stereo recovery algorithms depend on the type of primitive candidates on the image. Medioni and Nevatia[25] distinguish between two types of stereo techniques according to the primitive used: area based, and feature based.

Area based techniques use correlation among brightness patterns in the local neighborhood of pixels form both images. They make use of simplifications on the search space through the epipolar constrain and coarse-to-fine strategies. Regions are segmented through some characteristic of texture, size, form, etc.

This method requires the presence of texture within each correlation window, failing in the presence of repetitive patterns in the scene. It tends to be confused by the presence of discontinuities, is sensitive to absolute intensity, contrast and lighting, and has problems with depth fields that change quickly. Similar correlation algorithms have been commented for vergence control.

Feature based techniques rely on the extraction of a set of primitives of each image to accomplish the correspondence between these features using the rich information that they store. The primitives can be pixels that represent special characteristics of the scene as edge pixels, corresponding to sharp intensity changes [Marr-82], or zero-crossings of the Laplacian operator. This type of features presents some stable characteristics to photometrics variations. As typical features are used the contrast of the gradient, direction, radiance[27].

An additional step is based on in grouping pixels in primitives of greater abstraction level. The simplest case is the straight line which is invariant to the perspective projection. This representation is compact, contrasted, distinguishable, precise, and dense. This representation allows to easily model polyhedral objects, but has problems with curved contours, for whose polynomial approximations of higher order should be used [28]. Even so they can present additional problems due to partial occlusions between primitives. The line and curve segments have geometric characteristics incorporated as length, direction, curvature, average contrast, etc. The primitives can also be located with subpixel accuracy[29] increasing the robustness and accuracy of the matching.

3.3 The stereo algorithm

The developed stereo algorithm uses straight lines as features for modeling the elements of the scene giving a polyhedral approximation perfectly valid for the electric tower. As was previously mentioned this type of features stores a rich information: direction, size, sign and magnitude of gradient and figural information. This feature imposes, in this way, the figural continuity constrain.

Let $S^L=\{s_1^L,\dots s_n^L\}$ and $S^R=\{s_1^R,\dots s_m^R\}$ be the two sets of segments of left and right images respectively. The feature segmentation is accomplished by means a recursive edge detector[29]. The edge points are linked into straight segments. A second process links near collinear segments with small gaps due to noise in the segmentation. This linking process is necessary to facilitate the correctness for the epipolar constrain.

The algorithm generates a graph with all the possible correspondences enforcing local constrains as epipolar, direction, size, average gradient magnitude and sign, assigning a certainty of matching based on a weighted measure of the previous features. This certainty measure is not a probability as far the general condition $\sum_{\forall j} p(i,j)=1$ can not be assured because some hypothesis are eliminated in the pruning process. Matches below a threshold

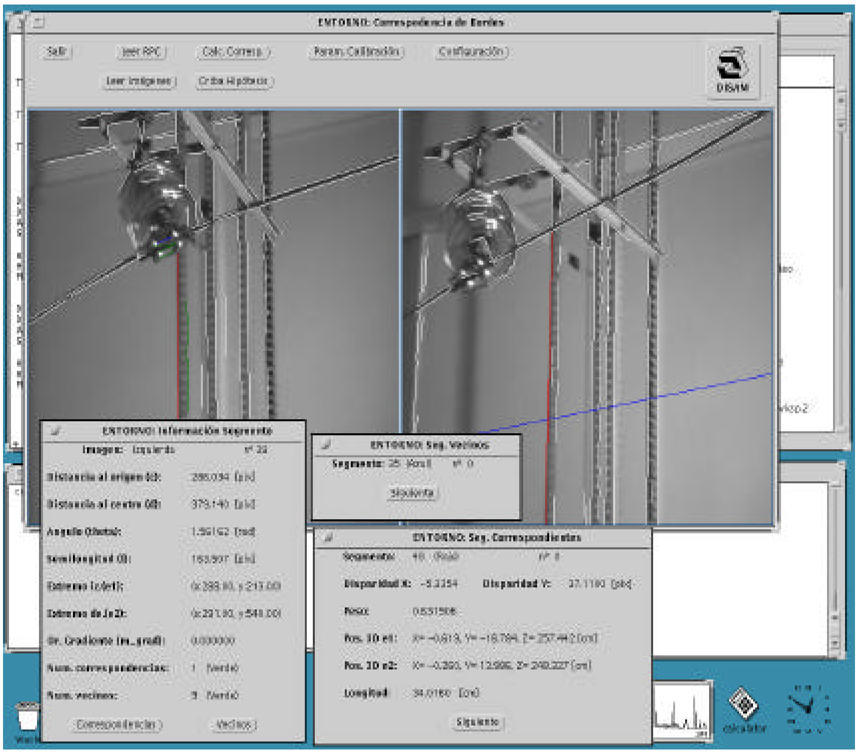

Figure 7 Stero matcher over the filtered segments of an electric tower.

are rejected. The epipolar constrain is considered over straight lines as an overlapping in the normal direction of epipolar lines of the middle points of segments.

Two graphs are generated with the correspondences for the left and right image edges. For each segment in the left image s_i^L a prediction graph is formed $H_{S_i^L} = \{s_{u_i}^R\}$ where u_i is a subset of S^R assigning a certainty measure $p^L(i,u)$. In the same way for each segment s_j^R, the graph $H_{S_j^R} = \{s_{v_j}^L\}$ is generated where v_j is a subset of S^L assigning a certainty measure $p^R(j,v)$. The intersection of the two graphs is obtained eliminating asymmetrical correspondences due to the epipolar constrain. The graph is pruned enforcing the uniqueness and continuity for neighboring segments (local verification). A neighborhood graph is generated with those segments in the proximity along the epipolar line $N_{S_i^L} = \{s_{p_i}^L\}$, $N_{S_i^R} = \{s_{q_i}^R\}$ where p_j, q_j are subsets of S^L and S^R respectively. For each of segments s_i^L in the left image, $s_j^R \in H_{S_i^L}$ (hypothetical correspondence), $s_k^L \in N_{S_i^L}$ (neighborhood of s_i^L), and $s_h^R \in \left(H_{S_k^L} \cap N_{S_j^R} \right)$ (neighborhood of the hypothetical correspondence), a compatibility function $c(i,j;k,h)$ is calculated. This compatibility function express the degree of support to the correspondence (s_i^L, s_j^R) from the neighboring possible correspondence (s_k^L, s_h^R). This function computes the distance, certainty, and size of the set $H_{S_k^L}$ (i.e. an alone correspondence set with higher certainty gives a bigger support)[3] giving a value in the interval *[-1,1]* (positive values represent excitatory support, and negative values represent inhibitory support).

A thresholded relaxation process implements the global continuity constrain, eliminating those correspondences with low support.

$$p^{t+1}(i,j) = p^t(i,j) + w \sum_k \sum_h c(i,j;k,h)\, p(k,h) \qquad (3)$$

A final global verification imposes the uniqueness and order along the epipolar line constrains together. A best mutual correspondence step between the two graphs gives the final correspondences. Figure 7 shows the stereo algorithm working over an electric tower at the laboratory.

4. COOPERATIVE VERGENCE –STEREO SYSTEM

The vergence of the focal axis constitutes the basic tool to obtain good results in the stereo correspondence algorithms because the disparity of the interest objects is reduced. In this way, objects located from different distances can be used. On the other hand, vergence gives also a measure of distance on the fixation point (intersection point of the focal axes on an object surface in the scene) through the geometrical triangulation obtained by the intersection from the focal axes. Therefore the vergence process arise as an adjustment mechanism of the sensor system geometry to the analyzed scene, providing coarse depth estimation[30].

Vergence facilitates the stereo correspondence, since the fixation point has zero disparity, and the neighboring areas will have small disparities. This reduces the number of candidates in the correspondence and therefore it reduces the possibility of false correspondences, especially when repetitive patterns exist in the scene. In the other hand the vergence system needs an approximate disparity map to start the fixation over an interest object avoiding local maxima or background objects. This interrelationship can be shown in figure 8.

The vergence produces a fixation point with a zero disparity, but more important is the depth estimation around this point (small disparities). This estimation is necessary as far the stereo matcher needs to fit the thresholds to impose the local and global constrains in order to reduce the size of the graph of correspondences. Furthermore, the calculation of the certainty measures $p(i,j)$ and compatibility functions $c(i,j;\, k,h)$ rely in the estimation of disparity for the object of interest. In this way the stereo matcher gives a precise depth map around the fixation point (object of interest), but can produce false correspondences far away of this point. This coarse depth map can direct the vergence system in order to fixate a new point.

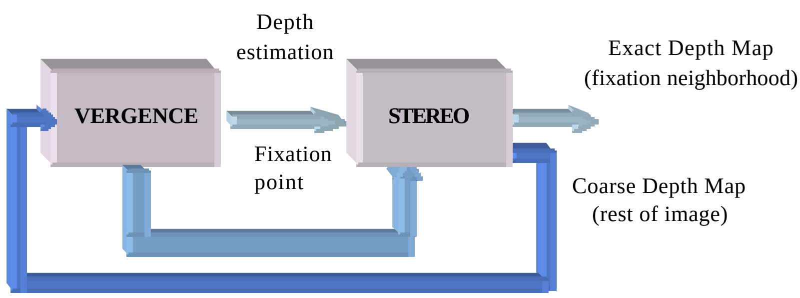

Figure 8 Cooperative algorithm using stereo and vergence subsystems.

5. CONCLUSIONS

The proposed system is made up of two subsystems: vergence control and stereo matcher. The vergence control is carried out through an algorithm based on a modified cepestral filter, trying to keep the vergence on the interest object. The sharp vergence changes produced by the variation of the interest objects are controlled through an estimation of the depth distance generated by a stereo correspondence system. The stereo matcher involves the generation of a correspondence graph that is prune using a relaxation process.

In some elements of the scene, those aligned with the epipolar plane, large errors in depth estimation are presented. These errors generate false correspondences and fixation points far from the interest objects of the scene. This kind of errors is presented in cables linked to the tower (narrow objects with an alignment close to the horizontal). A laser lighting system that simplifies the selection of the fixation point and the depth measure, assure an adequate vergence and depth extraction for this elements[5].

The overall system, including the manipulators and multimedia interface for teleoperated maintenance, may be considered as semiautomatic. In a first stage, a supervisory control in some degree has been implemented (for those standard task), although the final goal is to reach supervisory control next to fully automatic control. Some advantages of this way of operating, in contrast to a manual operation, are higher level of protection, safety for the worker and quickness.

ACKNOWLEDGMENTS

The work presented in this paper has been supported by electric utility IBERDROLA, S.A. under project PIE No. 132.198. The research work has been developed at University of Elche using the facilities of the Systems Engineering and Automation Department of the Polytechnic University of Madrid.

REFERENCES

1. Aracil, R. et. al. “3D Dynamic Non Structured Environment Modelling by Means Artificial Vision, for Robot Trajectory Planning”, Technical Report, 1995.
2. Aracil, R. et. al.. “Teleoperated System for Live Power Lines Maintenance”, Proc. of SPIE-The International Society for Optical Engineering. (International Symposium,on Intelligent Systems and Advanced Manufacturing.) Vol.2590. pp. 40-49, Octubre 1995.
3. Jiménez, L.M. Aracil, R., Sebastian J.M., Peñín L.F, Ferre M., Torres F. “Modelado de Entornos 3D para Planificación de Trayectorias de Robots Mediante Visión Estereoscópica y Control Activo de Parámetros Ópticos”, Proc. XVI Jornadas de Automática (Grupo de TrabajoVisión Artificial), San Sebastián, Spain, Septiembre 1995.
4. Aracil, et al..,“ROBTET: Robot for Live-Line Maintenance”. ICOLIM 1996.

5. Aracil, R. et. al. “Obstacle Avoidance for Teleoperated Robots for Live Power Lines Maintenance Using Artifical Vision”, Proc. Int. IFAC Conf. on Autonomous Vehicles, Sevilla, Spain, 1997
6. Ahuja, N., Abbot, A.L. “Active Stereo: Integrating Disparity, Vergence, Focus, Aperture, and Calibration for Surface Estimation”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp.1007-1029, October 1993.
7. Krotkov, E.P., *Active Computer Vision by Cooperative Focus And Stereo*, Springer-Verlag, New-York, 1989.
8. Nishihara, H.K. “Practical Rel-Time Imaging Stereo Matcher”, Opt. Eng., vol. 23, pp.536-545, Sept. 1984.
9. Nishihara, H.K. and Poggio, T. “Stereo Vision For Robotics”, R. Paul and M. Brady, editors, Robotics Research, The First International Symposium, pp.489-505, MIT Press, 1984.
10. Olson, T.J, “Stereopsis for Fixating Systems”, Proc. IEEE Int. Conf. on Systems, Man and Cybernetics, Charlottesville, Virginia, USA. October, 1991
11. Ballard, D.H., “Reference Frames for Animate Vision”, Proc. Int. Joint Conference on Artificial Intelligence, 1989.
12. Aracil, R., Jiménez, L.M., Reinoso, O., Fernández, C. “Estudio de Algoritmos de Cálculo de Disparidad para el Control de Convergencia en Sistemas Estereoscópicos”, XIX Jornadas de Automática, Madrid, Spain, 1998.
13. Coombs, J. D. “Real-Time Gaze Holding in Binocular Robot Vision”, Phd Thesis. University of Rochester, 1992.
14. Olson, T.J,, Coombs, D.J. “Real-Time Vergence Control For Binocular Robots”, Int. Journal of computer Vision, vol.7, no.1, pp.67-89, November, 1991
15. Mayhew, J.E.W. and Frisby, J.P. “Psychophysical and Computation Studies Towars A Theory Of Human Stereopsis”, Artificial Intelligence, 17, pp.349-386, 1981.
16. Faugeras, O. *Three-Dimensional Computer Vision*, The MIT Press, 1993
17. Yeshrun Y., Schwartz E. “Cepestral Filtering on a Columnar Image Architecture: a Fast Algorithm For Binocular Stereo Segmentation”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. PAMI-11, no.7, pp.759-767, July 1989.
18. Bogert B.P., M.J.R. Healy, J.W. Tukey “The Frequency Analysis of Time Series For Echoes: Cepestrum, Pseudo-Autocovariance, Cross-Cepestrum, and Shape Cracking”, Proc. Symp. Time Series Analysis, pp. 209-243, New York, 1963.
19. Kuglin, C.D., Hines, D.C., “The Phase Correlation Image Alignment Method”, Proc. IEEE Int. Conference on Cybernetics and Society, pp.163-165, 1975.
20. Person J.J, Hines D.C, Golosman S. Kugli C.D., “Video-Rate Image Correlation Processor”, SPIE v.119, Applications of Digital Image Processing, pp.197-205, San Diego, 1977.
21. Marr, D. , and Poggio, T. “A Computational Theory of Human Stereo Vision”, Proc. Royal Society London, Vol B 204, pp 301-328, 1979.
22. Marapane, S. , Trivedi, M. “Multi-Primitive Hierarchical (MPH) Stereo Analysis”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.PAMI-16, no.3 , pp.227-240, March 1994.
23. Marr, D. *VISION*, Ed. Freeman, 1982
24. Julesz, B. “Binocular Depth Perception of Computer-Generated Patterns”, Bell Syst. Tech. J., vol. 39, pp.1125-1162, 1960.
25. Medioni, G., Nevatia, R. “Segment-Based Stereo Matching”, Computer Vision, Graphics, and Image Processing, vol.31, no.1, pp.2-18, July 1985
26. Otha, Y., Kanade, T. “Stereo By Intra- and Inter-Scanline Search Using Dynamic Programming”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. PAMI-7, no.3, pp.139-154, March 1985.
27. Horn, B.K.P, *Robot Vision*, MIT Press, Cambridge, Mass., 1986.
28. Sherman, D., Peleg, S. “Stereo By Incremental Matching Of Contours”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. PAMI-12, no.11, Nov. 1990.
29. Reinoso, O. Sebastian, J.M., Aracil, R., “Image Restoration with subpixel accuracy in digital images”, Image Processing and Comunications. Vol.3, no.1,2, pp. 3-13. 1997
30. Bajcsy, R, “Active Perception”, Proc. of the IEEE, vol. 76, nº 8, pp.996-1005, August 1988.