Journal of Intelligent & Fuzzy Systems 21 (2010) 385–394
DOI:10.3233/IFS-2010-0459
IOS Press

An adaptive fuzzy system for the control of the vergence angle on a robotic head

Nikolaos Kyriakoulis[a,*], Antonios Gasteratos[a] and Spyridon G. Mouroutsos[b]
[a]*Laboratory of Robotics and Automation, Department of Production and Management Engineering, Democritus University of Thrace, University Campus Kimmeria, Xanthi, Greece*
[b]*Laboratory of Special Mechanics, Department of Electrical and Computer Engineering, Democritus University of Thrace, University Campus Kimmeria, Xanthi, Greece*

Abstract. An important issue in realizing robots with stereo vision is the efficient control of the vergence angle. In an active robotic vision system the vergence angle along with the pan and tilt ones determines uniquely the fixation point in the 3D space. The vergence control involves the adjustment of the angle between the two cameras' axes towards the fixation point and, therefore, it enables the robot to perceive depth and to compute obstacle maps. Vergence movement is directly related to the binocular fusion. Additionally, the decision for convergence or divergence is extracted either by motion affine models or by mathematical ones. In this paper, a new method for extracting the cameras' movement direction is presented. The movement decision is performed by an adaptive fuzzy control system, the inputs of which are the zero-mean normalized cross correlation (ZNCC) and the depth estimations at each time step. The proposed system is assessed on a 4 d.o.f. robotic head, yet it can be utilized in any active binocular system, since it is computationally inexpensive and it is independent to a priori camera calibration.

Keywords: Vergence angle, computer vision, mechatronics, fuzzy logic

1. Introduction

Among the contemporary mechatronics systems, robots are of the most complex in terms of structure and control. Also, they claim for an extensive sensorial system, which enables them with enhanced perception. In such a system, vision holds a dominant role and that's why it was utilized in robotic systems since their infancy [10]. Of special interest is the use of the binocular vision, that enables the biological systems to perceive depth and, thus, allows them a deeper knowledge of the scene. Vergence control supports the above capacity by turning the eyes, so that they are directed both towards the same point in the 3D space. Humans possess the medial and lateral recti muscles to rotate their globes so that an image pair is projected onto their foveae [2]. In a robot system the cameras play the role of the eyes, the servo motors this of the muscles, and the optic sensors correspond to the foveae. This principle was realized in [7], where the same landmark is identified in both images to enable the fixation of the active cameras of a human-like set-up to that point. A stereoscopic vision system controls the vergence angle by initially placing the target of interest at the image center as shown in Fig. 1. A plethora of different techniques have been applied for calculating and controlling the vergence angle, which can be classified into two main categories: disparity based and correlation based ones. The disparity (stereo matching) techniques deal with the problem of stereo correspondence. The correlation based methods utilize a measure such as normalized cross correlation (NCC), zero-mean normalized cross correlation (ZNCC) or sum of absolute differences (SAD), to max-

*Corresponding author. Nikolaos Kyriakoulis, Laboratory of Robotics and Automation, Department of Production and Management Engineering, Democritus University of Thrace, University Campus Kimmeria, GR-671 00 Xanthi, Greece. Tel.: +30 2541079359; Fax: +30 2541079343; E-mail: nkyriako@pme.duth.gr.

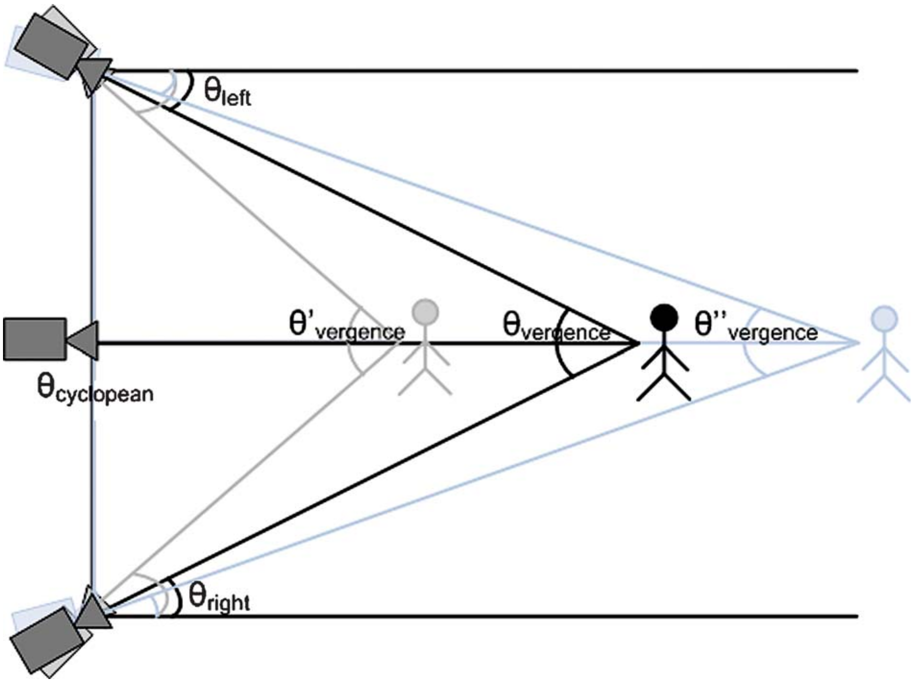

Fig. 1. The vergence angles in a stereoscopic vision system.

imize the stereo fusion index [5, 13, 14]. Furthermore the techniques for the control of the vergence angle are distinguished according to the images that are utilized, i.e. disparity estimation with log-polar images [4, 6, 12, 18] provides better results than with Cartesian ones. In [1] the cortically magnified visual cortex is used to match the entire image. Vergence control is based on the human's ability to rotate their eyes, a neuron-based procedure, mimicked in robotics in [16], where responses of energy neurons are used for vergence control and for disparity estimation. Furthermore, energy neurons have also been applied for the simultaneous gaze and vergence control [17]. In similar applications, the correct vergence angle can be reached by saccading towards the target in the periphery map and subsequently by the cameras to be guided through the visual cortex [11]. Moreover, a dynamic vergence control for direct estimation of the disparity can be achieved with a robust binocular fusion computation [4]. As the visual data is the dominant element for the vergence control, the followed strategy that is interpreted and processed leads to different results. A vergence control strategy based on Hering's law has been adopted for the motor movements by interpreting the visual data [15], with remarkable results.

In this paper, a new method is described for the control of the vergence angle. All the previous work has focused mostly on the optimal estimation of the 3D features between the stereo image pairs. We applied the proposed technique on a 4 d.o.f. robotic head [9]. The use of fuzzy logic on this scheme derives from the way the human eyes move towards a target, i.e. they converge when it approaches, and diverge when it draws away. The presented adaptive fuzzy system, is the optimization of the one presented in [11]. It has two inputs, the similarity and the depth, which determine whether the two cameras are aiming at the same target or not. If the two images are very similar it is believed that both cameras are focusing on the same target or scenery. The depth input, was selected due to the fact that it is easily manipulated, it is related with the cameras' angles and, finally, it is related to the disparity of the current scene. The similarity method chosen for the vergence control is the zero-mean normalized cross correlation as it has exhibited smoother results and does not require camera calibration [8]. The current vergence angle is read directly on the encoders and the depth is calculated accordingly. The output of the adaptive fuzzy system is used as feedback for the vergence control.

2. Mathematical formulation

2.1. Similarity measure

The similarity criterion is essential in vergence control. It should be able to provide accurate measures in a fast manner. For that purpose, a fusion index technique has been adopted as it exploits the fact that if a target is correctly verged, the stereo images are very similar. The goal of the vergence control system is to maximize the similarity. The index of the binocular fusion is computed using the zero-mean normalized cross correlation

$$\mathrm{ZNCC} = 1 - \frac{\sum_{(u,v)\in W} (I_r(u,v) - \overline{I_r})(I_l(u,v) - \overline{I_l})}{\sqrt{\sum_{(u,v)\in W} (I_r(u,v) - \overline{I_r})^2 \sum_{(u,v)\in W} (I_l(u,v) - \overline{I_l})^2}} \tag{1}$$

where I_r and I_l are the right and left images, respectively. $\overline{I_r}$ and $\overline{I_l}$ represent the mean values of the right and left images, respectively. The variables u and v are the image coordinates. The region in which the ZNCC is applied is the whole image: $(u, v) \in W$. ZNCC is invariant to the illumination changes and its range is normalized into [0 1]. ZNCC was selected as it has the smoother performance, even in the most difficult scenes. This measure is fast, accurate and robust in environmental alterations, as it compasses the correct vergence angle even with extremely low-resolution images, without making any topological rearrangement, such as log-polar mapping.

2.2. Depth estimation

The depth is calculated via triangulation of the 3D feature correspondence on the image pair. An appropriate vergence angle leads to high ZNCC values, i.e. either the stereo cameras are directed to the infinity or to the same object. The capacity of knowing the cameras' angles provides the better understanding of a captured image pair. Moreover, when the camera angles are known, the vergence angle is calculated simply geometrically (see Fig. 1). The mathematical equations for a symmetrical vergence system are:

$$Z = \frac{b}{2\tan\theta} \tag{2}$$

and

$$\theta_{\mathrm{vergence}} = \theta_{\mathrm{right}} + \theta_{\mathrm{left}} = 2\theta \tag{3}$$

where Z is the depth of the scene, and b is the baseline of the stereo head. $\theta_{\mathrm{vergence}}$ is the vergence angle, while θ_{right} and θ_{left} are the angle of one of the right and the left camera, respectively. The above formulae are only valid for a parallel camera setup, i.e. zero cyclopean angle. In order to ensure the symmetric operation of the cameras, the two respective independent degrees of freedom are controlled by the same absolute motion commands. The only difference in the movement is the sign, which results into symmetrical verge or diverge.

3. Fuzzy vergence system

We applied fuzzy logic to estimate the vergence angle in a simple and effective manner. The necessary data in our approach are the similarity of the two images and the depth of the scene. These two parameters are sufficient for computing the correct vergence angle, as the similarity determines whether the two cameras are aiming at the same target or not. The depth of the scene is easily extracted from the known angles, read from the encoders. From the formulae (2) and (3) it is obvious that the depth can be interpreted by means of the cameras' angles, which are the dominant features for the vergence control. In Fig. 2 the block diagram of the proposed vergence control is depicted. The control of the camera angles is performed by a classic scheme. The proposed fuzzy system calculates the feedback control signal. The images from the left and right cameras are used for the computation of the ZNCC and the readings of the encoders of the actuators of the camera angles are used for the computation of the depth (Z). Both ZNCC and Z are then fed into the fuzzy system. The output of the system is the amplitude and the sign of the angle applied to the actuators. The disturbance in the block diagram corresponds to potential noise including light reflections resulting to a global brightness dissimilarity between the two images of the stereo pair. However, the system is still capable to provide correct vergence angles since the ZNCC input utilizes the mean value of the respective images. As a result, the difference between the pixel's intensity and the image's mean value

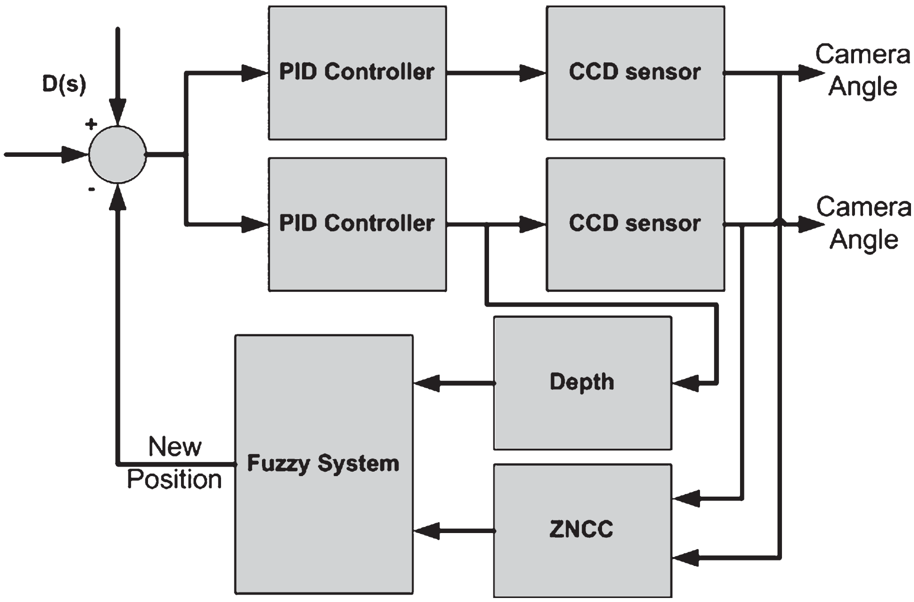

Fig. 2. Block diagram of the proposed automatic vergence system.

$(I_r(u, v) - \overline{I_r})$ is the same under any brightness level. On the other hand, the system is sensitive to cases where the disturbance is caused by an occlusion to one of the cameras, since the main algorithm prerequisite is that the two cameras have to operate simultaneously and to observe the same environment.

The design of a fuzzy system is highly depended on the application and the experience of the designer. Therefore, building such a system, relies on previous experience and exhaustive experimentation. This approach was followed also for the proposed system, where five Gaussian membership functions (MFs) are used for both inputs and five trapezoid MFs for the output, as they found to be more appropriate for the desired task. Apart from the type of the MFs, important role to the fuzzy system plays the possible adjustment methods, such as the defuzzification and the aggregation one. In the proposed system the aggregation method was set to sum. Other aggregation methods such as the maximum, the minimum, and the probabilistic OR (probor), were tested, but the sum provided a smoother output value in our tests. The defuzzification method was set to centroid, as it covers the output range more efficiently. The centroid defuzzification method provides the center value between the interacted rules, so in some cases the uttermost values of the output are not fired, reducing so the efficient range. We also assessed other defuzzification methods, namely: bisector, middle of maximum, largest of maximum, and smallest of maximum. The results acquired by the centroid exceeded the ones of the other methods and, thus, the centroid one was selected.

Key features to the design of a fuzzy system are the MFs and the rule base. In the fuzzy vergence system, apart from the first input, all MFs are normally distributed to their range. The membership functions are displayed in Fig. 3 and the rules interaction is shown in Table 1. The first input (Fig. 3a) is the similarity measure (ZNCC). The odd distribution of the MFs of the first input aims at distinguishing between the cases with a very high and a very low ZNCC. As the similarity between the two images is the criterion for the correct vergence angle, this kind of distribution highlights any alterations. Furthermore, as the ZNCC values around the correct vergence angle have very small differences, it is very difficult even for humans to identify the best angle. Thus, the arrangement of the MFs with small deviations and almost with no overlaps, enables the system to identify more efficiently the highest similarity. In the cases of high ZNCC, the system remains to the current state or makes small movements around the highest similarity point. On the other hand, in the cases of low ZNCC the system covers big distances until it reaches a point with higher similarity. The reason that three MFs are close to each other at the middle point, is that the ZNCC in most cases varies dramatically in that range. By having a denser distribution of the MFs in that range, instead of distributing them normally, we have

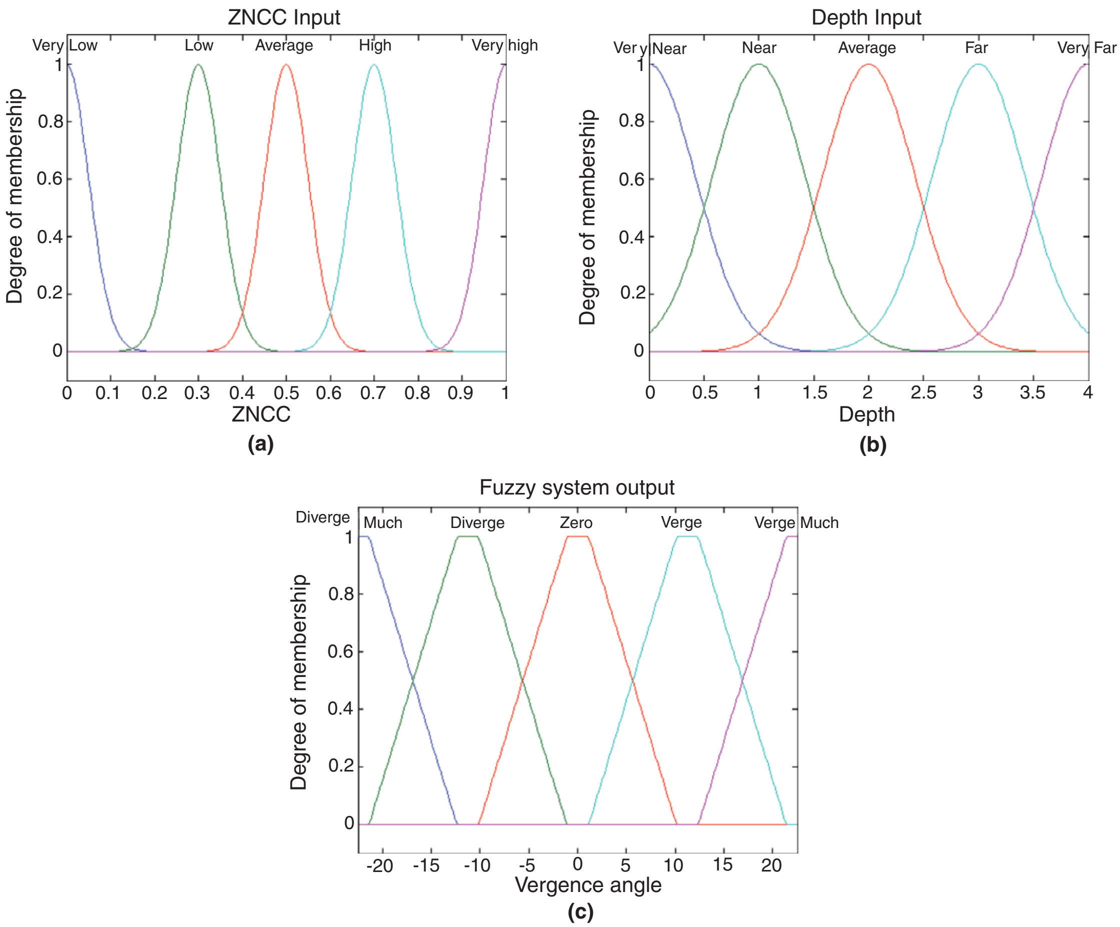

Fig. 3. Membership functions of the fuzzy vergence system for (a) input 1, (b) input 2 and (c) output.

Table 1
Rule base for the fuzzy kalman system

Depth	ZNCC				
	Very low	Low	Average	High	Very high
Very near	DM	DM	D	D	Z
Near	DM	D	D	D	Z
Average	DM	D	D	Z	Z
Far	VM	V	V	V	Z
Very far	VM	VM	V	V	Z

*DM = Diverge much, D = Diverge, Z = Zero, V = Verge, VM = Verge much.

increased more the system's efficiency. The second input (Fig. 3b) is the depth and is set to be [0 4] meters, due to the nature of the application in hand. In cases where the environment is uniform, the range can easily be adapted to larger values. The output range was set to [–22.5 22.5] degrees. The minus sign implies the vergence while the plus sign the divergence movement.

The fuzzy system is adaptive, as in every step the ranges of both inputs and the output are altered, depending on the values of the variables. The range is $[x - d\ x + d]$ for all variables, where x is the value of the variable at the current time step, and d is the range coefficient. The d values are 0.1 for the ZNCC input, 1.2 m for the depth (Z) and 10 degrees for the output. The d values are derived from the nature of the application in hand. For the ZNCC input the main criterion is whether the two cameras are fixated to the same target or not. Extensive experiments showed that the 1/10 of the ZNCC range is sufficient for the system to decide efficiently whether to verge or to diverge. Moreover, a small d value of the ZNCC enables the system to respond accurately in cases where the ZNCC is high, i.e. close to the desired vergence angle. For the depth input d is set to 1.2 m in order for the system to seek the correct vergence angle in a wider space. In cases where the d receives smaller values than 1.2 m, the system is restricted to search in fewer positions for the correct vergence angle resulting to much more iterations than with a higher d value. Finally, the output's range is altered by $d = 10$ in order to keep the number of iterations as low as possible. It was found that when the output range is reduced below 10 degrees, the d values for the ZNCC and the Z have to be significantly reduced in order for the system to respond accurately, yet increasing the computation time and the iterations needed. For example concerning the ZNCC input, if its value is 0.82 the adaptive range will not be [0 1] but [0.72 0.92] instead. In the cases where the ZNCC values exceed the maximum predefined range [0 1], there is a logical check, and the range limit is not violated, i.e. if the ZNCC is 0.99 the adaptive range will be [0.89 1]. The aforementioned restriction is applied to the inputs and the output. The adaptive range allows us to create a generic fuzzy system which is mostly dependent on the rule base and not on the variable values. The fuzzy rule base is shown in Table 1. The abbreviations DM, D, Z, V, and VM, are for Diverge Much, Diverge, Zero, Verge (Converge), and Verge Much, respectively.

4. Experimental results

The experimental setup is a 4 d.o.f. robotic head shown in Figs 4 and 5. For its movements (pan, tilt and vergence), four respective harmonic drive actuation mechanisms are used controlled by four flexible and compact DSP based PID controllers. The actuators' controllers are four respective high-performance intelligent drives, combining motion controller and PLC functionality in a single compact unit. A complete set of high level instructions permit to define and perform complex motion sequences from the host PC.

Two distributed computers are connected with the head. One dedicated to its control and the other to run the image processing to extract the control signals [3]. This approach was chosen due to the high data volume for imaging, which unless a dedicated system is devoted, it might late the control of the head. Thus, the system functions are separated into information processing and motion control. Image processing possess a high computational burden and it is recourse demanding. On the other hand, motion control requires a real-time operation system. This demand for high mul-

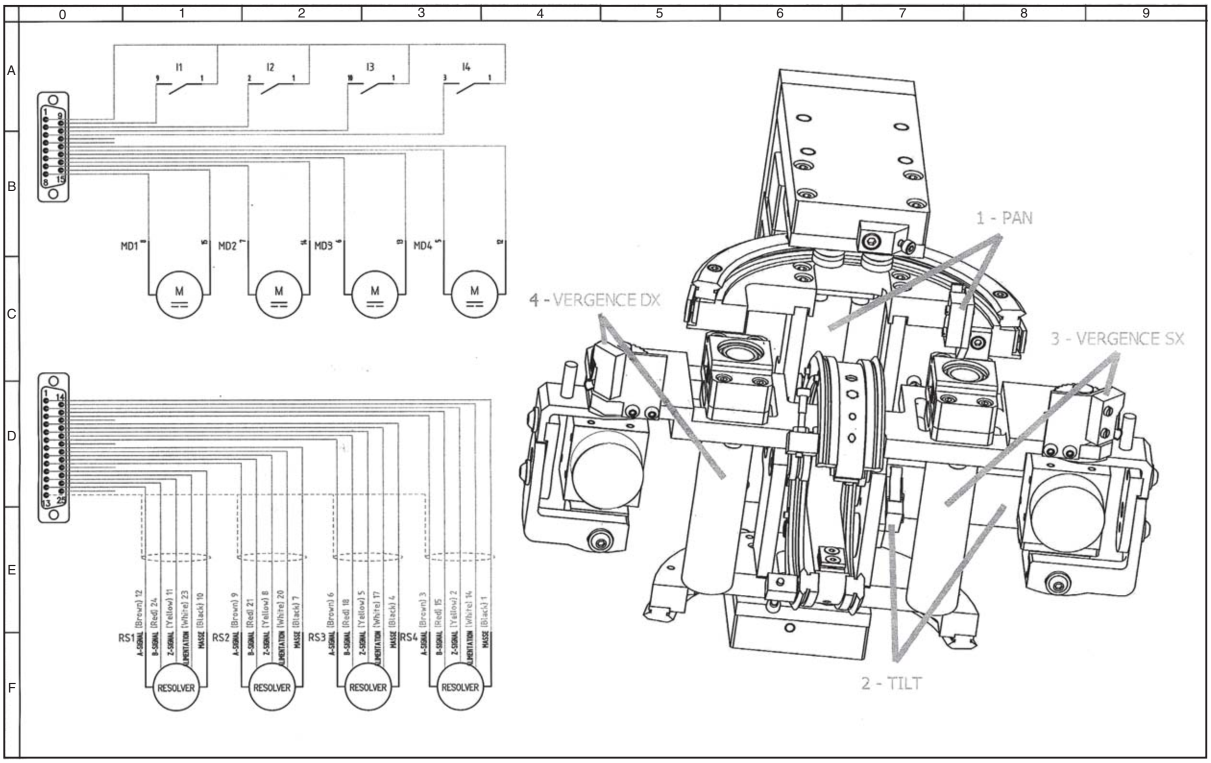

Fig. 4. The robotic head used for vergence control.

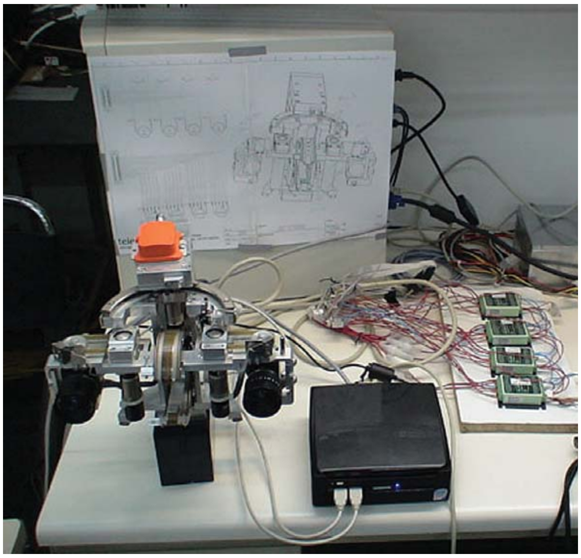

Fig. 5. The experimental setup: On the left-hand side the 4-d.o.f. robotic head; in the middle the dedicated PC for the image processing; on the right-hand side the four controllers for each of the d.o.f.; finally in the background the control PC.

timedia performance and real-time motion control led us to use a computer with Windows operating system for the image processing and one with RT-Linux for the control tasks. The computers are connected to each other with a high speed network in order to achieve real-time operation.

In order to evaluate the performance of the control scenario presented in the previous Section, several tests were executed. The tests include different vergence experiments. In every time step the ZNCC along with the depth were estimated. Extracting the vergence angle and the depth from the encoders, a new angle was calculated by our adaptive fuzzy system, which was read by the PID controllers and finally it was reached by the servo motors. The new vergence angle was read from the motors' encoders and the new ZNCC was calculated. The process continues perpetually as described above. The processing time of the fuzzy system in every time step is about 1.5 msec for a Pentium 4 PC, running at 1.8 GHz.

In the first experiment a human was put in a fixed position during the whole process. The distance from the robotic head was set randomly and as close to the cyclopean axis as possible. In the second experiment, some objects were put in front of the robotic head to a randomly selected position near to the cyclopean axis. In the last experiment a more difficult scene was selected, i.e. there were not a specific target for the stereo vision system and the system should aim to the infinity without any obstacle intruding directly into the field of view. In all experiments the system started from a random symmetrical vergence configuration and operated until the correct vergence one to be achieved. There was a restriction to the servo motors for not diverging further than the zero angle point. In case of violating the aforementioned restriction the system was programmed to return to its initial position. In the presented experiments the images are non rectified ones. Notwithstanding, the correct vergence angle was approximated after 4–5 iterations of the system. The fused images after several consequent iterations of the fuzzy vergence system are illustrated in Fig. 6. The sequences represent the vergence angles reached before the correct ones (with the highest similarity), which are illustrated for each row at the last subfigure.

For further evaluating and examining the adaptive fuzzy vergence system, every measurement was stored. Thus, all the experiments are evaluated not only by their visual results, but also by the record of the actuators' encoders, concerning the maximization of the similarity and the time intervals needed. The acquired results of our system were compared with a known and used technique for controlling the vergence angle presented in [4]. The decision of verging or diverging is taken by examining the relation between the current and the previous correlation index values. This approach is based on minimizing the given correlation function; the movement direction is determined by the sign of the subtraction between the previous correlation values and the current ones. If this sign is positive the chosen direction is maintained otherwise it is reversed. The comparative results are shown in Fig. 7. The abscissa expresses the time in seconds, while the ordinate respective ZNCC values. The blue lines correspond to the fuzzy system's steady state output [11], while the black ones to the steady state output of the technique presented in [4]. The red lines represent the computed ZNCC values at each time step before any correction has taken place. Finally, the green lines correspond to the output of the proposed adaptive system. It is obvious that adaptive range has increased the efficiency of the fuzzy system presented in [11]. For demonstration reasons, the go-between ZNCC values are not illustrated, but only the final ones. Thus, for a vergence angle with an initial respective ZNCC, the diagrams in Fig. 7 represent the final state achieved.

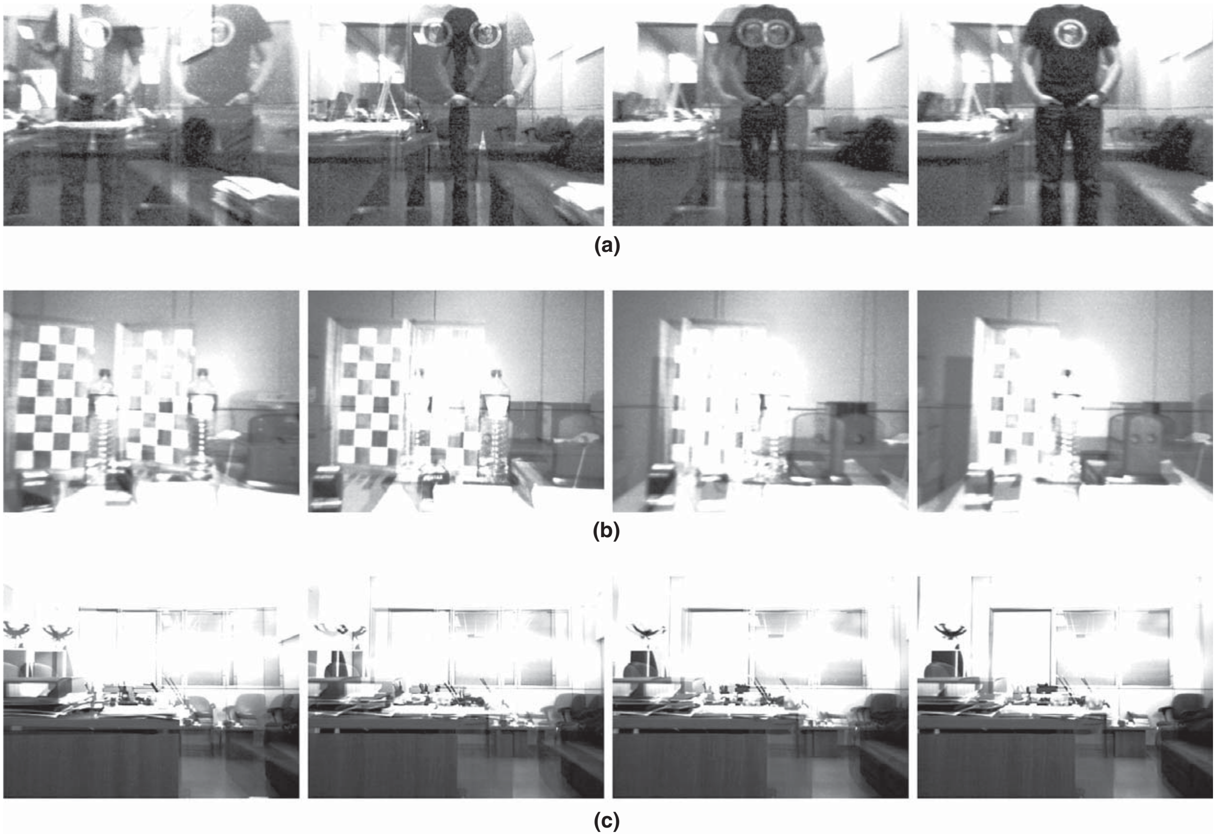

Fig. 6. Fused image sequences during the convergence of the system for several cases: (a) a human standing at a fixed position; (b) some objects were put in front of the robotic head and (c) there is absent of a specific target for the stereo vision system. From left to right, is displayed the index of the consequent vergence angles.

From Table 1, one can see that when a high ZNCC value is achieved, the system stands still. This is the reason of not reaching the highest possible similarity value. For example, in the Infinity experiment Fig. 7c and f, which had cluttered environment (Fig. 6c), and suffered from noise and occluded scenery, although the ZNCC values are quite high, the system responded accurately (the ZNCC is very high, about 0.92).

The demonstrated results of all the experiments can be further improved by filtering the noise and by storing all history positions. In any new measurement a check is made, with the form of a third input, whether the current or the old angle has a higher similarity. Thus, the system does not reach a position with low similarity twice, but instead it moves to a higher similarity angle, reducing the iterations and the time intervals needed.

5. Conclusion

A new vergence control, which uses an adaptive fuzzy system, is proposed. The adaptive fuzzy system decides whether to converge or diverge, depending on the given similarity and depth. The results show that the vergence direction was achieved during the whole process, while the amplitude of the exported vergence angle was the one which had provided the maximum similarity. The fuzzy system responded efficiently to all the experiments, considering the noise involved to the image processing. Its fast response renders it appropriate for real-time operation. To conclude fuzzy systems are a promising vergence control method due to their low response times and the high efficiency. Moreover, fuzzy logic incorporates the human way of thinking with simple “if… then” rules and, therefore, it provides better understanding of the vergence control procedure.

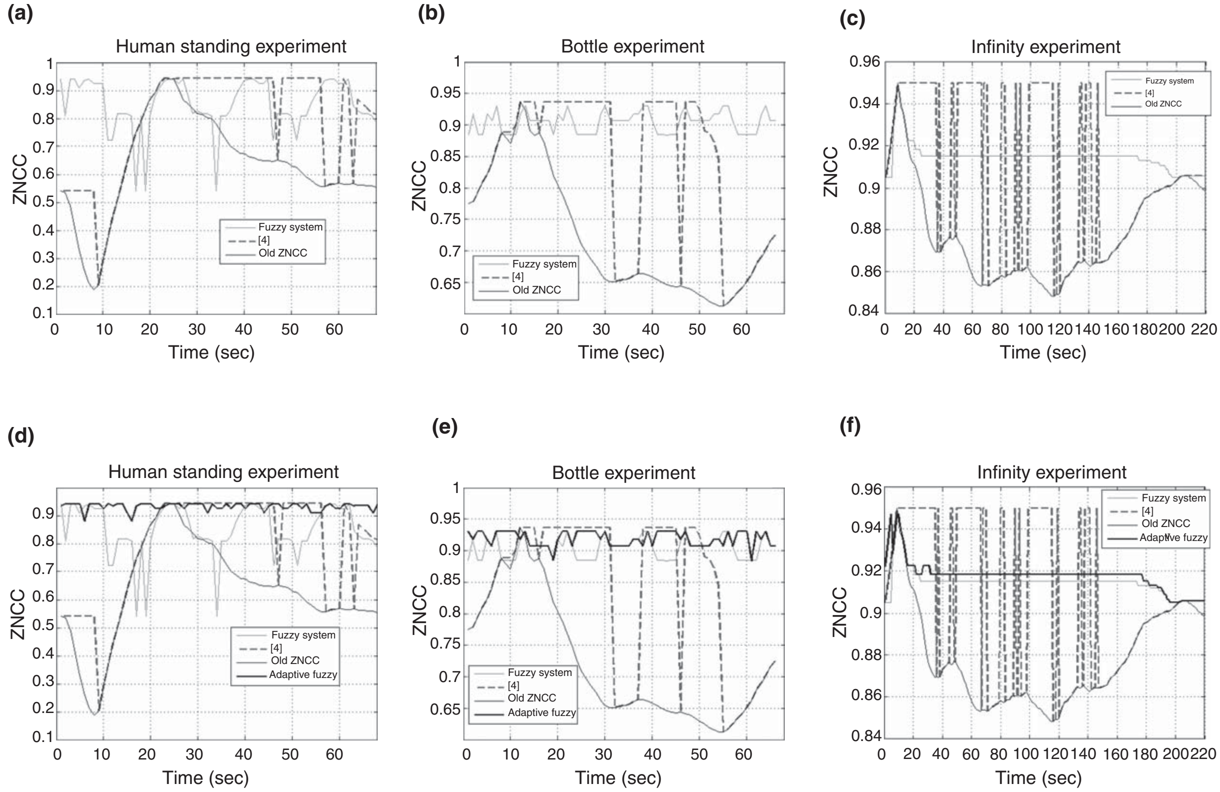

Fig. 7. The new ZNCC values acquired from the fuzzy system (gray thin lines) in relation with the old (black thin lines) and with the [4] ones (gray dashed lines) for the examined situations. The black bold lines represent the output of the adaptive fuzzy system.

Acknowledgements

This work was supported by the EC research projects: RESCUER, FP6-IST-2003-511492 and AC-ROBOTER, FP6-IST-2006-045530.

References

[1] E. Aho, J. Vanne, T.D. Hamalainen and K. Kuusilinna, Block-level parallel processing for scaling evenly divisible images, *Circuits and Systems I: Regular Papers, IEEE Transactions on* **52**(12) (Dec. 2005), 2717–2725.

[2] T.L. Alvarez, J.L. Semmlow and C. Pedrono, Dynamic assessment of disparity vergence ramps, *Computers in Biology and Medicine* **37** (2007), 903–909.

[3] A. Amanatiadis, A. Gasteratos, C. Georgoulas, L. Kotoulas and I. Andreadis, Development of a stereo vision system for remotely operated robots: A control and video streaming architecture, in: *Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2008. VECIMS 2008. IEEE Conference on*, July 2008, pp. 14–19.

[4] A. Bernardino and J. Santos-Victor, Correlation based vergence control using log-polar images, *International Symposium on Intelligent Robotic Systems*, 1996.

[5] M.Z. Brown, D. Burschka and G.D. Hager, Advances in computational stereo, *IEEE Trans on Pattern Analysis and Machine Intelligence* **25**(8) (2003), 993–1008.

[6] C. Capurro, F. Panerai and G. Sandini, Dynamic vergence using log-polar images, *International Journal of Computer Vision* **21** (1997), 79–94.

[7] S.-B. Choi, B.-S. Jung, S.-W. Ban, H. Niitsuma and M. Lee, Biologically motivated vergence control system using human-like selective attention model, *Neurocomputing* **69**(4–6) (2006), 537–558.

[8] D. Chrysostomou and A. Gasteratos, Comparison and evaluation of similarity measures for vergence angle estimation, *International Workshop on Robotics in Alpe-Adria-Danube (RAAD2007)*, 2007.

[9] A. Gasteratos, Tele-autonomous active stereo-vision head, *International Journal of Optomechatronics* **2** (2008), 144–161.

[10] B.K. Horn, *Robot Vision*, MIT press, Cambridge, Massachusetts, USA, 1986.

[11] N. Kyriakoulis, A. Gasteratos and S.G. Mouroutsos, Fuzzy vergence control for an active binocular vision system, in: *Cybernetic Intelligent Systems, 2008. CIS 2008. 7th IEEE International Conference on*, Sept. 2008, pp. 1–5.

[12] R. Manzotti, A. Gasteratos, G. Metta and G. Sandini, Disparity estimation on log-polar images and vergence control, *Computer Vision and Image Understing* **83**(2) (2001), 97–117.

[13] L. Nalpantidis, G.Ch. Sirakoulis and A. Gasteratos, Review of stereo matching algorithms for 3D vision, *International Symposium on Measurement and Control in Robotics*, 2007.

[14] H. Niitsuma and T. Maruyama, Real-time detection of moving objects, *Lecture Notes in Computer Science* **3203** (2004), 1155–1157.

[15] J.G. Samarawickrama and S.P. Sabatini, Version and vergence control of a stereo camera head by fitting the movement into the hering's law, in: *Computer and Robot Vision, 2007. CRV '07. Fourth Canadian Conference on*, May 2007, pp. 363–370.

[16] W. Stürzl, U. Hoffmann and H.A. Mallot, Vergence control and disparity estimation with energy neurons: Theory and implementation, *International Conference on Artificial Neural Networks (ICANN'02), Lecture Notes In Computer Science* **2415** (2002), 1255–1260.

[17] E.K.C. Tsang, S.Y.M. Lam, Y. Meng and B.E. Shi, Robotic gaze and vergence control via disparity energy neurons, in: *Cellular Neural Networks and their Applications, 2008. CNNA 2008. 11th International Workshop on*, July 2008, p. 3.

[18] A.X.J. Zhang, A.L.P. Tay and A. Saxena, Vergence control of 2 d.o.f. pan-tilt binocular cameras using a log-polar representation of the visual cortex, in: *International Joint Conference on Neural Networks* (2006), 4277–4283.