

## Research Article

# A Recursive Fuzzy System for Efficient Digital Image Stabilization

**Nikolaos Kyriakoulis and Antonios Gasteratos**

*Department of Production and Management Engineering, School of Engineering, Democritus University of Thrace, GR-671 00 Xanthi, Greece*

Correspondence should be addressed to Antonios Gasteratos, gasteratos@ieee.org

Received 14 March 2008; Accepted 23 May 2008

Recommended by Zne-Jung Lee

A novel digital image stabilization technique is proposed in this paper. It is based on a fuzzy Kalman compensation of the global motion vector (GMV), which is estimated in the log-polar plane. The GMV is extracted using four local motion vectors (LMVs) computed on respective subimages in the logpolar plane. The fuzzy Kalman system consists of a fuzzy system with the Kalman filter's discrete time-invariant definition. Due to this inherited recursiveness, the output results into smoothed image sequences. The proposed stabilization system aims to compensate any oscillations of the frame absolute positions, based on the motion estimation in the log-polar domain, filtered by the fuzzy Kalman system, and thus the advantages of both the fuzzy Kalman system and the log-polar transformation are exploited. The described technique produces optimal results in terms of the output quality and the level of compensation.

Copyright © 2008 N. Kyriakoulis and A. Gasteratos. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. INTRODUCTION

Digital video stabilization is the process, where the video signal is smoothened against unwanted oscillations while preserving the intentional camera movements. Almost any acquired image sequence is affected by noise and undesired camera jitters, caused by unstable holding and rough terrain. These unwanted positional oscillations of the image sequence affect the visual quality, which besides the aesthetic part is also crucial in many applications such as in robot vision or in video compression. High visual quality enables either humans or machines to easily watch and perceive the sequence, and thus meaningful results to be extracted. Several different image stabilization methods have been reported in the literature and they can be distinguished into three major categories; the technique where the unwanted fluctuations are mostly the rotational ones and the stabilization is implemented by servo motors, which compensates the pan and the tilt camera movements, respectively, is known as active image stabilization [1]. The image stabilization which is performed by electronic hardware is referred as electronic image stabilization [2]. Finally, when the unwanted oscilla-

tions are compensated by pure image processing techniques, the process is called digital image stabilization (DIS) [3]. A DIS system is built by two successive units: the motion estimation and the motion compensation one. The goal of the first unit is to compute the motion vectors, and eventually the GMV. The compensation unit follows the motion estimation and produces the vector to shift the current frame's position so that the output to be free from irregularities, preserving the desired global motion. An important feature affecting the performance of the DIS systems is the noise level. Apparently, the lower the noise is the smoother the results are. The GMV calculation has been realized by various techniques, such as phase correlation matching [4] and normalized cross-correlation [5]. A real-time DIS implementation that performs image matching of two successive images, by means of the Fourier-Mellin transformation has been reported in [6]. In [7], the GMV estimation is optimized by the exploitation of fuzzy logic. Kalman filtering has been utilized for the enhancement of the compensation of frame position [4, 8]. Apart from the matching techniques, optical flow ones have been adopted to estimate the motion in a sequence. The undesired motion

effects are calculated in [9] by estimating the rotational center and the angular frequency from the local translational motion definition by fine-to-coarse multiresolution motion estimation. In [10], the stabilization is accomplished by fixating at the central image region, while optical flow estimation optimizes this approximation. The LMVs determine the movement in a particle of the image, resulting in a better estimation of the indented camera movement and the undesired motion. A widely used technique is to compute the GMV via a series of LMVs. The computational cost of full-frame search algorithms implied the calculation of the global motion on subimages. The LMVs estimation on these regions has reduced the processing times to a high degree. The image sequence transformation to less computational intensive topological rearrangements has further reduced the processing and the computational resources.

In this paper, we transformed the Cartesian images into log-polar ones [11, 12] and there we computed the GMV from four LMVs in respective image regions. The resulting method achieves low processing times, efficient for real-time implementation. Due to the intrinsic attentional nature of the log-polar transformation, the motion estimation of the LMVs exhibits a space-variant distribution. Moreover, a fuzzy Kalman DIS technique is proposed. Kalman filter and fuzzy systems have widely been used in DIS applications. Recursive fuzzy systems provide optimal results. Prior smoothing of the imported displacements to the fuzzy system, either by Kalman filtering or another filter, has also provided efficiency to fuzzy systems [7]. However, in this work the recursiveness of Kalman filter is directly introduced to the fuzzy system, instead of expressing it as a standard discrete time-invariant system. The fuzzy inputs of the proposed system are expressed with the estimation-correction equations of the Kalman filter. Therefore, the intended camera movement is preserved more efficiently since it happens mostly in the foreground. Consequently, to the GMV estimation the fuzzy Kalman filter is utilized. In each time step, the estimated motion vector is the a priori measurement, while the output of the system is the a posteriori one. Finally, the correction is achieved through the previous measurements, which are used as the estimated ones. The fuzzy system was tested with several types of membership functions (MFs) and different aggregation and defuzzification methods. The measured fluctuations were not filtered further. The use of log-polar images for the motion field extraction issued fast and optimized results both for the stabilization of each frame and the visual quality of the video output, in all the tested situations. The whole operation exploits the advantages of the log-polar plane and the fuzzy Kalman system.

## 2. MOTION ESTIMATION

The motion estimation unit of the DIS system extracts the GMV. This unit distinguishes between the desired and the unwanted motion effects. The key feature is the accuracy of the intended camera motion estimation. Several motion estimation approaches were proposed in the past. Their main

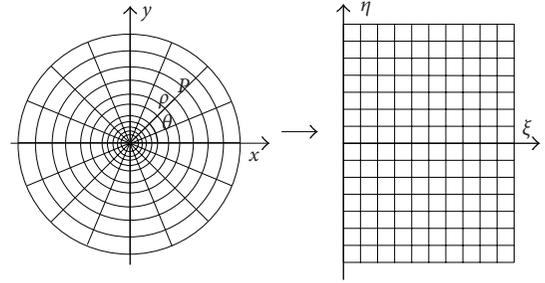


FIGURE 1: The log-polar mapping maps the radial lines and concentric circles into lines parallel to the coordinate axes.

categories are the block matching [13], the phase correlation [14], and the optical flow ones [15].

### 2.1. Log-polar transformation

Motion estimation is extremely demanding in terms of computation and resources. Subsampling of the images is often used in order to overcome this computational load. Therefore, a topological arrangement and notably a space-variant one, such as the log-polar, provides lesser volume of the image data without constraining the field of view or the image resolution at the fixation point. The log-polar transformation is based on the human's eyes projections of the retina plane to the visual cortex. It finds its origins into studies on the vision mechanisms of the mammals. The adoption of this topology into artificial vision systems exhibits several advantages as in visual attention, throughput rate and real-time processing. Many applications of the log-polar transformation have been reported, such as the time-to-impact estimation [11], wavelet extraction based on log-polar mapping [16], tracking [17], and disparity estimation and vergence control [18].

The mathematical model of the log-polar mapping can be expressed as a transformation between the polar  $(\rho, \theta)$  (retinal), the log-polar  $(\xi, \eta)$  (cortical plane), and the Cartesian plane  $(x, y)$  (image plane) as shown in Figure 1. Assuming that  $N_r$  is the number of cells in the radial direction and  $N_a$  is the number of cells in the angular direction, the mapping from the polar coordinates  $(\rho, \theta)$  to the log-polar coordinates  $(\xi, \eta)$ ; the log-polar variables  $\xi$  and  $\eta$  are defined as

$$\begin{aligned} \xi &= \log_{\alpha} \left( \frac{\rho}{\rho_0} \right), \\ \eta &= \eta \frac{N_a}{2\pi}, \end{aligned} \quad (1)$$

where  $\xi$  is each row pixel,  $\eta$  is each column pixel, and  $\rho_0$  is the radius of the fovea. The logarithmic basis  $\alpha$  is obtained from the foveal radius, the image radius  $\rho_{\max}$  and the radial resolution  $N_r$ :

$$\alpha^{N_r} = \frac{\rho_{\max}}{\rho_0} \quad \text{or} \quad \alpha = e^{(1/N_r) \ln(\rho_{\max}/\rho_0)}. \quad (2)$$

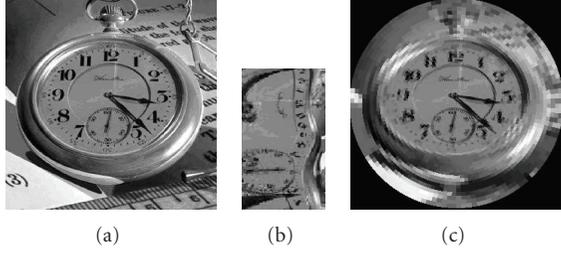


FIGURE 2: (a) Cartesian image, (b) log-polar image, and (c) reconstructed Cartesian from log-polar image.

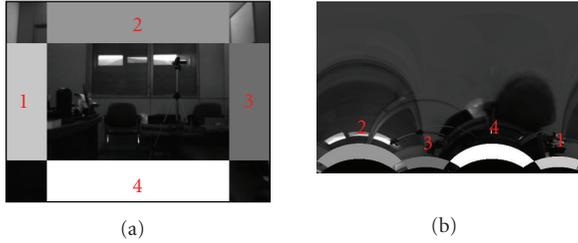


FIGURE 3: The selected subimages for motion estimation, in the Cartesian and log-polar plane, respectively.

The aforementioned mathematical formulation applied on the image in Figure 2(a) results to the log-polar image in Figure 2(b). In Figure 2(c), the reconstructed Cartesian representation of the log-polar image is shown.

## 2.2. Motion field extraction

The image motion is the projection of the real world 3D motion onto the two-dimensional image plane. This is expressed as either image velocities or image displacements on the  $x$  and  $y$  axes of the optical flow field. Optical flow techniques are divided into three main categories: the differential techniques, the frequency-based ones, and the matching methods [15]. The chosen calculation method is a differential one, that is, the classical Horn and Schunk optical flow model as modified in [19].

In order to reduce the computational load of the motion estimation, the horizontal and vertical axes displacements are computed on selected image regions located at the periphery of the image. On the Cartesian plane, these have a rectangular shape of  $440 \times 100$  pixels and  $100 \times 280$  pixels, respectively, as shown in Figure 3(a). Notwithstanding, the calculation of the LMVs was performed on the log-polar plane. The respective patches have an arch-like shape of dimensions 7353 pixels and 1893 pixels, respectively, as shown in Figure 3(b).

Yet, the motion estimation on the log-polar plane has some special features that should be taken into consideration, that is, the motion vectors are not transferred straightforwardly from the Cartesian to the log-polar plane due to the introduced fictitious gray-value curvature in the polar

image [12]. Having estimated the LMVs, the GMV was the average value of the four LMVs, as it provided better results for the tested image sequences. The displacements are then imported into the fuzzy Kalman system without further processing.

## 3. FUZZY KALMAN SYSTEM

The prediction-correction recursive equations of the Kalman filter were employed for the definition of the fuzzy inputs. The ground truth values of the fuzzy Kalman system are the displacements obtained during the optical flow technique, at the motion estimation phase. The use of the Fuzzy Kalman system equations are depicted into Figure 4 and are defined as follows.

*Prediction:*

$$\begin{aligned} P_k^- &= AP_{k-1}A^T + Q, \\ \hat{x}_k^- &= A\hat{x}_{k-1}^- + Bu_k. \end{aligned} \quad (3)$$

*Correction:*

$$\text{Input1}_k = z_{k-1} - \hat{x}_k^-, \quad (4)$$

$$\text{Input2}_k = \text{Input1}_k - \text{Input1}_{k-1}, \quad (5)$$

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}, \quad (6)$$

$$P_k = (I - K_k H) P_k^-, \quad (7)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H\hat{x}_k^-), \quad (8)$$

where  $k$  is the time index,  $z_k$  is the measurement value at the current time step, and  $\hat{x}_k^-$  is the a priori estimation of the frames positions.  $\hat{x}_k$  are the a posteriori estimated frame positions.  $P_k^-$  and  $P_k$  define, respectively, the a priori and the a posteriori error covariance matrices. The first input (4) is defined as the difference between the absolute frame translation and the a priori estimation of the stabilized frame position. The second input (5) indicates the rate of change of the first input at the current time step. The measurement values in each time index represent the frame's translation. The tuning variables  $Q$  and  $R$  for the process and the measurement noise, respectively, are set to a ratio of 10 ( $R/Q = 10$ ). Higher ratio yields to quicker responses, but the final output is not smooth enough, as the final frames position are close to the measured ones. High  $R$  values lead to low responses, though the high frequencies are cut off, providing smooth output. In order to provide a fast response the ratio was set to 10, although a ratio of 100 and higher introduced less error to the final output.

The key features in the designing of a fuzzy system are the shape of the MFs and the decision rules. In the proposed system, five MFs are used for each input and output, as they are efficient for the desired task. The construction of the fuzzy rules depends on the experience of the designer and the

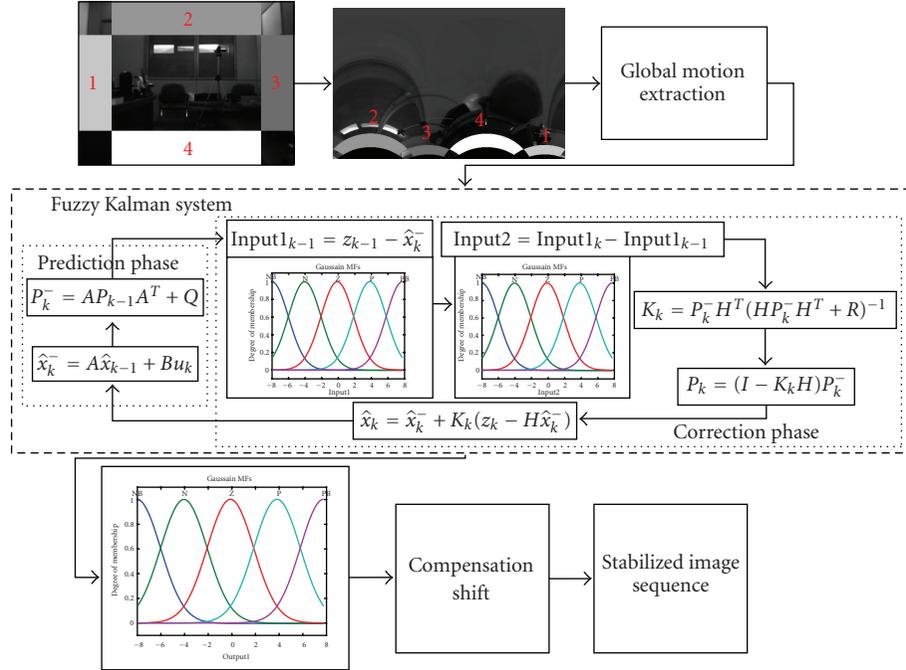


FIGURE 4: The block diagram of the fuzzy Kalman filter.

TABLE 1: Rule base for the fuzzy Kalman system.

		Input2				
		NB	N	Z	P	PB
Input1	NB	NB	N	N	Z	Z
	N	N	N	Z	Z	P
	Z	N	Z	Z	Z	P
	P	N	Z	P	P	P
	PB	Z	Z	P	P	PB

application used. In our task, there was a need of covering the range in order the final output to be smooth enough. Thus, the options are to distribute normally the MFs to their range or to import more MFs. More MFs lead to more fuzzy rules, and consequently to higher complexity. The selection of the type of the MFs is also crucial for the construction of a fuzzy system. The tested types of the MFs are Gaussian, trapezoid, and triangular ones. In all experiments, all the variables (inputs and output) had the same type of MFs. The two inputs and the output are normally distributed to their range in order to obtain, as it is mentioned, a smooth output. All the variables define the frame translations and are set to  $[-8 \ 8]$  pixels, as 8 pixels were the maximum absolute translation both on the horizontal and the vertical axis. The sign indicates the direction of the movement, that is, left or right and up or down. The rules interaction set is depicted in Table 1 and Figure 5 illustrates the fuzzy system for the Gaussian MFs. Important role to the fuzzy system play the possible adjustment methods, such as the implication, the defuzzification, and the aggregation ones. In the proposed system, the implication was set to product and

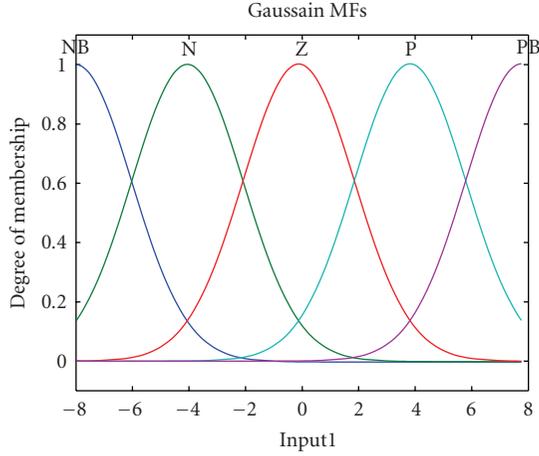
the aggregation method to sum, as it provided a smoother output value. The defuzzification method was set to centroid, as it covers the output range more efficiently.

#### 4. EXPERIMENTAL RESULTS

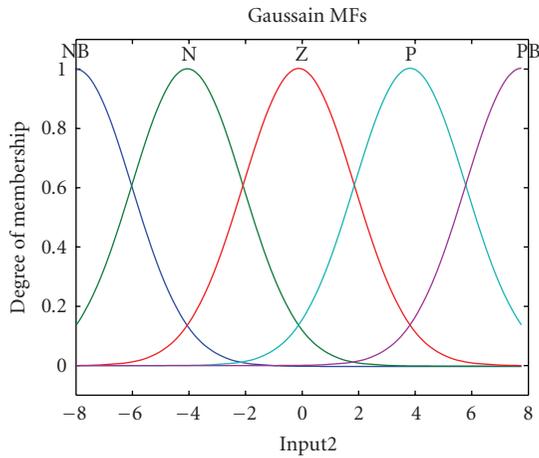
In order to evaluate the performance of the proposed system we performed several tests. These include different stabilization experiments captured by an active stereo vision head. The size of the acquired sequences is  $640 \times 480$  pixels. Some of the testing input videos were acquired, while an active image stabilization routine was running. All of these sequences suffer from high-frequency image jitters, produced intentionally by the user for testing purposes. They also suffer from high-illumination changes as well as from fluctuations caused by the servo motors. Further experiments were made, capturing video on a free course. These sequences suffer from motion blurred frames. The remedy to such sequences is a higher frame rate. As the acquired videos were tuned to 25 fps, the fast oscillatory movements during the course provoked loss of information to a high degree. The purpose of capturing such noisy and shaky sequences is to assess the proposed fuzzy Kalman system against complicated and challenging circumstances.

In order to compare the efficiency of our system the stabilization was assessed in four different combinations of image topologies as follows:

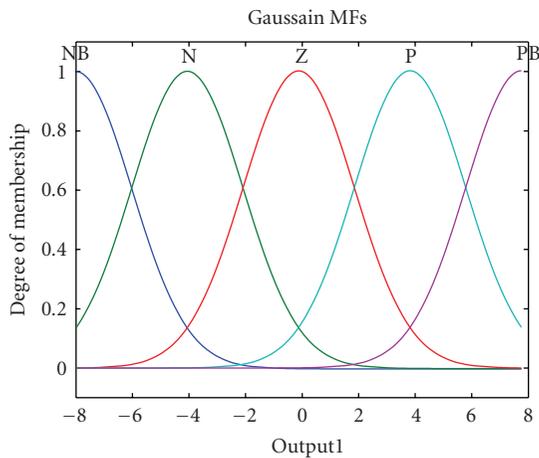
- (i) Cartesian image, full frame;
- (ii) Log-polar image, full frame;
- (iii) Cartesian image, subimages;
- (iv) Log polar image, subimage.



(a)



(b)



(c)

FIGURE 5: The membership functions of the fuzzy Kalman filter for Gaussian MFs.

The use of LMVs in Cartesian images provided better results than the full-frame ones. Table 2 summarizes the comparative results. In order to measure the performance of the proposed stabilization the mean square error (MSE),

TABLE 2: Error calculation table: the tested image topologies for motion estimation.

	Cartesian Full frame	Log-polar Full frame	Cartesian Subimages	Log-polar Subimages
MSE	2.311	1.872	2.151	0.751
LSE	0.0008	0.0000124	0.000006720	0.00000013
LMSE	1.23	1.68	1.51	0.522

TABLE 3: Error calculation table: the tested MFs.

	Gaussian	Trapezoid	Triangular
MSE	0.79777721	0.83456986	0.751706518
LSE	0.00000035	0.00000226	0.00000013
LMSE	0.60190373	0.66448906	0.52190102

the least square error (LSE), and the least mean square error (LSME) were calculated. The equations of these errors, as all the values are known, are defined as

$$\begin{aligned}
 \text{MSE} &= \frac{1}{m} \sum_{i=1}^m (f_i - z_i)^2, \\
 \text{LSE} &= (f_i - z_i)^2 = \min, \\
 \text{LMSE} &=, \frac{1}{m} \sum_{i=1}^m (f_i - z_i)^2 = \min,
 \end{aligned} \tag{9}$$

where  $f_i$  is the final stabilized frame position and  $z_i$  is the measured one from the motion estimation phase for every time index  $i$ . It is clear that the GMV extraction via LMVs in the log-polar plane provided the smoother output. The fuzzy Kalman system responded better by using triangular MFs. The visual results of the fuzzy Kalman system are demonstrated in Figure 6, while in Figure 7 the initial and the final frames' translation are shown for all the tested occasions. It is clear that the estimation of the GMV into the log-polar plane provides better performance.

Furthermore, these errors were also calculated for the efficiency of the different types of MFs. In Table 3, the comparative results for all the tested MFs are demonstrated. From Figure 8 and Table 3, it is clear that the triangular MFs provide a smoother output as they exhibit lower error cost in all the qualitative tests.

## 5. CONCLUSION

An image stabilization technique by means of a fuzzy Kalman system was proposed. The fuzzy Kalman system processes the GMV which is computed in the log-polar plane. The system provided a smoothly compensated output in all the tested image sequences. For the proposed fuzzy system, the triangular MFs proved to produce lesser errors. The use of log-polar images, along with the recursiveness of the Kalman filter, led to an optimum system, which not only stabilizes any fluctuations but also filters the noise during the process. To conclude, log-polar images are ideal for image stabilization,

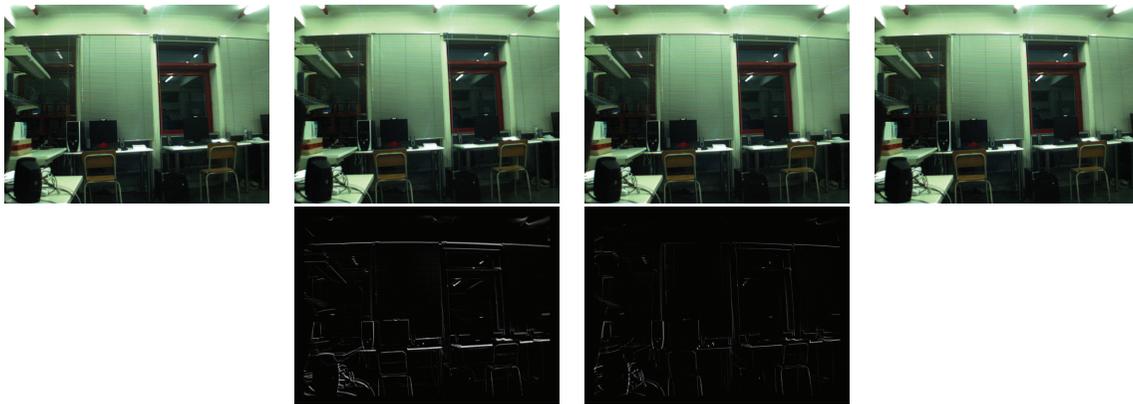


FIGURE 6: Two consecutive frames, the right column are the unstabilized while the compensated ones are depicted at the left one.

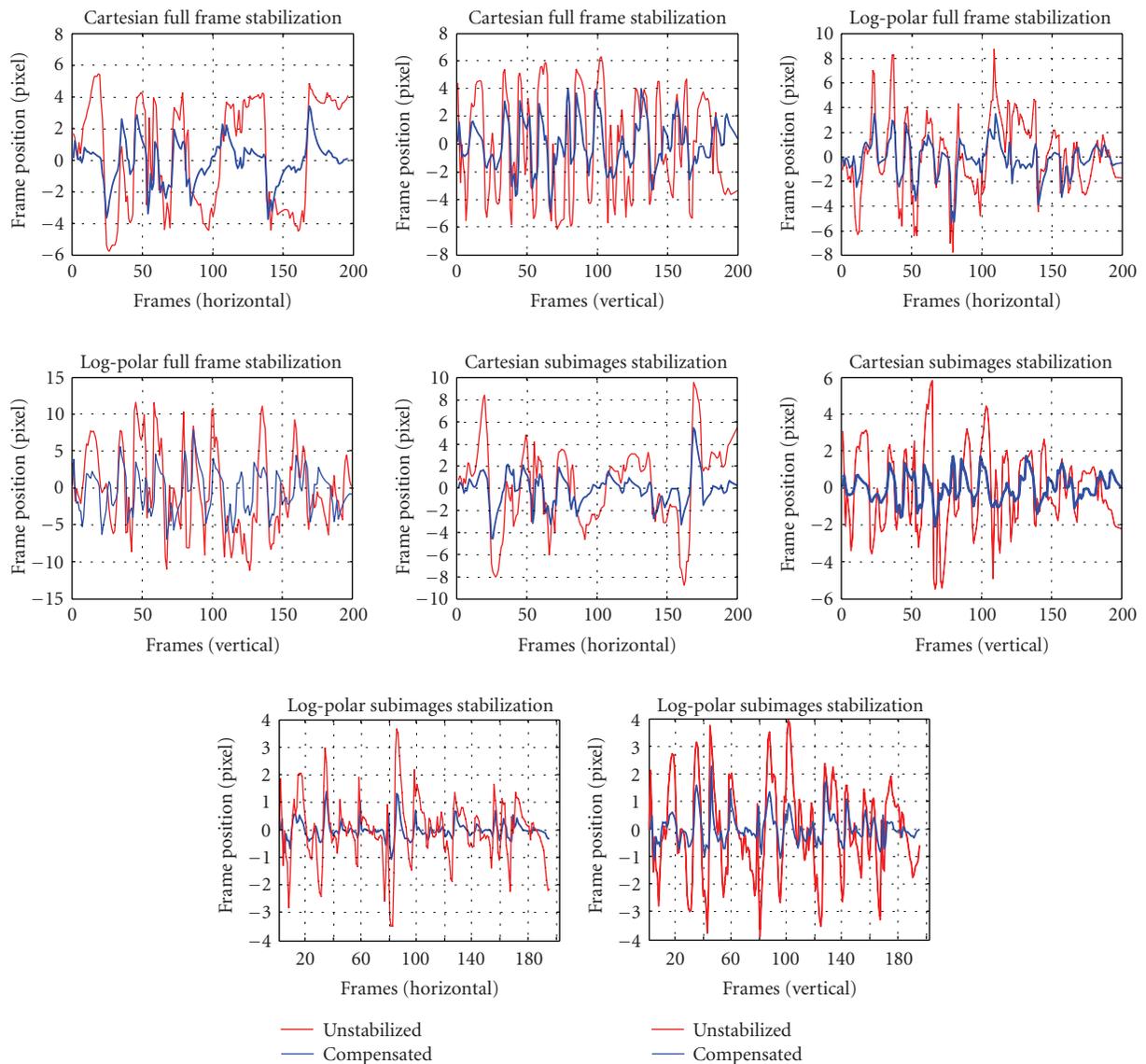


FIGURE 7: The absolute frame positions before and after the stabilization for all the tested topologies.

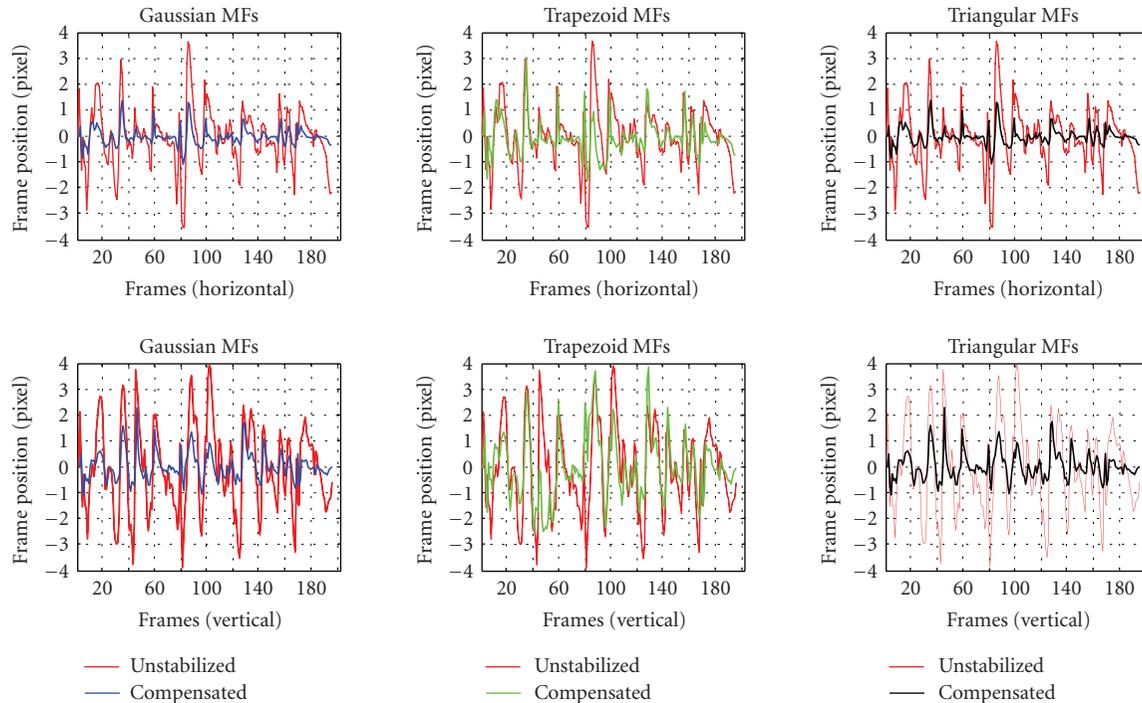


FIGURE 8: The red line presents each frame position before stabilization the blue, the green, and the black line the stabilized one for Gaussian, trapezoid, and triangular MFs, respectively.

as the errors are shorter. The proposed fuzzy Kalman system is a valuable and efficient tool for image stabilization.

## ACKNOWLEDGMENT

This work is partially supported by the EC research Project “ACROBOTER” FP6-IST-2006-045530.

## REFERENCES

- [1] F. Panerai, G. Metta, and G. Sandini, “Learning visual stabilization reflexes in robots with moving eyes,” *Neurocomputing*, vol. 48, no. 1–4, pp. 323–337, 2002.
- [2] C. Morimoto and R. Chellappa, “Fast electronic digital image stabilization for off-road navigation,” *Real-Time Imaging*, vol. 2, no. 5, pp. 285–296, 1996.
- [3] L. Xu and X. Lin, “Digital image stabilization based on circular block matching,” *IEEE Transactions on Consumer Electronics*, vol. 52, no. 2, pp. 566–574, 2006.
- [4] O. Kwon, J. Shin, and J. Paik, “Video stabilization using Kalman filter and phase correlation matching,” in *Proceedings of the 2nd International Conference on Image Analysis and Recognition (ICIAR ’05)*, vol. 3656 of *Lecture Notes in Computer Science*, pp. 141–148, Toronto, Canada, September 2005.
- [5] S.-C. Hsu, S.-F. Liang, and C.-T. Lin, “A robust digital image stabilization technique based on inverse triangle method and background detection,” *IEEE Transactions on Consumer Electronics*, vol. 51, no. 2, pp. 335–345, 2005.
- [6] J. R. Martinez-de Dios and A. Ollero, “A real-time image stabilization system based on fourier-mellin transform,” in *Proceedings of the International Conference on Image Analysis and Recognition (ICIAR ’04)*, vol. 3211 of *Lecture Notes in Computer Science*, pp. 376–383, Porto, Portugal, September 2004.
- [7] M. K. Güllü and S. Ertürk, “Membership function adaptive fuzzy filter for image sequence stabilization,” *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 1–7, 2004.
- [8] S. Ertürk, “Real-time digital image stabilization using Kalman filters,” *Real-Time Imaging*, vol. 8, no. 4, pp. 317–328, 2002.
- [9] J.-Y. Suk, G.-W. Lee, and K.-I. Lee, “New electronic digital image stabilization algorithm in wavelet transform domain,” in *Proceedings of the International Conference on Computational Intelligence and Security (CIS ’05)*, vol. 3802 of *Lecture Notes in Computer Science*, pp. 911–916, Xi’an, China, December 2005.
- [10] K. Pauwels, M. Lappe, and M. M. Van Hulle, “Fixation as a mechanism for stabilization of short image sequences,” *International Journal of Computer Vision*, vol. 72, no. 1, pp. 67–78, 2007.
- [11] M. Tistarelu and G. Sandini, “On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 4, pp. 401–410, 1993.
- [12] K. Daniilidis and V. Kruger, “Optical flow computation in the log-polar plane,” in *Proceedings of 6th International Conference on Computer Analysis of Images and Patterns (CAIP ’95)*, pp. 65–72, Prague, Czech Republic, September 1995.
- [13] J. S. Jin, Z. Zhu, and G. Xu, “A stable vision system for moving vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 1, pp. 32–39, 2000.
- [14] H. Foroosh, J. B. Zerubia, and M. Berthod, “Extension of phase correlation to subpixel registration,” *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, 2002.
- [15] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, “Performance of optical flow techniques,” *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994.

- [16] C.-M. Pun and M.-C. Lee, "Log-polar wavelet energy signatures for rotation and scale invariant texture classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 590–603, 2003.
- [17] G. Metta, A. Gasteratos, and G. Sandini, "Learning to track colored objects with log-polar vision," *Mechatronics*, vol. 14, no. 9, pp. 989–1006, 2004.
- [18] R. Manzotti, A. Gasteratos, G. Metta, and G. Sandini, "Disparity estimation on log-polar images and vergence control," *Computer Vision and Image Understanding*, vol. 83, no. 2, pp. 97–117, 2001.
- [19] Y. Lei, L. Jinzong, and L. Dongdong, "Discontinuity-preserving optical flow algorithm," *Journal of Systems Engineering and Electronics*, vol. 18, no. 2, pp. 347–354, 2007.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

