Autofocus System and Evaluation Methodologies:
A Literature Review

Yupeng Zhang,[1*] Liyan Liu,[1] Weitao Gong,[1] Haihua Yu,[1]
Wei Wang,[1] Chongying Zhao,[2] Peng Wang,[2] and Toshitsugu Ueda[3]

[1]Ricoh Software Research Center Beijing Co., Ltd.,
28th floor, Tengda Plaza, 168 Xiwai Street, Haidian District, Beijing 100044, China
[2]Wuhan Guide Infrared Co., Ltd.,
No. 6, Huanglongshan South Road, East Lake Development Zone, Wuhan, Hubei 430205, China
[3]Waseda University, Information, Production and Systems Research Center
Hibikino 2-2, Wakamatsu-ku, Kitakyusyu, Fukuoka 808-0135, Japan

The autofocus (AF) system has gained popularity in over 40 years. Embedded in a camera, it can bring the best focused image to the viewer within a few seconds, which frees users from focusing manually. An AF system usually consists of a motor, a lens, and the processing and control units. Within years of development, the AF system has become mature in terms of both technology and market, and many fast and accurate AF systems have been introduced and widely installed in compact cameras and digital single lens reflex (DSLR) cameras. On the other hand, the market for thermal infrared cameras has been increasingly growing in recent years owing to their decreasing cost and wide use by both civilians and the military. The AF system in those cameras adopted similar but more complex mechanisms. This article serves as a literature review of the state of the art of the AF system in both visible light digital cameras and thermal infrared cameras.

1. Introduction

The first mass-produced autofocus (AF) camera was the Konica-C35 AF, which was released in November 1977.[1] Nowadays, the AF system is embedded in almost every compact digital camera and digital single lens reflex (DSLR) camera, and some smartphones and tablets also incorporated the AF system. An AF system in a digital camera is a feedback control system normally composed of three parts: (1) a motor that drives a camera lens to move along the optical axis iteratively to search for the lens position of the best focus, (2) a group of lenses that converge light rays to the image sensor and (3) a processing unit that both carries out computation such as focus value per frame and issues control signals to the motor. Focus accuracy and speed are two important indicators for evaluating an AF system. They are affected by the selected algorithm for searching the best focus position, noise level, motor

performance, specification of the optics such as f-number, density of the lens, and whether or not the subject is moving.

AF can be achieved by actively, passively, or a hybrid of the two. The active method includes ultrasonic, infrared,[2] and time-of-flight (TOF) types. The camera emits a beam of light or sound and receives the reflected beam to estimate distance between the camera and the subject. For instance, Canon's "sure-shot" is an infrared-type AF. It uses a triangulation technique to estimate distance to the subject.[3,4] The passive method is categorized into phase detection and contrast detection AF. The former measures the phase difference between two captured images to estimate the focus position. The latter measures the sharpness of each frame to find the best focus position. Each method has pros and cons. For example, the active AF works under any illumination condition. However, it cannot "see" through windows and occlusions. Its accuracy is also inferior to that of the passive AF system.[1] The passive AF can detect and focus on subjects behind windows but may fail when the illumination is poor. In comparison with contrast detection, phase detection distinguishes between near and far focus. It is generally faster than contrast detection and adapts to focusing on moving objects. However, its accuracy is inferior to that of the latter.[5] Some cameras employ a hybrid-type AF, which combines active and passive AF, to ensure better performance instead of using active or passive alone, but they are more expensive and bulky.

Among the three types mentioned above, contrast-based AF has received considerable attention because of its good performance and low cost.[6,7] There are numerous papers and patents on contrast-based AF, some of which have already been used in consumer digital cameras or mobile phone cameras. In this paper, we review the literature on contrast-based AF from the early years to the state of the art. In addition, we contribute a section on the AF of thermal infrared cameras [also known as long-wave infrared (LWIR) cameras], because the LWIR camera has seen rapid growth in recent years in both consumer and military markets. Some high-end models incorporate the AF system, whose design is similar to that of digital cameras but more challenging.

2. Contrast-based AF

2.1 Focus measure function (FMF)

In an AF system, the FMF calculates the focus value for each frame. The conventional AF system searches for the peak of the focus values or the highest frequency component as the lens moves from out-of-focus to in-focus position.[8,9] For example, Fig. 1 shows some video frames captured during the AF process using an thermal infrared camera. The corresponding curves at the bottom are calculated at each frame using Gaussian focus measure (GFM). A good FMF should possess the following properties: (1) independent of the scene being captured; (2) fast and accurate; (3) has good reproducibility;[10] and (4) less affected by noise, which results in local maxima. Many FMFs were proposed in the 1970s–80s for microscopes.[11–15] Later, researchers found that some of the old FMFs are not suitable for digital cameras.[16] Therefore, many improved algorithms were introduced.
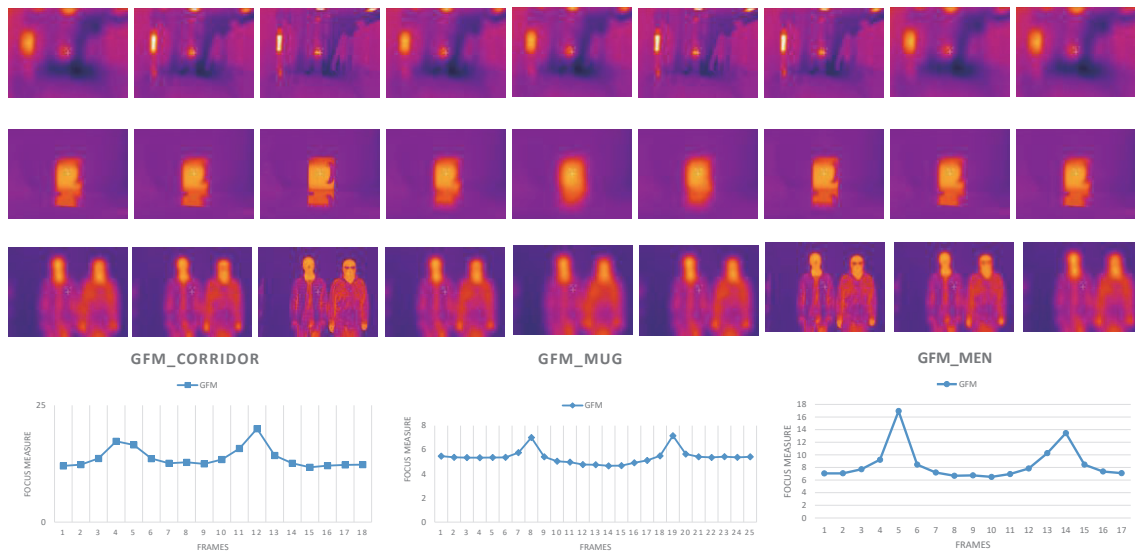
Fig. 1.   (Color online) Video frames captured during AF process using a thermal infrared camera. For each of the three scenes, GFM is used to calculate focus value at each frame. Two peaks corresponding to two in-focus positions are detected as shown in the curves.

A comprehensive study about FMF was published in 1985.[10]   The authors compared 11 different FMFs, and evaluated three different types of images. Results show that the squared gradient, Laplacian, and normalized variance outperform the other 8 FMFs for all three images. However, the number of images used for evaluation is small so it is still difficult to tell which FMF is superior.

Mir *et al.*[16] summarized the FMF into many categories: (1) first-order derivatives, such as the absolute gradient,[14] Prewitt operator,[17] Robert operator,[8] and Sobel operator;[8,17–20] (2) second-order derivatives, such as the Laplacian operator;[8,10,20–27] and (3) histogram-based,[11,12,18,20,21,28–30] (4) image-statistics-based,[8,10,12,17,18,20,21,26,28,29] (5) correlation-based,[31,32] and (6) data-compression-based[33] FMFs. They also divided the literature into AF methods for microscopy and digital photography and showed that some FMFs perform well on the former but poorly on the latter. They evaluated over 30 FMFs using a tremendous amount of real scene images (4303 in total) under many categories captured by a Canon DSLR. They found that the two first-order derivative methods, namely, Brenner *et al.*'s method[13] and the squared gradient function,[24–26,28] outperform others in terms of precision, recall, and the mean absolute error for the selected 4303 images. However, they did not show which method is less vulnerable to noise that usually increases under low light condition. Moreover, the AF speed was also not compared.

The same team in Ref. 16 conducted further research in 2015.[34] They expanded the image sets to 32 benchmarks for evaluation, a total of 5344 images. Additionally, 11 more sets were captured in a dark room to evaluate FMFs under low light condition. This is by far the largest image set used for evaluating an AF system as far as we know. However, they evaluated only

the squared gradient[24–26,28] and GFM,[18,35–37] and found that the latter is significantly better than the former under low light condition. This is because the GFM adds a smoothing effect to the input image before taking the first-order derivatives. This filters out a lot of noise such that it will not contribute to the focus measure. Other publications[38,39] also discussed the effectiveness of the GFM under low light conditions.

Another FMF that has been reported to be effective in reducing noise under low light condition is the frequency selective weighted median (FSWM).[40] Choi *et al.* proposed this filter as an FMF for the AF system.[41] It was based on previous research on weighted median filters.[42] The FSWM filter can extract high-frequency components from an image and reduce impulsive noise as well. The authors used 11 indoor and 3 outdoor scenes, a total of 2100 images for evaluating the proposed FSWM filter and compared with three other FMFs: one first-order derivative, one Laplacian, and one absolute gradient. As to the noise reduction ability, FSWM beat all the other three methods by effectively eliminating impulsive noise while keeping the useful high-frequency components such as edges and corners unchanged. However, the author did not compare the FSWM with GFM, which also demonstrates good noise reduction ability.

Some other methods have also been proposed to tackle noise problem under low contrast conditions. Many studies used variance-based methods employing discrete cosine transformation (DCT). Xu[43] *et al.* cited some traditional DCT-based methods that are vulnerable to noise and also some modified methods that are effective for noisy scenes.

A quantitative evaluation of the noise sensitivity of FMF was documented by Subbarao and Tyan.[35] They used standard deviation and root-mean-square error to evaluate noise sensitivity. Results showed that the best focus measure is dependent on both noise level and image texture.

2.2 Search algorithm

The search algorithm is the basis for all contrast-based AF systems. It directly relates to AF speed and accuracy. It can be divided into hill climbing search (HCS),[9] Fibonacci search,[19,44] curve fitting search,[44,45] binary search,[46] and a combination of the above-mentioned search methods. The conventional hill climbing algorithm without noise reduction results in a local maximum. It also depends on the FMF.

2.2.1 Global search

The simplest search algorithm is the global search, which means that the AF system measures the focus value at every lens step.[46] This is not efficient because it searches the entire lens range for the peak using the same speed. Some highly blurred frames captured during the AF process are not likely surrounded by a sharp frame, which is probably the in-focus position. The lens at these positions can move faster by taking large steps instead of taking the same steps.

2.2.2   Rule-based search

Therefore, a rule-based search algorithm was proposed by Kehtarnavaz and Oh.[46]   They divided the search range into coarse, middle, and fine regions. The fine region corresponds to the lens positions, which probably includes the global peak, so that the motor stops at every step to calculate the focus value. In the middle and coarse regions, the motor stops at three to four and seven to ten steps, respectively. The authors compared their rule-based algorithm with global search and binary search algorithms and showed the fastest speed in both the number of iterations and steps. Although the rule-based search is faster than the other two, it still requires a full sweep of the lens focus range.

2.2.3   Model-based search

The model-based method has also been reported. Chen *et al.* proposed a method by combing a discrete difference equation prediction model (DDEPM) and a bisection search algorithm to search the best focus position.[47]   The DDEPM can predict the trend of the focus value curve and locate the neighbors of the in-focus position quickly. Their method achieved real-time AF, which is 384.2 ms on average for 10 evaluations and showed good accuracy.

2.2.4   Coarse to fine search

The majority of the AF method[34,45–49] employs a coarse to fine search scheme. As mentioned above, this scheme can effectively reduce the time for searching the best focus position. The FLIR system has proposed a two-step searching algorithm.[45]   The first step uses a coarse but fast search method based on the low spatial frequencies of the image, which followed by a fine but slower search method based on the high frequencies of the image. Chen *et al.*[34] followed He *et al.*[48] and Li's[49] work but they employed two additional fine steps at the beginning to predict the direction where a peak most probably exists.

2.2.5   Curve-fitting-based methods

The curve fitting method fits the focus data to a curve such as a polynomial or Gaussian equation, then calculates the maximum on the curve and finds the corresponding lens position. During the AF process, it only needs three or four initial focus values to be calculated in order to locate the best focus position. Therefore, it reduces AF time. As mentioned earlier, the FLIR system uses a two-step search algorithm.[45]   During the fine search stage, they adopted a curve fitting method to locate the best focus position. Chen *et al.* used four initial lens positions[47] to predict the best focus position using a DDEPM. However, it is said that the curve fitting search algorithm is highly dependent on the acquired data around the peak. If the noise level is high, such as under low light conditions, local maxima appear. In this case, the algorithm may fail to locate the correct in-focus position.

2.2.6 Machine-learning-based methods

Machine learning has also been used for AF. Machine-learning-based methods do not require the lens to search the entire range but jump to the best focus position based on trained results. This reduces AF time significantly. The earliest learning-based searching algorithm was proposed by Park *et al.*[50] They used the depth from defocus (DFD) method to reduce the computational cost and AF time because this method searches the best focus position by referring to only two prefixed defocused lens positions compared with the depth from focus (DFF) method, which uses multiple positions. A multilayer neural network (MNN) was utilized to classify distance from the objects. However, the author did not mention how fast their method is compared with conventional methods such as the rule-based methods.

Chen *et al.*[51] used a well-trained self-organizing map (SOM) neural network to predict the best focus position in order to reduce the searching time. A frequency domain approach called "Discrete Wavelet Transformation (DWT)" was proposed in their paper to search for the highest frequency from the captured images, which corresponds to the best focus position. The input of the SOM network is three initial focus values. The output of the network is the lens position for the best focus. Then, a backward search is carried out to search for the best focus position more precisely. Compared with conventional full search methodologies, the SOM-based search algorithm increased the AF speed 2.5 times.

Similarly, Han *et al.*[52] also used a training-based method to reduce the searching time, but they used the focus value incremental ratio as the feature vector, which they said is less vulnerable to texture and illumination changes. This method also needs only three initial focus values to be calculated, so that the AF time is significantly reduced compared with those of conventional approaches, for example, 3.2 times faster than the rule-based approach.

A recent learning-based approach was proposed by Chen and van Beek. They introduced a supervised machine learning approach,[34] in which two decision tree classifiers are defined to decide the state of the focusing process and locate the best focus positions. They used two sets of feature vectors, and each set includes many different features. While their approach is superior to He *et al.*'s coarse to fine method[48] in AF accuracy, and obtained better accuracy even under low light conditions, it showed some decrease in AF speed. In addition, the authors of Refs. 33 and 52 all mentioned that the focus measure itself is not a good choice for the feature vector because this value is easily affected by texture and illumination changes.

2.3 Focus window

The selection of focus window also affects the AF accuracy and speed. The focus window being too large results in redundant data, thereby increasing the computational load. In contrast, a very small window may not contain the subject that needs to be focused on.[53] A focus window can be defined by users or automatically.[34,46,54,55] FLIR's AF system for a thermal infrared camera can choose a focusing window by analyzing input images[45] rather than routinely choosing the central area, which is sometimes not very informative. Lee *et al.* proposed an AF algorithm by dividing the entire image into multiple windows, each with 40

× 40 pixels. These windows are used in their two-step search scheme.[55] Compared with traditional methods, their method is very effective when multiple objects of different depths exist. Rahman and Kehtarnavaz proposed an AF approach for focusing on the human face.[54] Their method performs better than multiple-window AF.

2.4 Focusing on moving subjects

Focusing on a moving subject is challenging, especially when the subject moves fast. Fortunately, camera makers have already incorporated such a feature in some of their high-end products. Typical examples are the artificial intelligence (AI) servo of Canon's DSLR[56,57] continuous focus (AF-C) of Nikon's DSLR,[58] and trap focus[59] and a moving detection device to follow the movement of a subject and focus on it simultaneously[45] by FLIR systems. Zoom tracking, a method of continuous focusing on a moving subject along the camera's optical axis during zooming operation, was also described.[60–62]

3. AF System for Thermal Infrared Camera

Thermal infrared (long-wave infrared or LWIR) refers to the spectral band of approximately 7–14 μm. Cameras receiving this band of light do not require illumination of the subjects but "sense" the energy emitted directly from them. Therefore, the LWIR camera is very useful for night vision, surveillance, and military purposes.

There are many thermal imagers available on the market. Some high-end types employ AF feature. Basically, the thermal infrared camera can use similar FMFs and searching algorithms to locate the best focus position to a digital camera. A recent article about thermal imager AF by Srivastava *et al.* proposed a sharpness evaluation algorithm based on cumulative gradient measure.[63] They evaluated their algorithm under low contrast and noisy conditions as well as the effect of focus window on the focusing result. Cakir and Cetin used a cumulative probability of blur detection (CPBD) method to measure the amount of blur in an AF system for thermal infrared cameras.[64] The CPBD serves as an FMF, and the sharpest frame corresponds to the highest focus value.

One of the problems when designing an AF apparatus for the LWIR camera is the low level of energy penetrating to the sensor.[45] Therefore, the f number of the LWIR camera is always smaller than those of visible light cameras, normally between 0.8 and 1.2. This results in a shallow depth of field, which is much harder to focus than visible light cameras. Another issue is the temperature dependence of the optics, which also affects AF performance.[45] As a result, the design of the AF apparatus for the LWIR camera becomes much difficult. Cakir and Cetin[64] also addressed the difficulty when applying the CPBD to thermal imager because of its inherent noise problem. Their countermeasure is to modify CPBD algorithm to increase edge quantity rather than edge quality. Results showed that the modified CPBD outperformed conventional CPBD for LWIR imager. On the other hand, the optical lens of the LWIR camera usually adopts Ge as the material, which is highly transmissive to LWIR but has a much higher density (5.323 g/cm^3) than a digital camera lens. It means that the lens becomes very heavy,

which cannot be driven by conventional AF motors or driven at a low speed. Typical methods to solve the problem are to choose a lens with a low density (thus low weight), such as silicon (Si, 2.57 g/cm^3); select a motor with a high torque, speed, and accuracy; or move the sensor instead.

A thermal image is noisier than a visual image because it has low resolution, typically around 320 × 240 or 640 × 480. Therefore, antinoise is always required for an LWIR camera.

It should be mentioned that some exclusive properties of the LWIR camera, e.g., thermal difference, can be used for determining the focus window.[45] This cannot be achieved using visual cameras because they do not measure the temperature of subjects.

4. Evaluation Methodologies

An AF system is evaluated if the AF performance meets the required criteria such as accuracy, speed, and ease of implementation.[17] Many evaluation methodologies have been reported.

Shih mentioned the subjectivity of manually choosing the in-focus position in order to evaluate the accuracy of AF algorithms.[17] He compared the results of seven FMFs with manually selected in-focus images. As the manual selection is a subjective process, the author suggested collecting evaluation results from a group of observers for each FMF.

Chen and van Beek proposed a method of evaluating both accuracy and speed.[34] The speed is defined as the lens steps taken for AF and the accuracy is defined as the percentage where the peak is found.

Yousefi *et al.* evaluated the performance of four FMFs using 60 different AF sequences. For each FMF, they defined accuracy as the percentage of cases matching the true in-focus position and speed as the averaged time of all 60 sequences in millisecond.[21] However, they did not mention how the true in-focus position is defined.

Mir *et al.* introduced precision, recall, and mean absolute error (MAE)[16] for evaluating different FMFs. In their paper, ground truth is determined by letting a person view the captured images on the camera's display screen. This, however, may introduce subjectivity.

5. Conclusions

AF is an indispensable feature for DSLRs, compact digital cameras, and smartphone cameras. A good AF system should be fast, accurate, and adaptable to most scenes. It not only depends on the FMF, search algorithm, and selected focus window, but also the motor performance, properties of the optical lens, and processing capability. It is a feedback control system that requires optimization. In this paper, we mainly reviewed contrast-based algorithms and AF evaluation methodologies from the early years to the state of the art. It is clear that none of the reported algorithms can adapt to all scenes under a variety of illumination conditions, noise level, texture, and subject motion. Most of the commercial systems adopt gradient-based FMF and hill-climbing search with some modifications, which can adapt to most cases. In addition, we also reviewed AF for thermal infrared cameras, which is more challenging than AF for digital cameras. Although the current commercial AF cameras have reached a satisfactory

level, the future AF system could be faster, more accurate, and robust with the advent of new algorithms such as learning-based methods.

References

1   Wikipedia: https://en.wikipedia.org/wiki/Autofocus (accessed August 2017).
2   Y. Yoshida, S. Shinohara, H. Ikeda, K. Tada, H. Yoshida, K. Nishide, and M. Sumi: Proc. 1991 IEEE 34th Midwest Symp. Circuits and Systems (IEEE, 1991) 395–398.
3   P. J. Besl: Mach. Vision. Appl. **1** (1988) 127.
4   Wikipedia: https://en.wikipedia.org/wiki/Canon_AF35M (accessed August 2017).
5   R. Cicala: https://wordpress.lensrentals.com/blog/2012/07/(accessed July 2012).
6   S. M. Sohn, S. H. Yang, H. M. Park, S. W. Kim, Y. T. Cho, M. H. Sohn, S. H. Yi, K. T. Hong, K. H. Beak, and W. H. Paik: IEEE Trans. Consum. Electron. **52** (2006) 10.
7   V. Peddigari, M. Gamadia, and N. Kehtarnavaz: J. Imaging Sci. Technol. **49** (2005) 114.
8   J. H. Lee, K. S. Kim, B. D. Nam, J. C. Lee, Y. M. Kwon, and H. G. Kim: IEEE Trans. Consum. Electron. **41** (1995) 449.
9   K. Ooi, K. Izumi, M. Nozaki, and I. Takeda: IEEE Trans. Consum. Electron. **36** (1990) 526.
10 F. C. A. Groen, I. T. Young, and G. Ligthard: Cytometry **6** (1985) 81.
11 D. Mason and D. Green: IEEE Trans. Biomed. Eng. **22** (1975) 312.
12 M. L. Mendelsohn and B. H. Mayall: Comput. Biol. Med. **2** (1972) 137.
13 J. F. Brenner, B. S. Dew, J. B. Horton, T. King, P. W. Neurath, and W. D. Selles: J. Histochem. Cytochem. **24** (1976) 100.
14 R. A. Jarvis: Microscope **24** (1976) 163.
15 W. Lang: U.S. Patent 3,721,759 (1973).
16 H. Mir, P. Xu, and P. V. Beek: IS&T/SPIE Electronic Imaging. Int. Society for Optics and Photonics 2014: 90230I-90230I-11.
17 L. Shih: Proc. Soc. Photo. Opt. Instrum. Eng. **3** (2007) 1.
18 X. Y. Liu, W. H. Wang, and Y. Sun: J. Microsc. **227** (2007) 15.
19 E. Krotkov: Int. J. Comput. Vision **1** (1987) 223.
20 N. N. K. Chern, P. A. Neow, and M. H. Ang: Proc. 2001 ICRA, IEEE Int. Conf. Robotics and Automation (IEEE, 2001) 2791–2796.
21 S. Yousefi, M. Rahman, and N. Kehtarnavaz: IEEE Trans. Consum. Electron. **57** (2001) 476.
22 S. K. Nayar and Y. Nakagawa: IEEE Trans. Pattern Anal. **16** (1994) 824.
23 M. Subbarao and T. C. Wei: Proc. CVPR'92, IEEE Computer Society Conf. (IEEE, 1992) 773–776.
24 A. Erteza: Appl. Optics **15** (1976) 877.
25 A. Erteza: Appl. Optics **16** (1977) 2273.
26 R. A. Muller and A. Buffington: JOSA **64** (1974) 1200.
27 M. S. J. Tyan: Videometrics **2598** (1995) 89.
28 A. Santos, C. O. D. Solórzano, J. J. Vaquero, J. M. Pena, N. Malpica, and F. D. Pozo: J. Microsc. **188** (1997) 264.
29 L. Firestone, K. Cook, K. Culp, N. Talsania, and K. Preston, Jr: Cytometry **12** (1991) 195.
30 J. V. Daele, F. Dom, L. D. Buysscher, A. Oosterlinck, and H. V. D. Berghe: IEEE Computer Society Conf. Pattern Recognition and Image Processing (IEEE, 1979) 460–465.
31 D. Vollath: J. Microsc. **151** (1988) 133.
32 D. Vollath: J. Microsc. **147** (1987) 279.
33 I. H. Witten, A. Moffat, and T. C. Bell: Managing Gigabytes: Compressing and Indexing Documents and Images (Morgan Kaufmann, San Francisco, 1999) 2nd ed.
34 R. Chen and P. V. Beek: Pattern Recogn. Lett. **56** (2015) 30.
35 M. Subbarao and J. K. Tyan: IEEE Trans. Pattern Anal. **20** (1998) 864.
36 M. Gamadia and N. Kehtarnavaz: IEEE Image Proc. (ICIP) (IEEE, 2009) 2677–2680.
37 J. M. Geusebroek, F. Cornelissen, A. W. M. Smeulders, and H. Geerts: Cytometry Part A, **39** (2000) 1.
38 M. Gamadia, N. Kehtarnavaz, and K. Roberts-Hoffman: IEEE Trans. Consum. Electron. **53** (2007) 249.
39 M. Gamadia and N. Kehtarnavaz: IS&T/SPIE Electronic Imaging. Int. Society for Optics and Photonics (2009) 724408-1-724408-8.
40 K. S. Choi, K. H. Lee, S. J. Ko, and A.W. Morales: Proc. SPIE, 1999 Nonlinear Image Processing X (SPIE, 1999) 197–206.

41 K. S. Choi, J. S. Lee, and S. J. Ko: IEEE Trans. Consum. Electron. **45** (1999) 820.
42 L. Yin, R. Yang, M. Gabbouj, and Y. Neuvo: IEEE T CIRCUITS-II **43** (1996) 157.
43 X. Xu, Y. Wang, J. Tang, X. Zhang, and X. Liu: Sensors **11** (2011) 8281.
44 Y. Xiong and S. A. Shafer: Proc. CVPR'93, IEEE Computer Society Conf. (IEEE, 1993) 68–73.
45 M. Hellstrand: US Patent 7,110,035 (2006).
46 N. Kehtarnavaz and H. J. Oh: Real-time Imaging **9** (2003) 197.
47 C. M. Chen, C. M. Hong, and H. C. Chuang: IEEE Trans. Consum. Electron. **52** (2006) 1135.
48 J. He, R. Zhou, and Z. Hong: IEEE Trans. Consum. Electron. **49** (2003) 257.
49 J. Li: Opt. Eng. **44** (2005) 113201-1.
50 B. K. Park, S. S. Kim, D. S. Chung, S. D. Lee, and C. Y. Kim: Proc. SPIE, Int. Society for Optical Engineering (SPIE, 2008) 68170D-1-68170D-10.
51 C. Y. Chen, R. C. Hwang, and Y. J. Chen: Appl. Soft Comput. **10** (2010) 296.
52 J. W. Han, J. H. Kim, H. T. Lee, and S. J. Ko: IEEE Trans. Consum. Electron. **57** (2011) 232.
53 K. Gao, N. Liu, C. Sun, and G. Ni: Proc. SPIE (SPIE, 2009) 738339-1–738339-8.
54 M. T. Rahman and N. Kehtarnavaz: IEEE Trans. Consum. Electron. **54** (2008) 1506.
55 S. Y. Lee, Y. Kumar, J. M. Cho, S. W. Lee, and S. W. Kim: IEEE Trans. Circuits Syst. Video Technol. **18** (2008) 1237.
56 Ron. Donson: http://www.shutterfreaks.com/Tips/CanonAIServo.html (accessed August 2017).
57 S. Ono, T. Yoshida: US Patent 9,618,826 (2017).
58 Nikon: http://imaging.nikon.com/lineup/dslr/basics/16/03.htm (accessed August 2017).
59 F. Albu, V. Poenaru, and A. Drimbarean: US Patent 8,508,652 (2013).
60 J. S. Lee, S. J. Ko, Y. Kim, and A. Morales: 2002 Digest of Technical Papers. Int. Conf. Consumer Electronics (IEEE, 2002) 56–57.
61 J. A. Fayman, O. Sudarsky, and E. Rivlin: Proc. IEEE Int. Conf. Robotics and Automation (IEEE, 1998) 2783–2788.
62 P. Hoad and J. Illingworhth: 5th Int. Conf. Image Processing and Its Applications (1995) 291–295.
63 A. K. Srivastava and N. Kandpal: Proc. Int. Conf. Computer Vision and Image Processing (Springer, 2017) 377–387.
64 S. Cakir and A. E. Cetin: SPIE Security + Defence, Int. Society for Optics and Photonics (SPIE, 2015) 96480A-1–96480A-7.