System implications of implementing Auto-Focus on consumer digital cameras

Thomas Yeh[1], Nitin Sampat[2], Shyam Venkataraman[3] and Robert Kremens[4]
Questra Corporation
300 Linden Oaks, Suite 100, Rochester NY 14625
[1]tyeh@questra.com, [2]nsampat@questra.com, [3]shyam@questra.com, [4]rkremens@questra.com. Tel. (716) 381-0260

Abstract

In consumer digital cameras, some of the primary tasks in the image capture data path include automatic focus (AF), automatic exposure (AE) determination and auto-white balance. There are numerous algorithms used in implementing these tasks – auto-focus is implemented using maximum contrast, ranging or sonar; white balance using color gamut determinations and "gray value estimations", and auto-exposure using scene evaluations. We evaluate the system implications of implementing one of these tasks, namely auto-focus on an embedded system – a digital camera. These include, among other things, design approach, power consumption, leverage from film cameras, and component count. Commercially available digital cameras and their choice of AF implementation are discussed where appropriate. Such an evaluation will assist, we hope, anyone designing or building a digital camera *sub-system*.

Keywords: Digital Camera, Camera Design, Camera implementation, Auto Focus

Introduction

In a digital camera image-processing pipeline (see Figure 1), one of primary tasks the system must accomplish is in the image pre-processing stage (as shown in Figure 2). This stage must set up the system for the correct focus, exposure and white balance *before* the shutter is released. The design approach and components chosen to implement these pre-processing functions have implications on overall power consumption, shutter latency (time between the pressing of the shutter and when an image is captured) and final image quality. We will discuss these and related issues with respect to one image pre-processing function – auto-focus (AF).

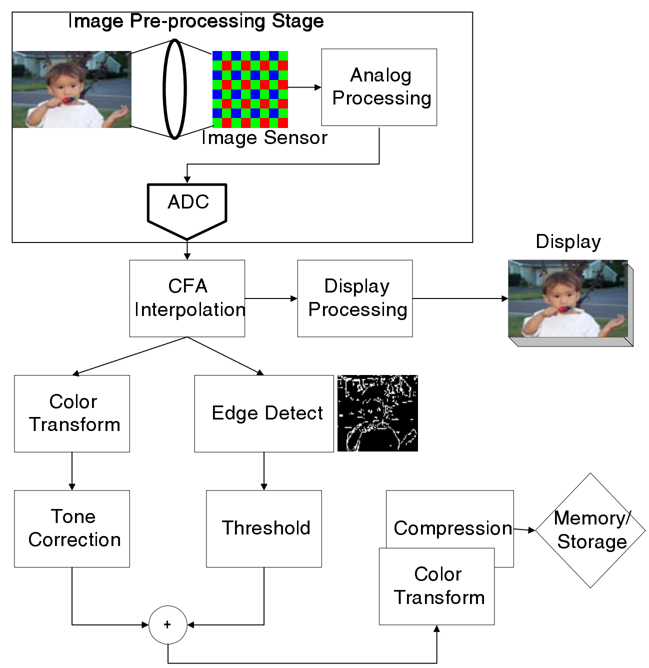

Figure 1: Image processing pipeline in a digital camera

Figure 2 is a view of the image pre-processing stage, expanded to show the details of the analog image processing chain found in a typical CCD sensor digital camera.

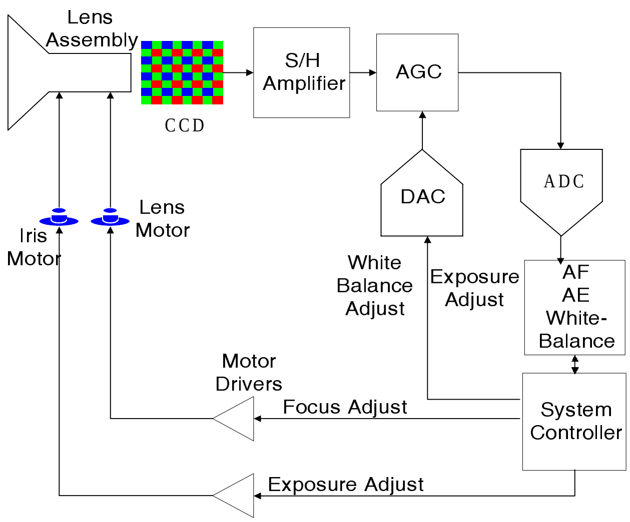

Figure 2: Expanded view of the image pre-processing stage

The Autofocus Sub-System – Background Information

Autofocus has been used in mid-range film cameras for quite some time. The use of AF is governed by two kinds of issues. On the consumer side, AF may not be used primarily because of cost. AF lenses are much more expensive than *fixed focus* lenses. On the professional side, AF may not be used more because of mistrust of the AF algorithm. Increasingly, however, most high-end cameras only come with AF lenses and their acceptance is improving – especially amongst photojournalists who are shooting sports events where manual focus may not be possible.

The low-end market, when trying to choose a lower cost alternative to AF will choose a fixed focus approach. A fixed focus lens with a wide depth-of-field (large range to object distance, that remains in acceptable focus) is used instead of an AF lens. This is typically a 35-mm wide-angle focal length lens for a 35-mm format camera. The tradeoff of this approach is that it does not allow selective focus control, and a wide-angle lens is not the optimum choice for applications like portrait photography.

Increasingly, both at the low end and high ends of the market the user/professional is demanding AF cameras and lenses. In camcorders, the existence of AF is assumed.

The basic idea behind AF is to determine the distance of the subject from the lens and move the motorized lens to correctly focus the subject onto the image plane of the optical system. The "high cost" generally associated with AF systems come from the cost of these motorized lenses.

There are numerous approaches that manufacturers have taken to implement AF. Amongst the largely popular ones, we can identify two approaches that stand out, namely:

1. Contrast-based algorithms:

These algorithms typically analyze the image and determine the highlight and shadow values. The difference between these values – the contrast – is at a maximum for optimum focus. This approach is analogous to *manual focusing* by the human eye. When a photographer focuses the image on the ground glass of the cameras, he is essentially maximizing contrast. An example of determining AF using this technique is to place 3 linear CCD's in the focal plane, in front of and behind the focal plane. The output of the CCD is analyzed for maximum signal variations (contrast). The disadvantage of this approach is that it requires high level of light and the operation is also slowed by the signal processing algorithms used. An alternative approach, which images the "focused" image through a lenslet array and determines the separation of the signals on a CCD array, overcomes the limitations of this approach (Minolta developed it in 1985). In low light or low contrast situations, these algorithms employ the use of projected beams (from an LED in the camera body) or the use of a pre-flash ("Flash-assist") to illuminate the scene.

In conventional cameras, this approach required finding an appropriate location for the sensors. In digital cameras, the advantage lies in the fact that the main image sensor itself can be used for focus sensing. One approach involves taking the image, computing its Fourier Transform and analyzing the frequency content for maximum frequency gain between frames. The image in focus will have the maximum frequency gain. This technique will also perform poorly in low-light conditions. However, image-processing techniques can be used to maximize the signal under low light conditions.

2. Ranging algorithms:

These algorithms involve scanning a scene with infrared light, which is emitted from a photodiode. The Infrared light that is reflected off the subject is used by another photo-diode to determine the distance from the cameras using triangulation. The use of Infrared allows for focusing in low and no light conditions. Some subjects, such as window glass or black objects may have very high or very low IR reflections, which generate false calculations of distance. Visible light is finding increasing use as it allows the use of visible laser that can be pointed to the subject of interest (much like handguns that have a laser pointing) and the photographer can see the subject that the beam is focusing on. This beam must be disabled when the photograph is taken (Kodak filed a patent on this approach in 1996). Many of the more recent patents extend on older patents (which generally used IR radiation) by simply changing the light to a laser source.

Used primarily by POLAROID, some ranging techniques employ the use of reflected ultrasonic sound to determined distance. A piezoelectric ceramic vibrator emits a "chirp signal" containing one or more ultrasonic frequencies. Simultaneously, a motorized lens starts focusing from its near field to infinity. The return echo, reflected from the subject, is detected and used to stop the focusing distance. The elapsed time for the return journey is proportional to the subject distance. The principle advantage of this approach lies in the fact that they can focus even in the dark. The disadvantage lies in that they will not focus through glass (or will give an erroneous reading because of the specular reflection from glass).

Review of Prior Art – Patent literature search

There are several reasons for reviewing patent literature. Prior to understanding a technical investigation or product development, a careful review of patents and other literature can help us learn what others have already achieved, to avoid repetition and reduce time-to-market. The patents are classified into class/subclass structure, which helps to cross-reference against a particular category as described in Table 1.

Category	Class	Subclass
Camera	354	288
Automatic Exposure Control	354	410+
Automatic Focus	354	400+
Camera Shutter	354	226+

Table 1: Patent Classification

Most of the AF patents filed are largely for film cameras or camcorders and relatively few for digital cameras. In the last 5 years, approximately 75% of the AF camera patents were assigned to Japanese camera manufacturers and approximately 25% to Eastman Kodak Company.

The largest group of patents concentrates on the "Maximum contrast approach". The patents all attempt to address the issue of "fast convergence" to an answer. Although the principle of this approach is well known (and well patented), the problem lies in the length of time it takes to converge to an answer. The newer patents exploit characteristics of lenses such as depth-of-field and propose a faster way to approach the problem. Some address the problem with low light (the limitations of this approach) by image processing techniques that amplify the low signal (while cleaning up the noise).

The newer projected beam approach patents concentrate on using lasers as their source instead of IR or red diodes. Also, other patents use optics (holographic elements) to project images on a wider area of the subject thereby enabling faster and quicker feedback. No one other than Polaroid Corporation seems to be using the *Sonar* approach.

Some example patent references for AF are listed at the end of this paper.

Autofocus Techniques Deployed in some Consumer Digital Cameras

The Kodak DC260 uses a visible light range measurement system. A red LED matrix projects an image consisting of a central spot and ring onto the scene (see Figure 3). As the distance to the subject increases, the size of the projected target increases. Two sensors measure the size of the target on a linear CCD or photodiode array. The input optic assembly for the detectors is tilted with respect to the optical axis of the camera.

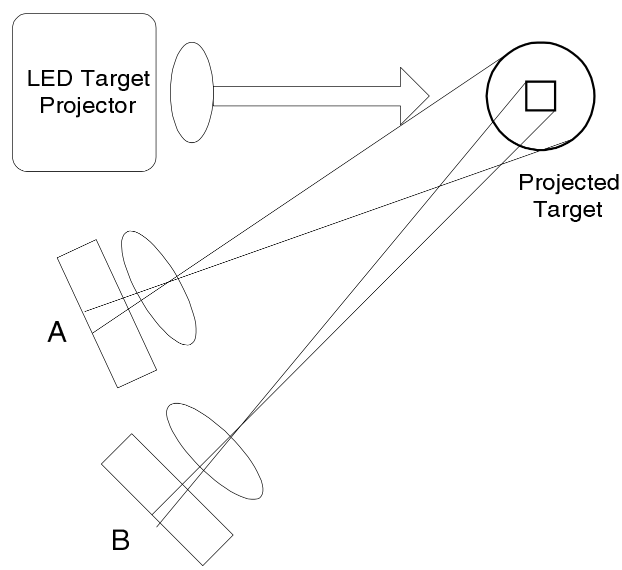

Figure 3: Visible ranging AF system in the DC 260. 'A' is a linear pixilated sensor, which measures the diameter of the projected target; 'B' is a similar detector, which measures the diameter of the inner projected image. In the DC260, the two sensors are part

The Kodak DC260 has three focus modes, selectable from the 'Advanced Focus Modes' menu from the capture menu tree. The camera uses a visible light multi-beam ranging system to determine the distance to the target. In 'Multi-Spot Auto Focus' mode, the camera uses three spots in the image plane to determine the average distance to the subject. In 'Single-Spot Auto Focus', the distance to the center of the image field is measured. These measurements and adjustments, along with ambient light metering and exposure calculation, are performed when the shutter button is depressed halfway.

Since it supplies its own light source, and is independent of the imaging system, the AF system can operate in complete darkness and still obtain accurate focus. This system is similar to, but not exactly like the IR autofocus system on many point and shoot film cameras, and Kodak's earlier DC120 digital camera.

The camera also has a manual focus setting that permits the camera to be set to fixed focus at any one of nine distances (20m, 10m, 5m, 3m, 2m, 1m, 70cm and 50cm). When in fixed focus mode, the delay between shutter press and shutter opening is slightly less than when either of the autofocus modes are enabled. By holding the shutter release in the half-depressed position, a shot can be taken almost instantly.

Independent methods to measure distance to the target, like in the DC260, entail the use of additional hardware over contrast detection focus systems that use the cameras lens and CCD to measure image contrast. The advantage of distance measurement schemes is their ability to operate in no light conditions, on subjects with inherently low contrast. Since the hardware for this system is independent from the image chain of the camera, it is also possible to make the system much faster. On the down side, the additional hardware adds cost and complexity, and can also be tricked by subjects that are highly reflective like glass or water surfaces.

The Casio QV-5000SX employs passive phase-difference detection AF, using a separate sensor pair. Similar AF sensor systems are in use on a number of middle to high-end autofocus point and shoot film cameras. Active AF has the advantage of being able to focus in low contrast or low light situations, since the light source for range determination is carried in the camera. Multi-beam active systems with multiple sensors also have fewer electronic and optical components than passive systems.

The QV-5000SX includes an easy to operate manual focus mode, which allows in-focus pictures to be taken when the AF system fails. The manual focus mode is accessed from the setup menu in 'Record' mode, and thereafter is actuated by the +/- button on the top of the camera. The rear panel LCD becomes active automatically in this mode so that a correct focus determination can be made. There are eight separate focus zones from infinity to 10 cm, so that the lens automatically switches from 'normal' to 'macro' mode.

The Casio uses standard phase detection AF ('passive AF') similar to that used in many of today's high-end point and shoot cameras. The phase detection sensor is located directly above the camera lens and consists of two plastic injection molded lenses in front of a two segment linear CCD array. Nearly identical sensors have been found on the Kodak DC260 as well as many film AF cameras. The linear CCD lies in the horizontal plane of the camera, so that the system is only sensitive to vertical lines. Since there is no auxiliary light or projected target (as there was on the Kodak DC260) in low light or low contrast scenes the camera will fail to AF.

The current trend in AF (either film or digital) is to use active infrared or NIR (near infrared) (range-finding). AF on low-cost cameras and passive (contrast-detection) systems on higher cost cameras, where a microprocessor and other necessary electronics are already available. Active systems have the advantage of rapid AF and accurate determination of focus even in zero contrast and no-light scenes, but are limited in range and suffer from interference from very bright lights. Passive AF can work at any range but requires some light and scene contrast to achieve focus lock. Some very recent APS film camera designs (notably the Canon ELPH series) use a combination of active (at short distances) and passive (at long distances), the passive system assisted at middle distances by a projected near infrared target.

Camera Product	Feature Focus
Kodak DC260	Visible light ranging AF - center weighted, AF 3 zone weighted, 9-zone manual focus. Can focus in low contrast and low light
Fuji MX-700	Contrast detect autofocus
Nikon Coolpix 900	Contrast detect AF, fixed focus at 30"

Table 2: Feature Comparison

Common Approaches for Automatic Focus & Their Implementation Issues

There are two approaches commonly used in implementing AF. For the sake of our discussion, we will refer to them as *ranging* and *contrast sensing*. The differences in the two approaches are characterized by the following design decisions:

- Power Consumption
- Shutter Latency (time from shutter press to image capture)

1. Power consumption issues:

The total power consumption for *ranging* AF system can be significantly less than that for contrast detection systems. In a contrast detection system, the entire camera imaging chain - image sensor, analog front end, and microprocessor-coprocessor chain must function during the AF cycle. Movement of a focusing lens group causes additional (and sometimes significant) energy expenditure. With the exception of the flash energy required, a contrast detection AF cycle is no different than a normal image capture cycle with regard to energy consumption. Since several AF cycles may be initiated for each captured frame, the total power consumed during the AF can be significant. Some digital still cameras that have been built use the video camcorder model and acquire (and focus) images continuously, thereby increasing the quiescent energy demand of the camera considerably. This model is especially prevalent in cameras without optical viewfinders, which of course must keep the image processing chain active for scene framing. In Figure 4, we show the relative power consumption of elements of a typical contrast detection AF system. The total power consumption listed is for the Nikon Coolpix 900 in continuous focus live view mode.

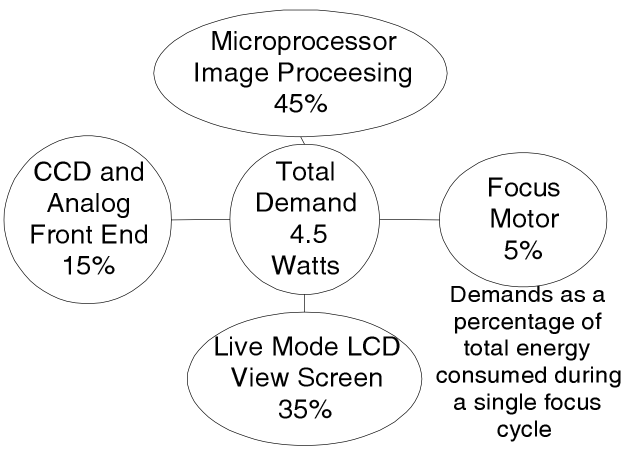

Figure 4: Power consumption breakup inside the Nikon Coolpix camera

Range finding active AF systems can be more energy efficient. Typical implementations use an LED-photodiode emitter-detector array and a simple analog circuit. Dedicated hardware or a simple 8-bit microprocessor can perform calculation of the focus condition. Many digital cameras (and all AF/AE film cameras) have a low computational power microprocessor to perform AF and AE calculations and user interface. By using this small auxiliary processor for AE and AF tasks, the main energy consumption components of the camera are powered off (CCD and power supplies, analog front end and image processing microprocessor) resulting in considerable energy savings. In active AF cameras, movement of the lens focus group consumes more energy than all other activities during focusing. Figure 5 shows the estimated power demand of an active autofocus camera, the Kodak DC260 in its normal (optical viewfinder) mode of operation.

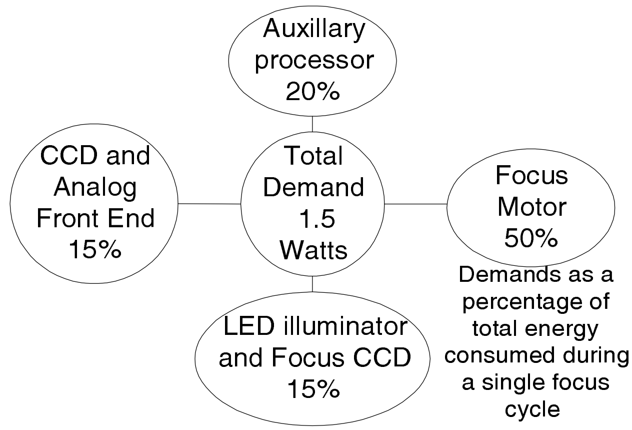

Figure 5: Power consumption breakup inside the Kodak DC-260 camera

2. Shutter Latency Issues:

Aside from power consumption, one of the most critical parameters for consumer acceptance of digital photography is the shutter latency, defined as the time from shutter depression to actual image capture. In modern autofocus film cameras, this shutter latency ranges from nearly zero to about 200 milliseconds. Current model, video camera-recorders typically require 1 to 2 seconds to achieve contrast-detected autofocus (older models taking as long as 6 seconds) which is totally unacceptable for still photography. Some early autofocus camcorder methods did not perform calculations on the image, but merely looked at the high frequency (contrast-related) components of the analog video signal.

Scores of patents have been issued to attempt to reduce the time required for autofocus using the contrast detection methodology. These patents can be broadly divided into algorithm improvement and data 'sparsing'. Algorithm improvements reduce autofocus time by reducing the number of steps required or the speed of convergence of calculations used to determine focus. Data sparsing methods are novel ways to reduce the amount of data processed, sometimes by more than a factor of 100. The ultimate data sparsing methods skip both the upper and lower portions of the image, and take only a small sample of the central portion of the image. This reduces the number of read out lines to no more than 30-50 out of the 1000 or more vertical lines of a typical megapixel CCD image sensor.

In spite of these advances, contrast detection autofocus, combined with auto exposure calculation at the time of shutter press, causes considerable shutter latency in all current digital cameras employing these methods.

Active AF methods can achieve focus lock in significantly shorter time than contrast detection methods. In film cameras using active AF, the time to achieve focus 'lock' is barely perceptible in even the least expensive models.

In light of the increased power consumption and longer shutter latency inherent in contrast detection AF methods, it is curious than manufacturers would choose this method at all. Given the short focal length and small apertures (high f/number) of most digital still cameras, which combine to produce great depth of field, the biggest disadvantage of active focus systems (less resolution in image space) is not a real issue. We wonder if the use of contrast detection auto focus systems - which are more applicable to camcorders with long focal length fast (small f/number) lenses - is just a carry over of a familiar technology to an unrelated application.

Focus Approach and Its System Implications

Using the ranging approach, the focus operation can be executed independent of other camera control and scene image capture operations. With dedicated focus control components, this approach allows the focus, exposure and white-balance controls be operated in parallel, reducing the shutter latency time for the user. The computational requirement for ranging focus approach is also a fraction of that needed for contrast sensing and places no impact to the memory or other imaging chain components.

Like film cameras that use the ranging approach, the location in the scene where the camera will focus is determined by the placement of the emitters and detectors.
Aside from power consumption, the main implications for the contrast sensing auto-focus approach are related to the successive acquisition of scene image and the calculation of scene contrast.

Prior to image acquisition, the camera must first determine the proper exposure for the scene, since as previously described, this focus technique is particularly sensitive to failure caused by low light level. This exposure may be only an initial set of settings for the lens iris and the gain value in the analog signal processing chain coupled to the image sensor. The purpose of this setting is to ensure that the signal level of the image data is within the capacity of the image sensor, and is of sufficient value to measure contrast. If the image sensor is also used to detect scene exposure, the acquisition of an additional frame may be required if the image sensor is operated continuously. This exposure setting may not be sufficiently accurate for use to capture the scene if the iris opening is adjusted during focus.

Similar to Video Cameras, the calculation of contrast requires the extraction of the high frequency content of the scene image. The algorithm moves the lens into the focus position, when the high frequency is at maximum. This requires the application of Infinite Impulse Response (IIR) or Finite Impulse Response (FIR) filters by the on-board microprocessor or by dedicated hardware. A bank of high pass filters tuned to the same frequency but with narrower pass band are commonly used in video cameras using dedicated hardware. However even with dedicated hardware this technique would still likely take excessive time to converge for still image application. Instead a single large context high pass filter coupled with limited number of high depth of focus zones may be preferred.

Since this focussing technique requires the capture of at least two image frames, the read-out time of the image data is a direct contributor to shutter latency. Image sensor with sub-sampled or fast-readout modes is useful to reduce acquisition time by skipping scan line and pixel data. This however has the effect of reducing the sampling rate and the effectiveness of the high pass filter. Ideally, only the pixel data outside the desired focus area of the scene is skipped. This would require a sensor with the ability to read only selective pixels of the sensor while skipping all others during AF operation.

Conclusions

Auto-focus is a very crucial component in the image-processing pipeline in a digital camera. AF implementation choices impact the shutter latency and power consumption. Table 3 enumerates the strengths and weaknesses of ranging and contrast sensing systems in digital cameras.

Auto focus system parameter	**Ranging**	**Contrast Sensing**
AF accuracy	Focus performance not impacted by lighting condition or scene content. Focus area in the scene limited to the location pointed to by the emitter. Potentially not functional through transparent barriers (Windows, etc.)	Requires sufficient scene lighting and contrast. Functional through transparent barriers.

Power consumption	Power only the low power dedicated components	Entire image chain used in AF calculations requiring higher power
Shutter Latency	Low	Moderate to High
AF versatility	Limited	Allow for high number of focus areas and discrimination algorithms (e.g. foreground / background, multiple objects, etc.)
Parts count	Dedicated emitters, detectors and possibly an additional microcontroller	No emitters or detectors but may require dedicated computational hardware

Table 3: Comparison of AF systems in digital cameras

References

1. Ray, S., Camera Systems, Focal Press, London and Boston, 1983.
2. IBM Intellectual Property Network - http://patent.womplex.ibm.com/.
3. U.S. Patent and Trademark Office - http://www.uspto.gov/.
4. Autofocus camera with power control for emitted beam, US Patent 5752115, Eastman Kodak Company, 1998.
5. Method and system for providing automatic focus control for a still digital camera, US Patent 5745175, FlashPoint Technologies Inc., 1998.
6. Auto-focusing device for camera, US Patent 5732288, Nikon Corporation, 1998.
7. Auto focus adjustment system and auto focus adjustment method, US Patent 5623707, Nikon Corporation, 1998.
8. Still video camera performing white-balance and auto-focus adjustment concurrently, US Patent 5619260, Asahi Kogaku Kogyo Kabushiki Kaisha, 1997.
9. Autofocus apparatus utilizing luma and chroma components to determine whether focusing is possible, US Patent 5604538, Canon Kabushiki Kaisha, 1997.
10. Multispot autofocus system having a radiation emitter which emits radiation that is divided into a plurality of discrete beams, US Patent 5569904, Eastman Kodak Company, 1996.
11. Auto-focus regulating method for video cameras and device thereof, US Patent 5477271, Samsung Electronics Company Ltd., 1995.
12. Camera auto focus system, US Patent 5416518, Sony Corporation, 1995.
13. Contrast-based autofocus mechanism, US Patent 5170202, Eastman Kodak Company, 1992.
14. Sonic autofocus camera having variable sonic beamwidth, US Patent 4490814, Polaroid Corporation, 1984.
15. Robert Kremens, Nikon Coolpix 900, Questra Technology Report, Digital Camera Series, Volume 1, Issue 6, 1998.
16. Robert Kremens, Kodak DC260, Questra Technology Report, Digital Camera Series, Volume 1, Issue 5, 1998.