やました　ゆういちろう
氏　　名　山下　雄一郎
研究科，専攻の名称　東北大学大学院工学研究科（博士課程）技術社会システム専攻
学 位 論 文 題 目　A Fundamental Study on Image Resolution Enhancement and Power Efficiency Improvement for Small Camera Module（小寸法カメラの高解像度化と電力効率向上の基礎的研究）
論 文 審 査 委 員　主査　東北大学教授　須川　成利　東北大学教授　長平　彰夫
東北大学教授　中村　健二　東北大学准教授　黒田　理人

論文内容要約

Almost 20 years has passed since CMOS image sensor (CIS) has adopted to the high-image-quality digital still camera (DSC) for consumers, while conventional cameras using silver-halide film had disappeared from the market. Now the large DSC has been transformed into a small camera module in the mobile device, i.e., smartphone, and it has been one of the primary means of the communication for the most of the people all over the world. While the market of the smartphone with small camera module expanded exponentially, firm demand supported a certain position of the DSCs in the market which is different to the case of film camera. It indicates the needs for the higher-quality imaging now only possible by DSCs, and also suggests the demands for image quality improvement of the small camera module. Among a few image quality indices, this research focuses on the image resolution, and intends to build the foundation on image resolution enhancement and power efficiency improvement of the small camera module employing CIS. The dissertation is organized into six chapters.

The introduction in Chapter 1 posed the future direction of the image sensor development based on the history of the imaging device and sensors. The history of the image sensor transitioning from image tube to CCD, and then, CMOS image sensor (CIS) told us that the innovations have been driven by the needs for the more pixels, higher frame rate, higher signal-to-noise ratio, and better color fidelity. The history of the imaging device confirmed the past movements such as the end of film-camera by the rise of digital single reflex camera and digital still camera, dismissal of the NTSC/PAL video camera by emergence of the HDTV format broadcasting and HD video camera, and the golden age of imaging by the mobile device enabled by the small camera module equipped with backside illuminated CIS. Those events indicate that the compact imaging device with acceptable or better image quality is the key enabler for the imaging innovation. Considering several expected scenarios among uncertainties, the improvement of the spatial resolution of small camera module was identified as the direction of the research. It is indispensable to increase the number of pixels of the CIS for improving the resolution of the image of the small camera module, but the camera module also has the restriction on the optical format of the pixel array and the acceptable power consumption. The requirement for more pixel with a constant optical format, due to the constraints from the nature of the small camera module suggested two key development directions: the pixel shrink, and the power efficiency improvement.

The chapter continued clarifying the challenges on the pixel shrink and reviewing the existing CIS power consumption studies, followed by the identification of the directions. The estimation based on the general physics pointed out the issue of increasing color cross-talk by the diffraction, caused by the micro-lens when the pixel was shrunk to around 0.5μm, and also the difficulty to suppress it even the light-pipe structure was used due to its ineffectiveness by the evanescent wave. With regard to the power efficiency improvement, the current situation where there was no systematic study on the power efficiency was confirmed, along with the reviews on the existing CIS FoM and the power efficiency FoM for ADC.

Finally, the challenges on the pixel shrink and power efficiency improvement were broken down to four sub-topics to be discussed on the thesis; 1) to clarify the factors defining the practical limitation of the spatial frequency of the camera; 2) a proposal and feasibility study on the image resolution improvement utilizing the sub-half-micron pixel (the pixel of the pitch and size comparable to or less than around 0.5μm); 3) to build a theory to realize the reduction of power consumption and the improvement of power efficiency for image sensors; 4) to realize the power efficiency improvement of CIS based on the theory.

Chapter 2 and 3 were dedicated to the study of high-resolution imaging by sub-half-micron pixel.

Chapter 2 first summarized the cause of resolution reduction to provide theoretical and quantitative fundamentals for the improvement of image spatial resolution on electro-optical system of small camera module. They were classified into two parts: the primary and secondary. The primary factors include the effect of diffraction, pixel aperture, sampling and color interpolation, which equally affect all the cameras, whereas the secondary factors are color cross-talk and lateral chromatic aberration, which mainly influence the small camera module.

The analytic MTF based on the diffraction optics quantitatively confirmed the benefit of using the sub-half-micron pixel, showing better numbers on the following indices, i.e., MTF_{200} (An amplitude of MTF curve at 200lp/mm), MTF_{400} (the MTF amplitude at 400lp/mm), and f_{SPR} (the maximum frequency without the influence of spurious). The improvement on analytical MTF was further compared with the slanted-edge based SFR MTF analysis with the image simulator. The benchmark revealed the good match between the analytical MTF curve and the MTF curve extracted from the simulated slanted-edge image, supporting the validity of analytical calculation as well as the image simulator.

The real improvement in the spatial domain, however, was unclear, since there existed no known correlation study between the values of MTF plot in the frequency domain and its visibility in the spatial domain. The structural resolution analysis (SRA) was proposed to resolve the situation, which substantiates the resolution limit directly in the spatial domain. The concept of SRA is based on the application of structural similarity index (SSIM) to compare the two images: the ideal simulated image depending only on the Rayleigh limit and Nyquist limit, and the simulated image including the effect of the primary and secondary factors. The result of analytical MTF and SRA with the primary factors were benchmarked to understand the correlation between the frequency domain and spatial domain, and it clarified that f_{SPR} and the resolution limit in SRA analysis showed strong correlation. This result gives us the following conclusion upon evaluating the spatial frequency improvement using MTF and SRA: SRA gives the same conclusion as stated by the Nyquist-Shannon sampling theorem while slanted-edge based SFR MTF analysis does not incorporate the effect of aliasing. In other words, slanted-edge based MTF tells us the relative improvement of image resolution if there is no repeating pattern in the actual image, whereas SRA provides the worst-case image resolution limit under the influence of the aliasing. This result also suggests that it is cost-effective to match the $f_{Rayleigh}$ and f_{SPR} to balance the benefit to the image resolution gain and the manufacturing cost of the CIS with sub-half-micron pixel and the fast lens for the small camera module. With the lens with F-number of 2.0, the pixel size of 0.8μm was optimum, and the pixel size of 0.5μm was enough when the F-number was 1.5.

The conclusion did not change under the pretense of the moderate photon shot noise, dark random noise, provided that the normalized performance indices of the pixel were constant, i.e., the pixel follows the scaling rule on its full well capacity and sensitivity. This result indicates the necessity of maintaining normalized pixel performance upon sub-half-micron pixel technology development.

The simulation with a sub-half-micron pixel did not degrade the image resolution when a large cross-talk existed, due to its inherently high pixel MTF. The simulation with all the secondary factors, however, revealed that the image resolution does not improve as expected due to the complication of color cross-talk and LCA. For example, simulations showed that sub-half-micron pixel of 0.5μm could achieve 700lp/mm with the lens F-number of 1.5 when only the primary factor included, while it provided only 200lp/mm when the complication of the secondary factor exists, and the further pixel shrink did not help to improve the response.

Since the effect of LCA increases in the small camera module and the cross-talk becomes more evident in sub-half-micron pixel generation, innovation is necessary to enhance the image resolution of the small camera module.

In Chapter 3, the concept of the preserved color pixel (PCP) was proposed as a solution to minimize the degradation of image resolution and saturation when the cross-talk and LCA exists. The supporting image processing algorithms: similarity-based blind cross-talk correction and adaptive normalized convolution by post-convolution-variation-minimization (PCVM) were developed as well. The feasibility was tested with image simulations.

The PCP arrangement consists of pixels called PCP and surrounding “buffer pixels” with sharing the color filter of the same color, and the pixels of the identical color plane constitute non-uniform sampling points. The idea to gain a benefit from PCP is straightforward: the buffer pixels absorb color cross-talk components from different color channels

so that the PCP preserves the true color, and the data of PCP are used to cancel the color cross-talk of the buffer pixels in image post-processing.

The similarity-based blind cross-talk correction was developed to restore the color information of the buffer pixels. This algorithm estimates the unknown cross-talk coefficient using the guide image created only by the signals of PCPs without prior information. It creates the system equations to calculate the coefficient based on the assumption that the pixels on similar locations are influenced by the same degree of crosstalk. The mathematical foundation was given in the thesis, together with its cross-talk model. The post-convolutional-variation-minimization (PCVM) algorithm was proposed as a method to select the optimum kernel upon the application of adaptive normalized convolution. It adaptively and locally selects the optimum convolution result from the set of candidates created with different kernels (applicability functions) based on the assumptions that 1) the optimum kernel on a location does not distort the original image after the convolution, and 2) the optimum kernel is also optimum to reconstruct the missing pixel information near the location.

The image simulation with the PVCM algorithm demonstrated a comparable SRA frequency limit as that of conventional Bayer CFA arrangement, whereas additional cross-talk correction process showed almost perfect recovery on the color component with degrading the SRA frequency limit from 700lp/mm to 400lp/mm, suggesting side-effect of the similarity-based cross-talk correction.

Compared to the results of simulations in Chapter 2 with the conventional Bayer arrangement with the presences of the secondary factors, e.g., 32% color cross-talk and 1.5um LCA, the PCP arrangement showed superior SRA frequency limit than the conventional approach when the pixel size was less than 0.5μm, which demonstrates the effectiveness of sub-half-micron pixel.

Chapter 4 and 5 were for the study of the power efficiency improvement for CIS, which is of primary importance to enable the high-resolution imaging with the small camera module using sub-half-micron pixels. Chapter 4 focused on the understanding of power efficiency, whereas an implementation example with 3D-stacked image sensor technology was exhibited in Chapter 5.

In Chapter 4, the power efficiency Figure-of-Merit (FoM) was introduced to understand the power efficiency, to benchmark the different design, and to identify the knob to improve it. It can be widely applied to most of the CIS as it requires only the pixel array size, frame rate, and total power consumption to derive the FoM index. The generalized CIS model was defined to describe different CISs using the common parameters. The model successfully provided the unified relationship of the power consumption, the number of rows and columns, and the frame rate for six different kinds of CIS read-out configurations. Moreover, the power-law power consumption model was introduced to express the power consumption of the block circuit with its operating frequency. Based on the two models, the number of rows, columns and the frame-rate of an image sensor were linked to the power consumption with a non-linear coefficient.

To derive the actual value of the non-linear coefficient, the power consumption of the image sensor was expressed in the mathematical equations with the process parameters (e.g., transistor gain: β) and the design parameters (e.g., transistor size), and the FoM was calculated for the various numbers of pixels between HDTV – 8K and frame rates between 60fps – 480fps with the range of candidates of non-linear coefficients from 1.0 to 2.0. The numerical analyses showed the optimum non-linear coefficient of 1.5 to deliver the nearly-constant value of FoM for the CISs made with the same process and design parameters.

The formulation of FoM also told that the power efficiency could be improved by increased transistor gain or reduced parasitic capacitance with the use of advanced technology, the reduction of operating frequency by employing sub-ranging column parallel ADC and its pipelining, and the increased parallelism factor with adapting 3D-stacked image sensor technology or the optimization of the layout. The findings were validated by reviewing FoMs of the 48 CIS studies reported from 1999 to 2018.

Chapter 5 focused on the feasibility study of the application of 3D-stacked image sensor technology for image sensors, and the demonstration of power efficiency improvement with it. The studies on the global shutter, single photon avalanche diode (SPAD) and Quanta Image Sensor (QIS) proved the readiness of the 3D-stacked sensor technology. Moreover, the 1.1μm-pitch, 33M pixels, 240fps, 3D-stacked image sensor successfully demonstrated the notable power efficiency improvement with 3-stage pipelined ADC in subarray column-parallel ADC readout configuration.

In the first half of this chapter, the several applications of 3D-stacked image sensor technology were explained including the author's work to confirm its effectiveness for the future of imaging. The history of "non-stacked" CMOS image sensor, and TSV-stacked CIS technology were briefly reviewed with their limitations, followed by the descriptions on 3D-stacked image sensor technology including the author's studies e.g., global electronic shutter CIS, LiDAR (Light Detection and Ranging) with single photon avalanche diode, and quanta image sensor (QIS) as tangible demonstrations enjoying the benefit from the technology. The application of 3D-stacking technology for the global shutter application realized unique features, such as excellent PLS, in-pixel ADC or very large full well capacity of 220Ke- for wide-dynamic-range function, but so far it lags the conventional global shutter CIS in terms of its pixel size or noise performance. The use of more advanced technology finer than 65nm CMOS, and the new readout architecture to enable compact implementation with low noise level is expected in the future. The SPAD chip employing 3D-stacking technology enjoyed the full benefit of the technology, i.e., better SPAD device performance due to the increased freedom of process tuning, and smaller and high-performance pixel using the advanced technology on the bottom-tier. The 3D-stacked LiDAR with 65nm CMOS technology realized 19.8μm pixel pitch with the 64-to-1 decision tree and shared TDC. This approach is, however, supposedly only effective for high-spatial resolution LiDAR, and economically not feasible for low-spatial resolution one due to the die-matching problem. The application to the commercial market needs careful considerations from the cost point of view. The 1.1μm pitch, 1M-pixel QIS successfully reproduced the image of 128x128 pixels with 8x8x8 cubicle summation using the cluster-parallel parallel readout architecture. The architecture defined a subarray of 256x16 pixels and it contained 8 column-parallel CDS, 8-to-1 multiplexer and one ADC operated at 4Msamples/s, which was enabled by the use of 3D-stacked sensor technology. The studies on above three different categories of sensing applications suggested that the use of 3D-stacking technology opens the door of the innovation by enabling the new design concept utilizing the increased layout freedom and the more advanced technology on the bottom-tier.

The second half of the chapter explained the importance of 3D-stacking technology for power efficiency improvement by enabling the new read-out architecture called "subarray column parallel ADC readout configuration," where the architecture was explained in detail with the quantitative analysis on power efficiency improvement. In the conventional non-stack approach, the shrink of the pitch of the column ADC was not straightforward since the use of advanced technology was not always possible. TSV-stacked sensor technology resolved the problem; although there remained the difficulty to shrink the analog circuit pitch. The 3D-stacked sensor technology eventually realized the subarray column-parallel ADC to place the analog readout under the pixel array area. The analog circuits could be small not only by the advanced technology, but also the reduced unused area – margin - by employing the smaller aspect ratio of the circuit layout. The demonstration by 1.1μm-pitch, 33M-pixel, 240fps CIS showed a similar power consumption compared to the previous design, despite the use of a pixel of 60% smaller pitch and the doubled frame rate. The estimate also showed that the use of 3D-stacked technology together with new 3-stage pipelined column ADC and 65nm CMOS process on the bottom-tier could have helped to reduce the power consumption by 60%. As a result, the fabricated CIS realized the world second best FoM with the data-rate of 8 billion pixels per second.

The 3D-stacked CMOS image sensor ADC configuration will be of primary importance to pursue higher spatial resolution imaging with sub-half-micron pixel while improving the power efficiency since it is so far only technology to increase the parallelism factor of the readout circuit either by pixel level or subarray level connection.

This research was dedicated to the fundamental study on enhancing the image resolution of the small camera module for mobile devices. The PCP concept has posed the possibility of resolution increase under the influence of the color cross-talk and LCA to researchers working on the imaging, and it is expected that this proposal will drive the trend of the pixel-shrinking down to sub-half-micron pitch. It is also anticipated that the power efficiency theory discussed in this thesis will contribute to resolving the inevitable side-effect of the sub-half-micron pixel approach, that is, increasing power consumption, by facilitating the discussion of the circuit designers and researchers based on the universal language of the CIS power efficiency.