DISSERTATION

A FREE SPACE OPTICAL INTERCONNECT BASED BIOSEQUENCE ANALYSIS SYSTEM

Submitted by

Tao Ao

Electrical and Computer Engineering

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2003

Colorado State University

July 31st, 2003

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY TAO AO ENTITLED "A FREE SPACE OPTICAL INTERCONNECT BASED BIOSEQUENCE ANALYSIS SYSTEM" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Advisor

Department Head

Abstract of A Free Space Optical Interconnect Based Biosequence Analysis System

Free space optical interconnects (FSOI) are of considerable interest for their potential of parallel high bandwidth optical data links. In order to demonstrate the feasibility of applying FSOI technology in complex information processing systems, a prototype FSOI based biosequence analysis system was constructed and tested. This thesis reports the design, setup and testing of the FSOI based biosequence analysis system, which has 64 optical channels that feeds optical signal to a custom designed biosequence processor chip. The system implements the basic local alignment search tool (BLAST) algorithm using a biosequence processor chip and a programmable logic device. The biosequence processor chip has a 8×8 comparison array, in which each basic comparison unit has a built-in optical input. A diffractive optical elements multiplies a modulated laser beam into an 8×8 spot array and a projection lens relays the spot array onto the optical inputs of the biosequence processor chip. Four major system development tasks were achieved:

1. System algorithm and architecture design;
2. Development of a FSOI processor chip for the biosequence comparisons;
3. Optics system design and setup;
4. Functionality and reliability tests
5. Biosequence analysis.

The FSOI biosequence analysis system successfully analyzed homologous and inhomologous biosequences. The performance of the proof of principle system is

compared with present electronic system. The potential of the FSOI system concept is also investigated.

Tao Ao

Electrical and Computer Engineering Department

Colorado State University

Fort Collins, CO 80523-1373

Fall 2003

Contents

Chapter 1

Introduction of Free Space Optical Interconnect

Modern information technology has a fundamental impact on almost every aspect of industry. According to the International Technology Roadmap of Semiconductors published by the Semiconductor Industry Association (SIA), the Moore's law has been valid for the last four decades.[1] Thus the number of transistors integrated on a single chip and the operation speed have increased dramatically. Although new materials, structures and manufacturing techniques are being constantly pursued and introduced, interconnects both inside and outside the IC chip are facing a serious challenge. [1][2][3][4] [5][6] David Miller concluded that the overall interconnect capacity inside an IC chip is limited due to the fact that the parasitic resistance and capacitance can not be reduced as the IC feature size scales down. [2][3][4][5][6] Free space optical interconnects may be an alternative for providing higher interconnect capacity.

In this chapter the reasons for developing free space optical interconnect (FSOI) technology are presented. The current status of FSOI technology, including device

technologies, architectures, and design issues, are introduced.

1.1 Reasons of free space optical interconnection (FSOI)

As the size of IC transistors shrinks, the top speed of the transistor is also increased. However, the speed of the metal interconnection lines can not be proportionately increased. The signal propagation delay caused by interconnect lines is much more significant than transistor delay. In the simplest transmission line model,the metal interconnect is considered as a lumped RC (parasitic resistance-capacitance) network. In order to send an electronic pulse from one end of the RC network to another end, the capacitance along the transmission line must be charged through the resistance. If two pulses are sent too close in time, they may not be distinguishable, i.e., the top transmission speed is limited. The transmission delay time is proportional to the lumped RC constant therefore it can be used to measure the delay time and the transmission speed. Because the charging process takes time and energy, the top speed that an electrical signal can travel through the line is limited. If the cross-sectional area of a metal interconnect is reduced to half of the original area, the capacitance would be reduced by half while the resistance is doubled. The overall RC time constant remains the same despite the fact that the cross-section of area is reduced, i.e., the top transmission speed of the metal line stays the same. Furthermore, as the density of the transistors increases, it would take more energy to connect all the transistors in a given area because the energy to drive each metal line remains the same and the number of the metal lines is increased. There are methods to reduce the resistance and capacitance by improvement of materials and

processing technology including copper lines and low-k materials. There also are methods to increase the overall space for interconnect lines in the IC chips. [1] However, it seems that many known methods of improving interconnect capacity will be exhausted without foreseeable replacements in about one decade. [1] In general, interconnect capacity inside the chip would not be able to match the demand for faster and smaller transistors.

The electronic signals travelling outside the IC chip face a harsher environment than on chip since the transmission length is much longer and the parasitic resistance and capacitance are much larger compared with those inside the chip. The transmission speed between the IC chips is significantly slower than that inside the chip. The printed circuit board (PCB) is currently the major means of assembling and interconnecting electronic components. It is unavoidable that multi-layer PCBs must be used to meet the demand for higher speed and density. The multi-layer PCBs are generally complex and expensive in terms of both designing and manufacturing. Besides, in order to provide sufficient I/O capacity at the interface of the IC, the packaging of the IC must include more and more I/O pins with less and less parasitic resistance and capacitance. Though the cost of manufacturing IC chips is dropping, the cost of high capacity IC chip packaging is increasing. Although the progress in these areas would drop the overall cost to some extent, there is no Moore's law in these areas, i.e., interconnect capacity outside the chip will not match the demand of the fast developing IC technology.

Free space optical interconnects (FSOI), on the other hand, have quite different properties. The most desirable properties of the FSOI technology are the parallelism capability and the 'bit-rate transparent' transmission path.

For FSOI the signals carried in modulated light beams (e.g. laser beams) that

travel through free space (e.g. air). The optical signals can be easily multiplied by using diffractive optical elements. Combined with integrated photodetector array technology, it is possible to construct a large number of optical inputs on an IC chip with complex logic design. Therefore, it is feasible to distribute the signals to a large number of individual receivers in parallel. Large optical fan-outs can be achieved this way. Furthermore, such optical inputs consume less power and chip estate than the electronic I/O bonding pads. On-chip optical outputs are also available with integrated semiconductor laser array and modulators.

In FSOI systems, there is no parasite-caused delay similar to that in the electronic transmission lines. The optical energy loss along the propagation path is related to the transmission loss of the transmitting medium, not the data rate being carried in the optical signal. The transmission speed of the interconnections is only limited by the speed of the optical transmitting and receiving devices in free space optical interconnects. The transmission path is 'bit-rate transparent', i.e. the optical interconnects are not sensitive to the bit-rate of the transmitting data stream. One may upgrade the transmission speed in the system without modifications of the optics.

In summary, free space optical interconnects promises to deliver board-to-board and chip-to-chip interconnects. Optical interconnects may provide considerably higher transmission capacity than current PCB technology because of the parallel optical fan-out and the 'bit-rate transparent' transmission path. This research work investigates the potential of FSOI technology by demonstrating a prototype biosequence analysis system. The prototype system developed in this project is not fast due to the limitation of photodetector design. However, the successful demonstration of the biosequence analysis system and the scaling calculations indicate

that further improvement on system speed and unit numbers can make FSOI based system comparable with the state-of-the-art electronic computers.

1.2 Technologies for FSOI

The concept of the free space optical interconnect (FSOI) technology was first proposed in early 1980's. [7][8] The FSOI technology has been constantly improved since then. Many relevant technologies, including new devices and processing techniques, have found applications in the FSOI systems. System architectures that are suitable for FSOI based systems have been introduced in several FSOI system prototypes. FSOI system demonstrations have shown impressive capabilities for board-to-board and chip-to-chip level data links. The following paragraphs review integration of active and passive photonic devices with silicon ICs.

The major objective of FSOI technology is to provide high speed, high density interconnects for electronic ICs. In order to achieve such objective, it is logical to place the optics and the electronics as close to each other as possible. Because the interconnections between the optics and the electronics are electronic based, these interconnection paths can not stretch too long otherwise they would become the bottleneck and this would defy the purpose of optical interconnection.

Much attention has been given to the integration of optics with electronics due to its potential in multiple areas including optical communication, optical storage and many other applications in which fast, reliable and low-cost electronic-to-optics (E/O) and optics-to-electronic (O/E) conversion are needed. The optical-electronic integrated circuit (OEIC) is the technology that can provide such devices.

The integration techniques are based on knowledge accumulated from research on semiconductor and optics device processing. Some photodetectors are fabricated

on the silicon substrate with processes compatible with complementary metal oxide semiconductor (CMOS) technology. Special processing technologies and substrates also allow optical devices homogeneously integrated with logic circuits. However, due to the fact that many different types of optoelectronic devices are not made of silicon or GaAs, heterogeneous integration is important for successful integration of optical transmitters, optical receivers, passive optical components and electronic logic circuits. Techniques for integration of vertical cavity surface emitting laser (VCSEL) and CMOS IC and integration of photodetectors with CMOS IC are available. [9] The following paragraphs briefly review an integrated optoelectronic array as an example of such integration techniques.

Honeywell is known for its distinctive achievements in VCSEL and smart pixel array (SPA) products and serves as an example of silicon IC integration with optoelectronic components. Honeywell's SPA technology heterogeneously integrates VCSELs, metal-semiconductor-metal (MSM) photodetectors, and micro lenses with Si-based IC. The integration process flow of the SPA is illustrated in Figure 1.1. [9] The first step manufactures VCSEL and MSM arrays on a GaAs wafer. The 850 nm VCSELs are patterned and insulated by ion implantation so that the top surface of the VCSEL wafer remains planar. The MSM photodetectors are made from a layer of undoped GaAs grown on top of the VCSEL layer. The MSM photodetector GaAs layer was removed from where the VCSELs are patterned. A rigid and transparent 'superstrate' layer is applied on top of the optical devices. This 'superstrate' layer is transparent to 850 nm light. Because the GaAs substrate is opaque to 850 nm light and the micro-optics components must be on top of the VCSEL and MSM photodetectors, the electronic contacts must be accessible from the bottom of the active optical devices. To accomplish this, the GaAs substrate is removed to expose the

bottom of the optical devices. After the substrate is removed, the 'superstrate' layer holds the optical devices. The micro-optics components can be directly fabricated on the 'superstrate' surface or be manufactured on separate wafers and aligned with the 'superstrate' surface after testing. Once the micro-optics is in place, the whole wafer is bump bonded to the Si IC.

The Honeywell integrated array is an example of a novel approach to optoelectronic processing. In order to take full advantage of such integrated devices, OEIC packaging must accommodate optical and electronic signals I/O and power supplies.

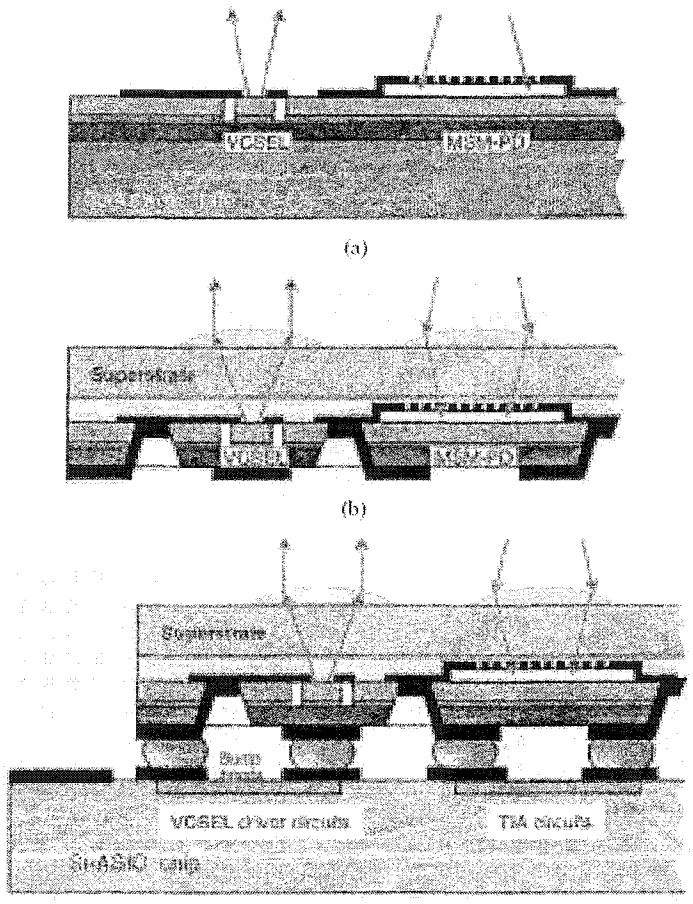


Figure 1.1: Process flow of Honeywell SPA [9]

University of California at San Diego (UCSD) and Honeywell have been working on the "three dimensional optoelectronic stacked processor" (3D-OESP) with DARPA support. [10][11] Based on the Honeywell SPA integration technology, the stacked OEICs have VCSELs as optical transmitters and metal-semiconductor-metal (MSM) photodetectors as optical receivers. As illustrated in Figure 1.2, the optical devices are on the top planar surface of the stacked devices. Macro optical components are used to steer the light beams to the mirrors. The light beams are reflected to the designated stacked processor by the reflection mirrors. The macro lens and the prism then project the light beams onto the photodetectors. The VCSEL and the MSM photodetector can operate at high speed (e.g. 250 Mb/s [11] or 1Gb/s [10]). The overall system transmission capacity is quite impressive due to the large number (1024 or more) of devices and the operation speed of each device.

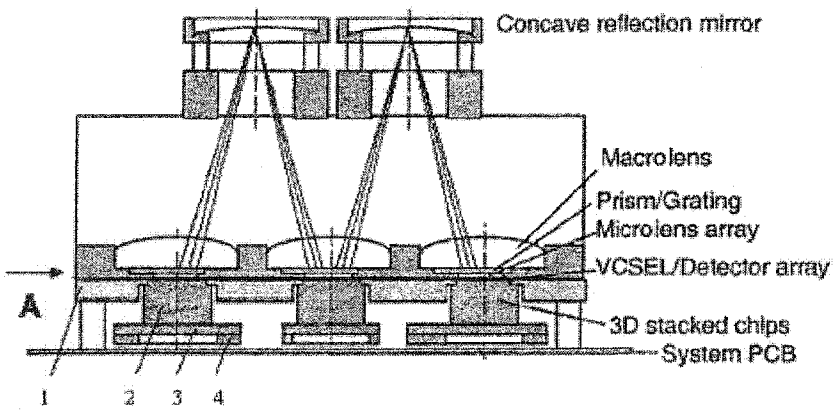


Figure 1.2: The 3D-OESP structure [11]

Because the positions of the optical devices are identical on each stacked processor, the mapping of the connections is fixed, i.e., the optical signal from a certain VCSEL only goes to one MSM photodetector on the corresponding position. The fixed mapping could limit the flexibility of the system, especially in complex sys-

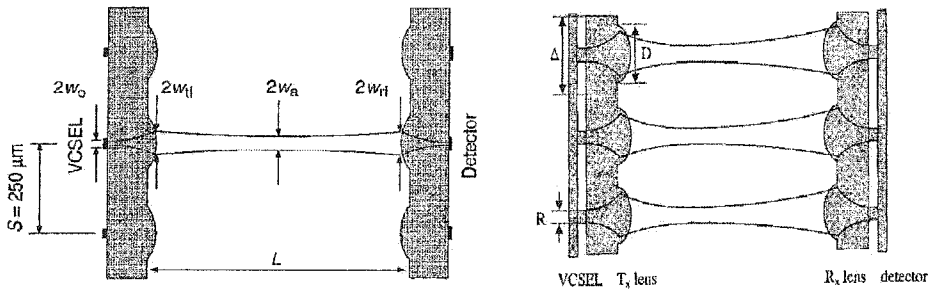
tems where the transmission pathways must be routed according to the need of logic design and timing constraints.

Other examples of optoelectronic device integration with Si IC for FSOI systems include multi-quantum well (MQW) modulators [12][13] and VCSELs. [14][15] In addition, many companies now have the potential to produce integrated optical transmitters and receivers for optical communication networks. If the integration technologies are openly available at reasonable cost, FSOI systems can employ planar arrayed optical transmitters and receivers to greatly simplify the system structure, construction complexity and overall, cost.

1.3 FSOI system structures

The structure of a FSOI system is determined by functionality and the technologies employed in the FSOI system. Several types of FSOI system have been reported. These system structures vary from straight-line 2-D point-to-point links to complicate multiple-stage crossbar rings. Each structure has its advantages and shortcomings.

The simplest optical interconnect in free space is the 'line-of-sight' type. The light beam from an optical transmitter is directly projected onto the photodetector surface of an optical receiver. The optical path between the transmitter and the receiver is a straight line without any optical component other than collimating lenses. One can build a multi-channel optical link by putting several 'line-of-sight' optical interconnects side by side and align their optical paths in parallel. As illustrated in Figure 1.3, two types of parallel optical link were reported. [20][21] Both systems use the simplest design using only collimating lenses. In order to maximize the transmission length, the waist of the Gaussian laser beam is designed to be in



(a) (b)

Figure 1.3: Two parallel optical link using 'light-of-sight' optical interconnect [16][17]

the middle of the propagation path. Figure 1.3(a) represents the schematic of a free space optical link between a VCSEL and a photodetector. [16] The VCSEL is a bottom emitter operating at 980 nm. The GaAs substrate is etched to the desired lens shape to form a microlens array. The VCSEL is flip-chip bonded on a microwave compatible board. The photodetector is prepared in a similar manner. The test result revealed that this type of optical interconnect is suitable for operation at 500 Mb/s over a distance of 5 to 10 mm. Figure 1.3(b) illustrates an array of microchannel optical links. [17] A 2×8 array of VCSEL in the configuration is commercially available with a 10 μm aperture and 250 μm pitch.

The 'line-of-sight' design is quite straightforward in design and implementation. It is useful for small prototype and proof-of-concept FSOI systems. However it is difficult to construct a large scale system based on such 'line-of-sight' links, because the transmitters and the receivers must be placed a short distance from each other.

In more sophisticated FSOI systems, it is common to use optical components to manipulate the light beams travelling in free space. The light beams can be steered, reflected and relayed such that the optical transmitters and receivers can be placed as desired without physical constraints of distance or position.

For chip-to-chip interconnects, the optical transmitters and receivers are usually placed on the same plane, i.e. the same printed circuit board. Based on their research on OEIC packing techniques (partially discussed in Section 1.2), UCSD has been working on such chip level interconnects for years. [10][11][18][19] As illustrated in Figure 1.4, each chip (a optoelectronic stacked processor OESP) has an interlaced 16×16 VCSEL/MSM photodetector array. A mirror is used to direct the optical signals from the transmitters on one chip to the receivers on the other chip. The optics components are in the center of the board and multiple capacitors are on both sides of the board. Each optical link can operate at up to 1Gb/s. With multiple chips on one board, the overall performance of the system could reach the 100 Gb/s range.

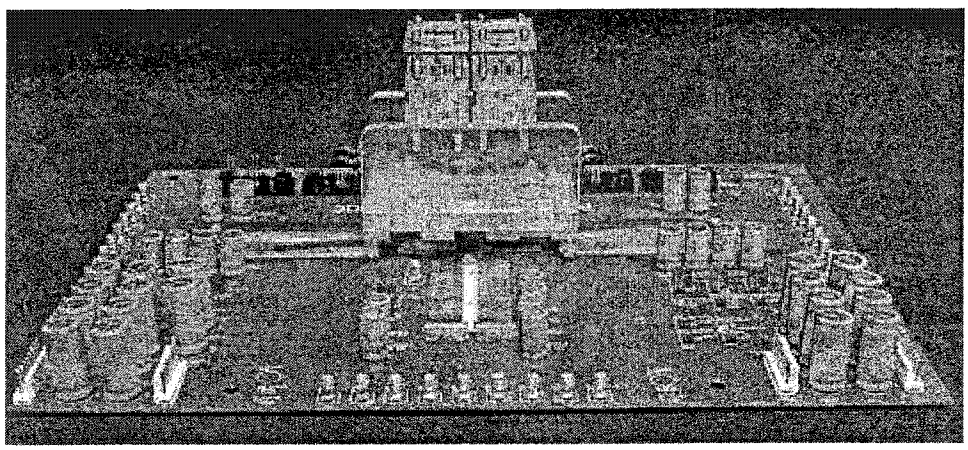


Figure 1.4: An assembled board level optical interconnect system by UCSD [10]

For board-to-board interconnects, the transmitters and the receivers are usually on different circuit boards. In many cases, these circuit boards are placed in parallel and fixed by mechanical means. Different types of beam steering and relaying techniques can be used in board-to-board FSOI systems.

Figure 1.5 illustrates a demonstration of board-to-board FSOI system using

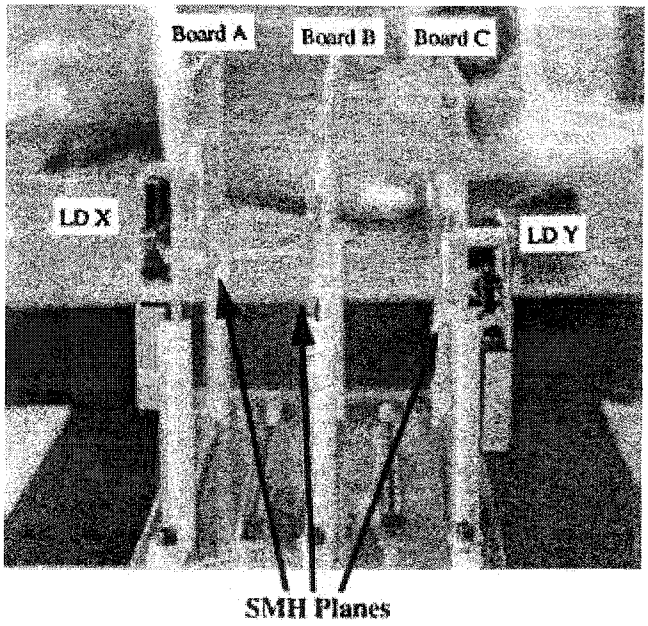


Figure 1.5: A board-to-board FSOI demonstration using SMH components [20]

substrate-mode hologram (SMH) components. [20] The SMH can be electronically controlled to multiplex inbound light beams and project them onto the desired optical receivers. The concept of the system can be used in optical clock distribution and optical bus implementations.

If the routing topology requires more complex structures, polarization beam splitters, prisms and liquid crystal (LC) components are often used to steer the optical beams. As illustrated in Figure 1.6, Nippon Telegraph and Telephone (NTT) designed a high capacity ATM switch based on board-to-board FSOI technology. [21] The ATM switch is a multi-stage routing network that consists of small-scale electronic ICs for ATM switching. Each printed circuit board has one complete stage, including application specific integrated circuit (ASIC) based packet switches, controlling logic and optical data link. Because the system is designed to provide very high capacity (100 Gb/s and more), the FSOI technology is needed to provide

data links between the adjacent switching boards.

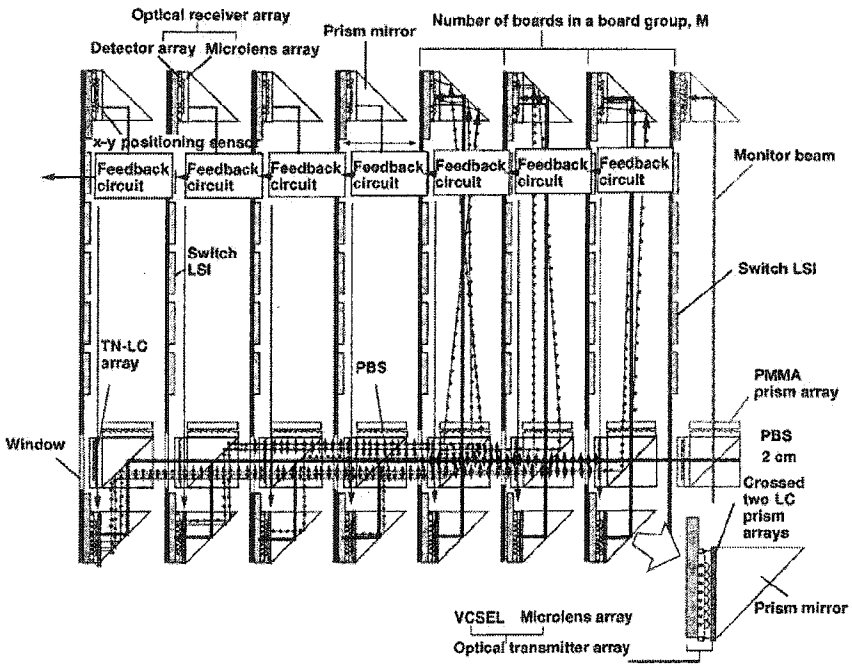


Figure 1.6: A board-to-board FSOI system for ATM switching [21]

The distance between the boards is around 30-50 cm. The optical transmitters are arrays of VCSELs. In order to realize the complicated routings required in the multi-stage routing network, sophisticated LC polarization-controller arrays are used to control the polarization of each light beam so that the prisms and the polarization beam splitters (PBS) can direct the light beams to their destined optical receivers.

Optical backplanes also can be constructed with FSOI technology. As a part of a bigger information processing system, the optical backplane functions as an optical data bus carrying all the high-speed data links while the rest of the system can access the data link.

A four-stage, unidirectional ring free-space optical interconnect system is illustrated in Figure 1.7. [22] The network topology of the system is essentially a crossbar

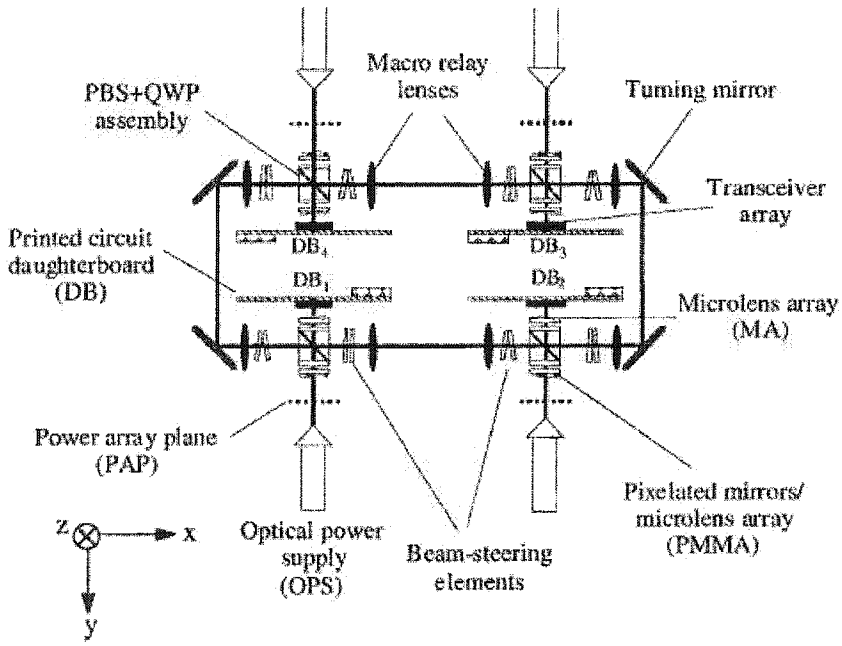


Figure 1.7: A FSOI based 4-stage unidirectional ring optical backplane [22]

structure that provides non-blocked data links from any given access point to all other access points.

The optical backplane uses CMOS –self-electro-optic-device-based (CMOS-SEED) devices as optical modulators. Functioning as an optical power supply distributor, a holographic spot array generator is used to generate an array of optical inputs from a single laser light source. The optical components are aligned using *in-situ* active alignment methods.

Just as the wavelength division multiplexing (WDM) technology is being widely applied in fiber optic telecommunication networks, similar wavelength based technique can be used to route the light beams in FSOI systems without complex polarization selection and controlling devices. In these systems, the essential components are the wavelength sensitive optical devices, including optical gratings, holograms

(essentially 2-D grating), interferometers, etc.

Researchers in University of Tokyo studied the feasibility of using micromachined tunable interferometers in the FSOI system. [23] As illustrated in Figure 1.8, the optical transmitters are tunable light sources and the optical receivers are tunable interferometers. Though all the transmitters share the same optical projection system, no additional routing mechanism is needed because each pair of optical transmitter/receiver has its own wavelength channel. The Fabry-Perot tunable interferometers can selectively filter light beams with a 2.8 nm wavelength difference.

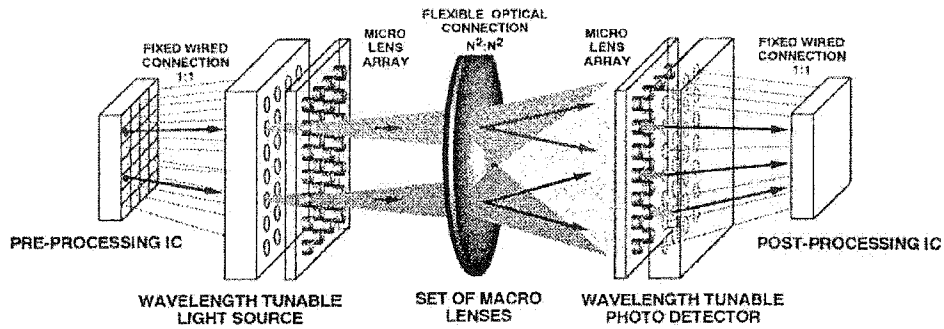


Figure 1.8: A FSOI system concept based on tunable wavelength light source and photodetector [23]

Optical gratings can project light beams with different wavelength to different positions on the focal plane. If the optical receivers are positioned on these focal spots then no additional routing is needed to distinguish optical links with different wavelengths. A wavelength FSOI system based on an optical amplitude grating is illustrated in Figure 1.9. [24]

Multiple parallel processing (MPP) system also benefit from wavelength FSOI technology. Nippon electronic company (NEC) reported such a MPP structure using VCSEL arrays operating at 4 different wavelengths. [25] Processing elements (PE) in one cluster use the same wavelength. Special wavelength-sensitive partial mirrors

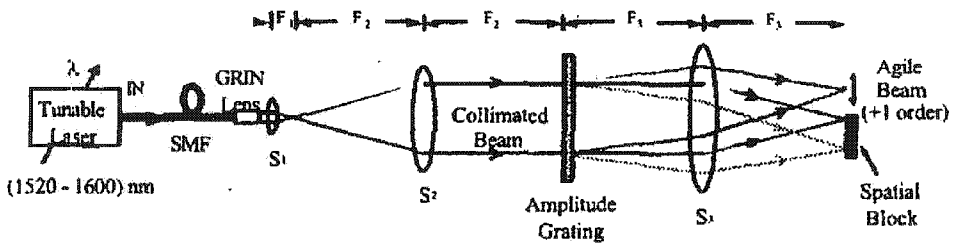


Figure 1.9: A FSOI system based on amplitude grating and wavelength selection [24]

are used for wavelength selection.

Many FSOI based systems, including high capacity telecommunication equipment and super computers, have been reported in recent years. [26][27][28]

UCSD proposed an ATM switch system based on the 3-D OESP technology. [26] The core of the ATM switch is a crossbar switch network, which performs the routings of ATM packets between the bus arbitration units (BAU). The crossbar switch is constructed on the 3-D OESP devices. Another example is the OCULAR-II, a massive parallel processing demonstration system developed by researchers at the University of Tokyo. [27] The system employs reconfigurable optical I/O to link the electronics data processors. Spatial light modulators (SLM) are used to provide the reconfigurable capability.

Researchers in University of Osaka and Nara National College of Technology proposed an optoelectronic parallel-matching architecture (PMA). [28] The communication among the processors is carried through the FSOI based data link. The demonstrator system use VCSELs as optical transmitters and CMOS photodetectors as receivers. Computer generated hologram (CGH) components are used to manipulate the light beams in the FSOI data link. The experimental operation speed on the prototype demonstrator is 15 MHz.

Colorado State University has been involved in FSOI devices and systems research for more than a decade. Recent achievements include hybrid integration techniques of VCSELs and integrated circuits, which are capable of at least 200MHz operation with considerable amount of optical output power. [29][30] A novel FSOI based ATM switch system that employs VCSEL array, integrated and stand-alone photodetector array and optical fan-out was proposed and constructed. [31][32]

1.4 FSOI alignment issue

One of the challenging issues that FSOI technology faces is the problem of aligning multiple light beams. The cause of the problem is the strict alignment tolerances due to the conflicting requirements of crosstalk, interconnect density, power budget and signal to noise ratio (SNR). The strict tolerances require special alignment techniques and the FSOI system may be vulnerable to external vibration if no alignment assurance mechanism is in place.

The simplest example of a FSOI misalignment scenario is the lateral displacement in the 'line-of-sight' type of interconnections. As stated in section 1.3, optical transmitters and receivers are directly aligned in the 'line-of-sight' interconnections. The size of the photodetectors is usually around tens of microns. The pitch of the photodetector array is usually around a hundred microns. Lateral alignment error within a few microns may greatly reduce the optical power received by the photodetector. Greater error could cause the light signal being completely off the photodetector surface. Since the received optical power is directly related with the SNR, the lateral misalignment could determine the quality of the optical link.

Eva M. Strzelecka studied the impact of the lateral displacement on bit error rate (BER) and insertion loss. [16] The system layout is illustrated in Figure 1.3(a).

The operation wavelength of the single-mode VCSEL is 980 nm. The aperture of the VCSEL is 3 μm. The diameter of the photodetector is 15 μm. The interconnection distance is 5 mm. The radius of curvature of microlenses is 360 μm and the thickness of the microlenses substrate is 500 μm. The diameter of the microlenses is much larger than that of the photodetector therefore the system can tolerate lateral misalignment greater than the dimension of the photodetector. The study indicated that the system performance deteriorated if the lateral alignment error was greater than 30 μm.

In more sophisticated FSOI systems, additional optical components may introduce more variables to the alignment. Many solutions have been proposed to solve this issue. Highly accurate optical designs can enhance the alignment accuracy of the FSOI systems. These designs have no active aligning mechanisms involved.

Daisuke Miyazaki reported a self-aligning optical connector for FSOI system. [33] The optical connector uses photosensitive resin to mark the correct position of the light beams. The photodetectors can be aligned with the mark instead of the light spots. The experimental alignment error is less than 20 μm. The alignment method might be useful in mass production of photonic devices.

Hugo Thienpont reported plastic microoptical interconnection modules for FSOI chip-to-chip and intra-chip data links. [34] The key component of the module is an optical bridge, in which precision semiconductor processing techniques are applied. The authors found 'the resulting requirements on tolerances less stringent than tolerances previous achieved by industry on similar geometries fabricated with injection molding" [34]

Instead of using highly accurate components, the introduction of redundant optical transmitter and receiver designs also increase the alignment tolerances of FSOI

systems. E. Bisaillon reported such a design. [35] The optical transmitter is an array of 3×3 960 nm VCSELs. The optical receiver is an array of 3×3 photodetectors with 70 μm diameter. The pitch of the VCSEL array is 250 μm and the pitch of the photodetector array is 125 μm. The spatial redundancy of the transmitter/receiver pair ensures that at least part of the photodetector array can receive sufficient optical power from the transmitter. The system is tolerant to 1.2 mm lateral misalignment at 20 cm interconnect distance. The maximum angular tolerance is 1.1 $^{\circ}$.

If one does not have the luxury of using an array of optical devices for a single data link, active aligning may be helpful. The concept of active alignment is nothing new to optics system designers: if the alignment error is measured in real time, then the correcting mechanism can instantly reduce the error to the minimum value. Similar means of alignment has been successfully adopted in optical compact disk drives (CD-ROM) and more recent digital video disc (DVD) drives.

An *in situ* misalignment error measurement system was demonstrated at McGill University. [36] The optical receiver plane has special bi-cell detectors (BCD) to measure the position of several light spots generated by alignment beams. If the system is perfectly aligned, then the alignment beams should fall evenly between two cells of the BCD. One can calculate the alignment error by measuring the difference between photocurrent from both cells. If multiple BCDs are deployed on multiple directions, the lateral misalignment errors can be measured.

Quadrant detectors are widely used as detectors for position measurement. If a Gaussian laser beam is projected on a quadrant detector, the position and the shape of the light spot determine the photocurrent of four photodectors. If quadrant detectors are carefully positioned in a FSOI system, one can tell if the system is correctly aligned and focused. Several institutes have reported FSOI systems with

quadrant detectors. [37][38][39][40] The measured alignment error is sent back to a micro controller, which in turn controls micro positioning motors to correct the error. Based on the fact that close-loop controlling has been successfully employed in many applications including CD-ROMs and other optical drives, the active alignment methods seem very promising.

A simple but effective alignment technique has been employed in the FSOI systems developed at Colorado State University. A cubic beam splitter in the optical path and a CCD camera make accurate visual alignment possible. Details of this technique is discussed in Chapter 5.

1.5 Summary

As the semiconductor industry pushes the IC technology for higher performance, the on-chip and off-chip interconnect capacity is going to meet the limit. As a possible solution to interconnect bandwidth shortage, FSOI technology has been under development for more than a decade. OEICs can provide active and passive optical components integrated with silicon electronic ICs. Many FSOI systems structures have been studied and reported in recent years.

A FSOI based biological sequence analysis system is presented in this dissertation. The following chapters will cover various development issues from background, design, assembly and tests. Chapter 2 presents the biological background and current status of the biological sequence analysis algorithms. Chapter 3 discusses the objective and system architecture of the biological sequence analysis system. Chapter 4 introduces the design and initial electronic test of the CMOS based biosequence analysis processor. Chapter 5 discusses the optics systems. Chapter 6 presents the functionality tests. Chapter 7 presents reliability tests and biosequence alignment

results. Chapter 8 concludes this thesis.

Bibliography

[1] *The National Technology Roadmap for Semiconductors*, Semiconductor Industry Association, 2001

[2] David A.B. Miller, "Physical reasons for optical interconnection", *International Journal of Optoelectronics*, Vol.11, No.3, p155-168, May-Jun 1997

[3] David A.B. Miller, "Why use optics for interconnects", *Conference Proceedings - Lasers and Electro-Optics Society Annual Meeting-LEOS, v1, 1997, Proceedings of the 1997 IEEE Lasers and Electro-Optics Society Annual Meeting, LEOS. Part 1 (of 2)*,p 192-193, Nov. 1997

[4] David A.B. Miller, "Motivations for optical interconnects to silicon chips", *Proceedings of SPIE - The International Society for Optical Engineering, v4089, 2000, Optics in Computing 2000*, p486, Jun.2000

[5] David A.B. Miller, "Optical interconnects to silicon", *IEEE Journal on Selected Topics in Quantum Electronics*, Vol.6, No.6, p1312-1317, Nov. 2000

[6] David A. B. Miller, "Rationale and Challenges for Optical Interconnects to Electronic Chips", *Proceedings of IEEE*, Vol. 88, No. 6, p728-749, June 2000

[7] P.R. Haugen, S. Rychnovsky,A. Husain, L.D. Hutcheson, "Optical Interconnects for High Speed Computing", *Optical Engineering*, Vol.25, No.10, p1076-1085, Oct. 1986

[8] Joseph W. Goodman, "Optical Interconnections in Microelectronics", *Proceedings of SPIE - The International Society for Optical Engineering, Optical Computing*, Vol. 456, p72-85, 1984

[9] Yue Liu, "Heterogeneous Integration of OE Arrays With Si Electronics and Microoptics", *IEEE Transactions on Advanced Packaging*, Vol. 25, No. 1, p43-49, Feb. 2002

[10] Guoqiang Li, Dawei Huang, Emel Yuceturk, Philippe J. Marchand, Sadik C. Esener, Volkan H. Ozguz, and Yue Liu, "Three-dimensional optoelectronic stacked

processor by use of free-space optical interconnection and three-dimensional VLSI chip stacks", *Applied Optics*, Vol. 41, No. 2, p348-360, Jan. 2002

[11] Sadik C. Esener, "Implementation and prospects of chip-to-chip free space optical interconnects", *Proceedings of IEEE Electronic Device Meeting 2001*, p541-544, 2001

[12] A.V. Krishnamoorthy, A.L. Lentine, K.W. Goossen, J.A. Walker, T.K. Woodward, J.E. Ford, G.F. Aplin, L.A. D'Asaro, S.P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. Dahringer, L.M.F. Chirovsky, D.A.B. Miller, "3-D integration of MQW modulators over active submicron CMOS circuits: 375 Mb/s transimpedance receiver-transmitter circuit", *IEEE Photonics Technology Letters*, , Vol.7, No.11, p1288-1290, Nov. 1995

[13] K. W. Goossen, J.A. Walker, L.A. D'Asaro, S.P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D.D. Bacon, D. Dahringer, L.M.F. Chirovsky, A.L. Lentine, D.A.B. Miller, "GaAs MQW modulators integrated with silicon CMOS", *IEEE Photonics Technology Letters*, Vol.7, No.4, p360-362, Apr. 1995

[14] Rui Pu, Chunjie Duan, Carl W. Wilmsen, "Hybrid integration of VCSEL's to CMOS integrated circuits", *IEEE Journal on Selected Topics in Quantum Electronics*, Vol.5, No.2, p201-208, Mar. 1999

[15] Rui Pu, E.M. Hayes, Carl W. Wilmsen, K.D. Choquette, K.M. Geib, H.Q. Hou, "Hybrid integration of VCSELs to foundry fabricated smart pixels", LEOS Summer Topical Meeting, p25-26, 1997

[16] Eva M. Strzelecka, Duane A. Louderback, Brian J. Thibeault, Geoff B. Thompson, Kent Bertilsson, and Larry A. Coldren, "Parallel free-space optical interconnect based on arrays of vertical-cavity lasers and detectors with monolithic microlenses", *Applied Optics*, Vol. 37, No. 14, p2811-2821, May 1998

[17] Rong Wang, Aleksandar D. Rakic' , and Marian L. Majewski, "Design of microchannel free-space optical interconnects based on vertical-cavity surface-emitting laser arrays", *Applied Optics*, Vol. 41, No. 17, p3469-3478, June 2002

[18] Sadik C. Esener, "Tightly integrated board level free space optical interconnects", *IEEE The 4th Pacific Rim Conference on CLEO/Pacific Rim 2001*, Vol. 2, p676 –677, 2001

[19] Xuezhe Zheng, Philippe J. Marchand, Dawei Huang, Osman Kibar, Nur S. E. Ozkan,

and Sadik C. Esener, "Optomechanical design and characterization of a printed-circuit-board-based free-space optical interconnect package", *Applied Optics*, Vol. 38, No. 26, p5631-5640, Sep. 1999

[20] Jang-Hun Yeh, Raymond K. Kostuk, "Free-space holographic optical interconnects for board-to-board and chip-to-chip interconnections", *Optics Letters*, Vol. 21, No. 16, p1274-1276, Aug. 1996

[21] Katsuhiko Hirabayashi, Tsuyoshi Yamamoto, Shinji Matsuo, and Shigeki Hino, "Board-to-board free-space optical interconnections passing through boards for a bookshelf-assembled terabit-per-second-class ATM switch", *Applied Optics*, Vol. 37, No. 14, p2985-2995, May 1998

[22] Yongsheng Liu, Brian Robertson, Guillaume C. Boisset, Michael H. Ayliffe, Rajiv Iyer, and David V. Plant, "Design, implementation, and characterization of a hybrid optical interconnect for a four-stage free-space optical backplane demonstrator", *Applied Optics*, Vol. 37, No. 14, p2895-2914, May 1998

[23] Hiroshi Toshiyoshi, Masakazu Kobayashi, Daisuke Miyauchi, Hiroyuki Fujita, "Design and Analysis of Micromechanical Tunable Interferometers for WDM Free-Space Optical Interconnection", *IEEE Journal of Lightwave Technolgoy*, Vol. 17, No. 1, p19-25, Jan. 1999

[24] Zahid Yaqoob, Nabeel A. Riza, "Smart free-space optical interconnects and communication links using agile WDM transmitters", *IEEE The 4th Pacific Rim Conference on CLEO/Pacific Rim 2001*, pII-5-II-6 Vol.2, 2001

[25] Mikihiro Kajita, Kenichi Kasahara, Tae Jin Kim, David T. Neilson, Ichiro Ogura, Ian Redmond, and Eugen Schenfeld, "Wavelength-division multiplexing free-space optical interconnect networks for massively parallel processing systems", *Applied Optics*, Vol. 37, No. 17, p3746-3755, June 1998

[26] James Rorie, Philippe J. Marchand, Jeremy Ekman, Fouad E. Kiamilev, and Sadik C. Esener, "Application of Networking Concepts to Optoelectronic Multiprocessor Architectures", *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 5, No. 2, p353-359, Mar./Apr. 1999

[27] M. Ishikawa, M.; M. Naruse, "Optoelectronic parallel computing system with reconfigurable interconnection", *IEEE The 4th Pacific Rim Conference on CLEO/Pacific Rim 2001*, pII-678 -II-679 Vol.2, 2001

[28] Keiichiro Kagawa, Kouichi Nitta, Yusuke Ogura, Jun Tanida, and Yoshiki Ichioka, "Optoelectronic parallel-matching architecture: architecture description, performance estimation, and prototype demonstration", *Applied Optics*, Vol. 40, No. 2, p283-298, Jan. 2001

[29] Rui Pu, Chunjie Duan, and Carl W. Wilmsen, "Hybrid Integration of VCSEL's to CMOS Integrated Circuits", *IEEE Journal of Selected Topics in Quantum Electronics*, Vol.5, No.2, p201-209, Mar./Apr. 1999

[30] Rui Pu, Carl W. Wilmsen, Kent M. Geib, and Kent D. Choquette, "Thermal Resistance of VCSEL's BOnded to Integrated Circuits", *IEEE Photonics Technology Letters*, Vol.11, No.12, p1554-1556, Dec. 1999

[31] J.R. Collington, W.A. Crossland, M.P. Dames, C. Duan, C.w. Wilmsen, VCSEL based high performance ATM switch, *1997 Digest of the IEEE/LEOS Summer Topical Meetings*, p29 -30, Aug 1997

[32] Tao Ao, Chunji Duan, Carl Wilmsen, VCSEL-based free-space optoelectronic ATM router, *Proceedings of SPIE - The International Society for Optical Engineering*, Vol. 4534, p122-128, Aug. 2001

[33] Daisuke Miyazaki, Shoji Masuda, and Kenji Matsushita, "Self-alignment with optical microconnectors for free-space optical interconnections", *Applied Optics*, Vol. 37, No. 2, p228- 232, Jan. 1998

[34] Hugo Thienpont, Christof Debaes, Valerie Baukens, Heidi Ottevaere, Pedro Vynck, Patrik Tuteleers, Guy Verschaffelt, Bart Volckaerts, Alex Hermanne, Mike Hanney, "Plastic Microoptical Interconnection Modules for Parallel Free-Space Inter- and Intra- MCM Data Communication", *Proceedings of IEEE*, Vol.88, No.6, p769-779, June 2000

[35] E. Bisaillon, D.F. Brosseau, T. Yamamoto, M. Mony, E. Bernier, D. Goodwill, D.V. Plant and A.G. Kirk, "Free-Space Optical Link With Spatial Redundancy for Misalignment Tolerance", IEEE Photonics Technology Letters, Vol. 42, No.2, p242-244, Feb. 2002

[36] G.C. Boisset, D.R. Rolston, B. Robertson, Y.S.Liu, R.Iyer, D.Kabal and D.V. Plant, "In situ Measurement of Misalignment Errors in Free-Space Optical Interconnects", *Journal of Lightwave Technology*, Vol. 16, No.5, p847-858, May 1998

[37] Tsung-Yi Yang, James Gourlay, and Andrew C. Walker, "Adaptive Alignment Packaging for 2-D Arrays of Free-Space Optical-Interconnected Optoelectronic Systems", *IEEE Transactions on Advanced Packaging*, Vol. 25, No. 1, p54-64, Feb. 2002

[38] Makoto Naruse and Masatoshi Ishikawa, "Analysis and characterization of alignment for free-space optical interconnects based on singular-value decomposition", *Applied Optics*, Vol. 39, No. 2, p293-301 Jan. 2000

[39] James Gourlay, Tsung-Yi Yang, Masatoshi Ishikawa, and Andrew C. Walker, "Low-order adaptive optics for free-space optoelectronic interconnects", *Applied Optics*, Vol.39, No.5, p714-720, Feb. 2000

[40] Makoto Naruse, Seiichi Yamamoto, and Masatoshi Ishikawa, "Real-Time Active Alignment Demonstration for Free-Space Optical Interconnections", *IEEE Photonics Technology Letters*, Vol.13, No.11, p1257-1259, Nov. 2001

Chapter 2

Overview of biological sequence alignment

With the latest achievements in molecular biology and other fields, life sciences have improved people's lives in various aspects such as medicine, food, environment, consumer products, etc. The life sciences have the potential of becoming one of the most important industries of the new millennium: tailored genetic drug, therapies for genetic related diseases, enhanced crops and many more.

Information technology has played a key role in processing and understanding the vast amount of biological related data. Bioinformatics, the biological information processing technology, involves the creation and development of advanced information and computational technologies for problems in biology (especially in molecular biology). A good example of a current bioinformatics research effort is genome sequencing, which has been pursued by multiple international research organizations and companies.

Section 2.1 provides the necessary background requires understanding the issues in bioinformatics. Section 2.2 introduces the biological sequence analysis algorithms.

Section 2.3 reviews several hardware based biological sequence analysis systems.

2.1 Biological sequence knowledge

2.1.1 A glimpse of molecular biology

In order to fully grasp the fundamentals of molecular biology requires considerable study and the aid of a few good textbooks as references. [1][2][3] A very brief summary is given here to aid the understanding of biosequene analysis. All organisms have a means of information record that determines their physical characteristics (*phenotype*). The information can be considered as the encoded blueprints (*genotype*) for the organisms. In most species, the codes are inside large molecules called deoxyribonucleic acids (DNA). The sequence of DNA is the encoded blueprint. The DNA sequence segments that can be transcribed are usually referred as a gene. Genetic information carried by a gene can be transcribed from DNA into ribonucleic acid (RNA) and then translated into protein. The sequential and spatial characteristics of protein distribution, i.e. when, where and how the protein is presented in an organism, determines the physical characteristics of the organism.

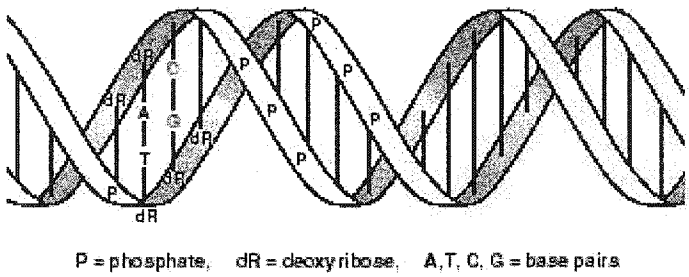


Figure 2.1: The double helix structure of DNA

As illustrated in Figure 2.1, DNA is a linearly arranged polymer composed of

4 types of *deoxyribonucleotides*,i.e. adenine, cytosine, guanine and thymine. These basic elements are usually referred as A, C, G and T. Two DNA strands are twisted around each other, forming the DNA double helix structure. Each of the basic elements in the strand is specifically paired with its complementary element: A with T, C with G. This unique feature ensures that both DNA strands carry exactly a complementary sequence. The chemical bonds between one basic element and its counterpart in the opposite strand are relatively weak, compared with the bonds between elements in the same strand. The two strands can separate under certain conditions (e.g. the presence of certain types of enzyme). Each of the strands can serve as a template for replicating new DNA strands or transferring the sequence to RNA strands. Furthermore, the double helix structure provides information redundancy for further error correction. These features of DNA make it ideal for nearly error-free information storage and propagation.

In different species, the length of DNA could vary from a few thousands elements to 10^{11} elements in some plants. The possibilities of those sequences are usually an astronomical number. Therefore DNA has enormous potential capacity for carrying genetic information. This capacity enables DNA to carry instructions for the highly sophisticated molecular mechanisms inside organisms.

An eukaryote cell is defined as a cell with *nucleus*. Observed through a microscope, the *nucleus* is the most notable part of an eukaryotic cell. [1] The nucleus contains the genetic material of the cell in the form of *chromatin*. Depending on the specific specie, chromatin contains long stretches of DNA surrounded by nuclear proteins. The nucleus is separated from the rest of the cell by a *nuclear membrane*. The nucleus can be considered as the container of DNA strands.

Parts of the gene carry information for constructing RNAs and regulating the

gene expression. Parts of the gene carry information for expressing proteins. The process of translating DNA sequences into proteins includes copying sequences into messenger RNA (mRNA) and assembling proteins in the ribosome.

Similar to DNA, RNAs are linear arranged polymers composed of four basic elements known as *ribonucleotides*. Though some viruses use RNA as gene carrier, RNAs usually play supportive roles in most organisms. Each of the RNA elements can bond to one specific DNA element. Therefore exact information transfer between DNA and RNA sequences is guaranteed.

The first step of synthesizing protein is to transfer one segment of the DNA sequence into mRNA (*transcription*). The DNA strands in the double helix would separate themselves from each other before the process starts, so that each individual DNA strand becomes accessible from the outside and the '-' strand is used as the template for RNA copying. Like an instrument used on a lathe to stop cutting at the proper position, RNA polymerase is a molecule that can be bonded to a certain location on DNA strands, and function as the indicator of the starting point and direction of RNA translation. The presence of polymerase starts a sequence duplication process, in which information in the DNA segment is to be transferred to its complementary mRNA strand (primary transcript). After some post-processing steps, the primary transcript becomes suitable for further use.

Using mRNAs as templates, proteins are assembled in ribosome. Consider the ribosome as a modern integrated mechanical processing center, it has processing instructions (mRNAs), means of transporting raw materials (tRNAs) and other highly sophisticated mechanisms. It is well known that the function of the ribosome is to produce proteins.

Similar to DNA and RNA, proteins have twenty types of basic elements called

amino acids. However, unlike the relatively simple DNA structure, proteins have complex three-dimensional spatial structures that may response to external condition changes. The versatile protein structures make proteins the foundation of organisms.

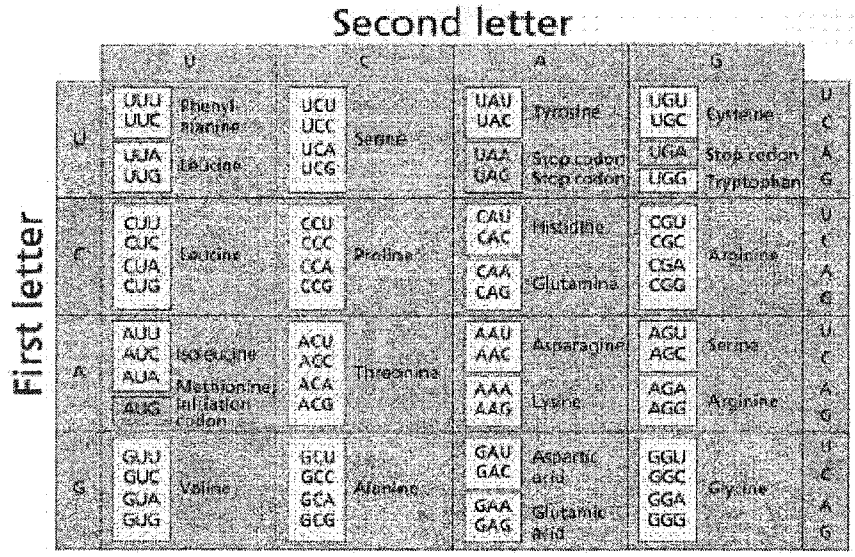


Figure 2.2: Function of the codons

Though more capable than any man-made devices, the fundamental principle behind protein synthesizing is surprisingly simple. The process of using mRNAs as a template to produce protein is called translation. The basic unit of translation operation is a codon, which is a 3-element word (in this case *deoxyribonucleotides,* the basic DNA sequence elements). Illustrated in Figure 2.2, some codons regulate the synthesizing process, e.g. codons used as start and stop indicators. Each of the remaining codons specifically corresponds to one type of amino acid. Since there are 64 (4^3) possible codons, one amino acid usually has more than one complementary codon. In many cases there are four quite similar codons for the same amino acid: the first two elements of these codons are the same. The third element can be any

element. This redundancy feature provides some error tolerances.

The decoding of the DNA sequence requires transfer RNA (tRNA) molecules. [2] tRNAs have two functions: to transport raw materials and to decode the codons. One end of a tRNA is an anti-codon, i.e., it binds to a certain type of codon. The other end of the molecule carries the corresponding amino acid. The anti-codon end binds to the mRNA, while the amino acid end is appropriately positioned for assembly. Proteins in the ribosome assembly line are long polymer chains. The ribosome adds new amino acids at one end of the protein and pushes the finished parts out of the assembly area. With assistance from ribosome and other molecules, freshly assembled proteins fold up into their designed structures. The dynamic process of protein folding currently is largely unclear. Despite the fact that computer simulations of such process are among the most demanding computation tasks, industrial giants including IBM have invested considerable research efforts into this field.

Some post-assembly processing (post-translational modifications) may be necessary for proteins to function properly. For example some additional chemicals, e.g. sugars, can be bonded to a certain place on the protein. Once the post-processing is done the proteins are ready to be moved to where they are needed. Various means of transportation bring proteins to where they can perform their designated functions.

The process of transforming encoded DNA messages into functioning proteins is remarkably simple but effective. The linear DNA sequences can create enormous 3-D protein structures, and the 3-D structures can provide enormous functions to support the organisms. Humans have yet to invent any technology comparable with the process that has been refined by several hundred million years of natural selection.

2.1.2 Evolution and gene: synopsis

With indisputable evidences, evolution is not just a hypothesis but a scientific fact. All the living beings on earth originated from a common ancestor, therefore they all share some commonalties. Genes play a vital role in evolution.

Futuyma [3] defined the evolution as follow:

> "In the broadest sense, evolution is merely change, and so is all-pervasive; galaxies, languages, and political systems all evolve. Biological evolution ... is change in the properties of populations of organisms that transcend the lifetime of a single individual. The ontogeny of an individual is not considered evolution; individual organisms do not evolve. The changes in populations that are considered evolutionary are those that are inheritable via the genetic material from one generation to the next. Biological evolution may be slight or substantial; it embraces everything from slight changes in the proportion of different alleles within a population (such as those determining blood types) to the successive alterations that led from the earliest protoorganism to snails, bees, giraffes, and dandelions."

There are three key factors in the evolution process: inheritance, variation and selection. Inheritance means the passing of the genetic materials from generation to generation. Variation means the gene changes that make the offspring different from their parents. Selection means that some of the variations may better suit the outer environment therefore have better chance of survival.

DNA is the major carrier of genetic sequences. As introduced before, DNA sequences can be duplicated in an almost perfect manner. This feature guarantees the inheritance.

Not all causes of variation are thoroughly understood. The simplest type of variation is the change of a single basic element in a sequence (point mutation). The possible ways of change are *insertion* of an additional element, *deletion* of an

existing element and *replacement* of an existing element with a new element. Other types of variation may involve larger numbers of basic elements: new sequence segments could be inserted, existing sequences could be rearranged, etc. Sex is also an important asset in sequence variation. Sexual reproduction blends the parents' characteristics and passes the combined genetic materials to the offspring.

The genetic variation may have different impact on the organism: positive, neutral or negative. Most variations are either neutral or negative. Changes that happen in some less important places in the sequence may have no significant effect. The redundancy in the codon translation is a good example: in many cases the amino acid translation is still correct even if the third element in the codon is wrong. If the first or second element of the codon is changed, the translation results in a different amino acid in the protein sequence. If the variation occurs on some less important spots, it has little or no impact on the 3-D structure of the protein. If the variation happens on vital parts of the sequence and the new amino acid is not similar with the old one, the protein probably will not function normally. In this case the variation would have a negative impact to the organism. A positive variation, i.e. a variation that the organism actually benefits from, is quite rare. Therefore, the crucial segments of the genetic codes are very conservative. The production of vital proteins must adhere to the original 'proven design' in order to function properly. Otherwise the organism will not be able to survive. Not surprisingly some segments in genetic codes of many species, no matter how apparently different they are, are exactly the same.

Evolution is the foundation of the genetic sequence analysis. The comparison between sequences from different species may help identify important parts of the genetic codes, plotting the evolution map, and revealing important historical events.

2.2 The genetic sequence analysis methods

The analysis of genetic sequences requires knowledge of molecular biology, mathematics and information technology. Statistical data could reveal the likelihood of whether several sequence fragments are biologically similar. The hidden Markov model is used to study the patterns inside the biological sequences. Various types of analytical methods have been developed to meet different analysis scenarios.

2.2.1 The pairwise alignment

The most common biosequences analysis method compares sequences and determines if they are related. One could find the similarities between sequences by putting similar sections side by side. It is known that some vital codes have been kept largely unchanged along the evolution path. [1] Aligned sequences could reveal those important sections. There are several issues concerning the alignment: (1) what kind of alignment shall be used; (2) how to evaluate the alignment; (3) how to find the 'best' alignment.

Based on the fact that the most common gene variation is single element mutation, it is sensible to assume that two related sequences have a lot in common, i.e., the pairwise alignment of one element in one sequence with its complementary element (either identical or similar) in the other sequence. Because of the presence of the insertion and deletion phenomenon, one must expect that some elements in one sequence may not have their counterparts in the other sequence (alignment gap). However the alignment gaps are relatively rare in vital parts of the sequence, otherwise the protein may not be able to function properly.

Figure 2.3 shows a classic example of pairwise alignment between two sequence segments named HBA_HUMAN (Top) and LGB2_LUPLU (Bottom). [4] The middle

line is a possible alignment: the letters indicate pairs of identical elements, and the plus signs indicate a pair of similar elements (i.e. have similar structure or function in biological or chemical aspects). The hyphen in the top line indicate possible alignment gaps. Although the example alignment could be used to evaluate the similarities between related segments, the alignment itself does not imply the segments in question are related.

```
GSAQVKGHGKKVADALTNAVAHV---D--DMPNALSALSDLHAHKL
++ ++++H+ KV   + +A  ++           +L+ L+++H+ K
NNPELQAHAGKVFKLVYEAAIQLQVTGVVVTDATLKNLGSVHVSKG
```

Figure 2.3: Example of pairwise alignment [4]

Several methods of pairwise alignment have been developed. One major concern of such methods is to distinguish biologically meaningful alignments from those random line-ups.

Based on biological facts and statistical data, scoring models have been established to guide the alignment process and validate the alignment results. The simplest concept behind the scoring models is to compare the likelihood that the sequences are related against that of the sequences being unrelated. Given one aligned element pair A, the probability that the two elements are related is P_R and that the probability of the two elements are unrelated is P_U. A scoring system can be established based on the odds ratios, defined as P_R/P_U.

The substitution matrix is an array of logarithms of odds ratios. Figure 2.4 shows a BLOSUM62 substitution matrix. Zeros in the matrix mean the chances are equal for a related alignment and a random alignment. Positive values indicate a higher probability for a related alignment. Negative values indicate a random alignment is more likely than a related one.

	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W	
C	9																				C
S	-1	4																			S
T	-1	1	5																		T
P	-3	-1	-1	7																	P
A	0	1	0	-1	4																A
G	-3	0	-2	-2	0	6															G
N	-3	1	0	-2	-2	0	6														N
D	-3	0	-1	-1	-2	-1	1	6													D
E	-4	0	-1	-1	-1	-2	0	2	5												E
Q	-3	0	-1	-1	-1	-2	0	0	2	5											Q
H	-3	-1	-2	-2	-2	-2	1	-1	0	0	8										H
R	-3	-1	-1	-2	-1	-2	0	-2	0	1	0	5									R
K	-3	0	-1	-1	-1	-2	0	-1	1	1	-1	2	5								K
M	-1	-1	-1	-2	-1	-3	-2	-3	-2	0	-2	-1	-1	5							M
I	-1	-2	-1	-3	-1	-4	-3	-3	-3	-3	-3	-3	-3	1	4						I
L	-1	-2	-1	-3	-1	-4	-3	-4	-3	-2	-3	-2	-2	2	2	4					L
V	-1	-2	0	-2	0	-3	-3	-3	-2	-2	-3	-3	-2	1	3	1	4				V
F	-2	-2	-2	-4	-2	-3	-3	-3	-3	-3	-1	-3	-3	0	0	0	-1	6			F
Y	-3	-2	-2	-3	-2	-3	-2	-3	-2	-2	2	-2	-2	-1	-1	-1	-1	3	7		Y
W	-2	-3	-3	-4	-3	-2	-4	-4	-3	-2	-2	-3	-3	-1	-3	-2	-3	1	2	11	W
	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W	

Figure 2.4: The BLOSUM62 substitution matrix [5]

As stated in section 2.1.2, the insertion and deletion of basic elements cause gaps in the sequences. In order to distinguish related sequences with gaps from unrelated random sequences, one must introduce some methods to evaluate the possible existence of gaps. The schemes of handing gaps varies from a simple gap penalty ,i.e. take off fixed score points for each gap, to complex equations.

Dynamic programming algorithms are designed to conduct different types of pairwise alignment. The general approach of such algorithms is to score all alignment possibilities with one of the substitution matrices and gap penalty rules. Figure 2.5 shows a scoring matrix used to record all the possible alignment scores. The sequence segments being aligned here are **HEAGAWGHEE** and **PAWHEAE**, which are shown on the top and the leftmost of the scoring matrix. Each value in the matrix represents the best score of a possible alignment, from the beginning of the two sequences to the two corresponding elements.

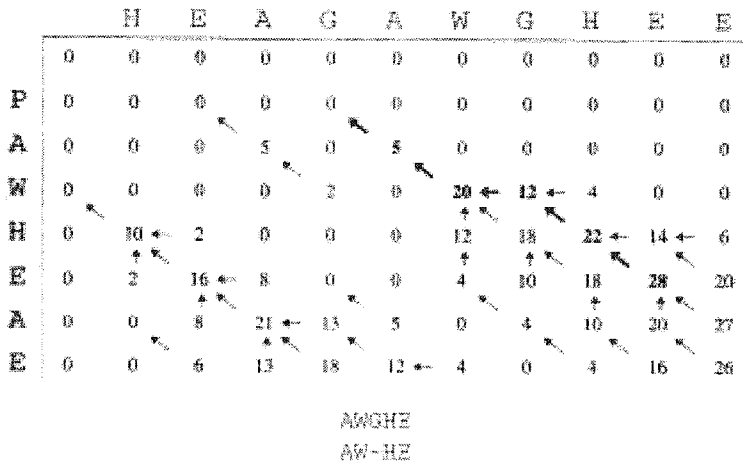


Figure 2.5: An example of scoring matrix for local alignment, Smith-Waterman algorithm [4]

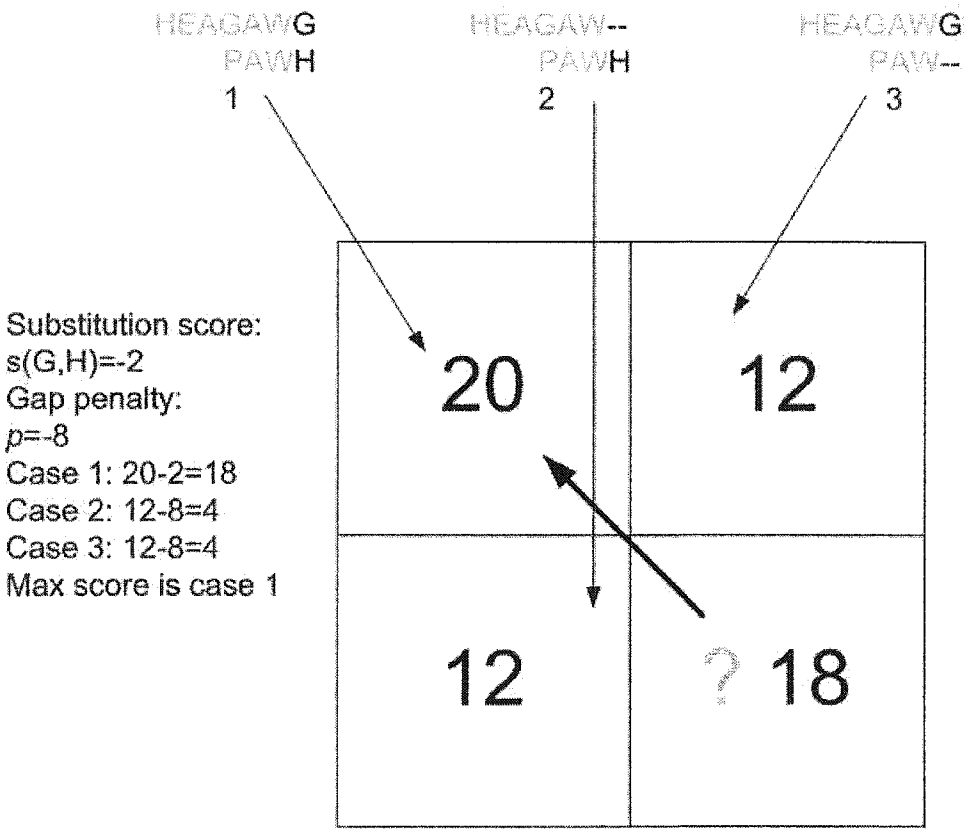


Figure 2.6: Three possible alignments for the two residues [4]

Figure 2.6 shows a small portion of Figure 2.5. In order to determine the best score for aligning **PAWH** and **HEAGWG**, one would look at several possibilities: (1) the residues shall be aligned to each other (the first situation in the figure) or (2) one of the residues shall be aligned to a gap (the second and third situation in the figure). In any case, the alignment in question shall be based on a previous alignment. In the case of aligning two elements, the substitution value shall be added to the value in the top left corner. In the case of aligning one element to a gap, the gap penalty shall be subtracted from the value in either top right corner or bottom left corner. The best alignment is the alignment with the highest score. Obviously the matrix would need to record the origin of the best score, as shown by the arrows in Figure 2.5. After the completion of the matrix, the arrows can be used to trace back the alignment process. The alignment result is generated in the trace back process.

The alignment principle can be used to fill up the scoring matrix starting from the top left corner. Additional column and line have been inserted at the top and the left of the matrix to facilitate the scoring process.

There are several types of pairwise alignment. The global alignment is designed to find out the best alignment between two complete sequences. The local alignment concentrates on locating similar fragments within the sequences in question. Other alignments can deal with more complicated situations.

Based on the discussion about evaluating different alignment possibilities, the alignment algorithms are quite straightforward. The Needleman-Wunsch global alignment algorithm is explained as follow: [4]

$$F(A_i, B_j) = \max \begin{cases} F(A_{i-1}, B_j) - p \\ F(A_i, B_{j-1}) - p \\ F(A_{i-1}, B_{j-1}) + s(A_i, B_j) \end{cases} \quad (2.1)$$

Assume the two sequences are A and B, the ith residue of A is Ai and the jth residue of B is Bj. The alignment score of Ai and Bj in the substitution matrix s is s(Ai, Bj). The gap penalty is a simple linear scheme that takes p points for each gap inserted. The scoring matrix is F. Equation 2.1 shows the scoring formula used for this global alignment.

All the elements in the sequence are not equally meaningful. The global alignment may not be able to distinguish the important fragments in the sequence. One may set a threshold for the score: the alignment shall stop once the score is below the threshold. The Smith-Waterman local alignment algorithm is designed to seek the most similar segments in the sequences by setting a scoring threshold at zero.

There are more complicated alignment situations. The protein sequences may carry a repeated pattern. One sequence segment may be completely encapsulated in other sequences. The gaps could be more than one element long. Some sophisticated pairwise algorithms have been developed to handle these facts. 2.2

Pairwise algorithms are usually relatively precise. However, the computational time and memory involved are proportional to the square or even the cube of the length of the sequences.

2.2.2 BLAST

Besides the Needleman-Wunsch algorithm for global alignment and Smith-Waterman algorithm for local alignment, dynamic programming algorithms have been ex-

panded to more complicated models. Markov chains and hidden Markov models have been utilized to develop faster and less memory consuming algorithms, including the most widely used basic local alignment search tool (BLAST) algorithm. [6]

In searching databases containing large numbers of sequences, only a relatively small number of sequence fragments would be found similar with the query sequence. These sequence fragments in the sequence database and the query sequence are likely to have some very similar (even identical) short words, which are called 'hits'. The BLAST program is designed to search a list of such hits between a sequence in question and a database with reasonable computing capabilities.

Similar to the Smith-Waterman algorithm, the original BLAST algorithm defines a maximum segment pair (MSP) as a pair of equal-length (therefore no gap is allowed) sequence segments with the highest alignment score. Though a MSP could be of any length, the MSP alignment score can not be improved by adding or removing additional neighbor characters. Since the biologically meaningful alignments usually share a high level of commonality, the alignment scores of those alignments would be significantly higher than random alignments. In order to minimize the search burden, only those alignments with scores higher than a threshold T are taken into further analysis process.

The classic implementation of the BLAST algorithm contains three major steps: (1) compiling a list of high scoring words, (2) scanning the database for the 'hits' and (3) expanding the hits.

BLAST assumes that some highly similar short strings are very likely to be found inside important alignments. Those strings can be used as 'seeds' to identify potential 'good' alignments. In the BLAST program the 'seeds' list contains some

short words, which would be generated based on biological facts. The computational time of this step is roughly proportional to the length of the list.

Once the word list is available, the database is searched for matches of the words. These hits are recorded. The search can then be extended in either direction of those hits to look for a longer alignment. Usually the alignment stops when the alignment score drops below a certain threshold. The extension process adds some small inaccuracies to the results.

By using the short words for the pre-screening, the BLAST program can efficiently generate satisfactory results. In order to search for short words in long sequences, one could use a large array to record all the occurrences or use a finite state machine to increase searching speed. For example, there are 20^4 (160,000) possible character combinations for protein short words with a length of 4 characters. Although in theory 20^4 pointers would be needed for a complete record array, in practice there usually would be only a tiny fraction of those 20^4 possible short words involved. [6]

Since there are only 4 characters in the DNA sequences, only 2 bits are needed to represent a character. Modern computers usually use 32-bit or 64-bit 'words' as basic operation numbers. Multiple characters can be compressed into one computer 'word' so that those characters can be compared at the same time. It is notable that DNA sequences may tend to have heavy concentrations of certain types of base. Besides, DNA sequences may have repeated sequence patterns that could cause large amount of alignment results without much interest. Special filters are added to remove those results.

The BLAST algorithm allows various modifications for different demands. Many improved versions of BLAST have been introduced for more specialized applications.

The gapped BLAST reduces the computing time of the BLAST programs by one third and enhances sensitivity to weak similarities. [7] Position-Specific Iterated BLAST (PSI-BLAST) combines "statistically significant alignments produced by BLAST into a position-specific score matrix, and searching the database using this matrix." [7] The program is also sensitive to weak similarities.

Today a biosequence database employing BLAST search engine is publicly available on National Institute of Health (NIH) sponsored websites. [8] BLAST 2.0 is provided with nucleotide and protein databases that contains newly found biological sequences. The BLAST algorithm conducts both local and global alignments that can assist identifying the function of unknown biological functions. PSI-BLAST can be used to further enhance the sensitivity of seeking weak similarities.

2.3 The hardware implementations of alignment algorithms

Due to improved sequencing techniques, increased competition and growing interesting in the genetic codes, the size of the biosequence databases, e.g. GenBank, has dramatically increased. Though the sequence analysis capability has been under constant improvement, it has been more and more challenging to analyze all the sequences with reasonable cost and time. One logical solution is to use special information processing hardware for highly demanding computational tasks.

The alignment algorithms have been put into parallel implementations custom hardware: either Application Specific Integrated Circuit (ASIC) or Programmable Logic Device (PLD) based systems that may operate independently or with general-purpose computers.

The biological information signal processor (BISP) is an ASIC implementation of modified Smith-Waterman algorithm. [9] Each BISP chip has 16 identical processing elements (PE), which can individually handle 16-bit (-32768 to 32767) score. The system was designed on CMOS technology that allows the PEs to run at a clock frequency of 12.5 MHz . The chip also has many customizable parameters that make the chip flexible to meet different demands.

The Smith-Waterman Algorithm-Specific ASIC Design (SWASAD) is a high performance DNA sequence alignment system based on an improved BISP Smith-Waterman algorithm. [10] SWASAD is designed to perform the Smith-Waterman algorithm at high speed (50 MHz) and low cost (both in term of design and fabrication). Similar to BISP, SWASAD also features an array of identical processing elements and adjustable parameters.

Mosanya et.al reported an FPGA implementation of the Generalized Profile Search algorithm, which is used for biological sequence processing. [11] A generalized profile is a modeled sequence or sequence segment, which contains the pattern of a family of sequences, i.e. a summary of how a certain type of sequences shall look like. The algorithm searches the sequence database for patterns that fit the generalized profile. The hardware implementation of such algorithm also includes an array of processing elements, which is considered to be suitable for dynamic programming.

Systolic Accelerator for Molecular Biological Applications (SAMBA) employs an array of 128 systolic processors for high performance comparisons based on the Smith-Waterman algorithm. [12]

Some hardware based bioinformatics systems are commercially available. These systems may use a combined scheme of customized hardware and off-the-shelf computer and network technologies.

DeCypher is a high-performance hardware suite that implements BLAST, Smith-Waterman and FrameSearch algorithms. [13] The system uses FPGA chips as the high speed processing core (the 'accelerator') and a desktop computer as controlling unit and interface. With the combined power of the accelerator and the general-purpose computers, DeCypher claims to provide affordable, flexible and high-speed analysis capabilities for researchers.

Bioccelerator, BioXL/P and BioXL/H are hardware biosequence analysis accelerators developed by Compugen. [14] BioXL/H is the latest version that employs Frame+ algorithm (an improved version of FrameSearch). The system uses Xilinx FPGA boards as the processing cores, high-speed ethernet as the communication port and a software package called GenCore for the front-end processing and interface. The system is accessible via a built-in world-wide-web interface so that the operation is easy for the end users. The system is also scalable from the basic 4-module model to the high-end 32-module model.

GeneMatcher2 is a combined Linux server farm and specialized hardware computing boards. [15] The hardware computing boards use ASIC instead of FPGA for the processing cores. Though the major selling point of GeneMatcher2 is its modified BLAST algorithm, the system can use multiple common algorithms including Needeleman-Wunsch, Smith-Waterman and their variations. Similar to BioXL/H, GeneMatcher2 also has a world-wide-web interface for common users.

All the systems introduced above are electronics based systems. These systems usually have impressive processing capabilities inside the ASIC or the PLD. Modern CMOS technology provides high-speed information processing capability. However, the processing unit only processes as much data as it receives from the interface, which would be running at considerably lower clock speed. The biosequence align-

ments require a high volume of data flow: the sequences in the databases as input and interim scoring and results as output. The bottleneck is the interface between the processing unit and the system memory that contains the biological sequence database and scoring data. This bottleneck issue, as mentioned in Chapter 1, is not uncommon in electronic based information processing systems. Optical interconnect may provide solutions to this issue.

2.4 Summary

Research in molecular biology provides background knowledge for the biological sequence analysis, which has drawn great interest in recent years. Based on statistical methods and biological facts, biological sequence analysis algorithms are designed for better understanding of biological sequences. Since these analysis algorithms usually require large amount of computation capability, several hardware based analysis systems have been developed to assist the analysis process.

Bibliography

[1] *Artificial Intelligence & Molecular Biology*, Lawrence Hunter, MIT Press, 1993

[2] *Principles of biochemistry*, Albert L. Lehninger, Worth Publishers, 1993

[3] *Evolutionary Biology*, Douglas J. Futuyma, Sinauer Associates, 1986

[4] *Biological sequence analysis*, R. Durbin *et. al*, Cambridge university press, 1998

[5] J.G.Henikoff, et. al, "Using substitution probabilities to improve position-specific scoring matrices", *Computer Applications in the Bioscience,* Vol. 12, p135, 1993

[6] Stephen F. Altschul, *et. al*, "Basic Local Alignment Search Tool", *Journal of Molecular Biology*, 215:403, 1990

[7] Stephen F. Altschul, et. al, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Research, Vol. 25 No.17, p3389-3402, Sep 1997

[8] http://www.ncbi.nlm.nih.gov/BLAST/

[9] Chow, E., et. al, "Biological information signal processor", *Proceedings of the International Conference on Application Specific Array Processors,* p144 –160, 1991

[10] Han, T.; Parameswaran, S. , "SWASAD: an ASIC design for high speed DNA sequence matching"

Proceedings of 7th Asia and South Pacific and the 15th International Conference on VLSI Design Automation Conference, p541-546, 2002

[11] Mosanya, Emeka et. al, "FPGA-based hardware implementation of generalized profile search using online arithmetic", *Proceedings of the 1999 ACM/SIGDA 7th International Symposium on Field Programmable Gate Arrays*, p 101-110, 1999

[12] P. Guerdoux-Jamet, and D. Lavenier, "SAMBA: Hardware Accelerator for Biological Sequence Comparison", *Computer Applications In the Biosciences*, vol. 13, p609-615, Dec. 1997.

[13] http://www.timelogic.com/decypher_intro.html

[14] http://www.cgen.com/products/hardware.htm

[15] http://www.paracel.com/products/gm2_specs.html

Chapter 3

System Overview

The major objective of the biological sequence (biosequence) analysis system is to explore the possibility of applying free space optical interconnect (FSOI) technology to complex information processing systems.

The system is a combination of a diffractive optical system, a customized integrated circuits (IC) chip with optical inputs, a programmable logic device board (in this case a field programmable gate array, FPGA) and a general-purpose computer for FPGA programming. BLAST was chosen as the biosequence analysis algorithm for the system. The algorithm was divided into modules that were to be implemented in the custom designed IC and the FPGA. The algorithm on the FPGA was developed by Mr. Klaus Hartinger.

3.1 Objectives

The current progress of free space optical interconnects technology and the biological sequence analysis techniques have been discussed in Chapter 1 and 2, respectively. FSOI technology has the potential of providing highly parallel, high speed data

interconnects to electronic hardware, including complex information processing systems. The FSOI based biosequence analysis system is a technology demonstration system that investigates such possibilities. The specific objectives are:

- Demonstrate the feasibility of a prototype FSOI biosequence analysis system
- Develop an architecture for the prototype system
- Apply available technologies to the prototype system

3.1.1 Objective 1: demonstrate the feasibility of a prototype FSOI biosequence analysis system

Fiber optics has become the major technology for the long distance telecommunication network. It has been more than a decade since the first FSOI system was proposed. Free space optics has not been introduced into the main stream of the information processing technology, although there is a prototype super computer using waveguide optical interconnect for clock signal distribution. [1] Despite the projection of making the technology commercially available in one decade [2], electronic engineers prefer to rely on the familiar electronic technologies rather than the unfamiliar FSOI concept. The successful construction and demonstration of the prototype in this research work enhances the credibility of FSOI technology and advances the overall research effort in this field.

3.1.2 Objective 2: develop an architecture for the prototype system

Massive parallelism increases processing speed of the biological sequence analysis system. Therefore, the system architecture for the biological sequence analysis sys-

tem may benefit from the FSOI technology, namely the large fan-out parallelism. Since there is no earlier report on FSOI based biosequence analysis system, this prototype will verify the concept and provide useful experience for future systems.

3.1.3 Objective 3: apply available technologies to the prototype system

There are multiple technological options for constructing a FSOI system, which can be optimized for speed, parallelism or efficiency. There are several relevant technologies to choose from. For example, for the light source one can choose a stand-alone VCSEL array, an VCSEL array integrated with custom IC chip or a single laser (VCSEL or edge emitter) with diffractive optical fan-out. For the photodetectors, one can choose a heterogeneously integrated photodetector or an on-chip CMOS silicon photodetector. In this research work, the main objective was to determine the proof of the concept with a limited budget. In the present FSOI based biosequence analysis system, the FSOI technology provides parallel optical inputs to a CMOS integrated circuit, which includes the photodetector array, receiving circuit and logic circuit. The biosequence sequence analysis process is partitioned between a FSOI based biosequence analysis chip and an FPGA. A computer is used for programming the FPGA.

The research on the FSOI based biosequence analysis system has been conducted with these objectives in mind.

3.2 System algorithm

3.2.1 Selection of the system algorithm

There were two candidates that might be suitable: the Smith-Waterman local alignment algorithm and the BLAST algorithm. The system algorithm for the FSOI based biosequence analysis system was chosen for its practical value and the feasibility of implementation.

The Smith-Waterman algorithm has been implemented in several research projects, as mentioned in Chapter 2. The advantages of using this algorithm are: (1) the hardware architecture of the algorithm has been well established and (2) the test result of the FSOI based system could more easily be compared with previous published data. However the hardware architecture requires relatively complex logic organized as identical function blocks, each of which may cost large chip real estate. Because one may not be able to put many such function blocks into one single chip, the processing chip may not have much processing power. The exchange rate between the processing chip and the rest of the system may not require high parallelism or data speed. The advantage of applying FSOI parallelism to the Smith-Waterman algorithm based systems are not obvious.

The BLAST algorithm and its variants are practically the standard biosequence local alignment algorithm. As introduced in Chapter 2, it has been implemented in several commercial products that use a hybrid parallel architecture for off-the-shelf computer clusters and custom designed electronic processing accelerators. Although the BLAST algorithm is more complex than the Smith-Waterman algorithm, the BLAST algorithm utilizes processing steps that benefit from parallel processing. FSOI systems may provide the parallelism in a simpler way compared with the

electronic systems. For this reason the BLAST algorithm was chosen as the system algorithm for the FSOI based biosequence analysis system.

3.2.2 The process flow of the BLAST algorithm

The process flow of the BLAST algorithm is illustrated in Figure 3.1.

The unidentified sequence is to be compared with sequences in the sequence database. The algorithm is composed of three function modules. Each of the modules corresponds to one step of the algorithm.

The preprocessing of the sequence module is illustrated in Figure 3.2. It first creates the word list, which is a data array stored in the temporary memory. Each entry of the array contains one of the possible permutations of the biosequence characters with a given length w. In the next step, the module shifts a window with length w along the unidentified sequence. The shifting window generates a word X with length w, which is compared with all the words in the word list. If the compared alignment score for X and a specific word Y is higher than a threshold T, i.e. 'similar' words, the module records the position p of word X in the unidentified sequence. The module then looks up the entry containing word Y in the word list and records position p in that entry. One word in the unidentified sequence may have several 'similar' words in the word list. Therefore for each word of the unidentified sequence, there may be several 'similar' words that have its position in the word list.

The 'hits' list generation module is illustrated in Figure 3.3. It creates the 'hits' list which is an indexed array that records positions of all the 'hits'. This module looks for matches, i.e. at least two of three matched characters, between the sequence from sequence database and the words list. These matches are 'hits'

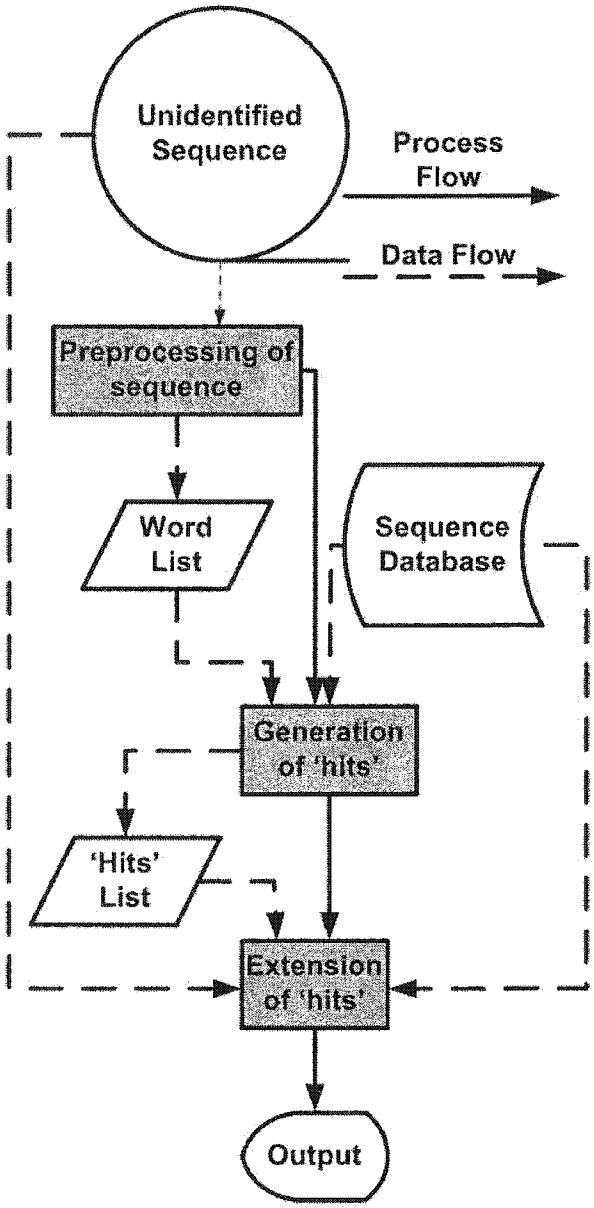


Figure 3.1: The process flow of the BLAST algorithm

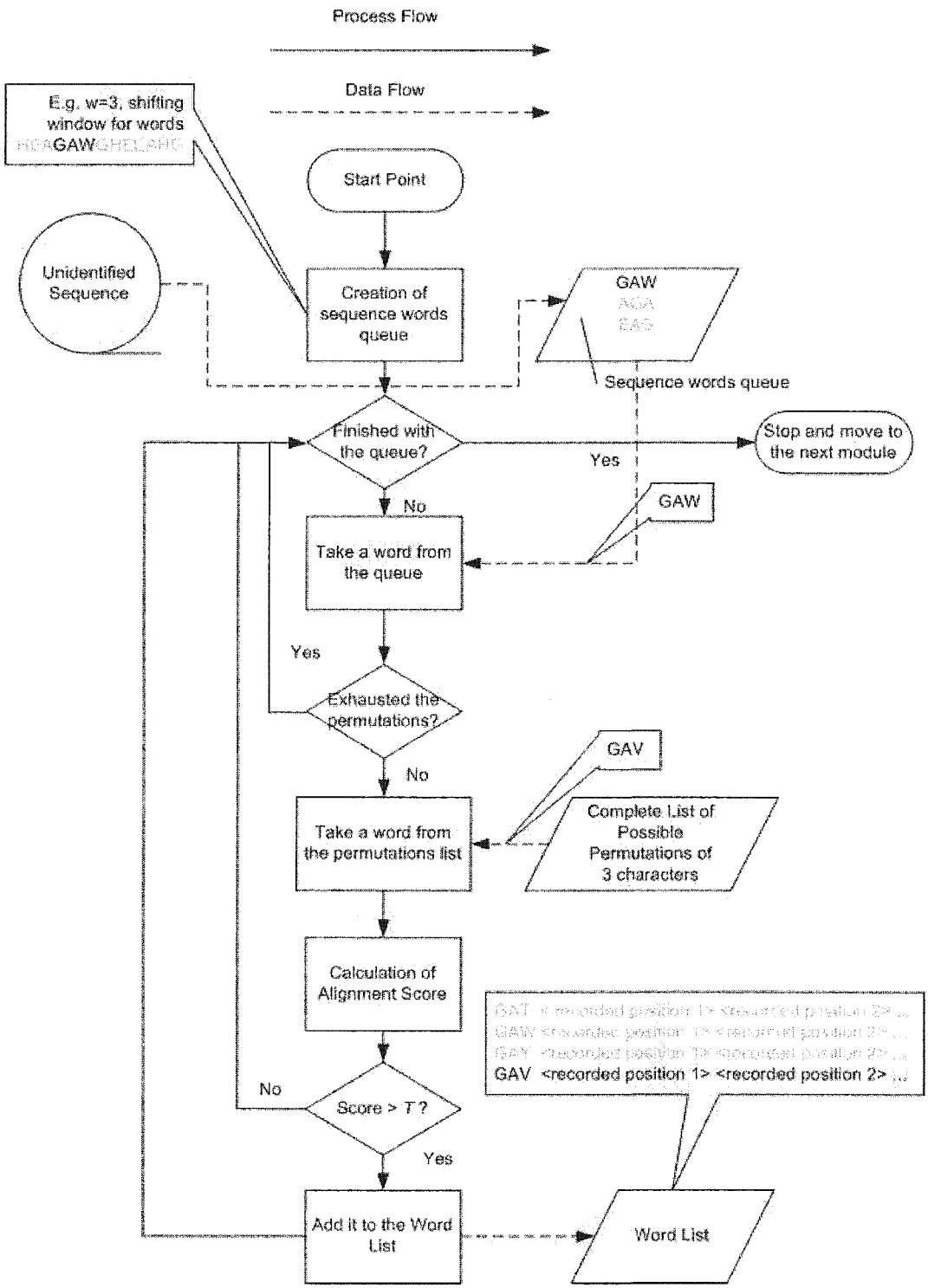


Figure 3.2: The preprocessing sequence module

and their positions in both sequences are stored in the 'hits' list.

The 'hits' extension module is illustrated in Figure 3.4. The module extends the length of the 'hits' in both directions. This is accomplished by examining the alignment score of each additional pair of characters. If the alignment score is positive or zero, then the characters are included into the extension. If the alignment score is negative, then the extension is stopped in that direction.

The alignment results are generated and displayed after the whole alignment process is finished.

3.2.3 The function partition of the algorithm

As stated in previous sections, the BLAST algorithm is quite complex. Realizing such an algorithm on a single custom designed CMOS chip might be possible but would require considerable resources that would be both costly and distract from the development of the FSOI technology itself. On the other hand, a state-of-the-art programmable logic device is capable of providing flexible and powerful logic processing capabilities at affordable cost. Therefore the algorithm modules are divided between a FSOI based processor chip and a programmable logic device.

The function of the FSOI based biosequence processor chip must be relatively simple in order to provide sufficient chip real estate for the optical inputs and to emphasize the FSOI technology instead of logic circuit design. The programmable logic device carries a large portion of the algorithm functions because of its flexible and powerful logic processing capability. The programmable logic device also manages the memory chips and communications between the FSOI chip. One can start the biosequence analysis process by pushing a button on the programmable logic device board. The development of FPGA programming is conducted on a personal

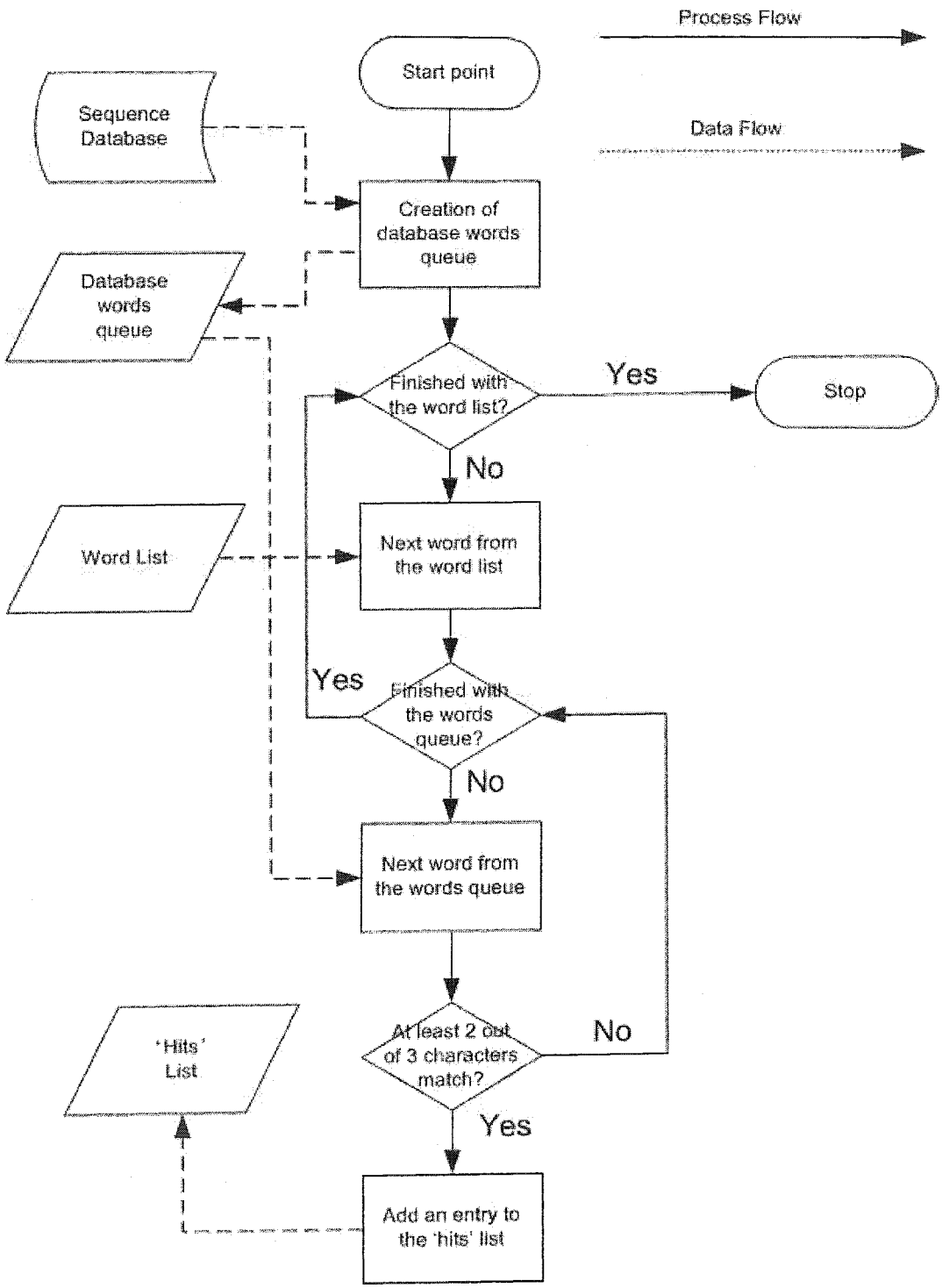


Figure 3.3: The 'hits' generation module

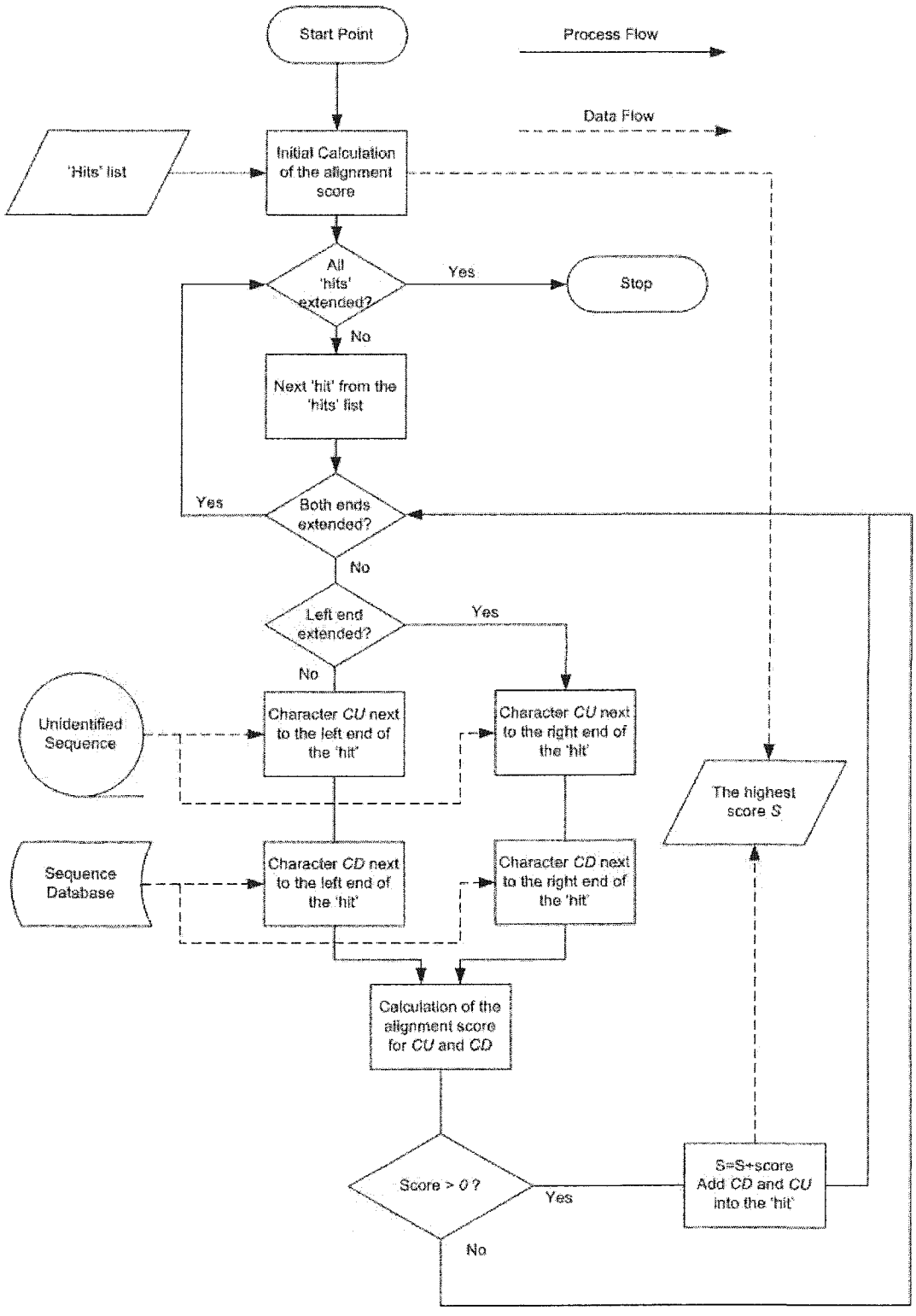


Figure 3.4: the 'hits' extension module

computer (PC) and the binary program file is downloaded via an USB cable. The function partition diagram is illustrated as Figure 3.5.

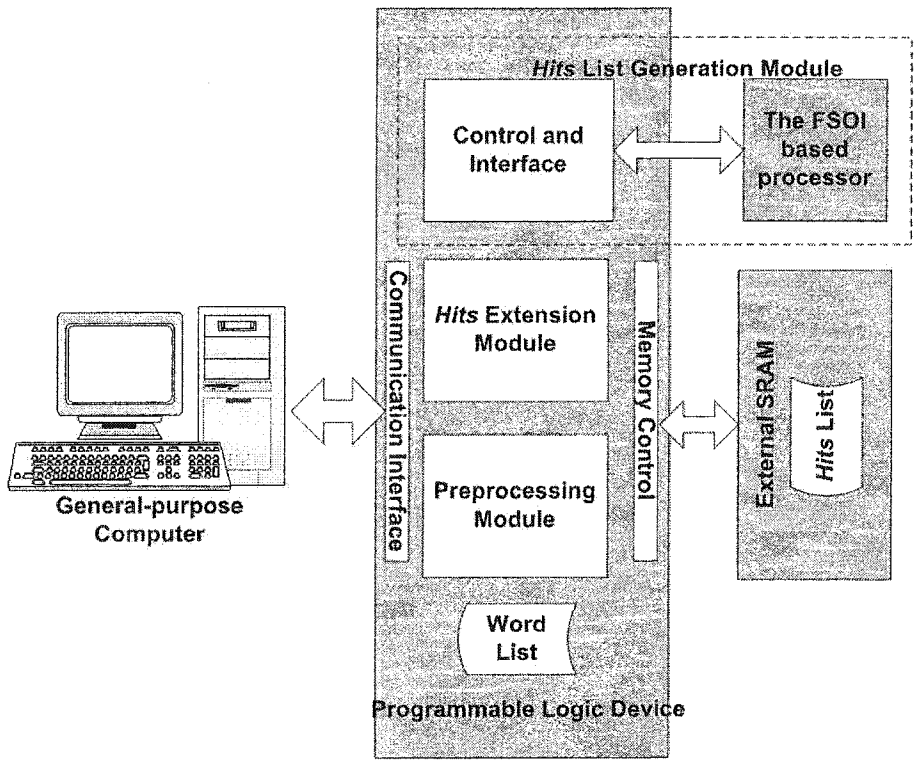


Figure 3.5: function partitions diagram

One frequently used function in the BLAST modules (shown in Figure 3.2, Figure 3.3 and Figure 3.4) is the character to character comparison, either for 'similar' word comparison between word list and unidentified sequence or for matches between the 'hits' and the sequence database. The implementation of a hardware parallel comparison process is relatively straightforward. The FSOI based processor chip was designed as a parallel character-to-character comparison chip.

The programmable logic device is commonly used in many applications. There are off-the-shelf development kits that include the programmable device, memory modules, communication ports, display and many other accessories. Development

software packages are also available to design, simulate and program these programmable logic devices. State of the art programmable logic devices, including field programmable gate array (FPGA) and complex programmable logic device (CPLD), are sophisticated programmable logic devices that can carry out complex logic operations. Compared with fully custom designed CMOS IC, these programmable devices have lower development and purchase cost. As stated in Chapter 2, several commercial biosequence analysis tools already have employed such devices for high-speed BLAST algorithm processing. There is no fundamental difficulties to implement the BLAST algorithm along with the FSOI base biosequence processor.

A AMD Athlon CPU based personal computer (PC) is used for the high demand device simulations and logic design compiling, design synthesis and binary file generation of the programmable logic device.

3.3 System layout

The individual components can not function as a system unless they are appropriately organized and connected. The complete system layout of the FSOI based biosequence analysis system is illustrated in Figure 3.6.

Several modules, including the personal computer, programmable logic device development board and FSOI based biosequence chip, have been described in section 3.2. The logic analysis system is used to monitor the signal flow inside the biosequence analysis system and is not involved in system operation.

The optics system is an important module for the optical interconnects. It multiplies the modulated optical signal from the laser system and projects the optical beams onto the FSOI based biosequence chip. Many issues, including power budgets, optics design and alignment, have to be taken into account to ensure reliable optical

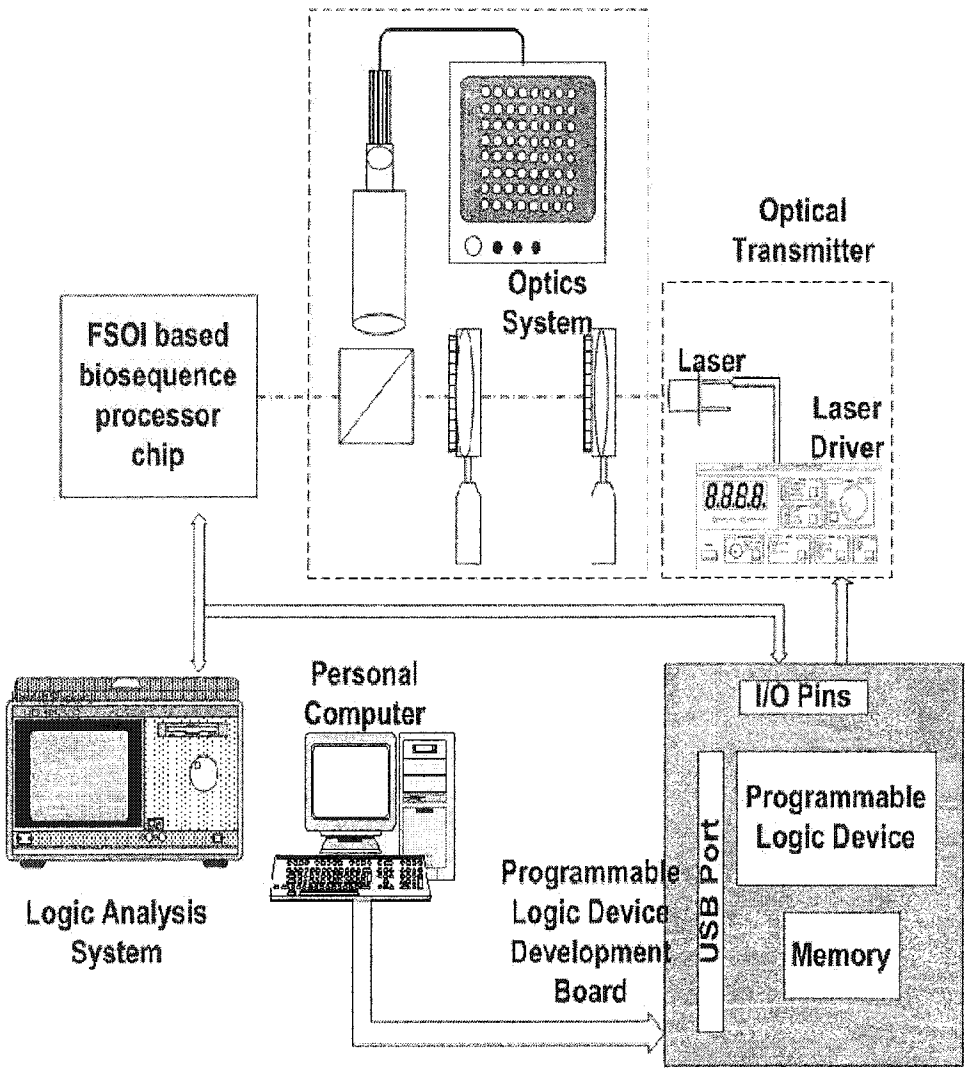


Figure 3.6: system layout of FSOI based biosequence analysis demonstration system

interconnects. Tradeoffs must be made to meet requirements raised by these issues, although these requirements may be conflicting.

The optical transmitter system consists of a laser driver and a semiconductor laser diode. The laser driver is a current source that can provide current modulated by external electronic signal. Depending upon the requirement of the optics system, the laser could be a single VCSEL, a VCSEL array or an edge emitter.

The logic analysis system is composed of a logic analyzer, a pattern generator and an oscilloscope. The logic analyzer is used to monitor the signal flow between the biosequence processor chip and the programmable logic device. It also can compare the monitored signal with stored signal waveform to detect any errors. The pattern generator generates a set of signals that are programmed in its memory. It is a stimulus source that can provide test signals to the FSOI based chip and the programmable logic device. If the FSOI based chip is faulty, a set of designed signals can help locate the source of error. The oscilloscope examines signals for debugging purposes. The logic analysis system can be programmed in such a way that all of its modules can work together to detect signals and patterns.

3.4 Summary

BLAST algorithm is selected for the FSOI based biological sequence analysis system because the algorithm is more suitable for a parallel processing system. The process flow and modules of the algorithm are discussed. The functions of the algorithms are partitioned between the biosequence analysis processor and the FPGA board.

Bibliography

[1] Ray Chen, et al, "Fully Embedded Board-Level Guided-Wave Optoelectronic Interconnects", Proceedings of IEEE, Vol. 8, No.6, p780-793, 2000

[2] Neil Savage, "Linking with Light", IEEE Spectrum, Aug. 2002, p32-35

Chapter 4

Development of the FSOI Based Biosequence Analysis Chip

The biosequence processor chip is one of the most important modules in the system since it contains the optical inputs that are the receiving end of the FSOI system. The chip also contains an array of optoelectronic comparison units that carry out the character to character comparison in parallel. The chip affects the design of the optics system and the programming of the programmable logic device.

The process flow of the processor chip and the desired specifications were determined before the design started. The CMOS processing technology type and the design tool were then determined. The photodetector design was selected. The actual design started with the analog design of the photodetector receiving circuit. The detailed logic design was then completed. The layout design was based on schematic driven layout (SDL) design technique. Standard library cells were used for some of the logic functions and other components. The design was constantly checked and verified with simulations. The fabricated chips were tested on an AMS test system.

4.1 Specifications and considerations

The biosequence processor is a CMOS chip that contains logic circuit, optical inputs, electronic inputs and outputs. The processor analyzes the biological sequences, which consists of characters representing amino acid. In the processor the characters are represented by 5-bit binary words. The chip receives one character from the optical inputs (identical character for all the optical inputs) and an array of characters from the electronic inputs. The logic circuit compares the character from the optical inputs with the characters from the electronic inputs and generates the signals indicating whether the character is identical with any of the characters. The comparison results are placed on the electronic outputs. The process chip should be able to work continuously, i.e. the chip should be ready to accept new characters for comparison after the results are generated. An external reset signal should be available to reset the whole chip operation.

The MOSIS foundry service is a long-established, cost-effective integrated circuit (IC) prototype fabrication service. Based on previous experience with its service, MOSIS was selected as the fabrication foundry. MOSIS provides a variety of semiconductor processes, including multiple sub-micron CMOS technology services that are suitable for the biosequence processor. Deep sub-micron (0.35-0.18 micron) fabrication services provided by the Taiwan Semiconductor Manufacture Company (TSMC) are high-speed, high-end and high-cost CMOS technologies. The project does not require such fancy technologies. AMI Semiconductor (AMIS) and Agilent (former HP) provide 0.5 micron technology that match the need of the project.

Mentor Graphics provides a versatile electronic design aid (EDA) tool package for IC designs. ASIC Design Kit (ADK) is a Mentor Graphics add-on package that provides libraries of several CMOS processes available via the MOSIS service. AMIS

0.5 micron library is included in the ADK package. Mentor Graphics software is widely used for both teaching and research purposes in the Department of Electrical and Computer Engineering (ECE dept.). ADK software is installed on the HP workstations on the Engineering Network Service (ENS) computer network.

AMIS 0.5 micron CMOS technology was selected because it is suitable for the processor chip, the design tools are available and the cost is appropriate.

4.2 Photodetector (PD) and PD receiving circuit

The photodetectors convert optical signals into electronic ones. There were two options for the photodetectors used for the optical inputs. Flip-chip bonded photodetectors are separately manufactured and bonded on the CMOS chip through flip-chip bonding process. They usually have high photo responsivity, fast response and established circuit models. However flip-chip bonding requires specialized processing that was not available under the project budget. Standard CMOS process provides structures including p-n junctions to form built-in photodetectors. The design of this type of photodetectors must adhere to MOSIS scalable sub-micron CMOS design rules. Although the high-speed Si-photodetector designs have been reported, there was no published data of their optoelectronic properties on AMIS 0.5 micron process. Therefore it would be difficult to design driving circuits for them. Less sophisticated photodetectors have been tested and used in previous research programs. [2][3] This design usually provides reliable operation and good responsivity, however, its operational speed is low (usually less than 50 kHz).

Based on the facts above, a simple photodetector (PD) design was selected. As illustrated in Figure 4.1, the PD has a p-n junction formed by an n-well and the p+ substrate. The rectangular shaped n-well collects the photons and generates

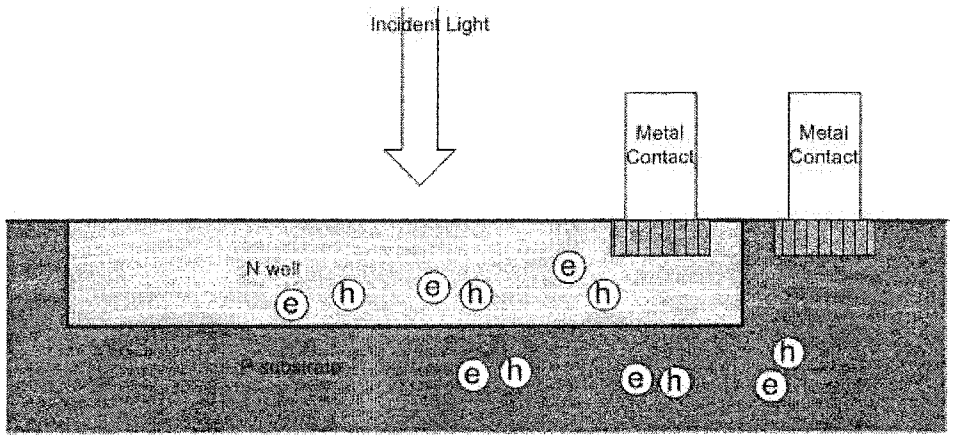


Figure 4.1: the cross-section view of the photodetector

electron-hole pairs primarily in the neutral region outside of the space charge layer. Since there is no strong electric field to attract and accelerate photon-generated electrons and holes, the carriers reach the collecting contacts by thermal diffusion. Because of the large photon collecting area and the depth of the n-well, the photodetector has good photo responsivity. However, the speed of the photodetector is slow since diffusion is the dominant means of collecting the carriers. Because of this, the falling edge of the optical signal is likely to have a long and slow-slope tail.

From previous research experience,[2][3] this type of photodetector is known to be reliably fabricated by standard IC process. The photo responsivity is usually around 0.3~0.4 A/W. The operation speed is usually less than 50 KHz.

To design a receiving circuit for the photodetector requires estimation of the minimum photo responsivity and optical power budget. The optical power budget has to be taken in account to estimate the minimum photocurrent that the Si photodetector can generate. Lens surface reflections, diffractive limitations of the optical components and efficiency of the diffractive optical element together greatly reduce the optical power detectable by the photodetectors. The optical power also

must be divided among 64 individual photodetector units. Furthermore, the need for manual alignment requires some type of mechanism that allows direct monitoring of the alignment. The beam splitter provides convenient aligning process at the cost of heavy (50%) power loss. Given the transmission efficiency of coated surface L_l=99.5% and that of uncoated surface L_u= 85%, the diffractive efficiency of DOE E_d=70%, the transmission efficiency of the beam splitter L_b=50%. For a simple optics system with 4 coated surfaces and 4 uncoated surfaces, the total efficiency η is:

$$\eta = (L_c)^4 \times (L_u)^4 \times \eta_d \times L_b = (0.995)^4 \times (0.85)^4 \times 0.7 \times 0.5 \approx 18\%$$

The calculation indicates that only a small fraction of the total laser optical power can reach the surface of the processor. Because the chip design must allow multiple optics system options, the worst case scenario, i.e. minimum optical power at the longest wavelength and the lowest photo responsivity, has to be considered. A VCSEL array (stand-alone or integrated) usually have enough power since each VCSEL only supplies one photodetector. In an optical fan-out system, the laser optical power is very important because only one laser is used in this case. Off-the-shelf 670 nm high power (~50 mW) edge emitters are readily available. Single mode 850 nm VCSELs have much less power and the Si PD photo responsivity is lower at 850 nm. The optical fan-out system with such a device is the worst case among all the possible options of the light source.

A single mode VCSEL may have 3.5mW power [1] therefore each photodetector unit should be able to receive at least 10μW power if a single mode 850 nm VCSEL is used. Given a relatively low (compared with simple p-n junction type photodetector) photo responsivity of 0.1 A/W at 850 nm wavelength and estimated minimum illumination of 10μW, the photodetector would be able to generate at least 1μA

photo current.

As illustrated in Figure 4.2, the photodetector is reversed biased and linked with a 1:16 current amplifier that amplifies the photo current. The change of photocurrent generated by the photodetector shifts the voltage potential at the input of a voltage comparator, which converts the current signal into a standard CMOS logic signal. In order to adjust the driver circuit to the appropriate operation point, a transistor with an external bias voltage applied to the gate has been added in parallel with the photodetector.

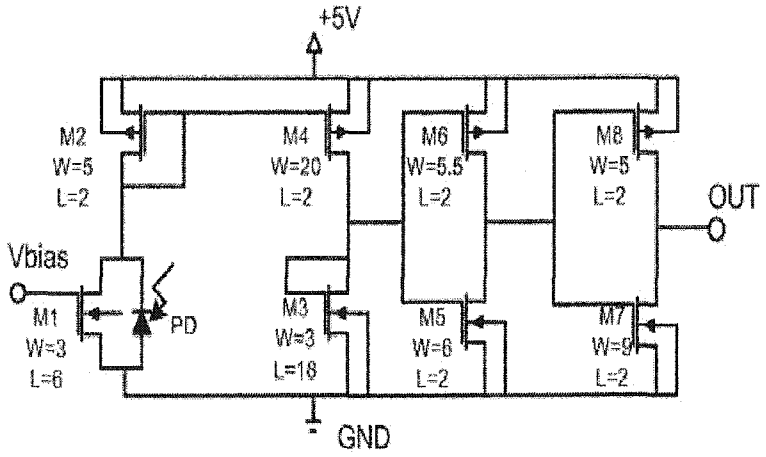


Figure 4.2: the schematic of the driving circuit

The MOS and BJT transistor library was obtained from the AMIS foundry service. The design and simulation were carried out with Avanti H-Spice software. The circuit netlist files were edited with standard text editors then saved to a hard drive. The netlist file includes the information of all the devices, nodes and their interconnections, the component library and the desired simulations. The H-Spice software loads the netlist file and runs the simulations specified in the netlist file. The simulation results are then displayed as figures and numbers in graphical windows. One can evaluate the simulation results for the performance of the driving

circuit.

The design focused mostly on the reliability. The circuit was expected to function with possible photo responsivity variations. The transistor that shunts the photodetector was designed to provide proper bias current for the driving circuit, which in turn controls the DC voltage potential at the voltage comparator. If the photocurrent is weak then the voltage potential is shifted closer to the threshold of the voltage comparator so that the circuit is more sensitive to photocurrent but more vulnerable to noise. If the photocurrent is strong enough then the bias can be shorted to ground to reduce noise.

The requirement for the operation speed of the receiving circuit was low because the photodetector was not a high-speed design. However, it is sensible to have some speed margin on the receiver circuit for the sake of possible improved generation. No obvious trade-off was found by having the top operation speed of the receiver circuit higher than that of the photodetector.

The parameters of the transistors were determined after multiple rounds of simulations, calculations and optimizations. The simulations indicated that the receiving circuit could operate with 1μA photocurrent signal up to 6 MHz frequency.

4.3 Logic design: basic unit and clock unit

The logic modules include two major parts: the comparison array and the clock unit. The comparison array is an array of identical basic units, each of which is a logic circuit that compares one character with the other character. The basic units perform the comparison functions described in the beginning of this chapter. Independent of each other, each basic unit operates under the control of a clock unit. The clocking unit receives an external clock that is sent into the processor

chip along with the character inputs. The data strings are presented on the inputs slightly ahead of the rising edge of the external clock signal. The clock unit uses the external clock to generate its own control signals with multiple instructions, including instructing the basic units to retrieve data from the inputs, to make the comparison, to put the results on the outputs and to reset the counter.

The inputs to the basic unit come from both the optical input and the electronic input. Because each character is represented by a 5-bit word, the basic unit compares the bits of the two words until all 5 bits are compared. If the words are not identical, the 'not match' result is represented as a zero. The 'match' result is represented as a one. Once the words are found to be not identical, the result is zero even though the comparison process continues until all 5 bits are compared. The state diagram of the basic unit is illustrated in Figure 4.3.

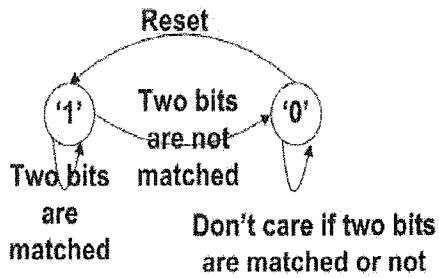


Figure 4.3: The state diagram of the basic unit

Since each of the eight inputs is shared by a row of eight basic units, an additional storage bit is necessary to store data from the electronic input. A control signal from the clock unit instructs when the basic unit should retrieve the input bit and store it. Since the eight outputs are also shared, the same control signal is also used to instruct the basic units to put the result on the outputs. The schematic of the basic unit is illustrated in Figure 4.4.

Because there are 64 basic units that are organized in an 8×8 array and the elec-

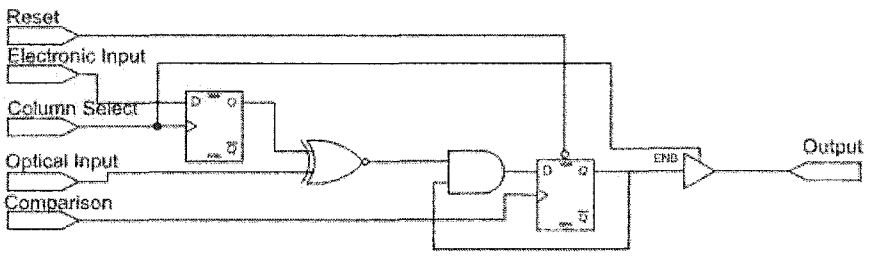


Figure 4.4: the schematic of the basic unit

tronic inputs and outputs are shared among the units, one of the primary functions of the clock unit is to regulate the sharing of the inputs and the outputs. As stated previously in this section, the rising edge of the external clock indicates that new data is available. It takes eight clock cycles for eight inputs to provide data for all 64 basic units. One additional clock cycle is used to make the comparison. One other clock cycle is reserved. Therefore, it takes ten clock cycles to provide one bit of data to each of the 64 basic units and make the comparison. After fifty clock cycles each of the basic units would have received 5 bits of data, i.e. a complete character and completed 64 individual comparisons. Starting from the fifty-first clock cycle, the final comparison results are put on the outputs.

In order to make the comparison process start over, the clock unit generates a reset signal on the sixtieth clock cycle. There is also an external reset input so that the user can interrupt and reset the comparison. The clocking scheme is illustrated in Figure 4.5. (Each grid in the figure represents one clock cycle.)

The schematic of the clock unit is illustrated in Figure 4.6. A 10-stage cascade counter continuously divides the clock into periods, each of which is 10 clock cycles long. Logic circuits generate comparison and internal reset signals. Additional buffers were added to provide sufficient driving capacity for the whole array and to reduce delays in the critical signal paths.

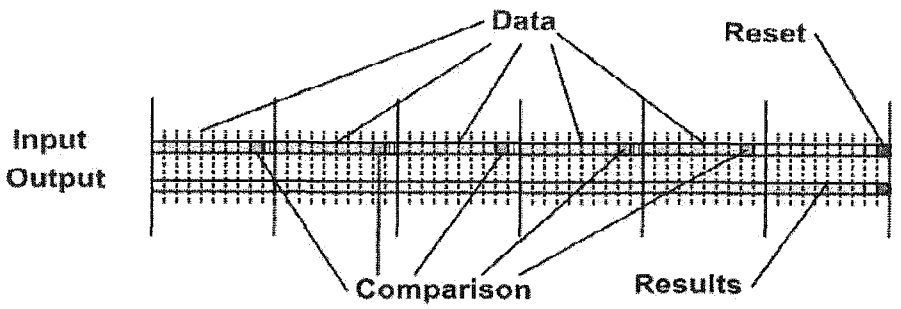


Figure 4.5: the clocking scheme of the clock unit

64 basic units are organized into an 8×8 array and the clock unit is connected to the appropriate signal paths. The complete logic design is illustrated in Figure 4.7.

The logic design was developed using the Mentor Graphics ADK DA (Design Architect) tool. The designs were simulated using AccuSim from the same ADK software package. The simulation indicated that the design should function as expected up to a clock frequency of at least 20 MHz. The simulation charts are presented in Appendix A.

4.4 Layout design: schematic driven layout (SDL)

The actual layout was designed based on the finished and simulated schematics. The ADK software package uses a technique called 'schematic driven layout' (SDL) to guide the layout design. The purpose of the concept is to design the layout based on verified schematic designs. The layout versus schematic (LVS) check ensures that the layout correctly represents the schematic without missing or adding any components and connections.

The SDL starts with components automatically placed in the design area. The logic connections are displayed in the form of virtual 'flying wires', which connect the components. The designer may move components, organize the layout and make

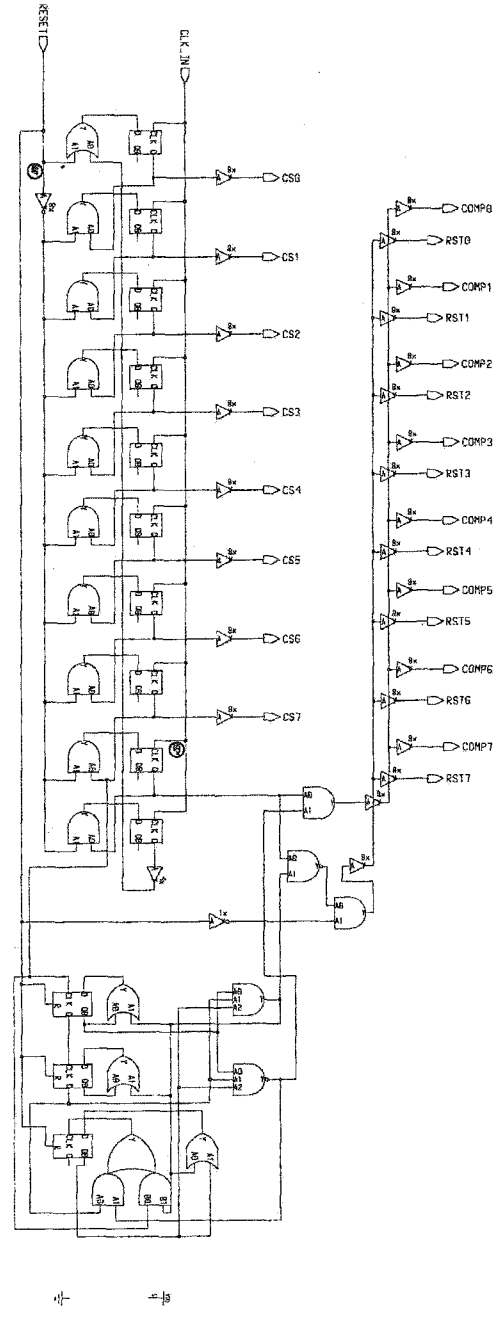


Figure 4.6: the schematic of the clock unit

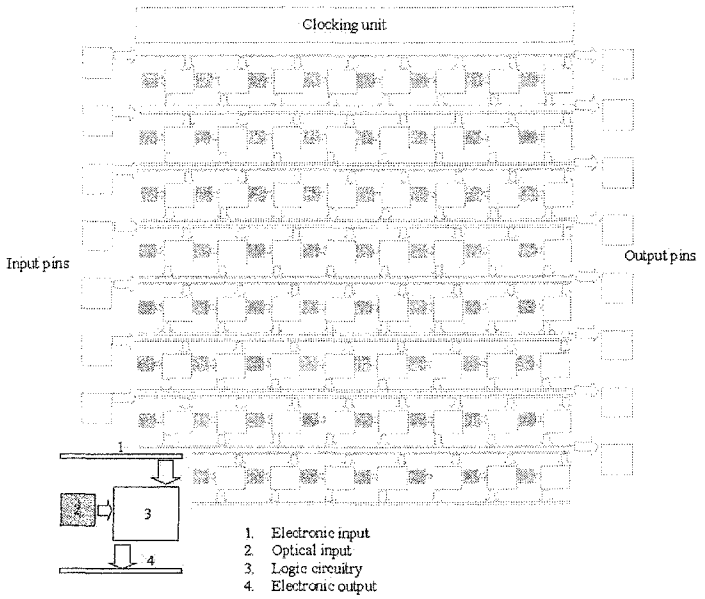


Figure 4.7: the complete schematic of the logic design

'flying wires' into metal connections.

The SDL provides two types of layout design. The transistor level design breaks the components in the schematics into individual transistors, even through the original schematics may use gate-level components. On another hand, only the gate-level components are used in the gate level design. The transistor level design usually uses less chip area and yields better performance. However it also requires a skillful designer and many work hours. The gate level design, on the contrary, takes shorter design time at the cost of area and performance.

The layout was designed using both design methods. The photodetector and its receiving circuit was customly designed at the transistor level. The rest of the basic unit and the clock unit were designed at the gate level.

The major concern in designing the layout of the photodetector and the receiving circuit is to ensure sufficient space for the photodetector and to reduce the overall

area. The design went through several generations. The final version of the layout is illustrated in Figure 4.8.

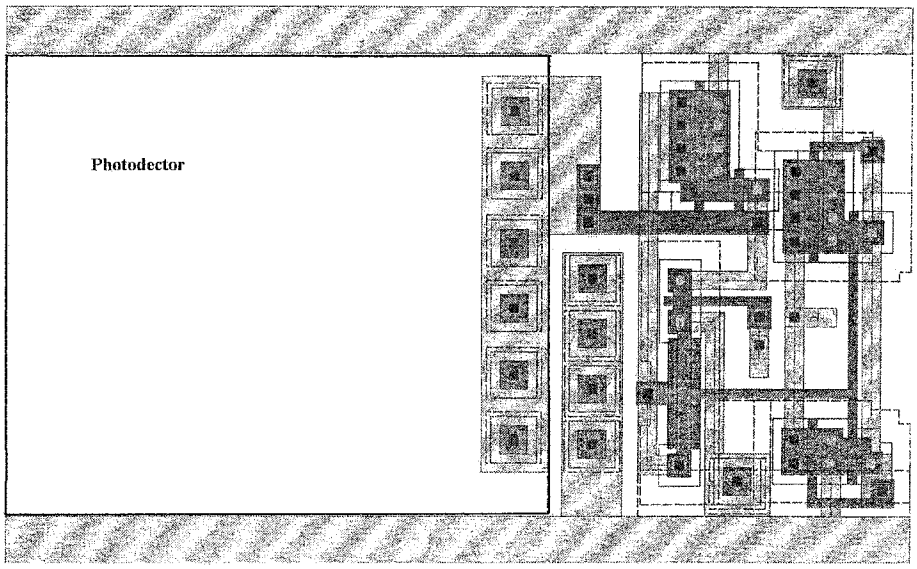


Figure 4.8: the layout of the photodetector and the PD driving circuit

As shown in the left part of the layout, the photodetector is roughly a square shape with dimension of 30×30 μm^2 (the maximum dimension that the power rails and clearance allow). With careful design, the PD driving circuit only takes relatively small area compared with the photodetector.

The layout of a basic unit is illustrated in Figure 4.9. The metal lines carry the input signal, output signal, control signals and power rails. The geometry of the metal paths was designed such that simply placing them side by side makes the connections between neighbor units. The dimension of a basic unit is 125×125 μm^2.

The layout of the clock unit was designed such that the signal paths can be connected with the basic unit array by simply aligning them in proper position. Although this design approach seemed complex, it reduced time and chance of error. The layout of the clock unit is illustrated in Figure 4.10. The majority of the clock

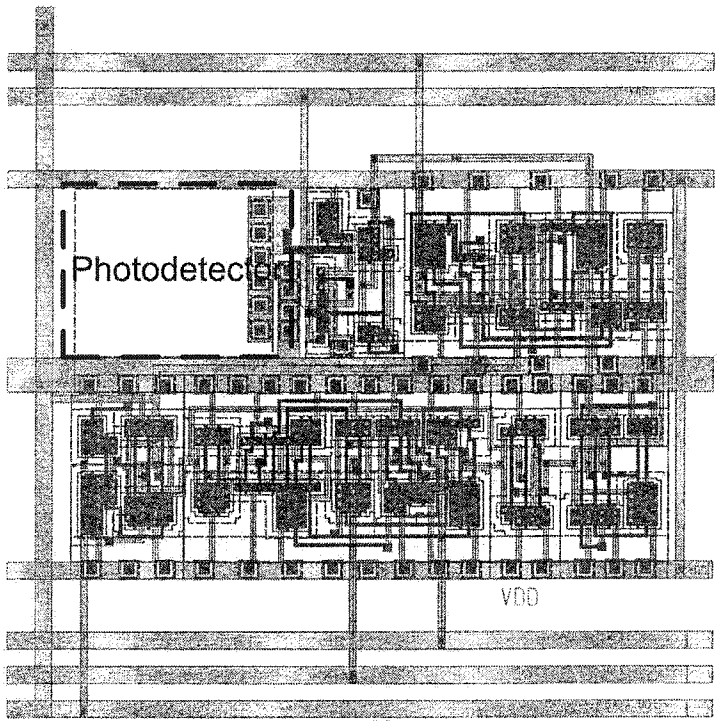


Figure 4.9: the layout of a basic unit

unit is in the bottom of the layout. The driving buffers are on the right side. The clock unit was then aligned with the basic unit array.

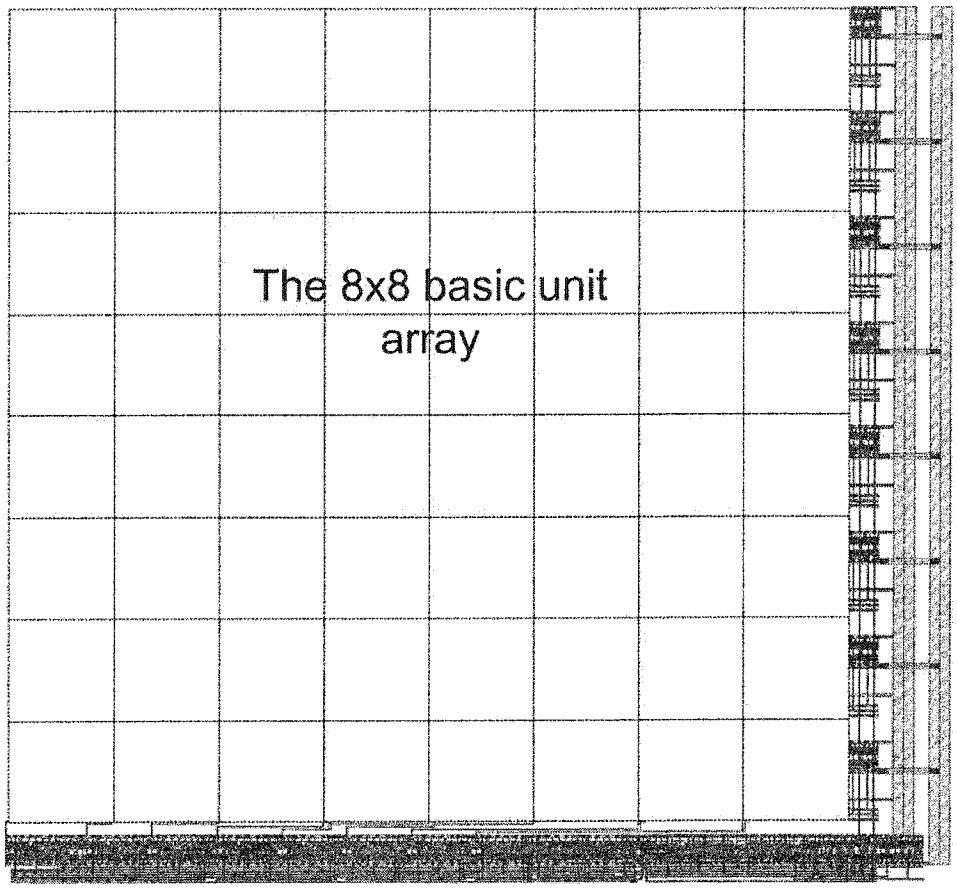


Figure 4.10: the layout of the clock unit

The completed the basic unit and the clock unit layouts must pass the LVS check to make sure that the layout is consistent with the schematic. Parasitic capacitance and resistance were calculated and added to the schematics in a process called 'back annotation'. The back-annotated schematics were then simulated to check if these parasitic parameters in the layout design would significantly affect the functionality and performance. Assembled layouts including the basic unit array

and the array with the clock unit (the 'core layout'), had also gone through this verification process.

After the verification process, the bonding pads were added to the schematics. Additional simulations were conducted to ensure that the pads were compatible. ADK then automatically generates a pad frame that encircles the 'core layout'. The layout design was completed with metal connections made between the pads and the 'core layout'.

4.5 Fabrication and initial electronic testing

The GDSII format design files were submitted to the MOSIS foundry for fabrication. After a few minor modifications the design files were accepted and queued into the MOSIS fabrication schedule. The fabricated chips were returned 3 months later.

Among 25 fabricated dies, 20 were packaged in LCC52 ceramic chip carrier. The remaining 5 were kept as bare dies, which were stored in a gel pack. Figure 4.11 illustrates a packaged and bonded chip.

The initial electronic testing was conducted in the LSI Logic Fort Collins test facility. With assistance from several LSI Logic test engineers, a SUN Solaris workstation controlled AMS semiconductor tester was used for the testing.

The concept of the AMS tester is to treat the chip under test as a 'black box'. The tester feeds test sequences into the chip and monitors the outputs from the chip. The outputs are compared with the expected outputs. If the outputs are consistent with the expected then the test can proceed with additional test sequences. If the outputs are not consistent, then special test sequences are used to determine the possible cause of error.

A series of test sequences were prepared before the tests were performed. The

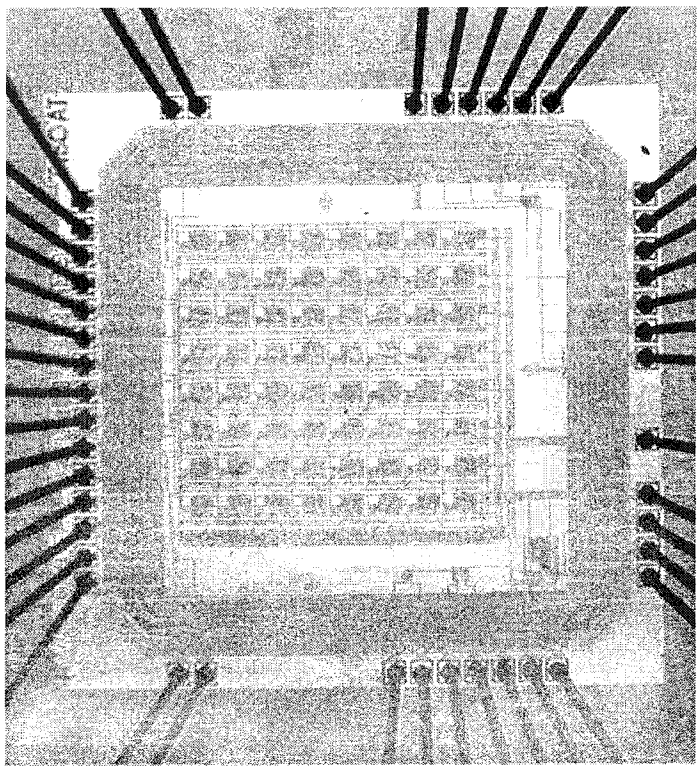


Figure 4.11: A packaged chip

simplest test sequences were merely '1's and '0's while other special test sequences were designed based on the logic design of the comparison circuit. A small C program generated many random test sequences at the maximum length allowed by the tester.

The LCC52 packaged requires special sockets that were not available at the time of testing. The problem was partially solved with a similar but different socket although sometimes the substitute was not able to provide stable connections.

The test was successful after the proper setup time and sampling time were found. The tests proved that the setup time (the time between when the data was presented on the inputs and the rising edge of the clock signal) should be no less than 60 ns. The sampling time (the delay between the rising edge of the clock signal and when the output signals stabilize) should be no less than 20 ns. The timing diagram is illustrated in Figure 4.12.

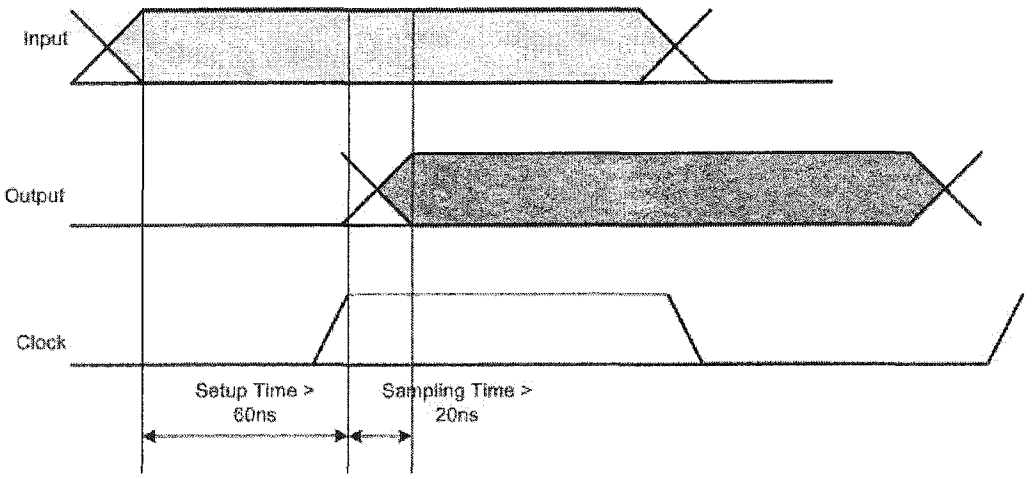


Figure 4.12: the setup time and the sampling time of the tested chip

A screenshot is illustrated in Figure 4.13. The left part of the screen shows the setup of the test signals. The right part of the screen shows the test sequence and the received response from the chip. Note that the circled part indicates that the tester has finished the test without finding any error.

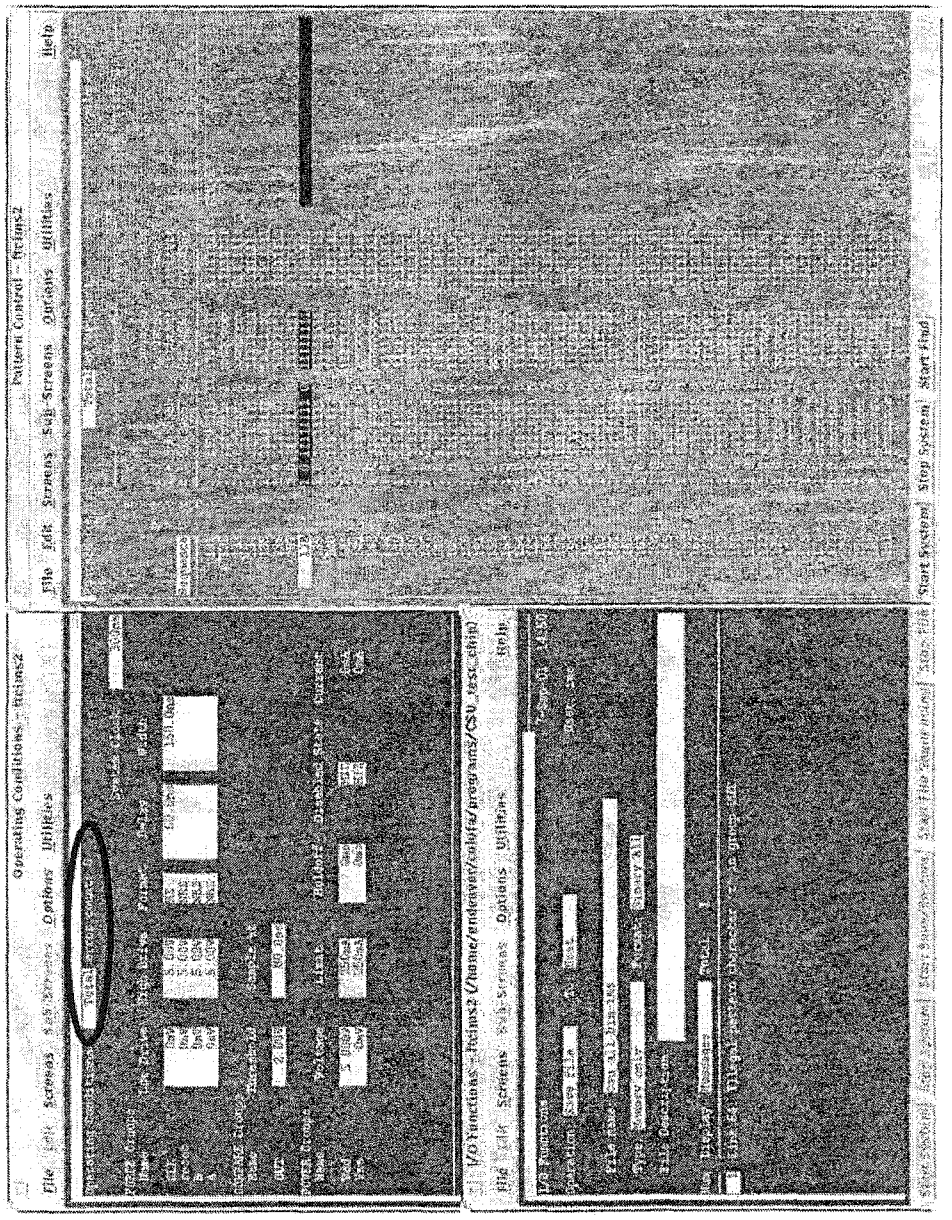


Figure 4.13: the screenshot of the AMS tester

4.6 Summary

The biosequence processor conducts the parallel character to character comparison in the FSOI based biosequence analysis system. It was determined that the AMIS 0.5 μm CMOS process was the most suitable technology for the processor. The processor contains two major parts: the comparison array and the clock unit. The comparison array is a 8X8 array of basic comparison units, each of which have an optical input, an electronic input and an electronic output. The photodetector for the optical input is a p-n junction type design that has good reliability and photo responsivity. The photodetector receiving circuit was designed with Avanti H-Spice and AMIS component models. The remaining schematics and layout design of the processor were conducted on Mentor Graphics ADK design software. Simulations verified the design. The processor was fabricated and packaged via the MOSIS foundry service. Initial electronic test confirmed the functionality of the processor.

Bibliography

[1] 850nm series VCSEL products, Luxnet Corp., http://www.luxnetcorp.com/products/vcsels.html

[2] F.R. Beyette, Jr., P.J. Stanko, E.M. Hayes, R.D. Snyder and C.W. Wilmsen, "An Optoelectronic Recirculating Sorter: Architecture and System Demonstration ", *Optical Engineering*, Vol.37, p312-319,1998

[3] C.W. Wilmsen, C. Duan, J.R. Collington, M.P. Dames and W.A Crossland, "VCSEL-Based Optoelectronic Asynchronous Transfer Mode Switch", *Optical Engineering*, vol. 38, pp. 1216-1222 ,1999

Chapter 5

Optics Design and Setup

The optics system is one of the most important modules of the FSOI based biosequence analysis system. It provides the optical signal source and the means of projecting optical signals onto the biosequence processor chip. A stable and well designed optics system is essential to the reliable operation of the optical interconnects.

In the FSOI based biosequence analysis system, the diffractive optical element (DOE) multiplies a single light source into a designed spot array. It is an optical fan-out mechanism that has found applications in previously reported FSOI systems. [1][2][3][4] In systems that employ DOEs, the quality of the DOEs directly determines the quality of the optics system. The DOE also brings some unique designing issues.

Conventional, i.e. non-diffractive, lenses are used to collimate the laser and to project the spot array onto the biosequence analysis chip. The design of optics system must consider both the optical quality and the availability of optical components. Analysis has been conducted to evaluate the optical quality and possible improvements.

5.1 Issues of diffractive optical element (DOE) fan-out optics

Diffractive optical elements (DOE) use diffraction phenomenon rather than refraction or reflection: they modulate the optical wavefront such that the interference of the modulated wavefront exhibits the desired patterns and optical quality. DOEs can perform functions that are difficult or impossible for refractive or reflective optical elements, including pattern generation from a single light source. Due to their diffractive nature, these elements are sensitive to the wavelength. Some DOEs also require a coherent light source. DOEs are suitable for monochromatic light sources including lasers. The common practice is to design a DOE for a specific laser wavelength.

Designing DOEs is a highly developed technique that requires experience and considerable computational capability. Although there is DOE design software available on the Internet, commercial design companies can provide high efficiency (>70%) DOE components with customer specified parameters. If the research funding allows, a custom-build DOE component would yield the best result with the least amount of optics design effort. However, due to the high cost of custom design and manufacture, one may have to use available components that were designed for some other applications.

Unlike refractive and reflective components, modelling a DOE with unknown structure is complex and difficult. Although it is relatively easy to estimate the behavior of the DOE with a single lens based on the far field pattern of the DOE, dealing with multiple lenses system is far more complicated.

One simplified common approach is to divide the optics into two parts: the

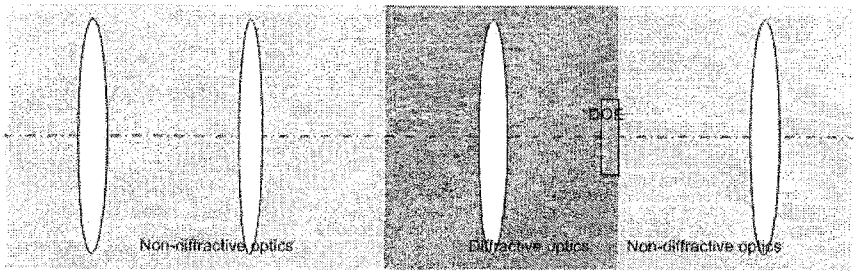


Figure 5.1: a simplified optics design method for systems with DOE components

diffractive part and the non-diffractive part, as illustrated in Figure 5.1. The diffractive part usually only contains the DOE component and one lens for the sake of simplicity. Scalar diffraction theory indicates that the far field, the observation distance $z >> 2D^2/\lambda$, where D is the aperture linear dimension, λ is the operation wavelength) distribution is essentially a Fourier transform of the DOE amplitude transmittance. Putting the DOE in front of the lens also forms a Fourier transform system of the DOE amplitude transmittance with the transformation result appearing on its image plane. One can observe the same pattern (with different scale) from the far field distribution and the lens induced distribution. Therefore any off-the-shelf DOE that can provide the desired far-field pattern is considered a good candidate.

Because the DOE is used as an optical fan-out in the FSOI system, a DOE that generates a spot array pattern is the most suitable for arrayed integrated optical receivers. For the spot array type of DOE, there is a rule-of-thumb that can satisfactorily determine the dimension of the projected spot array. In order to calculate the pitch of the spot array with a focal length of the lens f and the divergence angle (supplied by the manufacturers) α, the pitch of the spot d is:

$$d = 2 \cdot f \cdot \tan(\alpha) \tag{5.1}$$

E.g., for $f = 25$ mm and α=1.07 °, $d = 0.93$ mm.

The non-diffractive optics following the diffractive part of the system can be treated as normal geometric optics. The image plane of the diffractive part is treated as the object plane of the non-diffractive optics. In the case of the spot array, the object is the generated spot array on the image plane (the rear focal plane of the lens).

The method is effective for light sources with short coherent length (i.e. significantly shorter than the length of the optics system). Since laser diodes usually have coherent length around a few cm, the method is valid for the design of optics system using laser diodes as light sources.

5.2 The evaluation of different optical setups

In order to take full advantage of different types of optical system setup, one must evaluate these options for their feasibility and effectiveness.

5.2.1 850 nm VCSEL plus DOE

Vertical cavity surface emitting lasers (VCSEL) emit light that is normal to the wafer surface. The most desirable feature of using a single mode VCSEL in the FSOI system is that it generates a circular shape spot, which saves the need for a stop aperture or correction lens. The small divergence angle also helps to reduce collimated spot size.

The major disadvantages of 850 nm VCSELs are aligning difficulties and low power. The 850 nm emission is in the near infrared (NIR) region, within which small power light is invisible to the unaided human eye. An IR viewer or an IR

observation card can be used to assist the aligning process if the operator is in total darkness, but on average, aligning an IR system takes longer than aligning an optical system using a visible light source. A good single mode VCSEL can provide 2 to 3.5 mW of optical power. Based on the optical power budget calculations in Chapter 4, a single mode VCSEL is able to provide the minimum amount of illumination to trigger the output of the optical receiver. However, low laser power does not allow a margin for alignment error or performance improvement.

The first DOE was made by Rochester Photonics Inc. (now part of Corning) for University of Southern California. It has multiple patterns stacked side by side on one fused silica wafer. One of the patterns generates a 16×8 spot array.

Its characteristics were unknown by the time the component was received. The light source was an 850 nm VCSEL diode (supplied by Emcore) with 2.5 mW maximum power. The light beam was collimated by a short focal length (f=4 mm) lens. The laser and the collimating lens were packaged in a custom-machined holder, which was finely threaded to allow careful adjustment for best effect. The collimated laser beam generates a 2 mm diameter circular spot at 2 meters distance.

5.2.2 670 nm edge emitting laser diode plus DOEs

There were two available DOEs designed for red lasers. Both of them are 8×8 spot array generators, which were manufactured by New Span Opto-Technology Inc and Digital Optics Inc, respectively.

The New Span product was an off-the-shelf designed to operate at 650 nm wavelength and a 0.5° angular divergence. Although the manufacturer claimed that it has 5% intensity uniformity, actual test data illustrated in Table 5.1 indicates otherwise. The generated spot array is illustrated in Figure 5.2.

Table 5.1: The intensity distribution of the New Span DOE

4.39	8.14	11.92	6.91	5.70	7.60	5.67	4.61
5.30	4.26	10.21	4.96	9.18	5.07	5.96	3.30
5.92	8.29	3.86	5.75	8.07	12.06	4.04	4.26
4.44	10.61	8.95	10.41	9.82	5.47	4.78	5.20
4.93	6.39	7.47	11.05	8.12	10.29	9.69	6.12
4.56	4.59	11.60	10.09	7.23	6.54	9.37	5.77
3.28	6.56	4.78	9.69	5.97	9.84	2.99	4.81
4.29	6.02	6.78	3.55	4.71	10.31	6.86	4.34

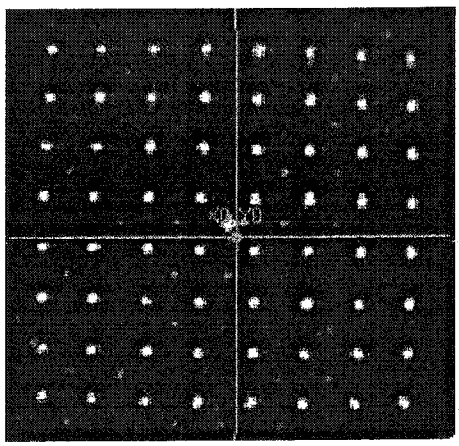


Figure 5.2: the spot array generated by New Span DOE

The Digital Optics DOE was custom designed and manufactured for a previous project. There was no documented data for the DOE, but tests revealed that it operates well with a 670 nm light source.

The 670 nm light source is a high power red edge emitting laser diode (Mitsubishi ML101). With 110 mA maximum driving current, it can generate up to 50 mW of optical power. The same collimating lens and lens holder were used. The collimated laser beam generates a 3×2 mm eclipse spot at 2 meters distance. Tests indicated that the collimated beam is sufficient for use of with the DOE. The 670 nm edge emitting laser and DOE was selected for the optics system.

5.2.3 The optics design

In order to keep the optics design as simple as possible, one must investigate all the possibilities.

One optics system design is illustrated in Figure 5.3. The VCSEL **A** generates a modulated laser beam, which is collimated by the collimating lens **B**. DOE **C** is placed on the front focal plane of a lens **D**. The lens **D** can be chosen from available lenses with a variety of focal lengths. DOE **C** multiplies the laser beam into a spot array. Beam splitter **E** is used to monitor the aligning process. The spot array eventually reaches the detector array of the biosequence processor chip **F**.

From Equation 5.1, the pitch of the spot array is proportional to the focal length of the lens. With a 0.5 ° DOE divergence angle, a lens with a 7.16 mm focal length can generate 0.125 mm pitch, which fits the pitch of the photodetectors on the biosequence processor chip. Obviously it would be difficult to obtain a lens with such a short focal length and sufficient aperture size. Furthermore, it is impossible to insert a beam splitter cube for monitoring the alignment.

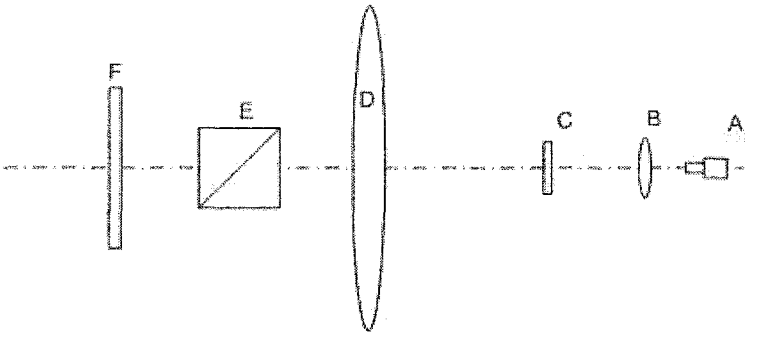


Figure 5.3: optics system with a single lens and a DOE

Another possible approach is illustrated in Figure 5.4. The design is essentially a two-part system. The laser beam generated by laser diode **A** is collimated by collimating lens **B**. The DOE **C** is put at the front focal plane of a front lens **D** to generate the spot array image. The projection lens **E** is used to shrink the spot array image to fit the photodetectors on the biosequence processor chip **G**. The beam splitter cube **F** is used to monitor the alignment process.

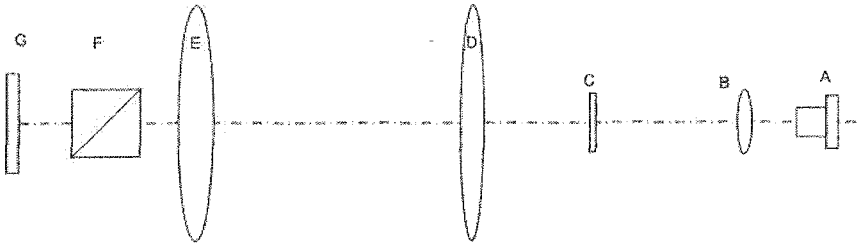


Figure 5.4: The optics system with two lens and a DOE

The key to this system is the projection lens **E**. Because it is impractical to generate the desired pitch with a single lens, a second lens is used to further reduce the pitch of the spot array. The reduction ratio is defined as the ratio of the spot array image size on the rear focal plane of the front lens **D** and the size of the photodetector array.

In order to control the length of the optics system, the reduction ratio must be

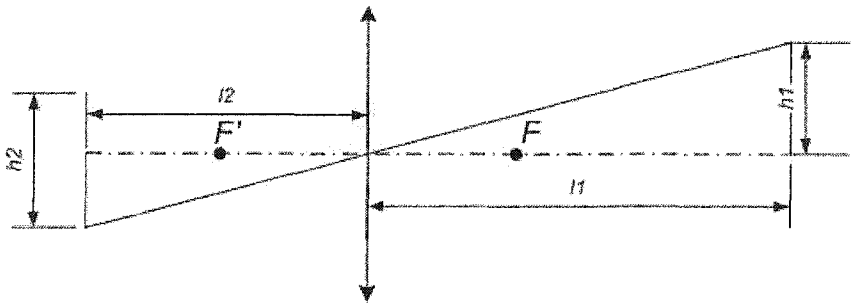


Figure 5.5: Object-image relation for a positive thin lens

within reasonable range. Because the pitch of the spot array is 125 μm, the focal length of the front lens must be quite short.

Newport KBX046 is a 25 mm focal length BK7 bi-convex lens, which is the shortest positive focal length available with a 25.4 mm diameter. Equation 5.1 indicates the pitch d of the spot array is 0.436 mm. Therefore the reduction ratio is 0.436/0.125= 3.488.

As illustrated in Figure 5.5, some rough estimation on the optics system is drawn by geometry optics rules. Given the object height $h1$, image height $h2$, the object length $l1$, image length $l2$ and focal length f, their relations are defined as follow:

$$\frac{h_1}{h_2} = \frac{l_1}{l_2} \tag{5.2}$$

$$\frac{1}{l_1} - \frac{1}{l_2} = \frac{1}{f} \tag{5.3}$$

The reduction ratio $|h_1/h_2|$is already known. Furthermore, l_2 must be large enough to accommodate the beam splitter with sufficient clearance. Because the projection lens and the beam splitter are installed on the optical rail, the minimum distance l_d between them is the distance between the mounting holes on an optical rail. l_d is 25.4 mm. The length l_b of the beam splitter is 12.7 mm. The refractive

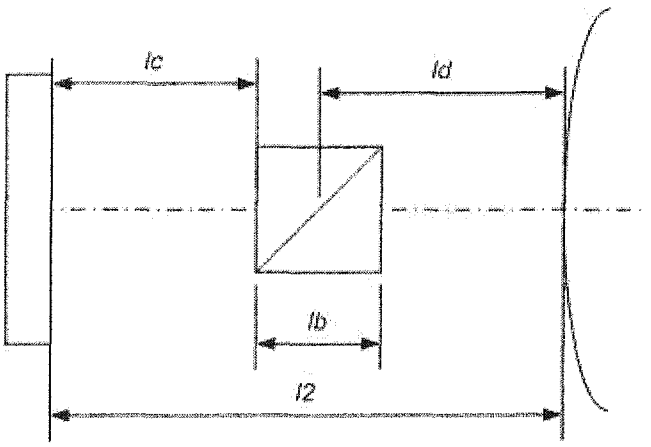


Figure 5.6: Calculation of the minimum optical length for l_2

index n of the BK7 glass (used in the beam splitter) is 1.51 at 670 nm. As illustrated in Figure 5.6, the minimum optical length of l_2 is determined by Equation 5.3.

$$l_2 \geq l_d + \frac{(n-1)\cdot l_b}{2} = 25.4 + \frac{0.51 \cdot 12.7}{2} = 28.634 \tag{5.4}$$

$$f = l_2 \cdot 0.771 \tag{5.5}$$

From Equations 5.2 and 5.3, the relation between the focal length f and l_2 is determined.

Assume the clearance l_c between the biosequence chip and the beam splitter is 10 mm, the focal length f must be longer than 29.8 mm.

Due to the limitation of the optical table and other instruments, the length of the whole optics system must be less than 350 mm. Several available lenses were evaluated using ORA CODE V optical design software. Newport KBX049 is a bi-convex lens with 38 mm focal length. Newport KBX052 is a bi-convex lens with 50 mm focal length. Newport KPX079 is a plano-convex lens with 38 mm focal length.

Table 5.2: Calculated basic parameters for different lenses

Lens	l_1	l_c
KBX049	169.426	23.023
KBX052	224.458	38.528
KPX079	166.605	29.503

The diameters of all of the lenses are 25.4 mm. Some of the software calculated basic parameters are listed in Table 5.2.

The detailed performance curves are included in Appendix B. KBX052 has the best ray aberration curve, which means that the spots on the edge would have more circular shape than the other two. KPX079 has the least field distortion, which means the spot array would be rectangular rather than star-like. Severe field distortion could cause the spots on the edge to be completely out of alignment with the photodetector. The photodetectors can still receive some optical energy even if the spots are comet-like. Therefore the field distortion performance is considered more important than the ray aberration, i.e. it is better to receive partial optical energy than to receive no energy at all.

KPX079 was selected for the projection lens.

5.2.4 Optical setup

An optical path must be established before the entire optical system can be setup. The collimated red laser is ideal for this role. The collimating lens holder that accommodates the edge emitter and the collimating lens were mounted on a 5-axis stage (a three-dimension rotation stage and a XY linear stage). Several apertures were used to align the laser beam such that the laser beam is aligned with the threaded holes of the optical table and parallel to the optical table surface. The

biosequence processor chip was then installed at the image plane and mounted on an extending arm from a XYZ linear stage. Judging from the light spot reflected back from the chip surface, the chip was adjusted to be perpendicular to the laser beam.

Once the chip was in place, the front lens was installed. The center of the lens was carefully aligned to the optical path. The beam splitter cube and the projection lens were mounted on the optical rail before the rail was installed on the optical table. The DOE was mounted on an extension arm attached to an XYZ linear stage.

The DOE was positioned on the front focal point of the front lens to generate the spot array on the rear focal plane of the front lens. The projection lens was moved to the approximate position where it was estimated to produce the desired reduction ratio.

A CCD camera mounted on a microscope was used to align the spot array and the on-chip photodetector. Once the camera caught the image of the chip, the longitudinal position of the chip was adjusted until one could observe a clear image of the photodetector array and the spot array. The projection lens was then slowly moved along the optical rail until the desired reduction ratio was achieved.

The transverse position of the chip was shifted until the spots were right on the photodetectors. Because there is no metal lines designed above the photodetectors, the photodetector actually 'shrines' when a light spot is right on a photodetector. This phenomenon was used to assist the alignment process.

The chip must be put into operation in order to verify the design of the optical inputs and judge the quality of the alignment. The electronic and optical test processes are discussed in Chapter 6.

The finished optics system setup is illustrated in Figure 5.7.

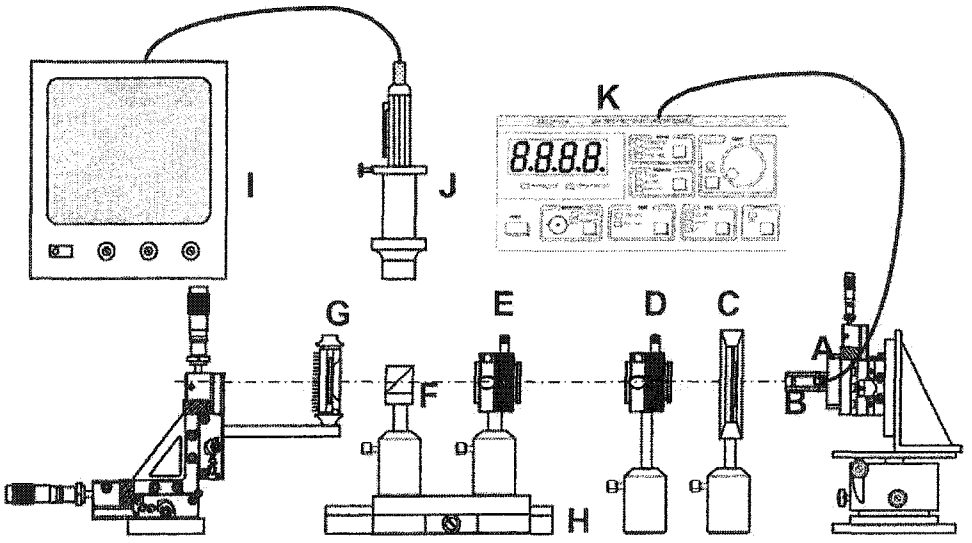


Figure 5.7: The finished optics system setup
A. red edge emitting laser; B. collimating lens; C. DOE; D. front lens; E. projection lens; F. beam splitter cube; G. biosequence analysis chip; H. optical rail; I. Video monitor; J. CCD camera with microscope; K. laser current source

The CCD image of aligned spot array and the photodetector array is illustrated as Figure 5.8.

5.3 Summary

The optics system uses a DOE pattern generator to yield the desired 8×8 spot array for the biosequence processor chip. A 670 nm edge emitter was selected as the light source. A two-lens system was designed to meet all the requirements for the optics system. The optics system was assembled on an optical table.

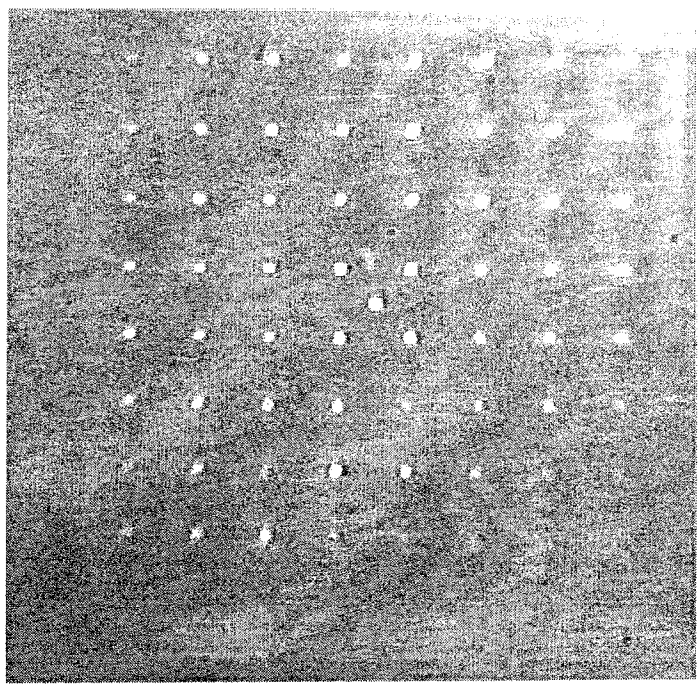


Figure 5.8: CCD image of the spot array aligned with the photodetector array

Bibliography

[1] Tao Ao, Chunji Duan, Carl Wilmsen, "VCSEL-based free-space optoelectronic ATM router", *Proceedings of SPIE - The International Society for Optical Engineering, Optoelectric and Wireless Data Management, Processing, Storage, and Retrieval*, Vol. 4534, p122-128, 2001

[2] Carl W. Wilmsen, Chunjie Duan, James R. Collington, Mark P. Dames, William A. Crossland, "Vertical cavity surface emitting laser based optoelectronic asynchronous transfer mode switch", *Optical Engineering*, Vol.38, No.7, p1216-1222, 1999

[3] John A. Neff, "Holographic optical interconnect elements in an FFT processor", *Proceedings of SPIE - The International Society for Optical Engineering, Diffractive and Holographic Device Technologies and Applications IV* , Vol.3010, p241-252, 1997

[4] D.W. Prather, S. Venkataraman, M. LeCompte, H. Bates, F. Kiamilev, "Three dimensional micro-optical architecture for chip level optical interconnects," *Proceedings of SPIE - The International Society for Optical Engineering*, Vol.4292, 2001, Optoelectronic Interconnects VIII, p197-207, Jan. 2001

Chapter 6

Initial System Tests

Once the optical system was assembled and the biosequence processor chip was in place, initial tests were conducted to evaluate the system functionalities. A HP16500 logic analysis system provided emulated inputs for the biosequence processor chip and monitored the signals between the biosequence processor chip and the programmable logic device.

The programmable logic device was introduced into the system for the first time. The signal timing between the field programmable gate array (FPGA) board and the biosequence analysis chip is crucial to successful data exchange between the two modules. A logic level converter board is also necessary for converting 3.3 V CMOS and 5 V CMOS logic signal levels.

6.1 Optical tests of the biosequence analysis chip

Once the optical system was in place, various optical tests were conducted to evaluate the operation properties of the optical interconnects. First the dc response of the optical inputs was tested to evaluate the alignment quality of the optical sys-

tem. Then the functionality test was conducted to verify the design of the build-in Si photodetector and its receiver circuit. The input clock frequency was gradually increased to determine the speed limit of the optical interconnects.

The layout of the complete optical test system is illustrated in Figure 6.1. The actual system setup is illustrated in Figure 6.2. The system was installed on a Newport optical table. The HP16500B logic analysis system is at the left side of the optical table. The optics system is in the center. The edge-emitting laser is at the right side of the optical table. The laser current driver is an ILX current source with external modulation signal input. A Panasonic CCD video camera is mounted on a magnifying lens system to monitor the aligning process between the photodetector array and the spot array. A power supply provided 5V DC power and tunable voltage bias to the biosequence analysis chip.

The HP16500 logic analysis system provided multiple emulation signals to test the functionality and performance of the biosequence analysis chip. It also monitored the output of the chip for functionality and performance tests.

6.1.1 Static functionality test

The purpose of the functionality test is to verify the design of the biosequence analysis chip. The first step of the test was to determine the working parameters for the chip, including the required laser output power, the external voltage bias for the receiver circuit, the clock frequency and the timing parameters. The HP16500B system provided emulated test signals. For example, if the electronic inputs were receiving '0' and the optical inputs were receiving '1', the comparison outputs should be '0', which indicate "no-match" comparisons. By providing various test inputs, one could determine if the optical signals are received correctly.

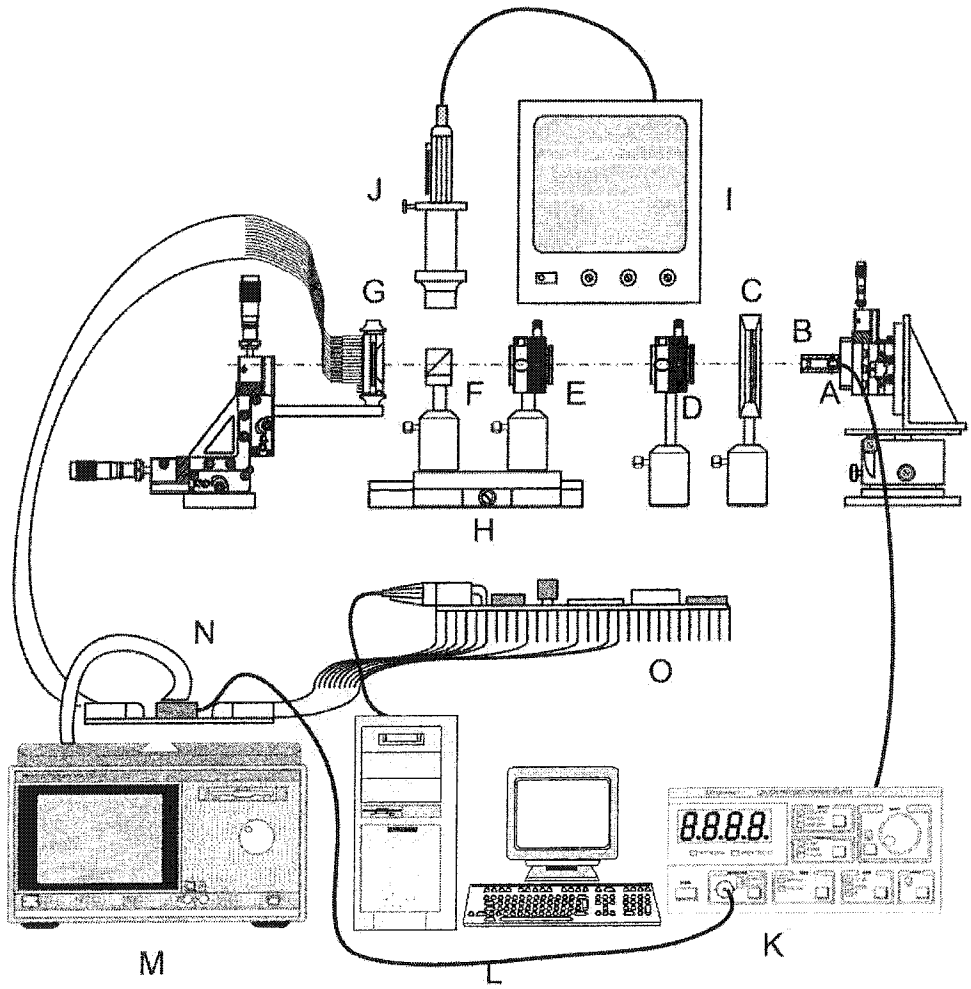


Figure 6.1: The layout of the optical test system
A. 670 nm edge emitting laser; B. collimating lens; C. DOE; D. front lens; E. projection lens; F. beam splitter cube; G. biosequence analysis chip; H. optical rail; I. Video monitor; J. CCD camera with microscope; K. laser current source; L. Computer; M. HP16500B logic analysis system; N. Logic level conversion board; O. FPGA board

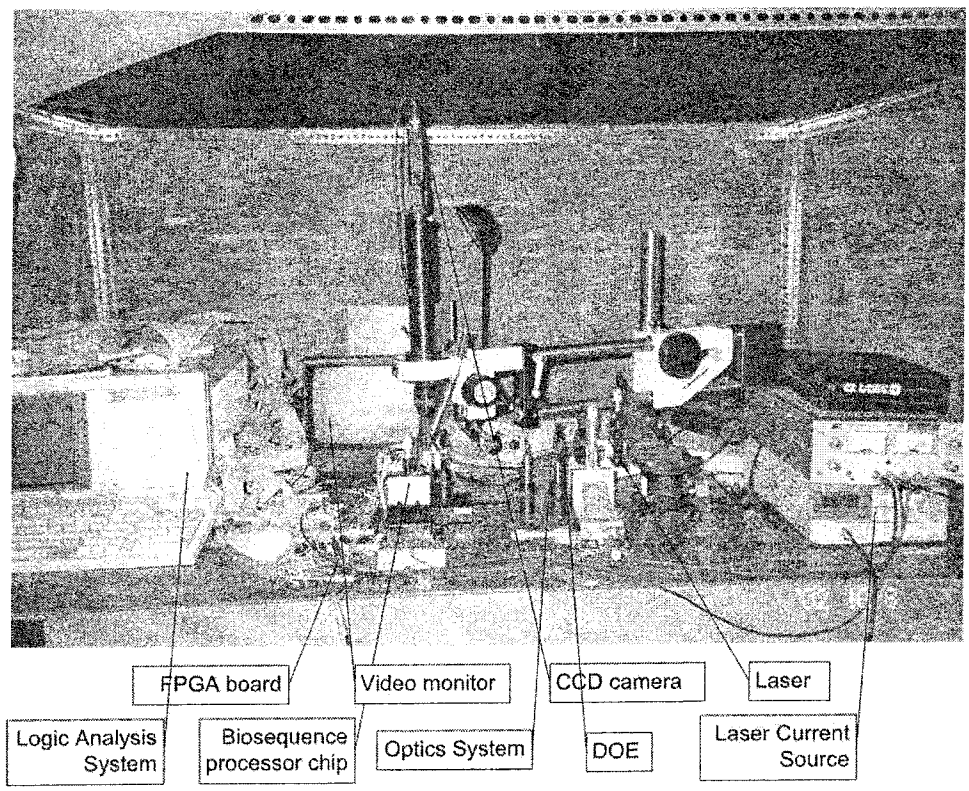


Figure 6.2: The picture of the optical test system

The bias voltage must be appropriate so that the photodetector receiver circuit would function. Because the properties of the photodetectors were unknown when the chip was designed, the receiver circuit was designed with flexibility between the photo responsivity and noise margin. Once the spot array and the photodetector array were roughly aligned, the laser was then turned off and the voltage bias to the chip was set to 0V. The outputs of the chip indicated that the optical inputs were not receiving any optical power because there were none. The bias voltage was slowly increased until the chip began to generate some random outputs. This indicated that the output of the photocurrent amplifier became very close to the threshold voltage of the comparator circuit. The voltage bias was then slightly decreased to stabilize the chip outputs to be '0'. The highest external bias voltage that generated a stable '0' output was about 1.0V.

Without any modulation, the laser driving current was set to approximately the maximum laser operating current to provide a high DC optical signal. The DC optical signal provides the comparison units with a stable '1' if the photodetectors receive the optical signal correctly.

The processor chip did not respond correctly the first time because the illumination of the chip was insufficient for precise alignment of the spot array with the photodetector array. Unless both the spot and photodetector arrays are clearly visible in the monitor, one could not tell if they are aligned well. Because the chemical mechanic polishing (CMP) processed chip layers are very flat and smooth, the chip behaves like a mirror, i.e., light reflected from the chip is highly directional without much scattering. The magnifying lens of the CCD camera only allows a small acceptance angle, i.e., the illumination light must be within a very limit range. This limitation imposed by the chip surface and the magnifying lens means that the illu-

mination light must be parallel with the optical path such that the smooth surface of the chip could reflect into the magnifying lens. A light source with fiber optic header was introduced to project the illumination light onto the chip. The fiber optic header was bonded on the microscopic lens so that the illumination beam could be reflected by the beam splitter cube and propagate parallel with the optical path. The improved illumination allowed the precise alignment. After some aligning efforts the chip outputs indicated that the receivers were getting a stable '1'.

The laser driving current was then slowly reduced while monitoring the output. Due to the size limitation of the power meter probe, the minimum laser power that maintains stable '1's on the receiver circuit was measured between the front lens and the projection lens. The measured power is 0.46 mW, which corresponds to a total laser power of 0.80 mW. Taking additional loss introduced by the beam splitter and the projection lens into account, each on-chip photodetector receives 4.4 μW of optical power. The corresponding laser driving current was 58.1 mA.

Figure 6.3 and Figure 6.4 are screenshots taken from the HP16500B system illustrate waveforms of static functionality test. The **B0_7** signals are the electronic inputs to the chip. The **CLOCK** is the clock signal and the **RESET** is the reset signal. Signal **A** is the modulation signal for the laser current source. In this case signal **A** was not connected with the laser current source. The **SPACER** is used to indicate different comparison stages. The **OUT0_7** signals are the outputs from the chip. Figure 6.3 illustrates multiple comparison cycles. The falling edge of the **RESET** signal indicates the beginning of each comparison cycle. If the optical inputs are functioning properly, then the outputs should be '1' in the first 10 clock cycles after the falling edge of the **RESET** signal. Then the outputs should drop to '0' for the remaining 50 clock cycles. The correct signal waveform was expected

to be very regular. The Figure 6.3 shows irregular outputs in the dotted rectangle that could be caused by poor alignment and insufficient optical power. Figure 6.4, on the contrary, shows output waveform exactly as predicted. Note that due to the resolution limit, the **CLOCK** waveform is shown as a black strip in the figures. There are 60 clock cycles between two adjacent **RESET** pulses and 10 clock cycles between two adjacent **SPACER** pulses.

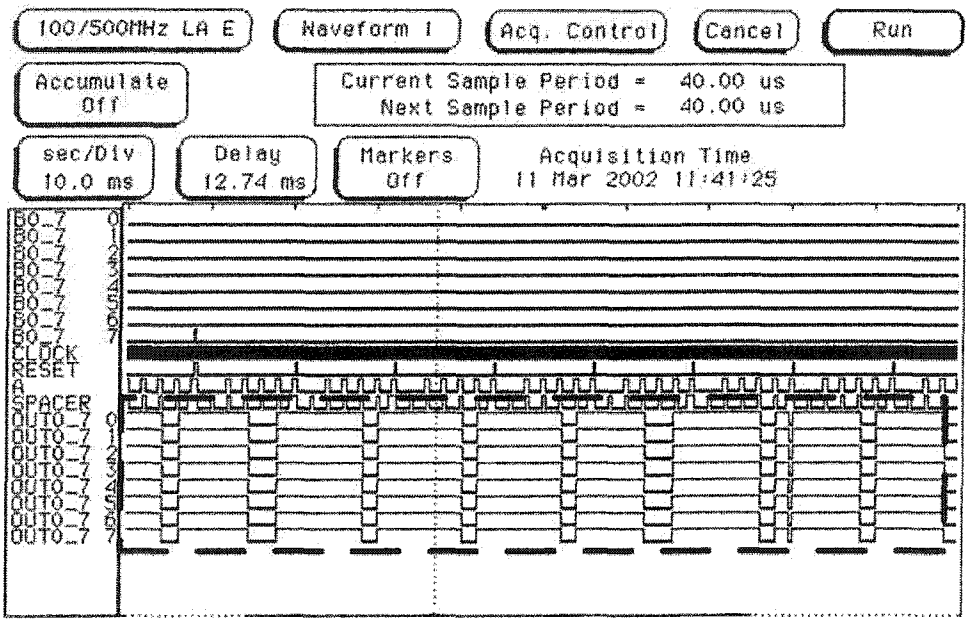


Figure 6.3: One set of the erroneous outputs of the chip under static tests

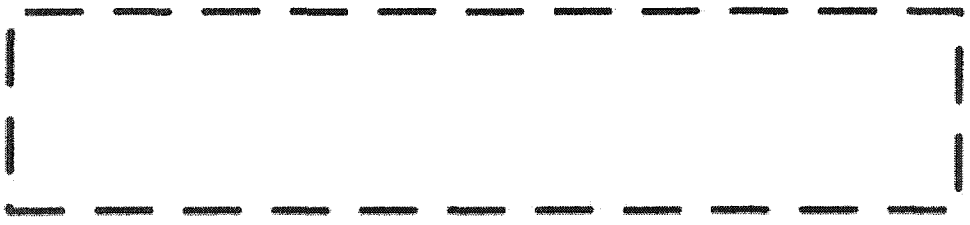


Figure 6.4: One set of the correct outputs of the chip under static tests

6.1.2 Dynamic functionality test

The dynamic working parameters, including the clock frequency and the timing requirements, were determined with the dynamic tests. Signal **A** was connected to the current source to modulate the laser. The optical signal frequency was 50 kHz, which corresponds to 500 kHz electronic clock frequency. The electronic inputs were '0' and the optical inputs were '1'. The predicted waveform should as the same as in Figure 6.4.

As illustrated in Figure 6.5, the initial dynamic test generated some erroneous results. Since the chip passed the static tests, the response speed of the photodetector was the first suspect as the cause of the problem. However because lowering the optical signal frequency to less than 10kHz did not solve the problem, it was unlikely that the photodetectors were unable to function with such low frequency. Further investigation suggested that the falling edge of the signal **A** might appear too early in the comparison cycle. The original timing was selected because the p-n junction type of photodetector was known to have a long tail that would turn a sharp falling edge into a gradually falling slope. It seemed that the 'tailing effect' was not sufficient to maintain the optical inputs long enough. The new signal **A** had a 1/10 duty cycle. If the optical inputs are '1's, an impulse would be presented immediately before the comparison starts, i.e. the 8^{th} clock cycle in each 10-clock-cycle long comparison cycles. The width of the impulse is exactly one clock cycle.

By doing so the photodetectors should generate sufficient photocurrent at the moment when the electronic and the optical inputs are compared. The modified timing scheme is illustrated in Figure 6.6. The outputs were all correct except the last column. It seemed that the optical signal did not last long enough for the last column of electronic inputs. More changes were needed to correct the error. Due to

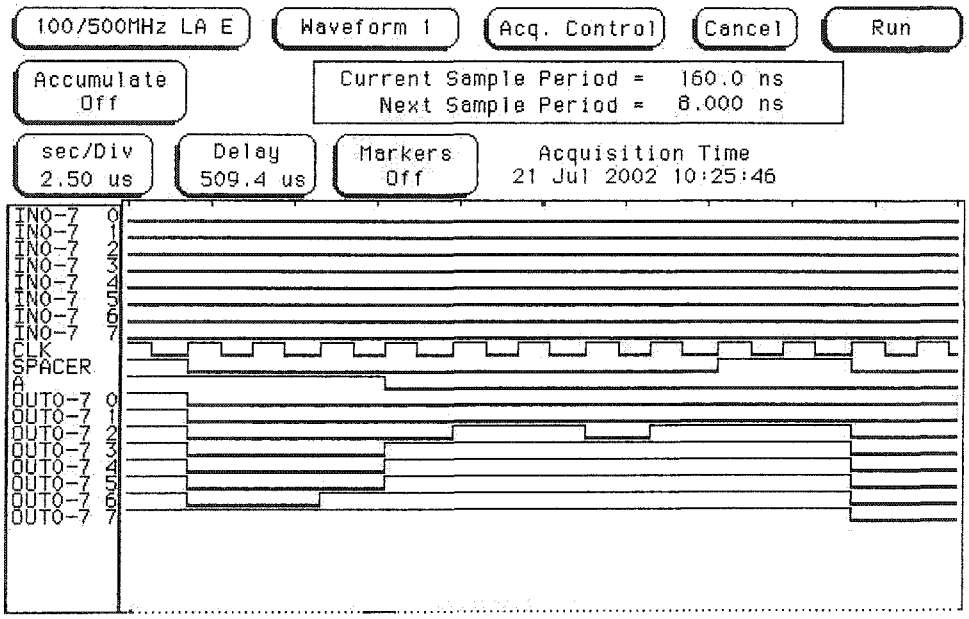


Figure 6.5: One set of the erroneous outputs of the chip under dynamic tests

the restriction of the HP 16500B system, the rising edge of the optical signal must be presented at the beginning of a clock cycle. The falling edge of the optical signal must be presented at the end of a clock cycle. To put the optical signal impulse to the right position appeared difficult.

The solution came with an inverted clock signal. By inverting the clock, the optical signal impulse became stable when the raising edge of the 9^{th} clock cycle is presented. With this timing scheme the chip generated the correct outputs. Figure 6.7 illustrats the screenshot of a set of correct outputs.

The new timing scheme also provided better performance. Originally it was estimated that the photodetectors could operate up to 50kHz. The dynamic tests revealed that the chip could operate with 100kHz optical inputs.

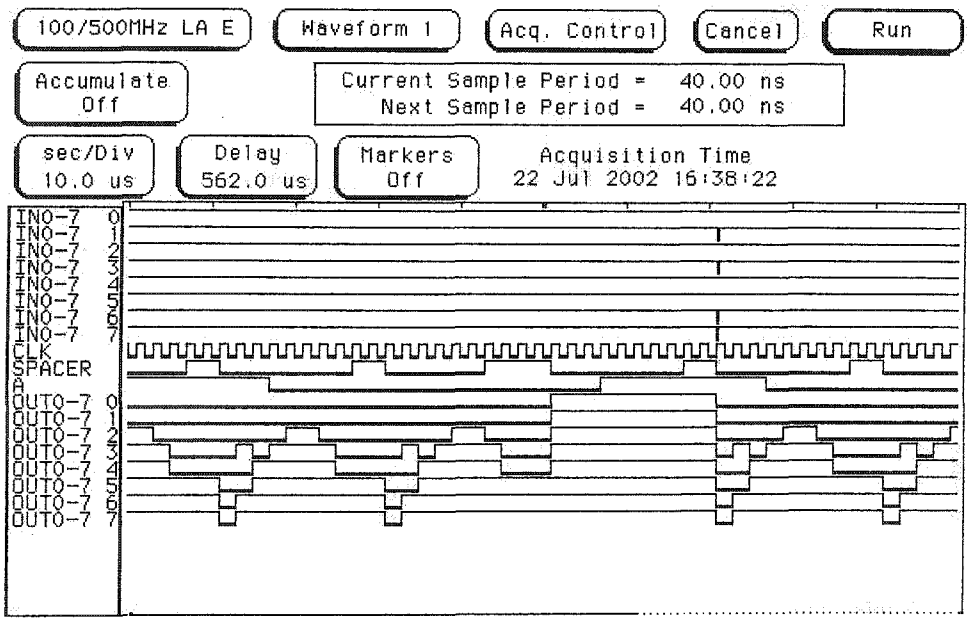


Figure 6.6: The outputs of the chip under the modified timing scheme

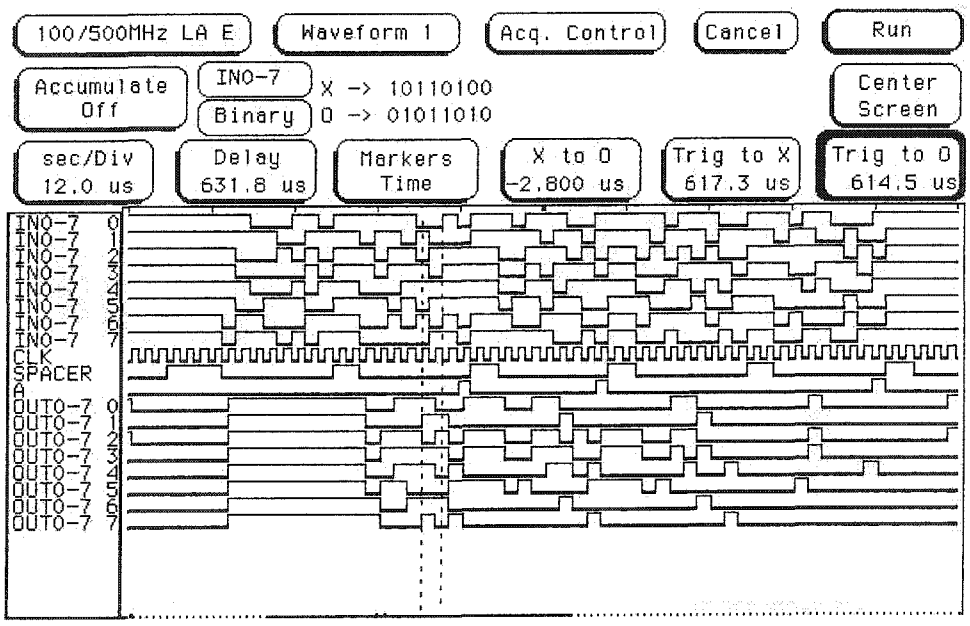


Figure 6.7: One set of correct outputs of the chip under dynamic tests

6.2 The FPGA boards

The field programmable gate array (FPGA) board carries a state-of-the-art FPGA chip, SRAM and flash ROM memories, multiple communication ports and other accessories including display, push button or switches. The FPGA board carries out two of three steps of the BLAST algorithm: it controls and exchanges data with the biosequence analysis chip and sends the result to the personal computer. Some heavy work of programming is required to fulfill all of these requirements. Some hardware design is also required to convert different CMOS logic levels between the FPGA and the biosequence analysis chip.

Two FPGA boards were used in the project. The first FPGA board was the XSA-100 prototyping board is supplied by XESS Inc. Its Xilinx XC2S100 Spartan-II FPGA has insufficient processing capability for the complete BLAST algorithm therefore the XSA-100 board was replaced. The second FPGA board was a TE-XC2Se Spartan IIe development platform, which was made by Trenz Electronic GmbH. The device used in TE-XC2Se board is a XilinX XC2S300E FPGA that has 300,000 gates, three times of the gate number in XC2S100 used in XSA-100 board. The board has 8 Mb flash memory, 4 Mb SRAM, monochrome LCD display and many other accessories. It provides up to 100 user programmable 3.3V CMOS logic I/O pins.

The logic functions of the FPGA boards was designed and programmed with Xilinx ISE integrated design environment. Mr. Klaus Hartinger carried out the algorithm design and programming of the FPGA. More details are available in his master thesis. [1]

6.2.1 The logic conversion

The FPGA board operates on 3.3V CMOS logic. The biosequence analysis chip operates on 5V CMOS logic. Tests revealed that the driving capacity of the FPGA chip was limited. The logic '1' generated by the FPGA chip was usually around 3V, which only leaves 0.5 V margin for 5V CMOS logic. Besides, the FPGA chip can not directly receive 5V inputs for a long period of time. Therefore, it was necessary to construct logic conversion mechanisms to translate the signal levels between the FPGA board and the biosequence analysis chip.

The key components for the logic conversion are integrated buffers. For example, a TI 74LVC244 octal buffer operates on 3.3V CMOS logic signal outputs but it can receive 5V signal inputs. A TI 74LVC07A hex open-drain buffer ICs also operates on 3.3 V but the output signals can be pulled up to 5 V with external pull-up resistors. 74LS244 and 74ACT244 octal buffers are also good alternatives because they use TTL logic, in which the threshold is 2V.

Two logic conversion mechanisms have been constructed. The first one is a logic conversion board. It uses two 74LVC07A for 3.3V to 5V logic conversion. The schematic and the printed circuit board (PCB) design were conducted using Protel99SE design software. PCB Express, Inc manufactured the PCB. The schematic design is illustrated in Figure 6.8. The PCB design is illustrated in Figure 6.9. The logic conversion board was hand soldered and connected with multiple connectors. It also has two dedicated connectors for HP16500B logic systems so that the probing is straightforward. The FPGA board, the processor chip and the logic conversion board are connected with ribbon cables.

The second logic conversion board is a simple prototype board with dual inline package (DIP) octal buffer ICs using two 74LS244 for 3.3V to 5V logic conversion.

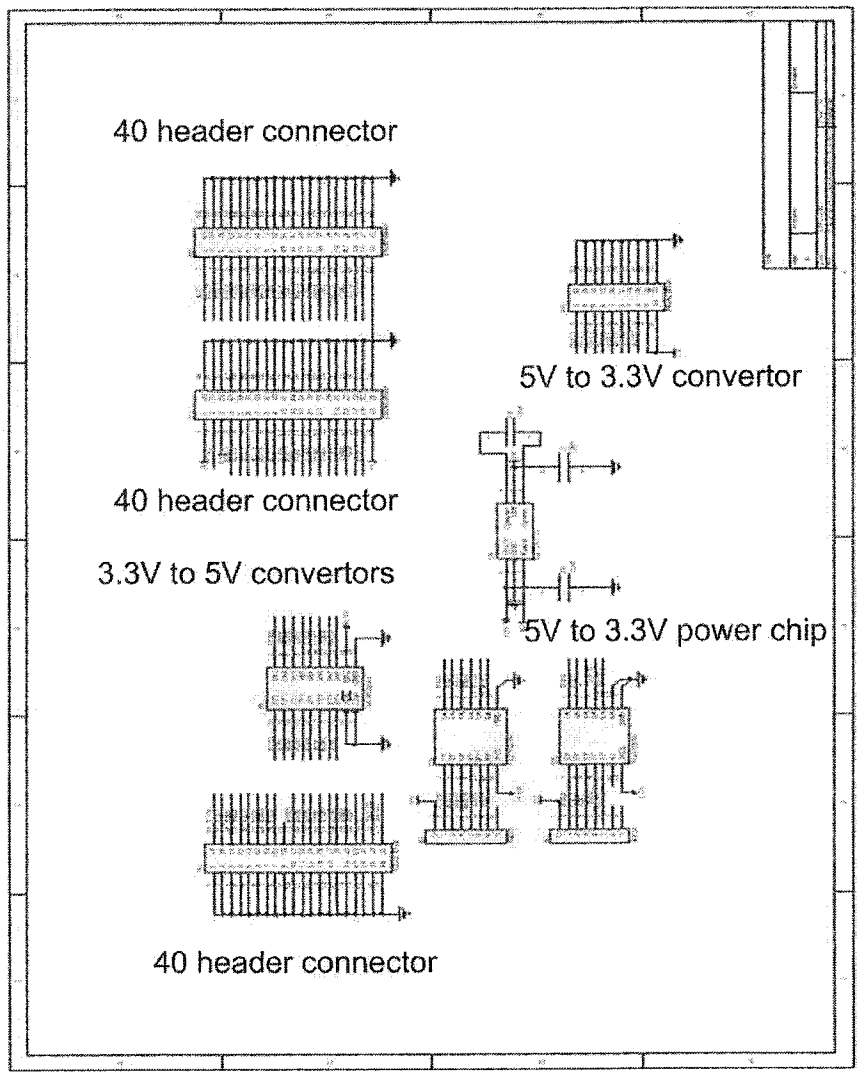


Figure 6.8: The schematic of the logic conversion board

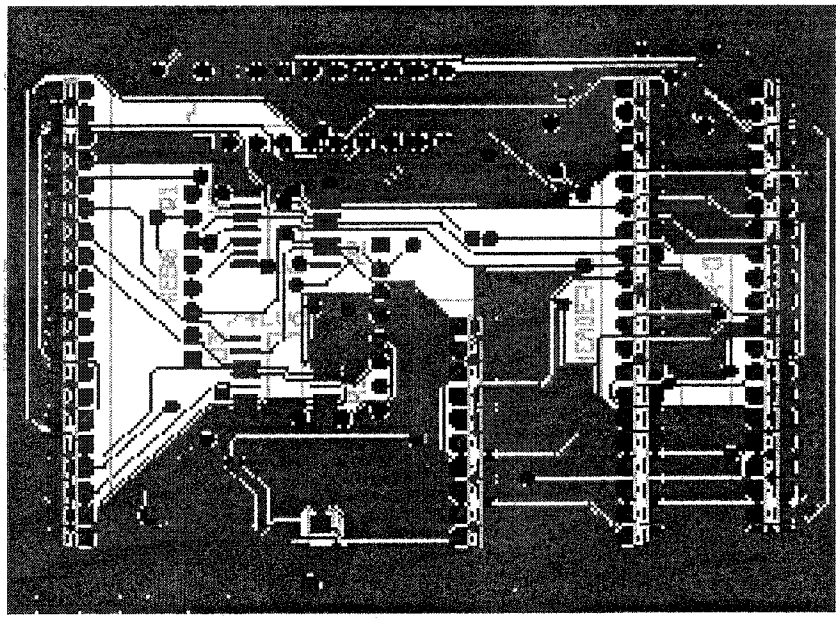


Figure 6.9: The PCB design of the logic conversion board

The prototype board offers more flexibility over the printed circuit board at the cost of lower speed.

6.2.2 The interfacing between the FPGA board and the biosequence chip

The interfacing between the FPGA board and the biosequence processor chip requires more than logic level conversion. The timing characteristics in the signals on both directions are also important.

The first issue is the clock distribution from the FPGA board to the biosequence processor chip. The maximum clock frequency is 1 MHz, which usually does not require impedance matching and other signal integrity measures. However, due to the fast internal operating speed of the FPGA, the duration of the rising and falling

edges of the clock is generally around 2~3 ns. These short edges require proper line terminations otherwise the signal becomes unreliable because of over- and undershoot caused by signal reflections on both the transmitting and the receiving end.

There are several conventional ways of reducing the over and undershoot, as illustrated in Figure 6.10.

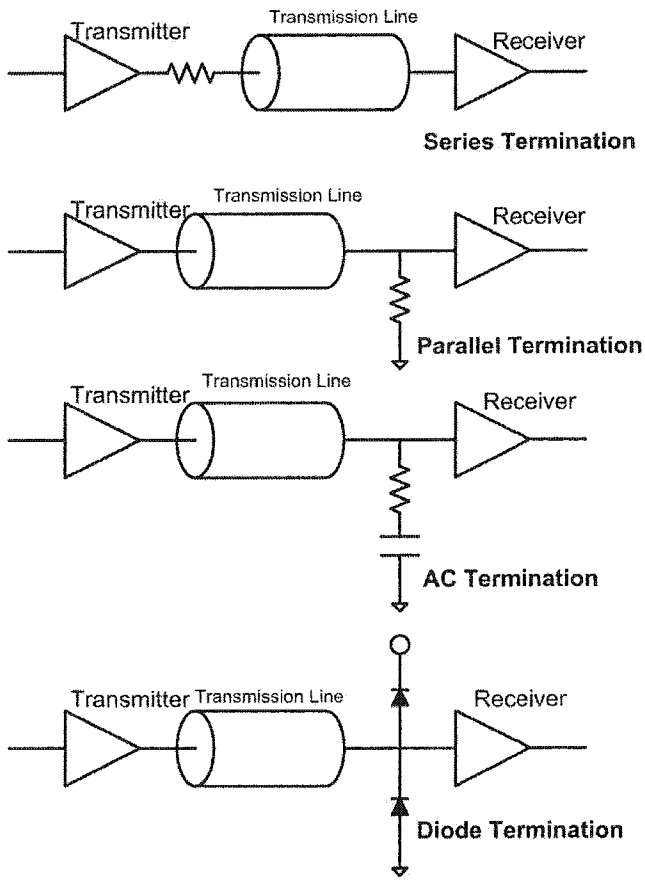


Figure 6.10: conventional ways of transmission line termination

The impedance matching is difficult for the 74LVC07 with pull-up resistors. If the impedance of the transmitting pin matches that of the transmission line, there is no reflection at the interface of the pin and the line. The series termination method calls for an additional resistor between the pin and the line such that the impedance

is matched. The method is impractical here because the output impedance may vary at different operation frequencies and logic states. The parallel termination is also impractical because the FPGA was found incapable of supporting an extra DC current drain. Many combinations of resistors and capacitors were tested on the actually translation board to look for the proper AC termination. As the theory predicted, the method was very sensitive to the capacitance value. A RC combination effective in the test became totally useless after the excessive length of the component pins was removed and the resistor and the capacitor soldered onto the printed circuit board. The diode termination method using fast Schottky diode was also tested but the result was not promising. In general, there is no good way to make a reliable line termination for the 74LVC07 design.

Line termination of 74LS244 is much easier. By using a 50 Ω resistor in series termination, the overshoot and undershoot were reduce from 30% to 1$\sim$2%. The amplitude of the terminated signal is around 3.5$\sim$4.1V. The signal quality is sufficient for reliable interfacing between the FPGA board and the biosequence processor chip.

Once the functionality of the processor chip was proven and reliable interface has been established, more performance tests concerning the processor chip and the optical interconnect were carried out as discussed in Chapter 7.

6.3 Summary

Initial system tests verified the functionality of the biosequence processor chip and the optics including the optical transmitter and the optics system. Issues for logic conversion were also discussed.

Bibliography

[1] Klaus Hartinger, Thesis for Master of Science, "An Optoelectronic System using BLAST implemented on a Field Programmable Gate Array to perform Biosequence Analysis", Department of Electrical and Computer Engineering, Colorado State University, Summer 2003.

Chapter 7

Complete System Operation

The basic functionality of the FSOI based biosequence processor chip was verified as presented in Chapter 6. Additional tests were conducted to evaluate the reliability and performance of the processor chip and the complete biosequence processing system. The system performed biosequence alignments with the BLAST algorithm to demonstrate its capability for biosequence analysis. The complete FSOI based biosequence processing system was shown to be operational with a BLAST algorithm implemented on the Te-XC2Se Spartan IIe FPGA development platform.

7.1 System operation requirements

The functionality tests of the FSOI based biosequence processor chip presented in Chapter 6 validated the logic design of the processor chip. The tests also proved that the chip had functional optical inputs and electric inputs/outputs. In order to prove that the complete FSOI based biosequence analysis system operates as expected, further tests have been conducted.

The first issue to be resolved was whether the FPGA board would be able to

reliably control and communicate with the biosequence processor chip. In the pervious tests the HP16500B Logic Analysis System provided emulated electronic input for the processor chip. The logic conversion board provided physical connection between the FPGA board and the processor chip. Now it must be shown that the FPGA board is able to send electronic data streams and control signals to the processor chip and receive comparison results from the processor. Furthermore, the reliability of such operations must be tested because biosequence analysis involves a large number of character-to-character comparisons.

As illustrated in Figure 7.1, the communication and control between the FPGA board and the processor chip requires every subsystem of the biosequence analysis system to function properly. For example, the optical data link contains the FPGA board, the logic conversion board, the laser current source, the laser, the optics system and the optical inputs. Failure in any of the components would cause erroneous comparison results from the processor chip.

The FPGA board can be programmed to send test data and check what it receives from the processor chip against the correct results it generates from the test data. Because everything in the biosequence analysis system is involved in such an operation, one can assume the FSOI design actually is exchanging information between the processor chip and the FPGA board if the received results are correct.

Once the reliability of the communication and control between the FPGA board and the processor chip was verified, actual biosequence analysis was carried out. The issue here was to compare the analysis results generated by the biosequence analysis system with the results from the on-line BLAST analysis tool. Because the biosequence analysis system uses a modified BLAST algorithm, one does not expect the results to be exactly the same. However, a working biosequence analysis system

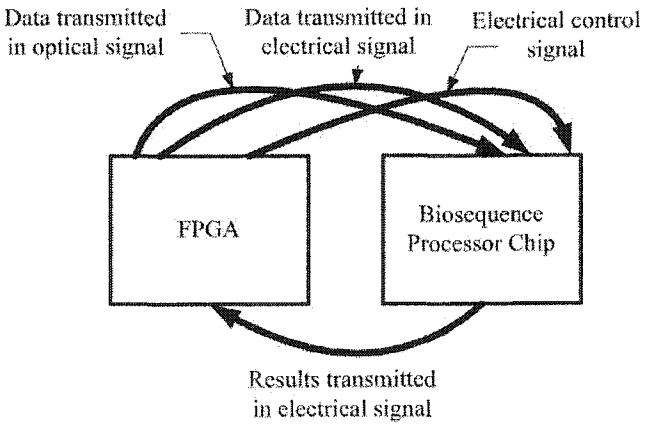


Figure 7.1: BER value with different laser power and external bias voltage

must generate sensible alignment results to be considered as functioning system.

7.2 Reliability tests

The complete FSOI based biosequence analysis system was assembled after the individual components and subsystems were checked out.

The reliability test of the biosequence processor chip was performed with the functionality test setup introduced in Chapter 6. The FPGA was programmed to generate random test data and examine the comparison results from the processor chip. The FPGA counted the overall number of comparisons and stopped after a preset number of errors (usually 1 5 errors) had occurred. The bit error rate (BER) could be calculated by dividing the number of errors with the number of overall comparisons. Because the comparison process involved the processor chip, the optical interconnects and the electronic inputs and outputs, the BER was considered a suitable figure of merit for measuring the reliability of the whole biosequence processing system.

The receiver circuit in the optical inputs was designed to operate with a range

of photodetector responsivitive and the receiver was designed to function with 1 μA of photocurrent at zero external bias voltage. If the photocurrent is less than 1 μA, the external bias can be increased to make the receiver circuit more sensitive to photocurrent and noise. Therefore the BER must be measured with various laser power and external bias voltage settings.

The electronic clock frequency for the BER was 166 kHz because this is the internal setting of the FPGA. Initial laser power was set at 0.8 mW, which was the minimum laser power needed for static functionality tests. Laser power was then increased until a reasonable BER value, i.e. sufficient to finish sequences alignment with only a few errors, could be obtained. Figure 7.2 plots measured BER of a single column of 8 comparison units for different laser power and external bias. The optical power in the figure is the laser output power converted from optical power measured at the front focal plane of the front lens.

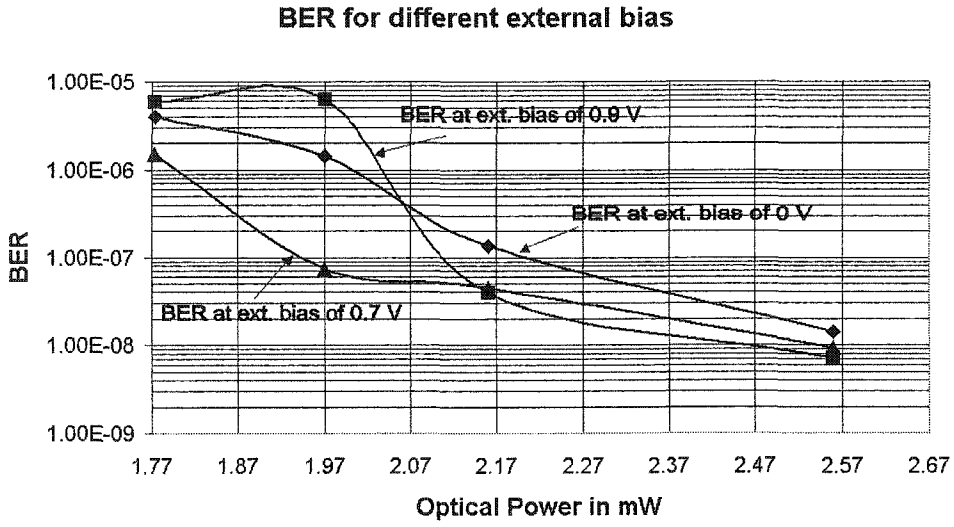


Figure 7.2: BER value with different laser power and external bias voltage

Figure 7.2 indicates that the laser power is the major factor in determining the

BER performance since the BER improves by two order of magnitude for a 45% increase in laser power. If the laser output power is higher than 2.1 mW, external bias voltage has only a secondary effect on the BER performance. However, external bias has a significant impact on the BER if the laser output power is below 2.0 mW. With low laser power, a moderate external bias yields the best BER performance because of the receiver circuit design. If the bias is too low then the laser power may generate insufficient photocurrent for the receiver to operate. If the bias is too high then the base voltage at the voltage comparator is too close to the threshold. Therefore at high bias the receiver circuit is vulnerable to the noise.

It is estimated that the BLAST comparison between two typical DNA biosequences with 200 characters requires less than 10^6 character-to-character comparisons. Since the BER is around 10^{-8} with 2.5 mW laser output power, the reliability of the optical interconnected biosequence processing system is sufficient for operation. Since the peak output power of the laser is 50 mW, there is a considerable power margin for better signal to noise ratio therefore better BER performance.

7.3 Biosequence Analysis

After the BER tests were conducted and the reliability of the biosequence processing system was confirmed, the complete system was put into operation for biosequence analysis. Since the National Center of Biological Information (NCBI) has an online PSI-BLAST biosequence analysis tool available for public uses, the alignment results from the biosequence analysis system were compared with results from the NCBI tool. However, the results were not entirely identical because the NCBI analysis tool uses a more complicated algorithm.

Initially short sequences were used to test the implemented BLAST algorithm.

The status of the system and the alignment result are displayed on the LCD screen of the Te-XC2Se FPGA development platform.

The test of the complete system was conducted in steps. The complexity of the alignments was gradually increased to ease the debugging of the whole biosequence analysis system. Some short (10 to 50 characters) dummy sequences and sequence fragments (50 to 100 characters) were used to validate the system. Once the the shorter ones had been successfully aligned, the system then analyzed longer sequences (100 to 200 characters) taken from real homologous sequences.

The first two sequences aligned were composed of all 'A's. The length of the sequences was 10 characters. Because the alignment of such short sequences was simple and straightforward, one could monitor the status of the system displaying from the LCD. If any error occurred the LCD display could help isolating the cause of error.

The second pair of sequences were of the same 10 characters long. However, only three characters in the middle of the sequences were identical. The remaining characters were randomly chosen. The system found hits around the identical characters, as designed in the algorithm.

Longer sequence fragments (50 to 100 characters)taken from a pair of homologous sequences "*Mus musculus*" and "*Cricetulus sp*" were next used to test the system. The former sequence is the genetic code for a common soluble protein found in mice and the latter a similar protein found in hamsters. The analysis system successfully aligned these sequence fragments.

After aligning the sequence fragments, the biosequence analysis system has successfully aligned the two complete sequences. Figure 7.3 illustrates the sequences aligned by the biosequence analysis system. The *Mus musculus* is shown with its

full length while only the aligned segments of the *Cricetulus sp* are shown in the figure. The dash signs indicates that fragments of *Cricetulus sp* that can not be aligned. The shadowed lines of letters and signs between the two sequences are the indication of the alignment: a plus sign indicates a pair of related characters, a letter indicates a pair of identical characters and spaces are for fragments that can not be aligned. Figure 7.4 is the alignment result from the NCBI online BLAST tool. One can observe the similarities between the aligned segments. The complete analysis of the biosequence alignment results is available in Mr. Klaus Hartinger's thesis.

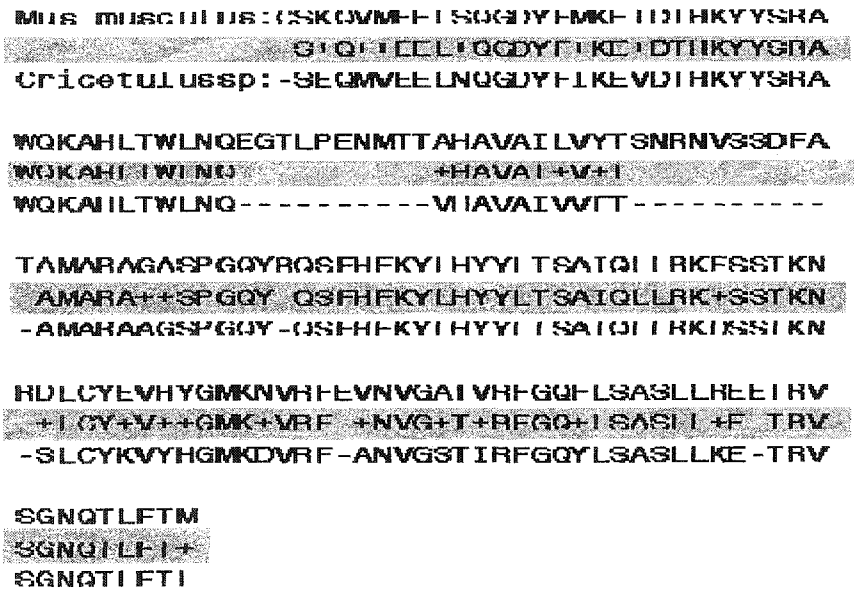


Figure 7.3: Sequences *Mus musculus* and *Cricetulus sp* aligned by the system

Another pair of homologous sequences *Tupaia* and *Cricetulus sp* are illustrated in Figure 7.5 and Figure 7.6. The similarities between the aligned segments are also indicated by the color lines.

Inhomologous sequences also have been tested with the biosequence analysis

```
SKQVMEELSQGDYFMKEIDTHKYYSRAWQKAHLTWLNQEGT
S+Q++EEL+QGDYF+KE+DTHKYYSRAWQKAHLTWLNQ
SEQMVEELNQGDYFIKEVDTHKYYSRAWQKAHLTWLNQAKA

LPENMTTAHAVAILVYTSNRNVSSDFATAMARAGASPGQYR
LPE+MT  HAVAI+V+T N NVSSD A AMARA  SPGQY
LPESMTPVHAVAIVVFTLNLNVSSDLAKAMARAAGSPGQYS

QSFHFKYLHYYLTSAIQLLRKESSTKNRDLCYEVHYGMKNV
QSFHFKYLHYYLTSAIQLLRK+SSTKN  LCY+V++GMK+V
QSFHFKYLHYYLTSAIQLLRKDSSTKNGSLCYKVYHGMKDV

RFEVNVGATVRFGQFLSASLLREETRVSGNQTLF
    NVG+T+RFGQFLSASLL+E TRVSGNQTLF
SIGANVGSTIRFGQFLSASLLKEGTRVSGNQTLF
```

Figure 7.4: Sequences *Mus musculus* and *Cricetulus sp* aligned by the NCBI BLAST

```
Cricetulus sp:CSKQVMEELSQGDYFMKEIDTHKYYSR
              C KQVMEELS GDYF KE+      YS
Tupaia:       CGKQVMEELSRGDYF-KEL-----YSQ

AWQKAHLTWLNQEGTLPENMTTAHAVAILVYTSNRNVSSDF
 W++AHL TWLN+              H+VA+LVY
-WIQAHLTWLNR----------HVVAMLVYT---------

AIAMARAGASPGQYRQSFHFKYLHYYLTSAIQLLRKESSTK
  AMA        QY  FHFKYLHYYLTSAIQLLRKE
--AMA-------QYEHAFHFKYLHYYLISAIQLLRKE----

NRDLCYEVHYGMKNVRFEVNVGATVRFGQFLSASLLREETR
   LCYEVH+GMK+V          L+RFGQF+SASLLREE++
---LCYEVHHGMKDV--------TIRFGQFISASLLREEVQ

VSGNQTLF
--------
```

Figure 7.5: Sequences *Tupaia* and *Cricetulus sp* aligned by the system

```
CSKQVMEELSQGDYFMKEIDTHKYYSRAWQKAHLTWLNQEG
C KQVMEELS+GDYF KE+  HK YS +W +AHLTWLN +
CGKQVMEELSRGDYFTKELKHHKNYSQIWHQAHLTWLNRAK

TLPENMTTAHAVAILVYTSNRNVSSDFATAMARAGASPGQY
 LP++MT  H VA+LVYT   NV +DF  AMA A  SP QY
ALPKDMTIPHVVAMLVYTLKSNVRADFIHAMASAARSPQQY

RQSFHFKYLHYYLTSAIQLLRKESSTKNRDLCYEVHYGMKN
  +FHFKYLHYYLTSAIQLLRKE   KN  LCYEVH+GMK+
EHAFHFKYLHYYLTSAIQLLRKERVMKNDTLCYEVHHGMKD

VRFEVNVGATVRFGQFLSASLLREETRVSGNQTLFTMVTC
V F+     T+RFGQF+SASLLREE +   NQTLFT+ TC
VSFKAYTGDTIRFGQFISASLLREEVQKFVNQTLFTISTC
```

Figure 7.6: Sequences *Tupaia* and *Cricetulus sp* aligned by the NCBI BLAST

system. As illustrated in Figure 7.7, the analysis of the two inhomologous sequences only yielded a single alignment of four characters. Considering the overall length of the sequences, it is clear that the alignment is not significant. The NCBI online BLAST tool also concluded that no significant alignment had been found in the inhomologous sequences.

The alignment usually takes about 1.5 seconds for the FSOI biosequence analysis systems to conduct the parallel character-to-character comparisons in the biosequence processor chip. The complete alignment process takes about 2.5 seconds.

The biosequence analysis system has been operational for a few months without much adjustments. The system setup is stable. The alignment had been conducted for multiple times and the results were consistent.

Mus musculus:
CSKQVMEELSQGDYFMKEIDTHKYYSRAWQKAHLTW
LNQEGTLPENMTTAHAVAILVYTSNRNVSSDFATAMAR
AGASPGQYRQSFHFKYLHYYLTSAIQLLRKESSTKNRDL
CYEVHYGMKNVRFEVNVGATVRFGQFLSASLLREETR
VSGNQTLFTMVTC
Ras-related C3 botulinum toxin substrate
MQAIKGVVVGDGAVGKTCLLISYTTNAFPGEDIPTAFD
NYSANVMVDGKLVNLGLWNTAGQEDYDRLRPLSYPQ
ADVFLICFSLVSPASFENVLAKWYPEVQHHCPNTPIILVG
TKLDLRDDKDRIQKLKEKKLTPITYPQGLAMAKEMGA
VKYLECLALTRRGLKTVFDEAIRAVLCPPPVKKRKRKC
LQL

Figure 7.7: Aligned complete sequences

7.4 Summary

The BER tests and the biosequence alignment results verify that the FSOI based biosequence analysis system performs BLAST algorithm with great reliability.

Chapter 8

Summary, Future work and Conclusions

A summation of the research results is first presented in this chapter. After the summary, future research on this topic are suggested. Finally the conclusions are presented and discussed.

8.1 Summary

The FSOI based biosequence analysis system has been successfully developed and tested. The system architecture was laid out and the algorithm functions were partitioned between the biosequence processor chip and the programmable logic device (in this case a field programmable gate array (FPGA) board). Several subsystems, including the FSOI based biosequence processor chip, the optics system, the FPGA board and other auxiliary components, have been designed and constructed. Multiple tests, including functionality tests and reliability tests, have been conducted to evaluate the design and the reliability of the FSOI and the biosequence analysis

system. Specifically, the milestones of this research work are listed as follow:

System algorithm and architecture: The BLAST algorithm is the most commonly used biosequence analysis algorithm today. The algorithm contains three major modules: sequence preprocessing, *hits* generation and *hits* extension. Realizing BLAST algorithm in a single custom designed CMOS IC is possible but costly and time-consuming. Therefore the algorithm functions are divided between a FSOI based biosequence processor chip and a FPGA board. System architecture was laid out to demonstrate a FSOI parallel processing system.

FSOI based biosequence processor chip: The FSOI based biosequence processor chip is a custom designed CMOS IC chip for parallel character-to-character comparisons in the biosequence analysis system. It features a comparison array with built-in optical inputs. The following tasks have been achieved in the development of the biosequence processor chip.

1. Design of the built-in photodetector and receiving circuit
2. Logic design of the basic comparison unit
3. Design of timing scheme and clocking unit
4. Layout of the basic unit and the clock unit
5. Initial electronic tests of the logic design

Optics system: The optics system in the biosequence analysis system multiplies a single modulated laser beam to a 8×8 spot array and projects the spot array onto the photodetector array of the biosequence processor chip. The following list summarizes work completed for the optics system.

1. Design and selection of the optics system layout
2. Calculation of the optics parameters
3. Selection of projection lens
4. Setup of the optics system

Tests: The tests on the biosequence analysis system include functionality tests, reliability tests and biosequence analysis. The functionality tests verified the FSOI system design. The reliability tests verified the stability of the optical/electrical data link and control signals. The biosequence analysis system analyzed multiple homologous and inhomologous biosequence pairs. The following list summarizes work completed in the tests.

1. Logic tests with emulated optical inputs
2. Logic tests with optical inputs
3. Bit error rate (BER) tests with different sets of optical power and bias voltage
4. Biosequence analysis with homologous biosequence pairs
5. Biosequence analysis with inhomologous biosequence pair

As summarized above, the biosequence analysis system has been successfully developed and tested. Experiences learned from this prototype system can be used for future development of FSOI based biosequence analysis systems.

8.2 Future Work

The prototype biosequence analysis system demonstrates the feasibility of applying FSOI technology to biosequence analysis hardware. It is possible to expand the

capability of the such a system based on current system design and technical know-how obtained in the system development. There are three factors that can be improved in order to expand the system capability: optical data speed, number of optical data links and on-chip logic.

The optical data speed is the bottleneck of the current generation of the FSOI based biosequence analysis system. The optical data speed is limited by the photodetector on the biosequence processor chip. Tests revealed that the maximum operation speed of the photodetector is 100 kHz and the electronic speed is 1 MHz The remaining system components can operate with higher speed. Electronic tests with emulated optical inputs indicated that the biosequence processor chip can operate at a clock speed of 20 MHz. Laser, laser driver and programmable logic device that are being used can match the 20 MHz speed. Therefore, the optical data speed can be increased by simply employing a faster photodetector in a new processor chip.

The key of such replacement is the information on the properties of the photodetectors. The full potential of a given type of photodetectors can be exploited only if a CMOS receiving circuit is specifically designed for the given type. As stated in Chapter 4, there are two options for the new photodetector design: external flip-chip bonded and built-in CMOS-compatible photodetectors. No change is required in the logic design of the biosequence processor. The optics system, the optical transmitter and the programmable logic device also are retained in the new design.

If the new photodetector operates with an 2 MHz optical signal then the speed of the biosequence processor chip will be increased twentyfold of the current speed without any modification to the other components of the biosequence analysis system.

There are 64 individual optical data links in the current system. More channels can be added by using a new DOE component in the optics system and enlarging the comparison array in the biosequence processor chip. Because of the added channels, a new lens with better abberation curve may be needed. A more powerful laser may also be required to fulfill the optical power budget.

At the present operation speed, the biosequence processor chip performs 1.33×10^5 character-to-character comparisons per second. If the optical operation speed is increased to 2 MHz, i.e. 20 MHz electronic clock frequency, then the processor chip can perform 2.13×10^7 comparisons per second.

Table 8.1 compares the performance of different FSOI based biosequence analysis systems and conventional computer. The electronic frequency of various FSOI systems is calculated for a given number of character-to-character comparisons per second.

Table 8.1: Quantitative comparison of FSOI systems and conventional computer

Number of Comparisons Per Second	Current FSOI System Design	Conceptual 16x16 FSOI System	Conceptual Full Optical 8x8 FSOI System	Conventional Personal Computer
1×10^6	938 KHz	422 KHz	281 KHz	28 MHz
5×10^6	4.69 MHz	2.11 MHz	1.41 MHz	140 MHz
1×10^8	93.8 MHz	42.2 MHz	28.1 MHz	2.8 GHz
1×10^9	938 MHz	422 MHz	281 MHz	28 GHz

The conceptual 16×16 FSOI System is an enlarged version of the current FSOI system. It has a 16×16 comparison array, 16 electronic inputs and outputs. The conceptual full optical 8×8 FSOI system has all optical inputs and outputs, which take only one clock cycle to receive or transmit data for all the units. The calculation assumes that the conventional computer CPU takes seven instructions to conduct one character-to-character comparison and each instruction on average takes four

clock cycles to execute.

Table 8.1 indicates that a current FSOI biosequence analysis system design running at 93.8 MHz electrical clock frequency is comparable with a state-of-the-art 2.8 GHz personal computer. An enlarged 16 $\times$ 16 comparison array running at 422 MHz is comparable with an array of 10 state-of-the-art 2.8 GHz computers operating in parallel. 93.8 MHz clock frequency is well within capability of the 0.5 μm CMOS IC technology. 422 MHz clock frequency is also achievable using better CMOS technology (e.g. 0.35 μm) The required optical frequencies are 8 MHz and 32 MHz, respectively.

Table 8.1 also indicates that a full-optical system will provide further performance improvement. Compared with the enlarged comparison array design, a full-optical system spends much less time on feeding the array with data therefore can make more comparisons in a second. 257 MHz clock frequency is also within capability of the 0.5 μm CMOS IC technology.

It is noteworthy that a state-of-the-art computer CPU requires deep sub-micron CMOS technology, which is much more expensive than the process required for the electronics of the FSOI based systems mentioned above. Though the cost of integration of high speed photodetector and laser must be taken into account, the performance improvement may justify such extra cost. If the same deep sub-micron process is applied in these FSOI based systems and the speed of the optical inputs/outputs matches, one may expect the performance of FSOI based systems would be much higher than its electronic counterparts.

8.3 Conclusions

This research project focuses on the application of the FSOI technology in a complex information processing system. A prototype FSOI based biosequence analysis system successfully demonstrated biosequence analysis of homologous and inhomologous sequences. We believe that it is feasible to employ FSOI in biosequence analysis hardware and similar systems.

Currently the FSOI system operation speed is not comparable with that of pure electronic systems. The bottleneck is the optical transmitter and receiver, which can be improved as the optical device and integration technology improves. Besides, the FSOI technology provides large parallelism that is transparent to data rate. The I/O speed of the FSOI systems may be brought to be comparable with that of the electronic systems. Therefore, the future of FSOI technology is promising.

Appendix A

The simulation results of the biosequence processor chip

Simulation charts Figure A.1, Figure A.2 and Figure A.3 were generated by the Mentor Graphics Accusim simulation software. The transistor model was obtained from the ADK software package. The netlist was generated and back-annotated using IC station software.

Figure A.1 illustrates the simulation of a basic unit. Two data streams were used to emulate the electronic and the optical inputs of the basic unit. **Output** is the output of the basic unit. **E-Input** is the emulated electronic input. **O-Input** is the emulated optical input. **CS**, **Reset** and **Cmp** are internal control signals generated by the clocking unit.

Figure A.2 illustrates the simulation of a clocking unit. It has two inputs: clock signal **Clock** and reset signal **Reset**. **CS7** $\sim$ **CS0** are the internal controlling signals generated by the clocking unit.

Figure A.3 illustrates the simulation of an assembled complete biosequence processor chip. It has 11 inputs: clock signal **Clock**, reset signal **Reset**, emulated

optical input **O-Input** and electronic inputs **IN7** $\sim$ **IN0**. **OUT7** $\sim$ **OUT0** are the output signals generated by the chip.

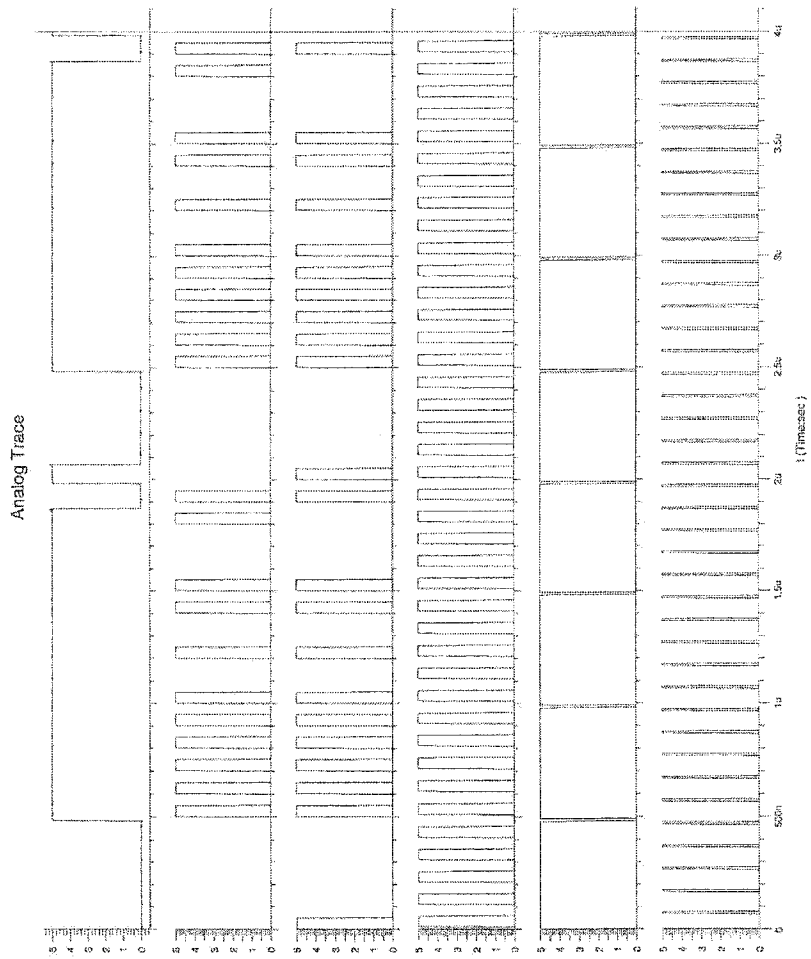


Figure A.1: Simulation chart of a basic unit. From left to the right, signals are: **Output E-Input O-Input CS Reset Cmp**

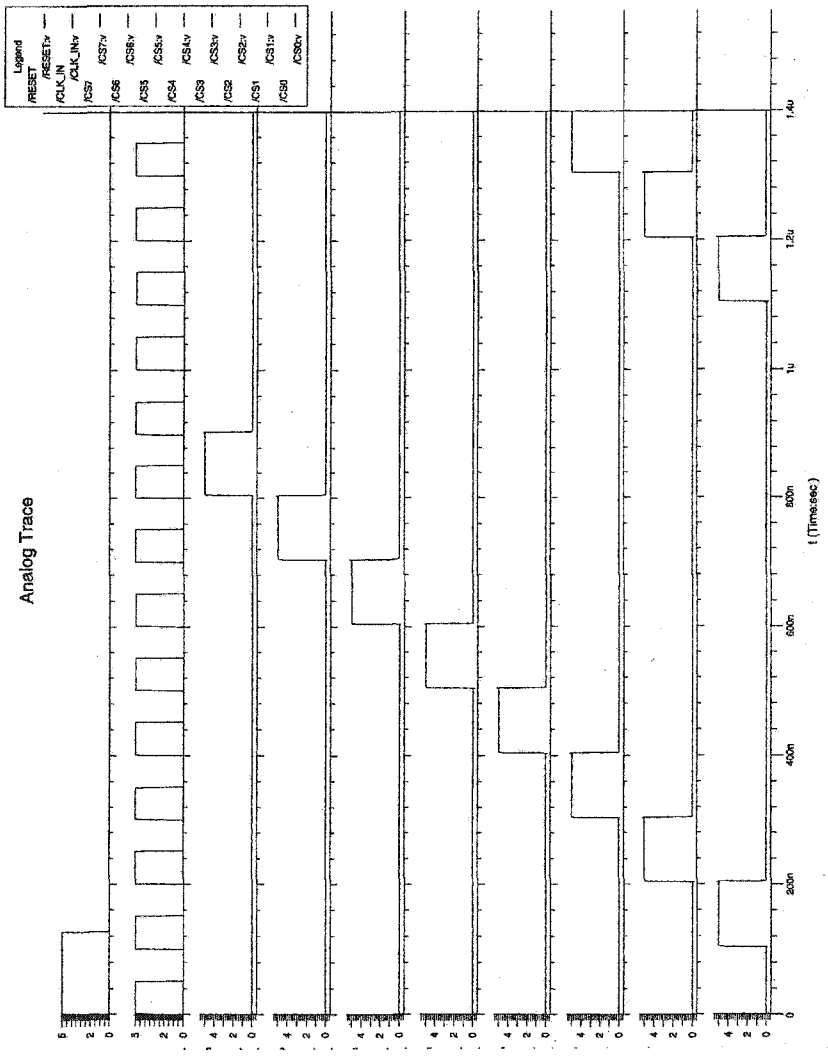


Figure A.2: Simulation chart of a clocking unit. From left to the right, signals are: **Reset Clock CS7 CS6 CS5 CS4 CS3 CS2 CS1 CS0**

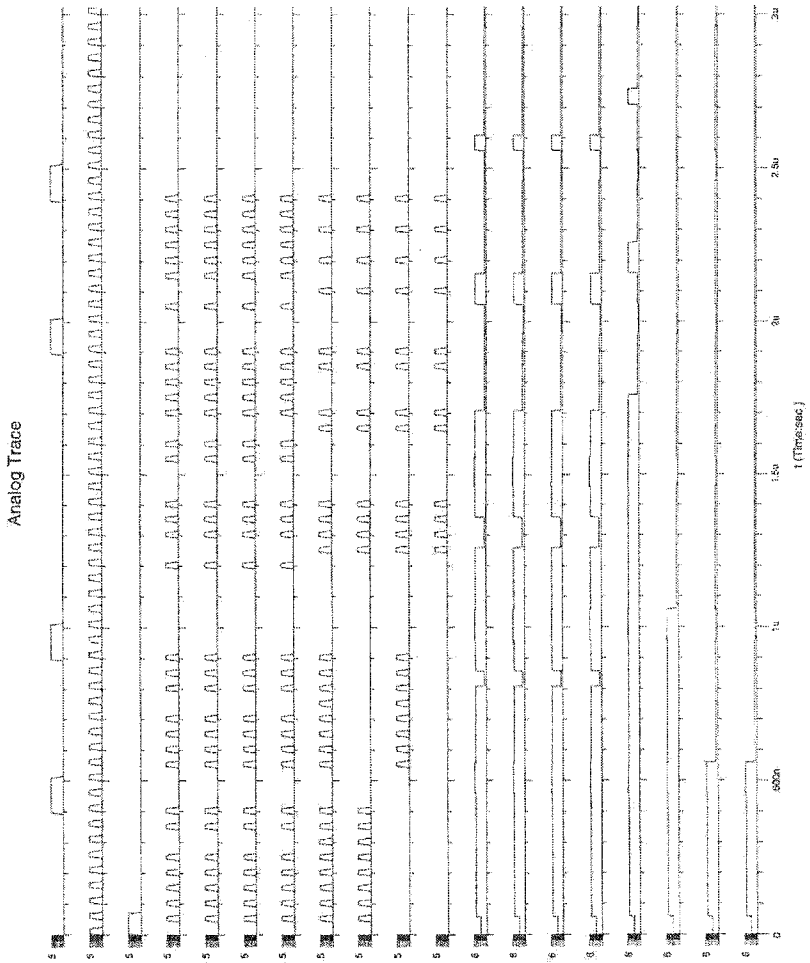


Figure A.3: Simulation chart of a clocking unit. From left to the right, signals are: **O-Input Clock Reset IN7 IN6 IN5 IN4 IN3 IN2 IN1 IN0 OUT7 OUT6 OUT5 OUT4 OUT3 OUT2 OUT1 OUT0**

Appendix B

The performance curves of three lenses

Figure B.1, Figure B.2 and Figure B.3 are simplified ray aberration curves for three candidates for the projection lens used in the optics system. The two curves illustrated in each of the figures are the ray abberations (the Y axis) versus the field (the X axis) for the X- and Y- directions. Figure B.4, Figure B.5 and Figure B.6 are simplified field distortion curves of these lenses. The curve illustrated in each of the figures are the field distortion (the X axis) versus the object height (the Y axis). Judging from the numbers shown on the charts, it is clear that KBX052 has the best ray aberration curve and KPX079 has the best field distortion.

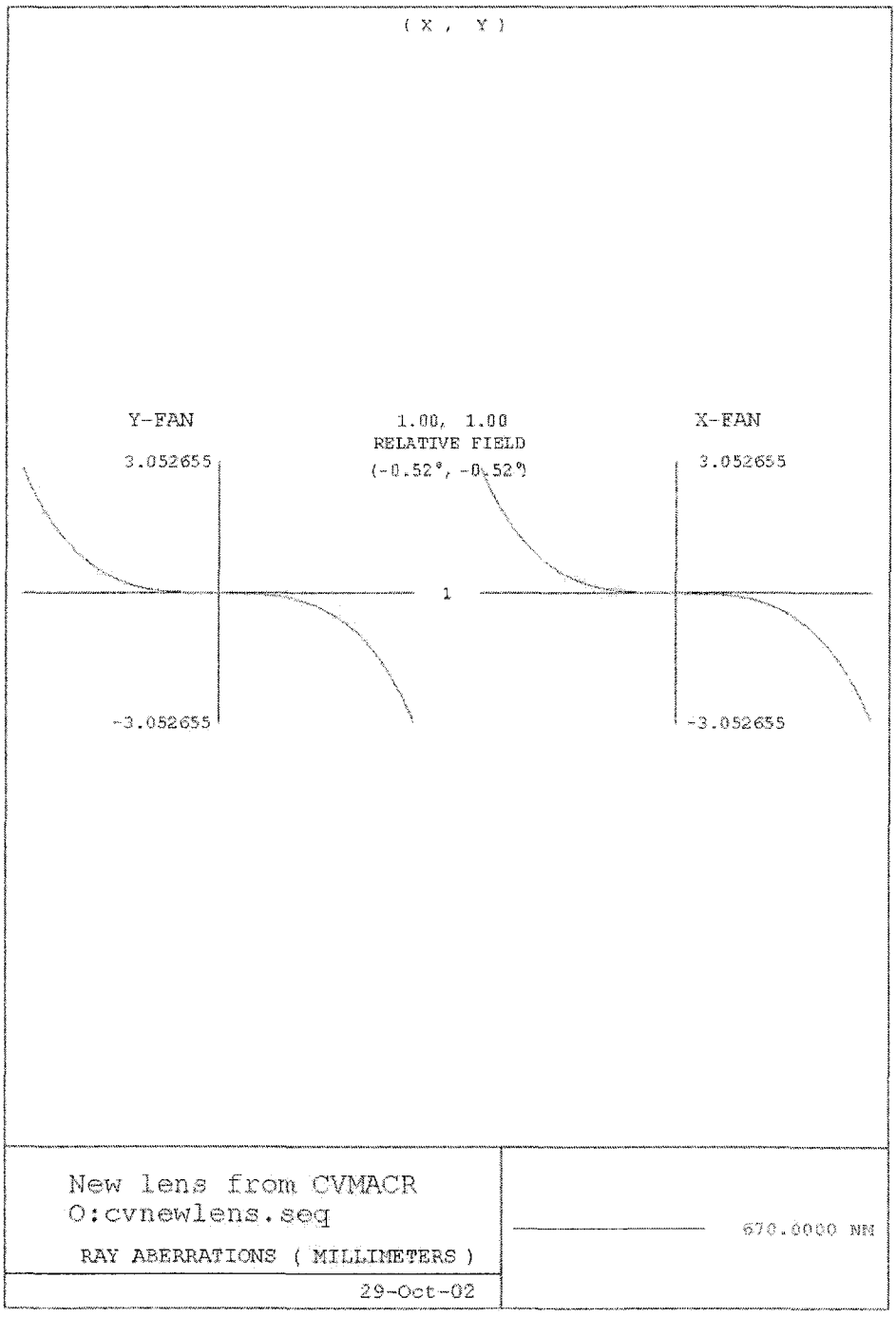


Figure B.1: Simplified ray aberration curves of KBX049

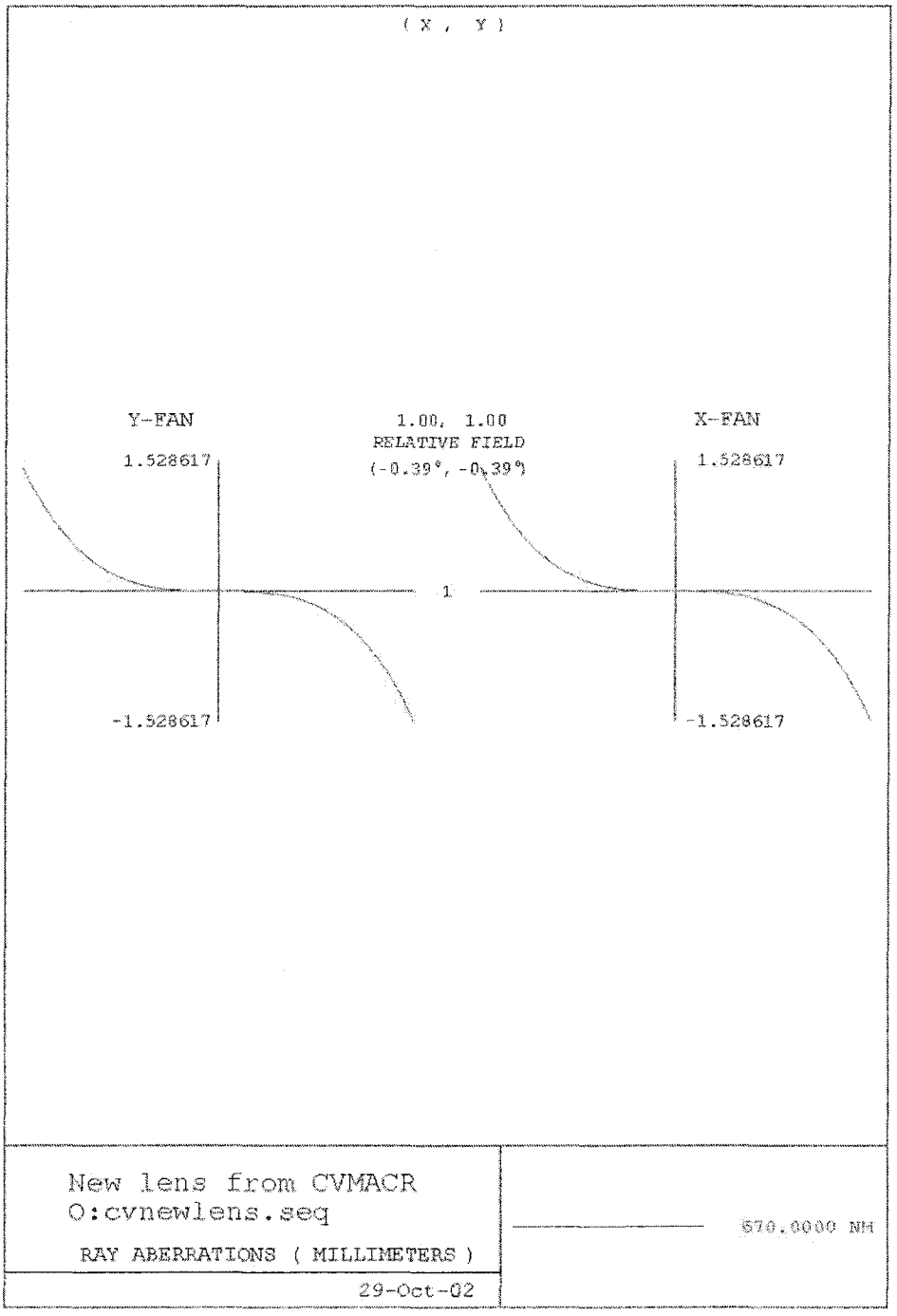


Figure B.2: Simplified ray aberration curve of KBX052

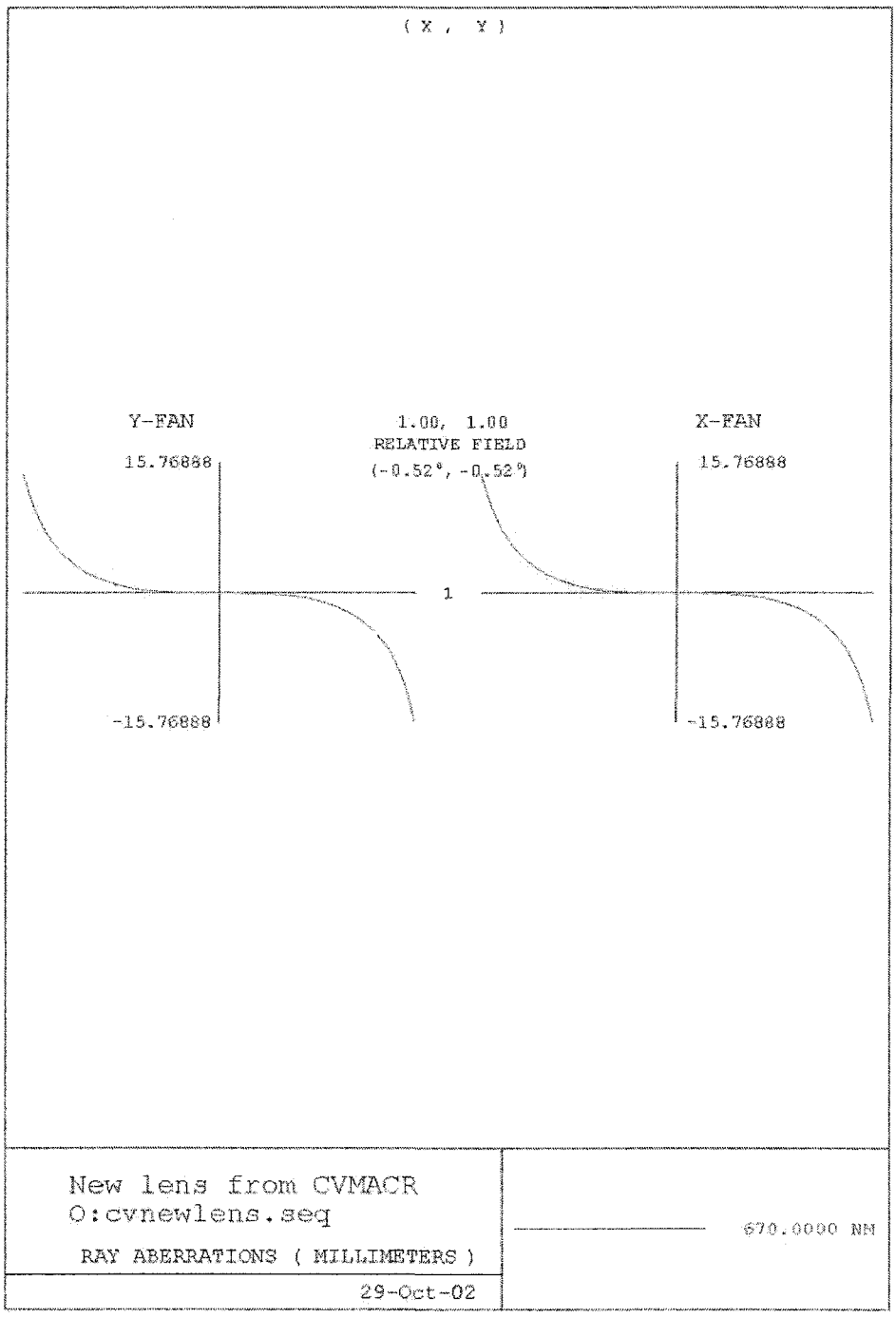


Figure B.3: Simplified ray aberration curve of KBX079

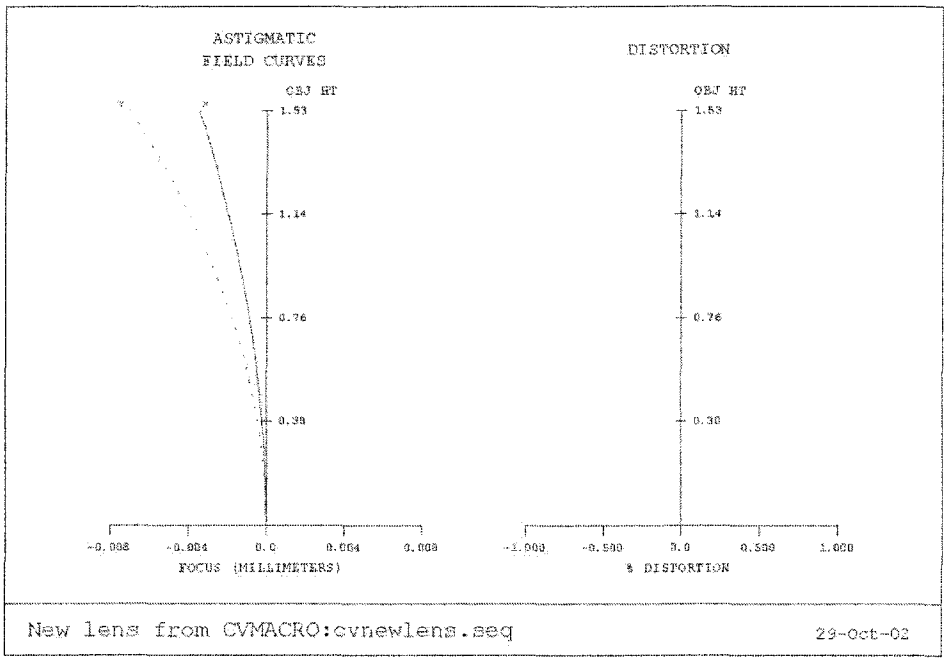


Figure B.4: Simplified field distortion curve of KBX049

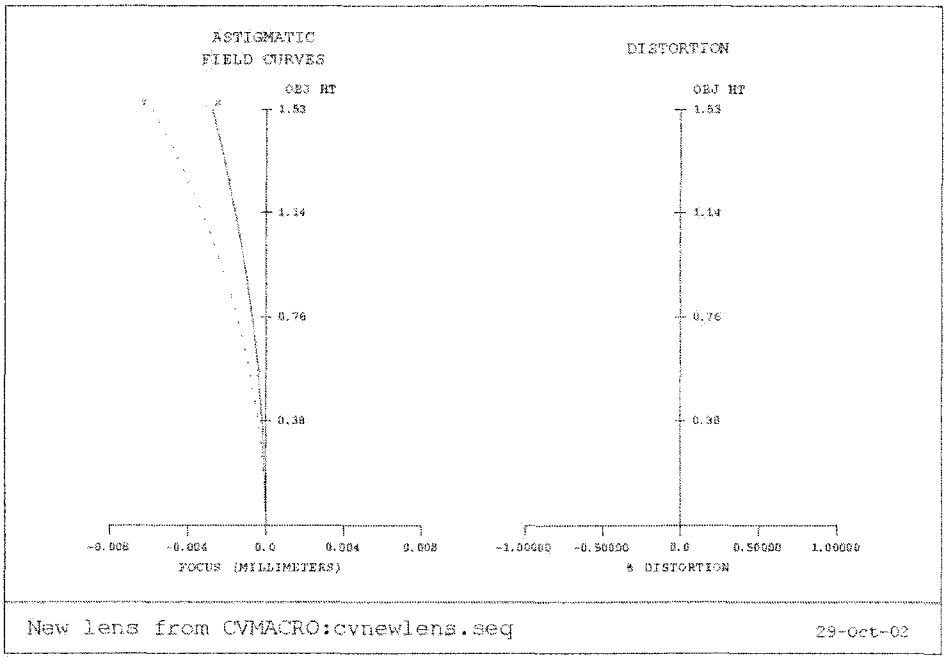


Figure B.5: Simplified field distortion curve of KBX052

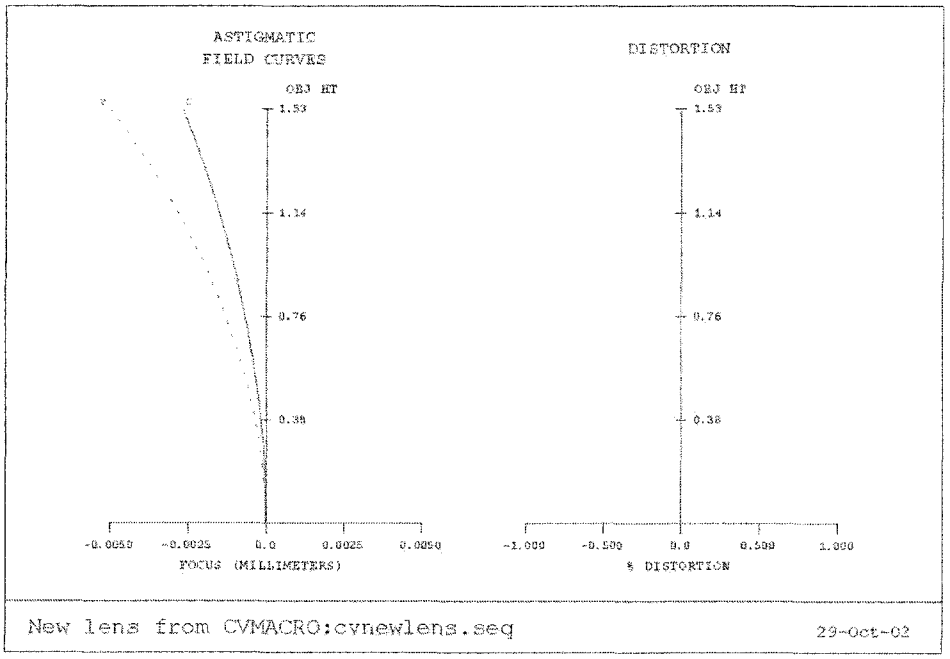


Figure B.6: Simplified field distortion curve of KBX079