The journal is partially supported by the Ministry of Scientific Research and Information Technology

Technical Secretary
- KATARZYNA SZTYLIŃSKA

Proof-reader
- HALINA MARCINIAK, IRENA SCHAROCH

Editorial Office
- Institute of Physics, Wrocław University of Technology,
  Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

Publisher
- Wrocław University of Technology,
  Wybrzeże Wyspiańskiego 27,
  50-370 Wrocław, Poland
  Optica.Applicata@pwr.wroc.pl
  www.if.pwr.wroc.pl/optappl
  tel. 48 71-320-23-93
  fax 48 71-328-36-96

Optica Applicata has been published since 1971 in a non-periodical form. Starting from 1973 it is published quarterly.

The Editorial Board invites the readers to send over any remarks or observations referring to the publications which will appear in the Optica Applicata. The correspondence of essential importance will be published in a separate column dealing with Letters to the Editor.

Review of matter

Diffraction theory and its application, quantum optics, problems in radiation coherence, light sources, holography and its application, scientific photography, methods of image reconstruction, optical application of Fourier transform, theory of optical systems, criteria of optical image evaluation, optical materials, technology of manufacturing optical elements, aspheric optics, optical properties of solids and thin films, lasers and their application, photo- and radiometry, problems in spectroscopy, nonlinear optics, optical data processing, optical measurements, fibre optics, optical instrumentation, interferometry, microscopy, non-visible optics, automation of optical computing, optoelectronics, colorimetry, optical detectors, ellipsometry and photoelasticity, optical modulation, optics of electron beams, biooptics, optometry.

Article and issue photocopies of this journal are available through University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106-1346, USA.

© Copyright by Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2009

Drukarnia Oficyny Wydawniczej Politechniki Wrocławskiej, Zam. nr 542/2009.
OPTICA APPLICATA

The quarterly of the Institute of Physics
Wrocław University of Technology, Poland
PL ISSN 0078-5466
Index 367729

HONORARY EDITOR IN CHIEF – MIRON GAJ
EDITOR IN CHIEF – WACLAW URBAŃCZYK
VICE-EDITOR – AGNIESZKA POPIOLEK-MASAJADA
EDITING COUNSELLOR – IRENEUSZ WILK

TOPICAL EDITORS

Fiber optics and optical communication, spectroscopy, lasers and their applications

Integrated optics, acoustooptics, microoptics, optical instrumentation, optical measurements, optical sensing

Nanooptics, plasmonics, optical imaging, optical computing, optical data storage and processing

Holography, diffraction and gratings, biooptics, medical optics, optometry, optical imaging, Fourier optics

Nonlinear optics, optical waveguides, photonic crystals

INTERNATIONAL ADVISORY BOARD

OLEG V. ANGE SKY, Chernivtsy University, Ukraine
YASUHIKO ARA KAWA, The University of Tokyo, Japan
IVAN GLESK, University of Strathclyde, UK
CHRISTOPHE GORECKI, FEMTO-ST, Besançon, France
ROMAN S. INGARDEN, Nicolaas Copernicus University, Toruń, Poland
EUGENIUSZ JAGOSZEWSK I (Chairman), Wrocław University of Technology, Poland
ROMUALD JÓZWICKI, Warsaw University of Technology, Poland
FRANCISZEK KACZ MAR EK, Adam Mickiewicz University, Poznań, Poland
BOLESŁAW KĘDZIA, Poznań University of Medical Sciences, Poland
MAŁGORZATA KUJAWIŃSKA, Warsaw University of Technology, Poland
NORBERT LINDELIN, University of Erlangen–Nürnberg, Germany
MIROSŁAW MILER, Institute of Photonics and Electronics of the ASCR, v.v.i., Prague, Czech Republic
JAN MISIEWICZ, Wrocław University of Technology, Poland
WOŁODZIMIERZ NAK WASKI, Technical University of Łódź, Poland
JAN PEŘINA, Palacký University, Olomouc, Czech Republic
BARBARA PIERŚCIONEK, University of Ulster, UK
COLIN SHEPPARD, National University of Singapore
CONCITA SIBILIA, Università di Roma “La Sapienza”, Italy
TADEUSZ STACEWICZ, University of Warsaw, Poland
TOMASZ WOLIŃSKI, Warsaw University of Technology, Poland
JAN WÓJCIC, Maria Curie-Skłodowska University in Lublin, Poland
PAVEL ZEMANEK, Institute of Scientific Instruments of the ASCR, v.v.i., Brno, Czech Republic
Contents

Optical communication

Multi-rate transmissions on spectral amplitude coding optical code division multiple access system using random diagonal codes
H.A. FADHIL, S.A. ALJUNID, R.B. AHMAD ............................ 225

Bandwidth analysis of multimode fiber passive optical networks (PONs)
G. STEPNIAK, L. MAKSYMIUK, J. SHUZDAK ............................ 233

Detecting faulty fiber with centralized failure detection system (CFDS) in fiber-to-the-home (FTTH) access network
M.S. AB-RAHMAN, B.C. NG, K. JUMARI ............................ 241

Influence of space charged particles on satellite optical communication system
R. HOU, S. ZHAO, J. XU, J. WU, Y. LI, S. ZHAN, L. SHI, S. FANG ............................ 251

Fiber optics

Light propagation in thermally expanded core fibers with graded-index
M.S. KOVAČEVIĆ, A. DJORDJEVICH, D. NIKEZIĆ ............................ 267

Multi-stage ring resonator all-pass filters for dispersion compensation
C. CHAIICHUAY, P.P. YUPAPIN, P. SAEUNG ............................ 277

Low cost incoherent pump solution for Raman fiber amplifier
A.M. ROCHA, B. NETO, M. FACAO, P.S. ANDRÉ ............................ 287

Photonic crystals

An investigation of the effect of intermediate layer in three-component planar photonic crystal waveguides
H. SHEN, H. TIAN, Y. JI ..................................................... 295

A 60° photonic crystal waveguide bend with improved transmission characteristics
M. KHATIBI MOGHADDAM, M.M. MIRSALEHI, A.R. ATTARI ............................ 307

Microwave properties of the generalized Fibonacci quasi-periodic multilayered photonic band gap structure
Y. TRABELSI, M. KANZARI, B. REZIG ......................................... 319

Image processing

Real time multiple planar volume clipping based on the programmable graphics process unit
F. YUAN ............................................................. 331

An image fusion algorithm based on polyharmonic local sine transform (PHLST)
L. SHANZHEHENG, H. JIUQIANG, B. LIU, Z. XINMAN ............................ 347
The feedback stability research of HDRI system

Efficient iris segmentation method with support vector domain description
R.N. Tang, J.Q. Han, X.M. Zhang ................................................................. 365

Sol–gel materials
Red-emitting LiEuW$_2$O$_8$ phosphor for white emitting diodes prepared by sol–gel process

Integrated design
Integrated design for large-scale opto-mechanical structure
A. Li, Y. Bian, Y. Liang, Y. Zhu ................................................................. 383

Optical measurements
Measurements of solution fluorescence – a new concept
S. Gebala ................................................................. 391

Vision optics
Biomechanical model of human eyeball and its applications
W. Śródką ................................................................. 401

Biomedical optics
What characteristics a clinical CSF system has to have?
E. Colombo, L. Issolio, J. Santillán, R. Aguirre ......................................... 415

Atmospheric optics
Coherence characterization of partially coherent flat-topped beam propagating through atmospheric turbulence
F.D. Kashani, M. Alavinejad, B. Ghafary ..................................................... 429
Multi-rate transmissions on spectral amplitude coding optical code division multiple access system using random diagonal codes

Hilal A. Fadhil*, S.A. Aljunid, R.B. Ahmad

School of Computer and Communication, University Malaysia Perlis (UniMAP), Kuala Perlis, 0200, Perlis, Malaysia

*Corresponding author: s0740810145@student.unimap.edu.my

In this paper, we study the use of a new code called random diagonal (RD) code for spectral amplitude coding (SAC) optical code division multiple access (OCDMA) networks, using fiber Bragg-grating (FBG). FBG consists of a fiber segment whose index of reflection varies periodically along its length. RD code is constructed using a code level and data level, one of the important properties of this code is that the cross correlation at the data level is always zero, which means that phase intensity induced phase (PIIN) is reduced. We find the performance of the RD code to be better than those of the modified frequency hopping (MFH) and Hadamard codes. It has been observed from simulation and theoretical results that considering the bit error rate (BER), the RD code performs significantly better than other codes. The ability of RD codes to support simultaneous transmissions at different bit rates is shown through simulated results of the BER and the eye patterns. 10 Gbps and 2.5 Gbps data transmissions have been successfully demonstrated together with FBG decoding scheme for canceling the code level from SAC-signal.

Keywords: optical code division multiple access (OCDMA), bit error rate (BER), SNR, phase intensity induced phase (PIIN).

1. Introduction

The success and extensive application of code division multiple access (CDMA) in the wireless area has renewed attention in exploring its application in the optics communication systems. Optical CDMA (OCDMA) has long been the subject of research because of its inherent ability to support asynchronous burst communications. Initially, it was employed for local area [1], then for access network applications [2, 3] and more recently for emerging networks such as generalized multiprotocol label switching [4]. The assumption of large code space does not always hold in practice, especially for the spectral-phase-encoded OCDMA [5, 6]. Because of optical hardware limitations and the code orthogonality required for low multiple access interferences, feasible spectral-phase-encoding solutions are limited to several well-known code families that contain relatively small number of codes. The spectral-phase-encoded OCDMA network exploits relatively simple all-optical pulse shaping to achieve optical
encoding and decoding. In this scheme, the entire pulse spectrum is divided into different spectral components called “chips”. The number of “chips” is chosen to be the length of the selected code. Across the spectrum, the phase is altered in each chip according to the phase code. The code set size is dependent on the code length. For the OCDMA scheme to be more realistic, it is desired to devise an optical code that can accommodate a larger number of simultaneous users with a low error probability for a given code length. A good set of codes is to obtain the maximum number of codes with maximum weight and minimum length with the best possible autocorrelation and cross-correlation properties [7]. To establish the OCDMA, we have to overcome the code orthogonality problem. Many researchers have proposed several codes such as prime code, optical orthogonal code, and so on. In this paper, we focus on random diagonal (RD) codes. In Section 2.1, we introduce the RD code construction, showing how the code has been developed theoretically, and discuss its properties. In Section 2.2, we focus on the proposed system scheme and design steps, and finally, draw conclusions is in Section 3.

2. RD code development

2.1. Construction of RD code

An \((N, W, \lambda)\) RD code is a family of \((0, 1)\) sequences of length \(N\) and weight \(W\), and \(\lambda\) is the in-phase cross-correlation which satisfies the following two properties: \(i\) zero cross-correlation will minimize \(\lambda\) and reduce phase induced intensity noise (PIIN); \(ii\) no cross-correlation in data level. The design of this new code can be preformed by dividing the code sequence into two groups, that is, a code level (segment) and data level (segment).

**Step 1, data segment**: let the elements in this group contain only one “1” to keep zero cross-correlation at data level \((\lambda = 0)\), which property is represented by the matrix \((K \times K)\), where \(K\) represents the number of users, these matrices have binary coefficient and a basic zero cross code \((W = 1)\) is defined as \([Y_1]\), for example, three users \((K = 3)\), \(y(K \times K)\) can be expressed as

\[
[Y_1] = \begin{bmatrix}
0 & 0 & 1 \\
0 & 1 & 0 \\
1 & 0 & 0 \\
\end{bmatrix}
\]

where \([Y_1]\) consists of \((K \times K)\) identity matrices. Note that for the above expression the cross-correlation between any two rows is always zero.

**Step 2, code segment**: the representation of this matrix can be expressed for \(W = 4\) as follows:

\[
[Y_2] = \begin{bmatrix}
1 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 0 \\
\end{bmatrix}
\]
where \([Y_2]\) consists of two parts: a weight matrix part and a basic matrix part; the basic part \([B]\) can be expressed as

\[
[B] = \begin{bmatrix}
0 & 1 & 1 \\
1 & 1 & 0 \\
1 & 0 & 1
\end{bmatrix}
\]

and the weight part \([M]\) = \[
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 0
\end{bmatrix}
\]

which is responsible for increasing number of weights. Let \(i = (W - 3)\) and \(M_i = \[
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 0
\end{bmatrix}
\]

where \(i\) represents the number of \(M_i\) matrix on \([M]\), as given by

\[
[M] = \langle M_1 | M_2 | M_3 \ldots M_i \rangle
\]

(1)

For example, if \(W = 5\), from Eq. (1) \(i = 2\), so that \([M] = \langle M_1 | M_2 \rangle\)

\[
[M] = \begin{bmatrix}
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0
\end{bmatrix}
\]

Notice that in order to increase the number of users simultaneously with the increase of code word length we can just repeat each row on both matrices \([M]\) and \([B]\); for the \(K\)-th user matrices \([M]\) and \([B]\) can be expressed as

\[
[M](j) = \begin{bmatrix}
0 & 1 \\
1 & 0 \\
0 & 1 \\
1 & 0 \\
0 & 1 \\
\vdots & \vdots \\
0 & 1
\end{bmatrix}, \quad [B](j) = \begin{bmatrix}
0 & 1 & 1 \\
1 & 1 & 0 \\
1 & 0 & 1 \\
0 & 1 & 1 \\
1 & 1 & 0 \\
\vdots & \vdots & \vdots \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
a_j_{1} \\
a_j_{2} \\
a_j_{3}
\end{bmatrix}
\]

where \(j\) represents the value for the \(K\)-th user \((j = 1, 2 \ldots K)\), and the value of \(a_j\) is either zero or one. The weights for the code part for both matrices \([M]\), \([B]\) are equal to \(W - 1\), so the total combination of code is represented by \((K \times N)\) where \(K = 3, N = 8\), as given by \([Z_1], [Z_2] = [Y_1 | Y_2]\)

\[
[Z_1] = \begin{bmatrix}
0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0
\end{bmatrix}
\]
From the above basic matrix $Z_1$, the number of users $K$ and the code length $N$, is given by $(K \times N)$ matrix. Notice that the code weight of each row is equal to 4, and the relation between $N$ and $K$ for this case can be expressed as

$$N = K + 5$$  

As a result, we can find that for $W = 5, 6, \text{ and } 7$, the code word length $N$ can be expressed as $K + 7, K + 9 \text{ and } K + 11$, respectively. As a result, the general equation describing the number of users $K$, code length $N$ and code weight $W$ is given by

$$N = K + 2W - 3$$

Many codes have been proposed for OCDMA systems, such as optical orthogonal code (OOC), modified frequency-hopping (MFH) codes, and Hadamard code, but the key point of RD code is that the RD code offers better performance than other codes in terms of the code length $N$ for the same number of users $K = 30$, as shown in the Table. A short code length limits the addressing flexibility of the codes, while a long code length is considered disadvantageous in implementation, since either a very wide bandwidth source or a very narrow filter bandwidth are required. The RD codes are neither too short nor too long.

### 2.2. Performance analysis

Figure 1 shows a setup for the proof-of-principle simulation for the scheme proposed. The performances of RD, MFH, and Hadamard codes are simulated by using the simulation software OptiSystem Version 6.0. A simple schematic block diagram consists of two users, as illustrated in Fig. 1. Each chip has a spectral width of 0.8 nm. The tests were carried out at a rate of 10 Gb/s for a 20-km distance with the ITU-T G.652 standard single-mode optical fiber. The attenuation $\alpha$ (i.e., 0.25 dB/km), dispersion (i.e., 18 ps/nm km), and nonlinear effects were activated and specified according to the typical industrial values to simulate real environment as close as possible. The performances of the system were characterized by referring to the bit-error rate (BER). As shown in Fig. 1, after transmission, we used a fibre Bragg grating (FBG) spectral phase decoder operating at the data level. The signals were decoded.
Multi-rate transmissions on SAC OCDMA system ... 229

by a photodetector (PD) followed by a 0.75 GHz low-pass-filter (LPF) and error
detector. The transmitted power used was 0 dBm out of the broadband source. The noise
generated at the receivers was set to be random and totally uncorrelated. The dark
current value was 5 nA, and the thermal noise coefficient was $1.8 \times 10^{-23}$ W/Hz for each
of the photodetectors. The eye pattern diagrams for RD, Hadamard and MFH codes
are shown in Fig. 2. The eye diagrams of Fig. 2 clearly show that the RD code system
gives better performance, having a larger eye opening. The eye pattern diagrams for
Channel 1 (10 Gbps) and Channel 3 (2.5 Gbps) are shown in Figs. 2c and 2d,
respectively, for RD code. Figure 2 also clearly shows the cross-talks experienced by
Hadamard and MFH codes [8]. The vertical distance between the top of the eye opening
and maximum signal level gives the degree of distortion. The more the eye closes,
the more difficult it is to distinguish between 1 s and 0 s in the signal. The height of
the eye opening at the specified sampling time shows the noise margin or immunity
to noise [9, 10].

The system’s SNR and the corresponding BER for RD code are given by:

$$
\text{SNR} = \frac{\left( \frac{2 \Re P_{sr} W}{N} \right)^2}{\frac{2 e B W P_{sr} \Re}{N} + \frac{B \Re^2 P_{sr} W K}{2 N \Delta V} (K - 1 + W) + \frac{4 K B T_n B}{R_L}}
$$

$$
\text{BER} = \frac{1}{2} \text{erfc} \sqrt{\frac{\text{SNR}}{8}}
$$
Fig. 2. Eye diagram of one of the Hadamard channels at 10 Gb/s after 10 km (a), and one of the MFH channels, at 10 Gb/s after 10 km (b), RD channel 1 at 10 Gb/s with BER of $6.4 \times 10^{-13}$ at a 20 km transmission (c), RD channel at 2.5 Gb/s with BER of $2.4 \times 10^{-18}$ after a 30 km transmission (d).

Fig. 3. BER against the number of users for RD, Hadamard, and MFH codes when $P_{sr} = 0$ dBm.
Figure 3 shows the relation between the number of users and the BER, for RD, MFH, and Hadamard codes, as plotted for different values of \( K \) (number of users). From this figure it is clearly seen that RD code gives a much better performance, i.e., (smaller BER) than MFH code and Hadamard code schemes. This is evident from the fact that RD code has a zero cross-correlation while Hadamard code has increasing value of cross-correlation as the number of users increases. However, a few code specific parameters are chosen based on the results published for these practical codes [7, 10, 11]. The calculated BER for RD was achieved for \( W = 7 \), while for MFH and Hadamard codes, \( W = 14 \), and \( W = 64 \), respectively.

3. Conclusions

A simple model using RD code is developed and employed to investigate the effects of SAC-OCDMA for different transmission rates on the OCDMA access links. The SAC-OCDMA is distinct because its requirement as to the light source is not strict and its encoder/decoder using FBGs is simple to be implemented. It has been shown that the RD code performs better than the system encoder with MFH and Hadamard, and the advantages of the proposed code are: shorter code length, no cross-correlation in data level (zero cross-correlation will minimize \( \lambda \) and reduce PIIN), data level can be replaced with any type of codes, more overlapping chips will result in crosstalk, flexibility in choosing parameters \( N, K \) unlike with other codes, such as MFH and Hadamard, and finally, RD code can support different transmission rates. The simulated result of one of the three RD coded carriers running at 10 Gb/s and 2.5 Gb/s over a communication-standard fiber shows a good quality transmission.

References


Received August 7, 2008
Bandwidth analysis of multimode fiber passive optical networks (PONs)

GRZEGORZ STEPNIAK*, ŁUKASZ MAKSYMIUK, JERZY SIUZDAK

Institute of Telecommunications, Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warszawa, Poland

*Corresponding author: stepniak@tele.pw.edu.pl

The analysis of frequency response in different branches of a multimode fiber based passive optical network is conducted. Both the theoretical and experimental results show that bandwidth of various network paths may be different.

Keywords: multimode fibers, frequency response, bandwidth, passive optical network (PON).

1. Introduction

The multimode fiber (MMF), due to its larger core size, has a number of advantages over single mode fiber (SMF) in access and local networks, where the cost and time of installation play substantial role. For the same reason passive optical network (PON) is expected to become the most popular technology in MMF networks.

A lot of work has been devoted to the investigation of MMF parameters, especially its bandwidth, with regard to the possible application of existing MMF LANs for higher bit rates (Gbit/s and 10 Gbit/s Ethernet) [1, 2]. Nevertheless, to the best of authors’ knowledge there are hardly any papers treating of the bandwidth performance of PONs incorporating MMFs. This matter is of utmost importance, as modal filtering in the network branching components may affect the node-to-node transmission parameters. In such a situation, to serve the users with equal quality, performance of the whole network has to be adjusted to the lowest available level.

In this paper, we extend the MMF link model [1] by adding a coupler/splitter module and thus we make it possible to analyze frequency response of any optical path within a MMF PON. The effect of splitter influence on the bandwidth is investigated both numerically and experimentally.
2. Theory

We employ the frequency domain analysis and matrix formalism to model the network components, which was first applied in [1]. Let us represent the mode power distribution with a vector [1]:

\[
A = \begin{bmatrix} A_1(\omega) & \ldots & A_M(\omega) \end{bmatrix}^T
\]

(1)

where \( M \) is the highest mode group number in the fiber and \( T \) denotes transposition. The elements of \( A \) are complex numbers – their modulus represents power in a mode group, whereas their argument – the group delay of a mode group for angular frequency \( \omega \) [1].

In our model, the network components – fibers, couplers/splitters and connectors are represented by matrices operating on vectors (1). The modulus of the element \( b_{i,j} \) of component matrix \( B \) represents part of the power of the \( j \)-th mode group that is transferred to the \( i \)-th mode group at this component whereas its argument represents possible delay (only for distributed elements such as fibers). To find the mode power vector at a certain node, the product of matrices on the path from the light source to that node must be computed. For instance, to find the frequency response of the optical path that contains (counting from the light source): a fiber, a connector and a splitter, with receiver in the \( (a) \) arm of the splitter, the mode power distribution vector at receiver has to be first computed

\[
A^{(R)}(\omega) = S(\omega) \cdot C(\omega) \cdot F(\omega) \cdot A^{(inp)}(\omega)
\]

(2)

where \( A^{(inp)}(\omega) \) is the input mode power distribution vector that depends on the launching conditions. We denote the splitter, connector, and fiber matrices with \( S, C, \) and \( F \), respectively. The splitter has a different matrix for each arm. Finally, the frequency response of the light path between transmitter and receiver may be expressed as (after [1])

\[
T^{(R)}(\omega) = \sum_{m=1}^{M} A_m^{(R)}(\omega)
\]

(3)

In the following, we shortly describe all the modeled elements.

2.1. Light source

The initial mode group power vector \( A^{(inp)} \) depends obviously on the light source type. In the following, results will be shown for two standard excitations [3]: overfilled launch (OFL) modeling LED source and restricted mode launch (RML) modeling vertical cavity surface emitting laser (VCSEL), but the method is not limited to them. Those excitations have been standardized to eliminate the dependence of measured frequency response on light source type or specimen [3]. Any kind of light launch can be easily introduced in the numerical model.
2.2. Fiber

The derivation of the matrix representation for the fiber, that takes mode mixing effects into account, is described in detail in [4]. Here, to simplify the formalism, we neglect mode mixing, as it usually has negligible influence in GI silica fibers [1]. The fiber matrix is then a diagonal $M \times M$ matrix with elements [1]

$$\tilde{F}_{m,m} = \exp(-\gamma_m L - j \omega \tau_m L)$$

where $\gamma_m$ and $\tau_m$ are the mode group dependent attenuation and the group delay, respectively, $j$ is the imaginary unit and $L$ is the fiber’s length.

2.3. Connector

Our approach to the connector modeling follows [1]. The connector transfer matrix can be found by finding the expansion coefficients of the input fiber modes in the basis of the output fiber modes.

2.4. Coupler/splitter

To obtain coupler matrices (one for each arm), the simplified wave equation for inhomogeneous coupler structure has to be solved repeatedly for all modes propagating in the couplers’ medium [5]:

$$\left\{ i 2k n_2 \frac{\partial}{\partial z} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k^2 \left[ n_1^2(x, y, z) - n_2^2 \right] \right\} \varphi_{\mu, \nu}(x, y, z) = 0$$

where $k$ is the wave constant, $n_2$ is the refractive index of the cladding, $n_1(x, y, z)$ is the refractive index value in the mesh point and $\varphi_{\mu, \nu}(x, y, 0)$ is the field distribution of the fiber $LP_{\mu, \nu}$ mode. Hence, two spatial amplitude fields $\varphi^{(\text{out})}_{\mu, \nu}(x, y)$ at far ends of both arms of the coupler are obtained. In the next step they are decomposed into orthogonal set of modes by calculating following integrals

$$a_{\mu', \nu', \mu, \nu} = \iint \int_\delta dxdy \varphi^*_{\mu', \nu'}(x, y) \varphi^{(\text{out})}_{\mu, \nu}(x, y)$$

Analogously to the connector case [1] the coupler power transfer matrix is obtained as follows

$$\tilde{S}_{m', m} = \frac{1}{2m} \sum_{\mu', \nu', \mu, \nu} \left| a_{\mu', \nu', \mu, \nu} \right|^2 \delta(m' - \mu' - 2\nu' + 1) \delta(m - \mu - 2\nu + 1)$$

where $m$ is the mode group number, $\delta$ is the Kronecker’s symbol, $2m$ stands for the number of modes in the $m$-th mode group. The decomposing procedure has to be repeated for two arms independently. This approach allows modeling of any coupler structure provided that its mechanical layout is known.
3. Results and measurements

We have examined various MMF PON structures of tree architectures both numerically and experimentally. We have calculated and measured frequency responses of various optical paths within such networks. The most important conclusion is that these responses are different for different paths for otherwise symmetrical network. This implies that the 3 dB bandwidths of various paths are not equal and may differ by a few tens of percent. The reason behind this is that splitters perform mode filtering differently in each arm (in the numerical model, the matrices are different in each of the splitters arms). If one of the arms of a multimode coupler is excited, the higher order modes (propagate near the cladding and may have higher angles in the geometrical approximation) tend to couple into the second arm of the coupler, whereas the lower order modes (propagate in the core center and have smaller angles) tend to stay in the first arm. Thus, the transmission matrix of a coupler with a geometry we consider, is typically different in its different branches. This effect is most pronounced in a single splitter network, whereas it tends to be averaged in more complex architectures.

According to our knowledge, the most popular couplers used in PON networks are side couplers, and only those are considered in this paper. They consist of two multimode fibers that are gradually brought close together, so that their modal fields overlap and energy transfer from the input arm to the lateral arm is enabled, and then they are gradually separated (example structure in Fig. 1). The coupling region can be additionally fused and tapered, to improve coupling efficiency. However, the exact geometrical layout of the coupler remains manufacturer’s secret. In our simulations, we consider the coupler structure presented in Fig. 1.

![Fig. 1. Geometry of the considered coupler. Exact parameters given in the Table.](image)

<table>
<thead>
<tr>
<th>Coupler No.</th>
<th>a [mm]</th>
<th>l [mm]</th>
<th>d [μm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.2</td>
<td>3.4</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>1.3</td>
<td>3.5</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>1.4</td>
<td>3.6</td>
<td>15</td>
</tr>
</tbody>
</table>
To show the effect of coupler filtering, we present results for a MMF (1 km long, 62.5 μm, NA = 0.275, graded index, with a profile flaw in the core center, and profile parameter selected to match the measurement data) which is connected to a 2×2 coupler either at its front or at its end. To get greater variety of results in calculations, we considered 3 realizations of coupler structure, of geometrical parameters given in the Table, that were numerically designed to give splitting ratio close to 50/50 at 850 nm wavelength. All possible to acquire frequency responses are plotted in Figs. 2 and 3 for OFL and RML, respectively. For comparison, results of measurements of analogous network incorporating 4 different commercial couplers from the same

![Fig. 2. Calculated frequency responses of a fiber interconnected with various couplers either on its input or output, in different arms. The dotted curve is the frequency response of the fiber itself. OFL.](image)

![Fig. 3. Calculated frequency responses of a fiber interconnected with various couplers either on its input or output, in different arms. The dotted curve is the frequency response of the fiber itself. RML.](image)
production series are depicted in Figs. 4 and 5 [6]. As we can see, the bandwidth is different in different coupler branches, and the variations can be even as high as 30% for RML and 15% for OFL. The dotted curve in each figure is the frequency response of the examined fiber itself (without coupler/splitter). It may be treated as the first approximation of the bandwidth in the network, however, it may not be considered as the upper or the lower bound.

Unfortunately, direct comparison of measured and calculated results is not possible. The calculation of a real commercial coupler matrix was not possible as the physical layout of the commercial couplers was not available, and it is evident that different
coupler types have different mode filtering properties. However, the proposed method allows modeling of any coupler structure, provided that the physical layout of the coupler geometry is known.

4. Summary

We introduced an extension of MMF link model [1], that allows numerical calculation of frequency response in MMF PON networks. It was shown that otherwise identical paths in MMF PON have different bandwidths. This follows from filtering properties of MMF couplers/splitters. The differences in bandwidth between the nodes are the highest for RML, which unfortunately is a typical launch in contemporary MMF systems. It is shown that the splitter may improve or decrease the bandwidth of a fiber itself, and that the bandwidth of optical fiber of equivalent length is a first approximation of the network bandwidth.

References


Received July 20, 2008
in revised form November 18, 2008
Detecting faulty fiber with centralized failure detection system (CFDS) in fiber-to-the-home (FTTH) access network

Mohammad Syuhami Ab-Rahman*, Boon Chuan Ng, Kasmiran Jumari

Computer and Network Security Research Group, Department of Electrical, Electronics and Systems Engineering, Faculty of Engineering and Built Environment, National University of Malaysia, 43600 UKM Bangi, Selangor, Malaysia

*Corresponding author: syuhami@vlsi.eng.ukm.my

A new technique for detecting any faulty fiber and identifying the failure location occurring in tree-based structured fiber-to-the-home (FTTH) access network with centralized failure detection system (CFDS) based on Visual Basic is proposed and experimentally demonstrated in the paper. CFDS is installed with optical line terminal (OLT) at central office (CO) to monitor the network system and detect any failure that occurs in multi-line drop region of FTTH access network downwardly from CO towards customer premises. CFDS enables the status of each optical network unit (ONU) connected line to be displayed on a computer screen with capability to configure the attenuation and detect the failure simultaneously. The failure analysis and information is delivered to the field engineers for prompt actions, and meanwhile the failure line is restored to stand-by line to ensure a continuous traffic flow. This approach has bright prospects of improving the survivability and reliability as well as increasing the efficiency and monitoring capabilities in FTTH access network. Besides, it is able to overcome the upwardly or downwardly monitoring issues with conventional fiber fault localization technique by using optical time domain reflectometer (OTDR). With CFDS database, the histories of scanning process and data can be recalled and further analysis can possibly be conducted.

Keywords: detection of faulty fiber, identification of failure location, fiber-to-the-home (FTTH), centralized failure detection system (CFDS), Visual Basic, downwardly.

1. Introduction

Fiber-to-the-home (FTTH) is a network technology that deploys optical fiber cable directly to the home or business to deliver triple-play (data, voice, and video) services with a high speed up to the customer premises [1]. Today, FTTH has been recognized as the ultimate solution for providing various communications and multimedia services, including carrier-class telephony, high-speed Internet access, digital cable television (CATV), and interactive two-way video-based services to the end users [2].
Since such architecture can accommodate a large number of subscribers, when a fiber break occurs in the feeder region, the access network is devoid of any function behind the breakpoint. The upstream or downstream signal after the breakpoint becomes unreachable [3]. Any service outage due to the fiber break can be translated into tremendous financial loss in business for the network service providers [1].

Conventionally, optical time domain reflectometer (OTDR) is used to identify a faulty fiber in FTTH access network upwardly from the customer premises side towards central office (CO). However, this approach would require much time and effort. Moreover, OTDR can only display a measurement result of a line at a time. Therefore, it becomes a hindrance to detection of a faulty fiber with a large number of subscribers and large coverage area in the fiber plant by using an OTDR. Besides, it is difficult to detect failure in a point-to-multipoint (P2MP) connectivity which is equipped with passive optical splitter by using an OTDR downwardly from CO because the Rayleigh back-scattered (RBS) light from different branches overlap each other in the OTDR trace and cannot be distinguished [1]. Therefore, CFDS has been introduced in this paper as a solution that is expected to reduce the capital and operational expenditures (CAPEX and OPEX) for FTTH access network.

2. Conceptual design for CFDS

This paper presents the development of a new technique for locating faulty fiber in FTTH access network with CFDS after taking into consideration the requirement for network monitoring capabilities, maintenance and repairing cost, restoration time, expandability, dependability, and redundancy. CFDS has potential for improving the survivability and increasing the monitoring capabilities in FTTH access network. This program is our second generation based on the Microsoft Visual Basic software. The first generation CFDS which was developed by using MATLAB software has been proposed in our previous paper [4]. It can reduce the time needed to restore the fiber fault to maintain and operate the optical access network more efficiently. However, it is quite inconvenient to use MATLAB software for the analysis due to its own limitation. Furthermore, this required much more steps and time for the data analysis. Therefore, we demonstrate in this paper a second generation based on the Microsoft Visual Basic software, with the complexity being reduced and a few additional features introduced. It has greater functionality as compared to our first approach. The program developed not only enables one to monitor the status for each optical fiber line and detect the failure location in FTTH access network downwardly, but also to determine the status, deployment, connection, and configuration in the network. It comes together with the additional features such as tracking the optical signal level (input power and output power) and losses (e.g., connection losses, splice losses, optical components/devices losses, fiber losses or attenuation) as well as monitoring the network performance.

CFDS can be broken down into three main parts to support its operation, including optical line measurement, interfacing OTDR with personal computer (PC), and data
Detecting faulty fiber with CFDS in FTTH access network

The whole operation process can be simplified in the flow chart, as depicted in Fig. 1. CFDS has the same features of the OTDR and computer-based OTDR emulation software for performing more OTDR trace processing functions, but with more flexibility used for optical communication link especially in the FTTH access network.

2.1. System architecture

To locate a fiber fault without affecting the transmission of services to other subscribers, it is essential to use a wavelength different from those of the triple-play signals for failure detection [5]. As illustrated in Fig. 2, CFDS uses a 1625 nm signal for failure detection control and in-service troubleshooting. The triple-play signals (1310 nm, 1490 nm, and 1550 nm) are multiplexed with a testing signal (1625 nm) from OTDR. The OTDR is installed with the OLT at CO and will be connected to a PC to display the troubleshooting result.

When four kinds of signals are distributed, the testing signal will be split up by the wavelength selective coupler (WSC) or wavelength division multiplexing (WDM)
coupler, which is installed before the splitter. The WSC only allows the testing signal at 1625 nm to enter the taper circuit and reject all unwanted signals (1310 nm, 1490 nm, and 1550 nm) that contaminate the OTDR measurement. The downstream signal will go through the WSC, which is in turn connected to a splitter before it reaches the ONU s at different residential location. The distance between the OLT and ONU is about 20 km. On the other hand, the testing signal which is demultiplexed by WSC will be split up again in power ratio 99:1 by using directional coupler (DC) to activate the microprocessor system. The 99% 1625 nm signal will then be configured by using a splitter, with each of its outputs being connected to a single line of ONU. The operation of optical switch is controlled by microprocessor system that is activated by 1% of 1625 nm signal. With the method described in this article, no any expensive additional equipment or devices are required and it also enables the upwardly or downwardly monitoring issues to be overcome by means of conventional techniques.

CFDS is interfaced with the OTDR to accumulate every network testing result to be displayed on a single PC screen for further analysis. The analysis result is sent to field engineers or network service providers through the mobile phone or Wi-Fi/Internet computer using wireless technology for prompt action. Anywhere, the traffic from the failure line is diverted to stand-by (protection) line to ensure a continuous traffic flow. After the restoration/maintenance process, the traffic is switched back to the normal working line.

2.2. Experimental setup and network testing
To verify the concept and benefits of CFDS, we conducted two experiments through a point-to-point (P2P) network testbed, mainly focusing on the identification of the faulty fiber and locating failure. The fixed connection (FC) connector and optical attenuator are the two optical devices used in the experiments. The FC connector is used to establish connection between two fibers under test, whereas the optical attenuator was used to represent the breakpoint in an optical fiber line in the first

Fig. 2. The system architecture for CFDS.
Detecting faulty fiber with CFDS in FTTH access network

experiment. It visualized the actual breakpoint of an optical line at that distance in real conditions. The optical attenuator was also used to reduce (attenuate) the optical signal level (optical power) in a line tested in the second experiment. The characteristics of the optical fiber line under working (good/ideal) and non-working (failure/breakdown) conditions are measured at an early stage by using an OTDR. Then, the measurement results for each line are saved in the OTDR and then transferred into the PC. After completing the transferring process, the results need to be recorded in database and then loaded into CFDS for further analysis.

3. Execution display for CFDS

One of the basic functionalities of CFDS is to detect failure in FTTH access network. The CFDS consists in: i) plotting optical signal level (dB) versus distance (km) graph, ii) checking the status of each optical fiber line, and iii) displaying line details. CFDS also tracks the optical signal level and monitors the network performance. There are other functionalities of CFDS that are not described in detail nor demonstrated in this paper. The CFDS provides a convenient way to solve the particular upwardly or downwardly measuring issues with OTDR and capability of fiber fault localization in FTTH access network.

3.1. Analyzing line status

Figure 3 shows the ability of CFDS to specify a faulty fiber and failure location among a number of optical fiber lines in FTTH access network through measuring the fiber attenuation, connection losses, splice losses, and optical components/devices losses. In the first experiment, it is assumed that there is a fiber fault occurring in an individual subscriber’s infrastructure among eight subscribers. OTDR is an instrument that is used to measure the fiber attenuation, locate fault, measure splice loss and fiber

![Fig. 3. Eight graphs are displayed in the Line Status form. A failure message is displayed to show the faulty line and failure location in FTTH access network.](image-url)
uniformity or the attenuation coefficient throughout the installed fiber length. The measurement results are recorded in the OTDR and then transferred into the database in PC.

Every eighth network testing results are displayed in the Line Status form for centralized monitoring. CFDS uses an event identification method to differentiate the mechanism of the optical signal in working and non-working condition. The loss in reflective fault event is representing the condition of a tested line. A failure message Line x FAILURE at z km from CO! is displayed to inform the user if it has detected any fiber fault in the network.

---

**Fig. 4.** An example of working line in the Line’s Detail form. The optical power level in line 1 is decreasing 0.595 dB at a distance of 15.1969 km.

**Fig. 5.** An example of failure line in the Line’s Detail form. The line 8 undergoes failure at 15.1867 km when the attenuation in an optical attenuator is 13 dB, which shows the breakpoint occurring at that distance in the real condition.
3.2. Displaying line’s detail

The developed program is able to identify and present the parameters of each optical line such as the line’s status either in working condition (normal operation) or non-working condition (breakdown), the magnitude of attenuation as well as the location, and other details (breakdown location, line’s parameter such as return loss, crosstalk, etc.) are shown in the computer screen. The advantage of this feature as compared to the OTDR and computer-based emulation software is CFDS displayed every status for the testing line in the Line’s Detail form which display onto one screen board. A Good condition or Decreasing y dB at z km message displays at the line’s status panel in a working condition (see Fig. 4). However in the non-working condition, a failure message Line x FAILURE at z km from CO! displayed to show the exact failure location in the network as illustrated in Fig. 5. It is flexible and easily to use for those who are inexperienced in the optical fiber testing by just reading the information gain from the messages.

3.3. Tracking optical signal level and losses

This feature enables CFDS to compare every line’s optical power and losses. In the second experiment, optical attenuator is used to reduce the optical power in the

![Fig. 6. The summary of optical signal level among eight lines.](image1)

![Fig. 7. The summary of system losses among eight lines.](image2)
tested line. The attenuation of optical attenuator is set to 0 dB and increased by 2 dB for every following test. CFDS tracks the input power, output power, and losses for each line in the Optical Power Comparison form, as shown in Fig. 6. Figure 7 gives the comparison for losses among the eight lines in the Losses Comparison form.

3.4. Monitoring network performance

At the same time, CFDS stores the analysis results in database for further processing and queries. All kinds of additional information can be easily accessed and queried later. We can evaluate the performance of the network via the summarized daily and monthly network performance graph, as illustrated in Figs. 8 and 9. In the daily network performance form, we can evaluate the daily network performance which may

Fig. 8. The summary of daily network performance. The daily network performance (%) versus time (hour) plot is recording the network performance (in unit %) every 4 hours.

Fig. 9. The summary of monthly network performance.
require some prompt action. The network performance can be monitored by CFDS 24 hours a day and 7 days a week. In the monthly network performance form, the graph is able to give an overview to show the daily network performance for the respective month. It clearly shows the best cast (higher performance), worst cast (lowest performance) as well as the average performance for the respective month.

4. Discussion

The optical signal level gives a visual representation for the network deployment and connection. CFDS observes the optical fiber line’s attenuation characteristics and losses through events identification method. CFDS can track all losses and attenuation in the network for the preventive maintenance and network performance monitoring purposes. When a fiber fault occurs, the field engineers may determine sharply the breakpoint before it has restored the stand-by line and repaired for post-fault maintenance through CFDS. With CFDS database, the histories of scanning can be recalled and further analysis can be conducted. Through in-service monitoring with CFDS, the field engineers can view the service delivery and detect any breakdowns as well as other circumstances which may require some prompt action before it turns into a big trouble and causes a tremendous financial loss.

5. Conclusions

Locating a fiber fault within FTTH access network becomes more and more important due to the increasing demand for reliable service delivery. CFDS can help any network service providers and field engineers to monitor the status and detect the failure location in tree-based structured FTTH access network downwardly. It is a cost-effective way to detect the failure location within FTTH access network with CFDS to improve the service reliability and reduce the restoration time and maintenance cost. It should be mentioned that CFDS is not limited to scenario with a single event (single condition), but could be applied to more complicated network configurations. In future research activity, we aim to develop the third generation of this program based on underlying software code (such as C or C++ language) which will be focused on adding extra new features and larger database. Finally, we highlight the possibility of modifying our strategy so that it would be applicable to the long haul optical communication link with various topologies or other fiber-to-the-x (FTTx) schemes.

Acknowledgements – The present work was conducted in Photonic Technology Laboratory, Institute of Micro Engineering and Nanoelectronics (IMEN), Universiti Kebangsaan Malaysia (UKM), Malaysia, funded by the Ministry of Science, Technology and Innovation (MOSTI), Malaysian through the National Science Fund (NSF) 01-01-01-SF0493.
References


Received September 4, 2008
in revised form October 28, 2008
Influence of space charged particles on satellite optical communication system

RUI HOU*, SHANGHONG ZHAO, JIE XU, JILI WU, YONGJUN LI, SHENGBAO ZHAN, LEI SHI, SHAOQIANG FANG

Telecommunication Engineering Institute, Air Force Engineering University, Xi’an 710077, China

*Corresponding author: houruirachel@yahoo.com.cn

Although research of satellite optical communication system has been carried out for many years, there is scarce literature to comprehensively analyze the influence of space environment on satellite optical communication system. Different kinds of particles and fields exist in the space environment, including high energy charged particles, solar radiation, plasma environment, space fragment, etc. The Influence of space charged particles on satellite optical communication system was investigated in detail, which mainly related to single event upset (SEU), total dose effect and plasma environment. For SEU analysis the relation between single proton upset rate and satellite orbit was analyzed in detail. The reliability index of equipment based on SEU was proposed, the numerical calculation results have proved that the SEU effect was relatively less and corresponded to higher reliability of SRAM/MOS equipment under lower orbit altitude and inclination. For plasma environment analysis there is no obvious influence of plasma on laser signal transmission. But charging and discharging processes on satellite surface would lead to the malfunction of satellite communication system. The influence of charged particles and its related plasma on satellite optical communication system was investigated, which would be helpful for the design and the improvement of performance of satellite optical communication system.

Keywords: satellite optical communication, charged particle, single event upset, plasma.

1. Introduction

The satellite communication technique based on microwave could not meet the requirement of space information transmission. The satellite optical communication system has advantages compared with microwave communication system [1], it has larger information capacity, good direction, it is smaller and lighter, and so on. It has become a possible information transmission scenario in the space. A great deal of literature and experiments have been developed to explore and improve the satellite optical communication system, which mainly focus on the improvement of BER (bit error ratio) system. The transmission medium was considered to be free space. Actually, there are different kinds of particles and fields in the space environment [2],
including high energy charged particles, solar radiation, plasma environment and space fragment. The laser communication terminal is a complex integrative equipment, which has higher requirements of satellite optical communication system compared with microwave’s. According to statistics, nearly 40% of satellite malfunctions were induced by space environment. Although the research of satellite optical communication system had been carried out for many years, there was scarce literature to comprehensively analyze the influence of space environment on satellite optical communication system, which we believed to be interesting and important. It would be helpful for the design and the performance and stability of satellite optical communication system. The charged particles are common in space, which were considered to be first investigated.

2. Single event upset effect

2.1. SEU prediction

The single event upset (SEU) is caused by a single particle deposition, which results in “soft error”. The high energy charged particles in space (proton, heavy ion, and neutron) shorten micro-electronic equipment, and produce large numbers of electron–hole pairs inner equipment. The instantaneous upset of electronic equipment would appear under the action of inner electric field equipment. It induces large scale and super-large scale electronic circuits causing mistakes or temporary invalidation. Single event upset rate is defined as the SEU occurrences in unit time, being generally represented by 1/(bit·day).

The research on SEU is always performed through satellite experiment and ground simulation. The latter is completed by a heavy ion accelerator, which is convenient to be realized. There are different methods to complete SEU prediction, in which some errors exist. Two methods with less errors, differential energy spectrum and figure of merit (FOM) have been used to evaluate SEU. Further, the effects of SEU on the reliability of satellite optical communication system are discussed.

2.1.1. SEU prediction through differential energy spectrum

The satellite is always shielded by an aluminum shell. The differential energy spectrum [3] of inner particles $f(E)$ can be expressed as:

$$f(E) = f'(E') \frac{S(E')}{S(E)} \exp(-ct)$$

(1)

$$E' = R^{-1}[R(E) + t]$$

(2)

$$c = \frac{\eta (A^{1/3} + 8.6)^2}{27} \times 5 \times 10^{-26}$$

(3)
where \( f'(E') \) is the differential energy spectrum of satellite surface (cm\(^{-2}\)·d\(^{-1}\)·MeV\(^{-1}\)); \( E \) is the particle energy inner satellite (MeV); \( R(E) \) is the transmission distance of particle with energy \( E \); \( S(E) \) is the held back energy of particle; \( A \) is the mass of particle; \( t \) is the thickness of aluminum shell; \( \eta \) is the Avogadro constant.

The SEU rate \( R_h \) (1/(bit·day)) could be expressed through linear energy transition (LET) spectrum and SEU cross-section area with different LET values

\[
R_h = \int_0^\infty f_e(L) \sigma_0(L) dL
\]

\( L \) is LET value of heavy ion (MeV·cm\(^2\)·mg\(^{-1}\)); \( \sigma_0(L) \) is the SEU cross-section area of heavy ion; \( f_e(L) \) is the equivalent differential energy spectrum \([4]\) of heavy ion, which can be expressed as

\[
f_e(L) = \frac{f(L)}{2\pi} \int_{\theta_c}^{\pi/2} \cos(\theta) d\Omega = \frac{f(L)}{2} \cos^2(\theta_c) =
\]

\[
= \begin{cases} 
\frac{f(L)}{2} \left( \frac{L}{L_0} \right)^2 & L \leq L_0 \\
\frac{f(L)}{2} & L > L_0 
\end{cases}
\]

\( f(L) \) is the differential energy spectrum of heavy ion; \( L_0 \) is the LET threshold of SEU; the critical angle at which SEU occurs is \( \theta_c = \arccos(L/L_0) \).

According to experimental results, the SEU rate \([4]\) of proton \( R_p \) could be expressed as

\[
R_p = \int_0^{\infty} \sigma_0 \left[ 1 - \exp \left( -\left( \frac{L-L_0}{W} \right)^8 \right) \right] \varphi(L) dL
\]

\( \varphi(L) = 2.4 \times 10^{-6} \times \int_0^{\infty} f(E_p) \left( 0.134 + \frac{9}{E} \right) \exp \left[ 0.134 + \frac{9}{E} \right] L dE \)

where \( W \) and \( S \) are Weibull parameters; \( \varphi(L) \) is the effective differential LET function of proton.

The final SEU rate, under the action of both proton and heavy ion, can be expressed as follows:

\[
R = R_p + R_h
\]
2.1.2. SEU prediction through FOM

The FOM method can also be used to predict SEU rate, which can be calculated as

$$ R = R_p + R_h = (C_p + C_h) F = (C_p + C_h) \frac{\sigma_{hL}}{L_{0.25}^2} $$

(9)

where $C_p$ and $C_h$ are the orbit SEU coefficients of proton and heavy ion (upsets/(bit·day)); parameters $F$ could be achieved through experimental results for heavy ion $\sigma_{hL}/L_{0.25}^2$; $\sigma_{hL}$ is the SEU saturation cross-section (cm$^2$/bit); $L_{0.25}$ is the LET value corresponding to the 25% of SEU saturation cross-section, $L_{0.25} = L_0 + 0.2881/S_W$.

There are two kinds of equipment (enhanced and un-enhanced), which are characterized by the same shielding conditions. The SEU rate $R_1$ and parameter $F_1$ of equipment 1 can be calculated by differential energy spectrum method, then the SEU coefficient can be obtained from $C = R_1/F_1$. For equipment 2, the parameters $\sigma_{hL}$ and $L_{0.25}$ are obtained, the parameter $F_2$ can be calculated. The SEU rate of some orbit can be calculated as $R_2 = CF_2$. For the satellites on the same orbit, the SEU coefficient changes with the shielding thickness. $C_t$ is defined as the SEU coefficient of shielding thickness $t$, $C_t$ is the SEU coefficient of shielding thickness $t_1$, then

$$ C_t = 2 C_{t_1} - 0.5 C_{t_1} \log(t) $$

(10)

<table>
<thead>
<tr>
<th>Devices</th>
<th>Satellites</th>
<th>Type</th>
<th>$L_0$ [MeV/mg/cm²]</th>
<th>$W$ [μm²/bit]</th>
<th>$S$ [μm²/bit]</th>
<th>$\sigma_{hL}$ [μm²/bit]</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HM6516</td>
<td>UOSAT-2</td>
<td>Hitachi, 16K</td>
<td>5</td>
<td>14</td>
<td>1.9</td>
<td>180</td>
<td>1.195×10⁻⁸</td>
</tr>
<tr>
<td>2164</td>
<td>CRESS</td>
<td>MOS</td>
<td>0.487</td>
<td>4.95</td>
<td>1.42</td>
<td>170</td>
<td>2.62×10⁻⁷</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Orbits</th>
<th>$\sigma_p$</th>
<th>$\sigma_h$</th>
<th>$C$</th>
<th>$R_p$</th>
<th>$R_h$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Earth orbit (LEO) (500 km/52°)</td>
<td>103.31</td>
<td>3.62</td>
<td>106.93</td>
<td>1.23×10⁻⁶</td>
<td>4.33×10⁻⁸</td>
<td>2.09×10⁻⁶</td>
</tr>
<tr>
<td>Sun synchronous orbit (870 km/98.9°)</td>
<td>436.16</td>
<td>38.96</td>
<td>475.12</td>
<td>5.21×10⁻⁶</td>
<td>4.66×10⁻⁷</td>
<td>5.68×10⁻⁶</td>
</tr>
<tr>
<td>Geostationary Earth orbit (GEO) (35790 km/0°)</td>
<td>0.11</td>
<td>75.61</td>
<td>75.72</td>
<td>0.13×10⁻⁸</td>
<td>9.04×10⁻⁷</td>
<td>9.05×10⁻⁷</td>
</tr>
<tr>
<td>High Earth orbit (HEO) (20000 km/63.4°)</td>
<td>0.10</td>
<td>113.19</td>
<td>113.29</td>
<td>1.195×10⁻⁹</td>
<td>1.35×10⁻⁶</td>
<td>1.35×10⁻⁶</td>
</tr>
<tr>
<td>Highly elliptical orbit (200×36000 km/28.5°)</td>
<td>878.56</td>
<td>57.80</td>
<td>936.36</td>
<td>1.05×10⁻⁵</td>
<td>6.91×10⁻⁷</td>
<td>1.12×10⁻⁵</td>
</tr>
</tbody>
</table>
2.2. SEU rate analysis

In the satellite optical communication system, SRAM and CMOS equipment saves important control bit, whose SEU-induced errors would lead to the malfunction of satellite optical communication system. According to some experimental results [6, 7], under the shielding of a 3.3 mm aluminum shell, the SRAM device HM6516 and MOS device 2164 on satellite have been discussed, whose parameters are listed in Tab. 1.

2.2.1. SEU rates of different orbits

According to the parameters from Tab. 1, the SEU rates of unenhanced HM6516 on different orbits are shown in Tab. 2.

It can be seen from Tab. 2 that the orbits of SEU induced by proton mainly include LEO, Sun synchronous orbit, and highly elliptical orbit; the orbits of SEU induced by heavy ion mainly include GEO and HEO.

2.2.2. SEU rate of LEO

Although the space radiation and flux of high energy proton and heavy ion are lower, the SEU of LEO is an important factor that affects the reliability of equipment on satellite. The MOS device 2164 was analyzed, the SEU rates on different orbits were calculated, as listed in Tab. 3. The typical altitudes of orbits were selected as 500 km and 1000 km, the proton SEU rates of different orbit inclinations were calculated, see Tab. 4.

It can be seen from Tabs. 3 and 4 that the SEU of proton increases with an increase in the orbit altitude; the SEU rate at 0° inclination increases much faster. For 500 km and 1000 km orbits the SEU rates reach maximum at 30° and 20°, respectively.

\[
\text{Table 3. Proton SEU rates of MOS 2164 at 0°, 30° and 60° inclinations.}
\]

<table>
<thead>
<tr>
<th>Orbit altitude [km]</th>
<th>0° (C_p) (R_p)</th>
<th>30° (C_p) (R_p)</th>
<th>60° (C_p) (R_p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.01 2.62\times10^{-9}</td>
<td>0.1 2.62\times10^{-8}</td>
<td>0.07 1.83\times10^{-8}</td>
</tr>
<tr>
<td>300</td>
<td>3.1 8.12\times10^{-7}</td>
<td>2 5.24\times10^{-7}</td>
<td>0.85 2.27\times10^{-7}</td>
</tr>
<tr>
<td>400</td>
<td>72 1.89\times10^{-5}</td>
<td>10 2.62\times10^{-6}</td>
<td>4.7 1.23\times10^{-6}</td>
</tr>
<tr>
<td>500</td>
<td>225 5.90\times10^{-6}</td>
<td>40 1.05\times10^{-5}</td>
<td>18 4.72\times10^{-6}</td>
</tr>
<tr>
<td>600</td>
<td>1000 2.62\times10^{-4}</td>
<td>200 5.24\times10^{-5}</td>
<td>85 2.23\times10^{-5}</td>
</tr>
<tr>
<td>700</td>
<td>300 7.86\times10^{-5}</td>
<td>420 1.10\times10^{-4}</td>
<td>143 3.75\times10^{-5}</td>
</tr>
<tr>
<td>800</td>
<td>553 1.81\times10^{-5}</td>
<td>690 1.81\times10^{-4}</td>
<td>302 7.91\times10^{-5}</td>
</tr>
<tr>
<td>900</td>
<td>1000 1.45\times10^{-4}</td>
<td>900 1.81\times10^{-4}</td>
<td>545 1.43\times10^{-4}</td>
</tr>
<tr>
<td>1000</td>
<td>1600 4.19\times10^{-4}</td>
<td>1500 3.93\times10^{-4}</td>
<td>700 1.83\times10^{-4}</td>
</tr>
<tr>
<td>1100</td>
<td>3000 7.86\times10^{-4}</td>
<td>2000 5.24\times10^{-4}</td>
<td>1000 2.62\times10^{-4}</td>
</tr>
<tr>
<td>1200</td>
<td>3260 8.54\times10^{-4}</td>
<td>3300 8.65\times10^{-4}</td>
<td>4950 1.30\times10^{-3}</td>
</tr>
</tbody>
</table>
Table 4. Proton SEU rates of MOS 2164 at 500 km and 1000 km.

<table>
<thead>
<tr>
<th>Orbit inclination [°]</th>
<th>500 km</th>
<th>1000 km</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(C_p)</td>
<td>(R_p)</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>2.62(\times)(10^{-9})</td>
</tr>
<tr>
<td>10</td>
<td>0.5</td>
<td>1.31(\times)(10^{-7})</td>
</tr>
<tr>
<td>20</td>
<td>14</td>
<td>3.67(\times)(10^{-6})</td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td>1.05(\times)(10^{-5})</td>
</tr>
<tr>
<td>40</td>
<td>31</td>
<td>8.12(\times)(10^{-6})</td>
</tr>
<tr>
<td>50</td>
<td>22</td>
<td>5.76(\times)(10^{-6})</td>
</tr>
<tr>
<td>60</td>
<td>17</td>
<td>4.45(\times)(10^{-6})</td>
</tr>
<tr>
<td>70</td>
<td>16</td>
<td>4.19(\times)(10^{-6})</td>
</tr>
<tr>
<td>80</td>
<td>16</td>
<td>4.19(\times)(10^{-6})</td>
</tr>
<tr>
<td>90</td>
<td>16</td>
<td>4.19(\times)(10^{-6})</td>
</tr>
</tbody>
</table>

Fig. 1. Variations of SEU rate of proton with orbit altitude.

Fig. 2. Variations of SEU rate of proton with orbit inclination.
The numerical fit of discrete data in Tabs. 3 and 4 would directly present the variations of SEU, which are shown in Figs. 1 and 2. It can be seen from Fig. 1 that when orbit altitude changes from 200 km to 900 km, the SEU rate reaches its maximum at 30° inclination, the second maximum place is at 60°, the 0° inclination is minimum, but it changes fast. When the orbit altitude is higher than 1100 km, the SEU rate reaches maximum at the 0° inclination and minimum at 60° inclination. It can be seen from Fig. 2 that for a 500 km orbit, when the inclination is between 0° and 30°, the SEU rate of proton changes greatly; when the inclination is between 30° and 60°, the SEU rate varies slowly; when the inclination surpasses 60°, the SEU rate almost does not change. For a 1000 km orbit, when the inclination is between 0° and 20°, the SEU rate increases with an increase of orbit inclination; when the inclination is over 20°, the SEU rate decreases with an increase of orbit inclination.

2.3. Equipment reliability analysis

According to the long term measurements, the SEU is considered as a random event $\xi$, with the probability of SEU of a single memory bit being the Poisson distribution,

$$p(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, ...$$

where $k$ is the number of times that SEU happens ($k = 0$ represents the probability that no upset happens); $\lambda = RT$ is the mean value of SEU; $P_0 = e^{-RT}$.

The SEU reliability index of equipment $P_R$ is defined as

$$P_R = 1 - e^{-RT}$$

For enhanced equipment $P_R$ can be expressed as

$$R_r = f R_0$$

$$P_R = 1 - e^{-TR_r} = 1 - e^{-Tf R_0}$$

where $T$ is the lifetime of satellite; $f$ is a parameter related to the enhanced equipment. The relations between the SEU of LEO satellite and inclination and altitude through the lifetime of satellite are shown in Figs. 3 and 4.

3. Plasma environment

3.1. Plasma frequency

The plasma environment is also related to charged particles. The plasma frequency $\omega_p$ is determined by the density of free electrons in plasma. Collisions are caused by the movement of charged particles, therefore the electromagnetic wave would be...
absorbed by plasma. When the frequency of incident wave $\omega > \omega_p$, it could transmit in the plasma; if $\omega < \omega_p$ the incident wave would be totally reflected.

Because the mass of positive ion is greatly larger than that of electron, here we only consider the movements of electrons. The density of electrons is $N$; the mass of electron is $m$; velocity is $v$; charge quantity is $e$, then the movements of electrons satisfy

$$\frac{\partial N}{\partial t} + \nabla \cdot (Nv) = 0 \quad (15)$$

Because the magnetic force is much lower than electric force, which could be neglected [8], then

$$m \frac{dv}{dt} = m \left[ \frac{\partial v}{\partial t} + (v \cdot \nabla)v \right] = -eE_i \quad (16)$$

The electric field of incident wave is $E_i = E_e(x e^{-j\omega t})$ and $j = j_0 e^{-j\omega t}$, then

$$j = i \frac{n_0 e^2}{m_\omega} E_e = \sigma_e(\omega) E_e(\omega) \quad (17)$$
The electric conductivity $\sigma_c(\omega) = i[(n_0 e^2)/(m \omega)]$ is pure imaginary number, which shows a phase difference of $\pi/2$ between $j$ and $E_e$. There is no energy attenuation, the average power is $\bar{P} = 0$. The complex dielectric constant is $\varepsilon' = \varepsilon + i[\sigma_c/\omega]$, then

$$\varepsilon' = \varepsilon_0 - \frac{n_0 e^2}{m \omega^2} = \varepsilon_0 \left(1 - \frac{n_0 e^2}{m \varepsilon_0 \omega^2}\right) = \varepsilon_0 \varepsilon'_r \quad (18)$$

It can be seen from Eq. (18) that the plasma frequency $\omega_p$ is determined by $[1 - n_0 e^2/(m \varepsilon_0 \omega^2)] = 0$, $\omega_p = n_0 e^2/m \varepsilon_0$. When the frequency of electromagnetic wave $\omega > \omega_p$, the real part of $\varepsilon'$ is larger than 0, the incident wave could transmit into plasma and be used in satellite communication, except for some attenuation. When $\omega < \omega_p$, the incident wave is totally reflected at the interface between free space and plasma. For satellite optical communication system, whose frequency of laser signal is $10^{14}$ Hz or even much higher, and the plasma frequency is generally between $10^6$ Hz and $10^7$ Hz, therefore $\omega \gg \omega_p$, there would be no total reflection of laser signal.

### 3.2. Effects of plasma on laser signal transmission

The effects of plasma on electromagnetic wave will be used to analyze the effects of plasma on laser signal transmission. The transmission equation of electromagnetic wave in plasma [9] is

$$\nabla \times \nabla \times E = -\varepsilon_0 \mu_0 \hat{\gamma} \frac{\partial^2 E}{\partial t^2} \quad (19)$$

Here, we only consider the interactions between plane wave and single layer plasma. Assume that the electromagnetic wave transmitted along $z$ axis, so

$$E = E_0 \exp(j \omega t - \hat{\gamma} z) \quad j \equiv \sqrt{-1} \quad (20)$$

where $\omega$ is the frequency of electromagnetic wave; $\hat{\gamma}$ is the transmission constant; the dispersion relation of electromagnetic wave in plasma is

$$\hat{\gamma}^2 = -\frac{\omega^2}{c^2} \hat{\varepsilon}_r \quad (21)$$

where $c$ is the light velocity. The transmission constant [10] of electromagnetic wave in plasma is

$$\hat{\gamma} = j \frac{\omega}{c} \sqrt{\hat{\varepsilon}_r} = \alpha + j\beta \quad (22)$$

where the real part $\alpha$ is attenuation coefficient; the imaginary part $\beta$ is phase constant. In partially ionized plasma the density of neutral gas is greatly larger than that of plasma, therefore the collisions of electrons and plasma could be neglected. Besides,
the mass of ion is also greatly larger than that of electron, whose movement could be neglected, so

$$\tilde{\varepsilon}_r = 1 - \frac{\omega_p^2}{\omega^2 + \nu_{en}^2} - j \frac{\omega_p^2 \nu_{en}}{\omega^2 + \nu_{en}^2}$$

(23)

where $\omega_p$ is plasma frequency, $\omega_p = \sqrt{n_0 e^2 / m \varepsilon_0}$; $\omega$ is the frequency of electromagnetic wave; $\nu_{en}$ is the collision frequency between electrons and neutral gas. The plasma frequency is selected as $\omega_p = 10^7$ Hz, collision frequency $\nu_{en} = 10^8$ Hz; the relation between complex dielectric constant and frequency of electromagnetic wave is shown in Fig. 5.

The initial transmitting power of electromagnetic wave is assumed to be $P_0$; the power transmitted into plasma is $P_i$; the power reflected by interface is $P_r$, and it satisfies

$$\frac{P_r}{P_0} = \left| \frac{1 - \sqrt{\tilde{\varepsilon}_r}}{1 + \sqrt{\tilde{\varepsilon}_r}} \right|^2$$

(24)

The power transmitted into plasma is

$$P_i = P_0 - P_r \approx P_0$$

(25)

The power that transmitted into plasma at the place $z$ is

$$P(z) = P_0 \exp(-2\alpha z)$$

(26)

The thickness of plasma is $d$, the power of electromagnetic wave transmitted through plasma is

$$P_i = P_0 \exp(-2\alpha d)$$

(27)
So, the power absorbed by plasma is

$$P_a = P_0 - P_l$$

(28)

The relation between transmission constant $\tilde{\gamma}$ and electromagnetic frequency $\omega$ could be obtained using Eqs. (22) and (27), which is shown in Fig. 6. It can be seen that the real part of transmission constant $\tilde{\gamma}$ decreases fast with an increase of electromagnetic frequency, which is about $10^{-14}$ at the frequency of $10^{14}$ Hz. According to Eq. (27), there is nearly no attenuation of laser signal through plasma.

What is worthy of attention is that the resonance [11] of electromagnetic wave and plasma was mentioned in literature. When the frequency of electromagnetic wave is around the plasma frequency, the resonance of electrons in plasma would be enhanced. There is some peak value of absorption coefficient. When collision frequency, frequency of electromagnetic wave and plasma frequency are getting close to each other, two kinds of resonances happen instantaneously, the peak value is higher. In
order to illustrate this phenomenon, the plasma frequency was selected as $3 \times 10^9$ Hz. The relation between attenuation coefficient and electromagnetic frequency is shown in Fig. 7. Obviously, there is a peak value of attenuation coefficient at $3 \times 10^9$ Hz, and it increases with collision frequency. Certainly the above situation only occurs at the lower frequency of electromagnetic wave, for laser signal it will not happen.

### 3.3. Charging and discharging effects

The effects of plasma on laser signal transmission in satellite optical communication system have been analyzed, and it has been proved that the plasma imposed no obvious attenuation on laser signal. But we cannot say that the plasma has little influence on satellite optical communication system. The charging and discharging process induced by plasma occurred on the satellite surface and inner system, which is also the reason for the satellite malfunction. A schematic diagram of charging process is shown in Figure 8.

In order to analyze the charging process induced by plasma, the satellite surface is divided into several equivalent planes. The resistor of the $i$-th unit is denoted by $R_i$, capacitance by $C_i$, the electricity flown in by $I_{Ti}$, voltage by $\varphi_i$, the resistance of metal unit is denoted by $R_0$, capacitance by $C_0$, the electricity flown in by $I_{T0}$, voltage by $\varphi_0$; $\varphi_i$ and $\varphi_0$ satisfy [12]

\[
\begin{align*}
    \frac{d \varphi_1}{dt} &= \sum_{i=0}^{n} \frac{I_{Ti}}{C_0} + \frac{I_{T1}}{C_1} - \frac{\varphi_1 - \varphi_0}{R_1 C_1} \\
    \vdots \\
    \frac{d \varphi_n}{dt} &= \sum_{i=0}^{n} \frac{I_{Ti}}{C_0} + \frac{I_{Tn}}{C_n} - \frac{\varphi_n - \varphi_0}{R_n C_n} \\
    \frac{d \varphi_0}{dt} &= \sum_{i=0}^{n} \frac{I_{Ti}}{C_0}
\end{align*}
\]

(29)

The back scattering electricity is neglected, the surface electricity of the $i$-th unit [13] is

\[
I_{Ti} = \left[ J_{p0}(1 + f_{pD}) + J_{e0}(f_{eD} - 1)e^{\varphi_i/T_e} \right] A_i
\]

(30)

\[
I_{T0} = \left[ J_{p0}(1 + f_{pM}) + J_{e0}(f_{eM} - 1)\left(1 - \frac{\varphi_0}{T_e}\right)e^{\varphi_0/T_e} \right] A_0
\]

(31)
Influence of space charged particles on satellite optical communication system

\[ J_{k_0} = 2.7 \times 10^{-14} n_k T_k^{1/2}, \quad k = p, e \]  \hspace{1cm} (32)

\( f_{e_D} \) and \( f_{p_D} \) are the second radiation coefficients of nonmetal material; \( f_{e_M} \) and \( f_{p_M} \) are the second radiation coefficients of metal material; \( A_i \) and \( A_0 \) are the surface areas of the \( i \)-th unit and metal unit; \( T_e \) is electron temperature. Equation (29) can be solved by the Runge–Kutta method, the surface charging voltage of any unit can be solved by ODE23 order in MATLAB.

The second radiation electricity of electron [14] is

\[ I_{T_2} = J_{e_0} A_i \frac{e^{\phi/T_e}}{f_{e_D}} = I_{T_1} f_{e_D} \]  \hspace{1cm} (33)

Based on Eq. (29), the charging processes of satellite surface of 100 and 300 eV plasma are shown in Figs. 9 and 10. The 100 eV plasma would impose no obvious charging process on satellite surface. But the 300 eV plasma would cause the charging

Fig. 8. Schematic diagram of surface and inner charging/discharging process.

Fig. 9. Surface charging process of 100 eV plasma: initial voltage 0 V (a), initial voltage –300 V (b).
voltage of satellite surface to get close to –800 V, which would cause arc discharging. It has been proved that the middle energy plasma like 300 eV could effectively disturb the normal functions of satellite optical communication system.

4. Conclusions

Nearly 40% of satellite malfunctions were induced by space environment, mainly including high energy charged particles, solar radiation, plasma environment and space fragment. The charged particles are common in space, which might cause single event upset and total dose effect. This is also related to plasma in space. For SEU analysis the relation between single proton upset rate and satellite orbit has been analyzed in detail. The reliability index of equipment based on SEU has been proposed, the numerical calculation results have proved the SEU effect to be relatively less and correspond to higher reliability of SRAM/MOS equipment under lower orbit altitude and inclination. For plasma environment analysis there is no obvious influence of plasma on laser signal transmission, but charging and discharging processes on satellite surface would lead to the malfunctions of satellite communication system. Although research into satellite optical communication system has been carried out for many years, there is scare literature to comprehensively analyze the influence of space environment on satellite optical communication system, which would be helpful for the design and the performance and stability of satellite optical communication system. The charged particles and related plasma are first investigated, other space environments like solar radiation, temperature variation and space fragment need to be further investigated.

Acknowledgements – This work was supported by the National High Technology Research and Development Project of China (863); (Grant No.2007AA01Z294).
References


Received September 23, 2008
in revised form November 23, 2008
Light propagation in thermally expanded core fibers with graded-index

MILAN S. KOVAČEVIĆ1,2*, ALEXANDAR DJORDJEVIĆ2, DRAGOSLAV NIKEZIĆ1

1Faculty of Science, Department of Physics, University of Kragujevac, R. Domanovica 12, 34000 Kragujevac, Serbia

2City University of Hong Kong, MEEM Department, Tat Chee Avenue, Kowloon, Hong Kong

*Corresponding author: kovac@kg.ac.yu

Light propagation is analyzed in thermally expanded core (TEC) fibers with graded-index profile. Used as power mixers among others, their core structure at the boundary between the heated and non-heated regions is represented by linear taper. Ray optics is used as the transverse taper dimensions are large relative to the wavelength of propagating light. Trajectories of meridian rays are derived analytically. Numerical results presented show ray trajectories as functions of the position within the taper and taper slope. These are modulated sinusoidal functions whose amplitude and period rise with the taper radius. Both, bound and leaky rays have been examined.

Keywords: optical fibers, geometric optics approach; thermally diffused expanded core fibers; ray trajectories; linearly taper model.

1. Introduction

Because of the complexity of Maxwell’s and derivative equations describing light propagation through optical waveguides, simpler but approximate solutions are often sought. Asymptotic theory, with ray optics as its main component, has been found to be adequate for most aspects of light propagation in step- and graded-index optical waveguides (e.g., [1]). A comparison of wave and ray techniques for solving graded-index (GI) optical waveguide problems is presented in [2]. As a wide range of problems involving GI waveguides cannot be solved analytically, techniques of ray optics elaborated in [2] are often attempted. Ray optics has attracted the attention of researchers because its applications for various types of optical fibers have made it possible to calculate ray-path parameters and analyze the ray temporal dispersion. Application of geometric (ray) optics to investigate the ray dispersion, light power acceptance properties of multi-step fibers, and coupling losses for multi-step index plastic optical fibers have been reported [3].
Tapered dielectric structures maximize the coupling of light into optical fibers and integrated-optic devices or waveguides in general. With the advancement of optical communication networks and optical switching and fiber sensing systems, the thermally diffused expanded core (TEC) fibers [4–6] have gained in prominence. TEC fibers have an expanded mode field diameter obtained by heating a step-index single-mode fiber locally at high temperature (of approximately 1300 °C to 1650 °C) and diffusing the germanium dopant into the core. The expansion rate of the core depends on the heating temperature and duration and dopant intensity in the fiber core. When a TEC fiber is applied to low-connection loss connectors and laser diode modules, the enlarged mode field diameter is an important factor which is related to fabrication time and cost. Theoretical and experimental analysis of TEC fiber characteristics has been reported [7]. Methods for analyzing the propagation, modal and coupling characteristics of same fibers have been elaborated [8, 9].

In this paper, the propagation is described of meridian rays in TEC fibers using a geometric optics approach. Ray trajectories in the expanded core region of the fiber are estimated. Using the model of linear taper with small angles, the analytical solution for trajectories of meridian rays is obtained. This solution describes the ray path in terms of sinusoidal functions with amplitudes proportional to the taper core radius. Attention is given to rays that remain bound to the core region for a specified length of taper.

In terms of the structure of the paper, we first present briefly the refractive-index profile of a TEC fiber under isotropic thermal diffusion in order to explain the linear taper model. We then use the general equation for the ray path in the medium of refractive index \( n \) in order to obtain an equation of ray trajectories. Finally, the propagation of meridian rays in TEC fibers is described and numerical simulation of the ray trajectories is presented.

2. Linearly taper model for TEC fibers

Longitudinal view of a graded-index tapered core of thermally diffused expanded core fibers is shown in Fig. 1 in references [8] and [9]. Following the formalism used in ref. [9], the variation of the refractive index profile under conditions of isotropic thermal diffusion has the form

\[
\begin{align*}
n^2(r) &= n_1^2 \left[ 1 - 2\Delta f(r) \right] \\
\end{align*}
\]

(1)

where \( \Delta = (n_1^2 - n_2^2)/2n_1^2 \) is the index profile height, \( n_1 \) is the maximum core refractive index (on the taper axis), \( n_2 \) is the refractive index of the cladding layer, \( f(r) = 1 - (a^2/A^2)\exp[-(r^2/A^2)] \) while \( a \) is the radius of the input end of taper, and \( A \) is radius of the larger end of taper. A feature of TEC fibers is that the dopants in the core are thermally diffused into the radial direction as the core expands. Variation of the refractive index profile of a TEC fiber as a function of heating time is given in Fig. 2 in ref. [8].
We assume that the thermally diffused expanded core fibers have a slow change in non-uniformity along the fiber. If \( \alpha \) is the slope of the taper, the variable taper’s radius is defined as \( A(z) = a + \alpha z \). In this manner, the refractive profile could be approximated by

\[
n^2(r, z) \simeq \begin{cases} 
    n_1^2 \left[ 1 - 2\Delta \frac{a^2 r^2}{(a + \alpha z)^4} \right] & r < a + \alpha z \\
    n_1^2 \left[ 1 - 2\Delta \right] = n_2^2 & r > a + \alpha z
\end{cases}
\]  

(2)

where the approximation \( \exp(-nx) \simeq (1 - nx) \) was used. In this manner, the TEC fiber considered is modeled as linearly tapered graded-index fiber as shown in Fig. 1. For tapered graded-index fibers, the cladding in the region \( r > a + \alpha z \) is a homogeneous medium with refractive index \( n_2 = n(a, 0) = n(a + \alpha z, z) \). For the taper of length \( L \), the core radius increases linearly from \( a \) at \( z = 0 \) to \( A = (a + \alpha L) \) at the larger end of the taper. It is assumed that both \( \Delta \) and \( \alpha \) are much smaller than unity. It is also assumed that \( A \) is much larger than the wavelength of light propagating in the fiber, allowing us to use ray-optics in studying the propagation of light through the tapered region of the fiber.

3. Rays in graded-index linearly tapered fiber: analytical solution

The general vector form of equation for the ray path in the medium of refractive index \( n \) is [10]

\[
\frac{d}{ds} \left( n \frac{d\mathbf{R}}{ds} \right) = \nabla n
\]

(3)

where \( s \) is the distance measured along the ray path and \( \mathbf{R} \) is the position vector for a point on the ray path as shown in Fig. 2. This equation can be regarded as a generalization of Snell’s law. It can be derived in more than one way [11]. The ray equation describes the beam trajectory in terms of the position vector \( \mathbf{R} \) along the ray measured from some starting point. In this analysis, propagation properties of multimode
Gaussian index profile linear taper are investigated using geometric optics on which the Eq. (3) is based.

Since the transverse dimensions of the tapers are assumed to be large compared with the wavelength of light, the ray optics approach is sufficiently accurate, yet simple for the description of light propagation within the taper. Consider a tapered dielectric waveguide with geometry shown in Fig. 1. This waveguide serves as a simplified model for TC fibers with graded-index refractive profile. The parameter $\Delta \ll 1$ is the index profile height, $n_1$ is the maximum core index and $r, z$ represent the cylindrical radial and longitudinal coordinates, respectively. This waveguide is tapered in such way that the core radius increases linearly in the $z$ direction from its initial value of $a$ to a final value of $\Delta$ over a length $L$. The core-cladding boundaries of the taper form a plane defined by $A(z) = a + \alpha z$ where $\alpha = (\Delta - a)/L$ is the slope of the taper. It is assumed that the taper angle is small so that $\tan^{-1}\alpha \approx \alpha \ll 1$. Using this model in the ray equation, ray trajectories of meridian rays are investigated. In the vector form (3), ray equation is independent of any particular choice of coordinate system.

In Cartesian coordinates, it can be expressed as

\[
\frac{d}{ds} \left( n \frac{dx}{ds} \right) = \frac{\partial n}{\partial x} 
\]

\[
\frac{d}{ds} \left( n \frac{dy}{ds} \right) = \frac{\partial n}{\partial y} 
\]

\[
\frac{d}{ds} \left( n \frac{dz}{ds} \right) = \frac{\partial n}{\partial z} 
\]  

For applications involving optical fibers, ray equation in cylindrical coordinates must be known. The transformation from Cartesian coordinates $(x, y, z)$ to cylindrical coordinates $(r, \varphi, z)$ is accomplished by the following transformations

\[ x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = z \]
\[ x = r \cos \varphi \]
\[ y = r \sin \varphi \]
\[ z = z \]
\[ r = (x^2 + y^2)^{1/2} \]
\[ \varphi = \arctan(y/x) \]

The partial derivatives of \( n \) with respect to \( x \) and \( y \) may be expressed as
\[ \frac{\partial n}{\partial x} = \frac{\partial n}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial n}{\partial \varphi} \frac{\partial \varphi}{\partial x} = \frac{\partial n}{\partial r} \cos \varphi - \frac{\partial n}{\partial \varphi} \frac{\sin \varphi}{r} \] (6)
\[ \frac{\partial n}{\partial y} = \frac{\partial n}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial n}{\partial \varphi} \frac{\partial \varphi}{\partial y} = \frac{\partial n}{\partial r} \sin \varphi + \frac{\partial n}{\partial \varphi} \frac{\cos \varphi}{r} \] (7)

and the derivatives of \( x \) and \( y \) with respect to \( s \) become
\[ \frac{dx}{dr} = \frac{\partial r}{\partial s} \cos \varphi - r \frac{\partial \varphi}{\partial s} \sin \varphi \] (8)
\[ \frac{dy}{dr} = \frac{\partial r}{\partial s} \sin \varphi + r \frac{\partial \varphi}{\partial s} \cos \varphi \] (9)

Using Eqs. (5) to (9), the ray equation can be derived in cylindrical coordinates. Listed below are \( r \), \( \varphi \) and \( z \) in cylindrical coordinates of the ray equation:
\[ \frac{d}{ds} \left( n \frac{dr}{ds} \right) - nr \left( \frac{d\varphi}{ds} \right)^2 = \frac{\partial n}{\partial r} \] (10a)
\[ \frac{d}{ds} \left( nr^2 \frac{d\varphi}{ds} \right) = \frac{\partial n}{\partial \varphi} \] (10b)
\[ \frac{d}{ds} \left( n \frac{dz}{ds} \right) = \frac{\partial n}{\partial z} \] (10c)

For the meridian rays \((d/(d\varphi) = 0)\) of the TEC fibers with refractive index profile given by (2), the Eqs. (10) are summarized as follows:
\[ \frac{d}{ds} \left( n \frac{dr}{ds} \right) \approx - \frac{2n_1 a^2 \Delta r}{(a + \alpha z)^4} \] (11a)
It is noted that the \( z \) component of \( \nabla n \) is of order \( \alpha \Delta \), and its \( r \) component is of the order \( \Delta \), which is much larger than the \( z \) component for the tapers with small slope \( (\alpha \ll 1) \). Hence, the right-hand side of Eq. (11c) is set to zero in order to determine the solution for the ray trajectories. After this assumption is made, this equation becomes

\[
\frac{d}{ds} \left( n \frac{dz}{ds} \right) = 0 \tag{12}
\]

Integrating both sides of Eq. (6), it is obtained that

\[
n \frac{dz}{ds} = \text{const} = \gamma \tag{13}
\]

Equation (7) is Snell’s law derived from ray optics. For the case of \( \Delta \ll 1 \) the assumption \( \gamma \approx n_1 \) can be used. Using the result (13) and noting that \( \frac{d}{ds} = (d/dz)(dz/ds) = (\gamma/n)(d/dz) \), the first equation in the set (11) becomes

\[
\gamma^2 \frac{d^2 r}{dz^2} - \frac{1}{2n} \frac{\partial n^2}{\partial r} = 0 \tag{14}
\]

Substituting for \( n(r, z) \) from Eq. (2), Equation (14) reduces to

\[
(a + \alpha z)^4 \frac{d^2 r}{dz^2} + 2n_1^2 a^2 \Delta \frac{\gamma^2}{r} = 0 \tag{15}
\]

In order to solve (15), a new variable \( \rho \), such that \( \rho = (a + \alpha z)/a \) is introduced. Then, in terms of \( \rho \), Eq. (9) becomes

\[
\rho^4 \frac{d^2 r}{d\rho^2} + \omega r = 0 \tag{16}
\]

where \( \omega = \frac{2n_1^2 \Delta}{\alpha^2 \gamma^2} \). For solving Eq. (16), we used the transformation \( \rho = 1/t \). Equation (16), in terms of \( t \), is then expressed as

\[
\frac{d^2 r}{dt^2} + \frac{2}{t} \frac{dr}{dt} + \omega r = 0 \tag{17}
\]
Next, we state that \( r = u/t \), where \( u \) is a new function of \( t \). Then, Eq. (17) becomes

\[
u'' + \omega u = 0 \tag{18}
\]

The solution of (18) is readily available as \( u = c_1 \sin(\sqrt{\omega} t) + c_2 \cos(\sqrt{\omega} t) \). After some arithmetic, the solution of Eq. (15) is obtained as

\[
r(z) = C_0 \frac{a + \alpha z}{a} \sin \left[ \frac{2 n_1^2 \Delta}{\alpha^2 \gamma^2} \frac{a}{a + \alpha z} + \theta_0 \right] \tag{19}
\]

where \( C_0 \) and \( \theta_0 \) are constants of integration. These constants can be determined by using the initial ray conditions at \( z = 0 \) for \( r \) and \( dr/dz \). If we assume that at \( z = 0 \), \( r = 0 \) and \( dr/dz = \alpha_0 \), the constant \( C_0 \) and \( \theta_0 \) are obtained in the following form:

\[
\theta_0 = -\sqrt{\omega} \equiv -b
\]

\[
C_0 = \pm b \frac{a \alpha_0 - b \cos(b^2) + b \sin(b) \sqrt{\sin(b^2)}}{b \alpha \left[ b \cos(2b) - \sin(2b) \right]} \left[ \sin(b) - \frac{\sin(b^2)}{\sqrt{\sin(b^2)}} \right]
\]

The solution (19) together with \( \varphi = \varphi_0 = \text{const} \) completely describe the ray trajectories of meridian rays in the tapered region of TEC fibers.

The solution (19) indicates that the ray trajectories of meridian rays are described by sinusoidal functions. Their amplitude and period increase as rays propagate from smaller to larger end of the taper. It is emphasized that (19) describes ray trajectories of meridian rays in the core of the fibers only. If a ray reaches the core-cladding boundary of the taper, it will enter and remain in the cladding region travelling straight along and away from the boundary. Such rays are considered leaky. They contribute to the radiation loss of the fiber. In order for rays to remain bound to the core over the entire length of the taperer, the condition \( |r| < (a + \alpha z) \), \( z \leq L \) must be satisfied. This condition is met if \( L \leq z_0 \) where the \( z_0 \) is the smallest positive solution of \( |r(z_0)| = a + \alpha z_0 \).

Figure 3 shows trajectories of meridian rays with the slope of 0.0004, \( a = 4 \mu m \), \( A = 8 \mu m \), \( n_1 = 1.50 \), \( n_2 = 1.48 \), \( \Delta = 0.27 \) and \( L = 10 \text{ mm} \) for \( \varphi = 0 \). It should be noted that the ray in Fig. 3 is fully bound to the core region throughout the entire length of the taper.

Figure 4 shows trajectories of a meridian ray with the slope of 0.0004, \( a = 4 \mu m \), \( A = 15 \mu m \), \( n_1 = 1.492 \), \( n_2 = 1.48 \), \( \Delta = 0.000804 \) and \( L = 10 \text{ mm} \) for \( \varphi = 0 \). For rays to be bound to the core throughout the taper, it is required that Eq. (19) satisfy the condition \( |r| < a + \alpha z \). This condition is met if \( L \leq z_0 \), where \( z_0 \) is the smallest positive solution of \( |r(z_0)| = a + \alpha z_0 \). Figure 4 shows also that the ray is bound for a portion of the taperer. It leaves the core and enters into the cladding region at \( z_0 \approx 500 \mu m \), it is thus considered to be a leaky ray.
Fig. 3. Trajectory of meridian rays in graded-index TEC fibers (slope 0.0004, $a = 4 \mu\text{m}$, $A = 8 \mu\text{m}$, $n_1 = 1.50$, $\Delta = 0.27$ and taper length $L = 10 \text{ mm}$).

Fig. 4. Trajectory of meridian rays in graded-index TEC fibers (slope 0.0004, $a = 4 \mu\text{m}$, $A = 15 \mu\text{m}$, $n_1 = 1.492$, $\Delta = 0.00804$ and taper length of $L = 10 \text{ mm}$).

Fig. 5. Trajectory of meridian rays after heat treatment of 10 h (TEC fiber with $D = 3.9 \times 10^{-16} \text{ m}^2\text{/s}$, $n_2 = 1.46$ and $\Delta = 1.25\%$).
Following the formalism used in the study [8, 9], we studied the ray trajectory in the TEC fiber as a function of heating time. After the heat treatment of 10 h for $D = 3.9 \times 10^{-16} \text{m}^2/\text{s}$, for the fiber with $a = 2 \mu\text{m}$, $n_2 = 1.46$ and $\Delta = 1.25\%$, the ray trajectory of meridian rays is shown in Fig. 5. After the heat treatment of 10 h, the maximum value of $A$ is around 7.494 $\mu\text{m}$. It can be observed in Fig. 5 that the rays are fully bound to the core region throughout the entire length of the taper.

For a TEC fiber with $D = 3.9 \times 10^{-16} \text{m}^2/\text{s}$, $a = 2 \mu\text{m}$, $n_2 = 1.46$ and $\Delta = 1.25\%$, Fig. 6 shows the calculated trajectory of meridian rays after a 6 h heat treatment. It is obvious that the ray in Fig. 6 is not fully bound to the core region throughout the entire length of the taper. These results indicate that, for specific initial conditions, the taper length depends on the duration of the heat treatment. As the duration is increased, larger length of the taper becomes possible. It should be mentioned that rays analyzed entered the taper region at small angle with respect to the axis as the approximation of weakly guiding fiber was used.

### 4. Conclusions

A graded-index thermally expanded core (TEC) fiber is analyzed. Ray optics is used as the transverse taper dimensions are large relative to the wavelength of propagating light. A linear taper model for analysis of TEC fiber is proposed. This model describes the behavior of meridian rays in tapered region of TEC fibers. For small angles consistent with the approximation that the fiber is weakly guiding, an analytical solution for the trajectory of meridian rays is obtained. The solution describes the ray trajectory by a sinusoidal function whose amplitude and period rise with the taper-core radius. Both bound and leaky rays have been examined. Ray trajectories of meridian rays have been calculated for two sample cases of TEC fibers. The proposed model may be used to determine conditions for rays to remain bound to the core region throughout the taper length. The function of TEC fiber suggests a possible application of the taper as a power mixer.

![Fig. 6. Trajectory of meridian rays after heat treatment of 6 h for a TEC fiber with $D = 3.9 \times 10^{-16} \text{m}^2/\text{s}$, $n_2 = 1.46$ and $\Delta = 1.25\%$.](image)
Acknowledgements – The work described in this paper was supported in part by a grant from City University of Hong Kong (Project No. 7002313) and in part by Serbian Ministry of Science, through Project No. 141023.

References


Received August 7, 2008
in revised form September 25, 2008
Multi-stage ring resonator all-pass filters for dispersion compensation

CHINDA CHAICHUAY1, PREECHA P. YUPAPIN2*, PRAJAK SAEUNG3*

1Department of Mathematics and Computer Science, Faculty of Science, King Mongkut’s Institute of Technology Ladkrabang, Bangkok 10520, Thailand
2Advanced Research Center for Photonics, Department of Applied Physics, Faculty of Science, King Mongkut’s Institute of Technology Ladkrabang, Bangkok 10520, Thailand
3Electronics Technology Division, Department of Physics, Faculty of Science, Ramchamhaeng University, Bangkok 10240, Thailand

*Corresponding authors: P. Saeung – prajak_neer27@hotmail.com; P.P. Yupapin – kypreech@kmitl.ac.th

This paper describes group delay time property of the multi-stage ring resonator all-pass filters (RRAPF) in either cascading single stages or using lattice architectures. The present analysis is restricted to directional couplers and waveguides characterized by various parameters, and careful design of these parameters can optimize the group delay response. The extra phase shifters of each single stage have been adjusted to yield a broadband group delay. By increasing the number of filter stages, a larger bandwidth over the dispersion can be obtained. This device is able to provide dispersion compensation to systems such as the high speed dense wavelength division multiplexer (DWDM) for the optical fiber communication system.

Keywords: ring resonator, all-pass filter, group delay, quadratic dispersion.

1. Introduction

The basic type of an autoregressive moving average (ARMA) planar waveguide filter is a single ring resonator connected to one coupler which provides no path back to the input port. This filter is called an all-pass or, in the absence of loss, unit transmittance networks [1], because the magnitude of their transmission factor is unity on the whole spectrum, independent of wavelength. Although lossless all-pass filters do not display magnitude filter characteristics, their phase response is frequency dependent. Therefore, they can be configured for group delay equalization and dispersion compensation [2–4], polarization mode dispersion compensation [5], and other applications based on their phase-frequency characteristics such as band-pass filtering when used in conjunction with other optical components. There have been growing interests in tunable dispersion compensators (TDC) for high-speed
wavelength division multiplexed (WDM) networks. This is because the chromatic dispersion of transmission path could be changed frequently in a dynamically reconfigurable WDM networks. The TDC based on a ring resonator all-pass filter is one of the key components in these networks. Optical ring resonator all-pass filters (RRAPF) can be realized using multi-stage ring resonator in either cascading single stages or using lattice architectures [6]. In this paper, multi-stage RRAPF for dispersion compensation is proposed and analyzed. Desired group delay shape, which has a larger value and is sharper, can be tuned by the amount of power coupling to the ring.

2. Transfer functions of ring resonator all-pass filters

2.1. Cascaded ring resonator all-pass filters

The architecture of single ring and three stage cascaded RRAPF is illustrated in Fig. 1, whose every stage is constructed by one ring resonator and one 2×2 optical coupler. The insertion loss of the coupler \( \gamma \) and \( \kappa_i \) is the coupling factor of the \( i \)-th coupler. When a coherent source is input into a device, the coupling intensity for the throughput path in each stage is denoted by \( c_i \) and for the cross path it is \( -j s_i = -j \sqrt{1 - \kappa_i} \), where \(-j\) represents the \(-\pi/2\) phase shift. As to the transmission of light along the ring resonator (the closed pass), we can represent as \( x z^{-1} \), where \( x = \exp(-\alpha L/2) \) is the one round-trip losses coefficient, and the \( z^{-1} \) is the Z-transform parameter, which is defined in terms of normalized angular frequency \( \omega \) as

\[
z^{-1} = \exp(-j\omega) = \exp(-j\beta L)
\]

where \( \beta = kn_\text{eff} \) is the propagation constant, \( k = 2\pi/\lambda \) is the vacuum wave number, \( n_\text{eff} \) is the effective refractive index of the waveguide and the circumference of the ring is \( L = 2\pi R \), here \( R \) is the radius of the ring.

When all rings have the same circumference, a device we call a uniform cascaded RRAPF. Therefore, each single stage has the same periodic resonant responses in the frequency domain with the free spectral range (FSR) between two resonance peaks given by

\[
\text{FSR} = \Delta f = \frac{c}{n_g L}
\]

where \( n_g = n_\text{eff} + f_\omega (d n_{\text{eff}} / df)_{f_\omega} \) is the group index of the ring waveguide, \( f_\omega \) is the center frequency and \( c \) is the velocity of light in vacuum. The optical resonators resonate at a high order mode. At the \( f_\omega \), the perimeter of the ring is an integer number of guide wavelengths, and this integer \( M_r \) is the order number of mode and \( f_\omega = M_r \cdot \text{FSR} \). Using the scattering matrix with Z-transform or signal flow graph technique as in [7, 8], we can express the transfer function for single RRAPF by

\[
S(z) = \frac{E_o}{E_i} = \frac{c - xz^{-1}}{1 - cxz^{-1}}
\]
By evaluating $S(z)$ at $z = \exp(j\omega)$ and defining the phase delay by $S(\omega) = |S(\omega)| \exp(j\theta(\omega))$, we obtain from (3) the relative intensity transfer

$$|S|^2 = \frac{c^2 + x^2 - 2cx \cos(\omega)}{1 + (cx)^2 - 2xc \cos(\omega)}$$  \hspace{1cm} (4)

and the phase delay is given by

$$\theta(\omega) = \tan^{-1}\left[\frac{x(1 - c^2) \sin(\omega)}{c(1 + x^2) - x(1 + c^2) \cos(\omega)}\right]$$  \hspace{1cm} (5)

The resonances for Fig. 1a occur at frequencies where $\cos(\omega) = 1$, that are at $f = M_c c / n_{\text{eff}} L$. The minimum transmission of $|S|^2$ at resonance is

$$|S|_{\text{res}}^2 = \frac{(c - x)^2}{(cx - 1)^2}$$  \hspace{1cm} (6)

and if the coupling coefficient reaches the critical value of $\kappa = \kappa_c = 1 - x^2$, the intensity reaches zero, and there is no transmission, i.e., the fractional loss around the ring is exactly the same as the fractional loss through the coupler.
2.2. Lattice ring resonator all-pass filters

Figure 2 illustrates the two stage lattice RRAPF and lattice RRAPF 2×2 arrays which the improved group delay can be obtained. For simplicity, the waveguide is considered lossless so that \( x = 1 \), the transfer function of Fig. 2a can be expressed as

\[
S = \frac{E_o}{E_i} = \frac{c_1 - \exp[-j(\omega - \phi)]}{1 - c_1 \exp[-j(\omega - \phi)]}
\]  

(7)

where \( \phi \) is the phase delay shift resulting from the upper ring resonator which is given by

\[
\phi(\omega) = \tan^{-1} \left[ \frac{(1 - c_2^2) \sin(\omega)}{2c_2 - (1 + c_2^2) \cos(\omega)} \right]
\]  

(8)

3. Group delay of ring resonator all-pass filters

The filter’s group delay is defined as the negative derivative of the phase of the transfer function with respect to the angular frequency as follows [6]:

\[
\tau_n(\omega) = -\frac{d\theta(\omega)}{d\omega} = -\frac{d}{d\omega} \tan^{-1} \left\{ \frac{\text{Im}[S(z)]}{\text{Re}[S(z)]} \right\} \bigg|_{z = \exp(j\omega)}
\]  

(9)

where \( \tau_n \) is normalized to the unit delay of the waveguide \( T \). The absolute group delay is given by \( \tau_g = T\tau_n \). Thus, we substitute (5) into (9) and given that \( \tan^{-1}[g(x)]/dx = g'(x)/(1 + g^2(x)) \), the normalized group delay of Fig. 1a is given explicitly in terms of \( c \), \( x \) and \( \omega \) as follows:

\[
\tau_n(\omega) = \frac{x(1 - c^2) \left[ x(1 + c^2) - (1 + x^2) c \cos(\omega) \right]}{\left[ x(1 + c^2) - (1 + x^2) c \cos(\omega) \right]^2 + \left[ (1 - x^2) c \sin(\omega) \right]^2}
\]  

(10)

Equation (10) is a periodic group delay response in frequency domain, which exhibits sharp peaks at \( \omega = 2M\pi \). The value of (10) at resonance where \( \cos(\omega) = 1 \) is

\[
\tau_n\bigg|_{\omega} = \frac{x(1 - c^2)}{x(1 + c^2) - (1 + x^2)c}
\]  

(11)

The normalized group delay as a function of normalized angular frequency \( \omega \) for a lossless waveguide is given by

\[
\tau_n\bigg|_{x = 1} = \frac{1 - c^2}{1 + c^2 - 2c \cos(\omega)}
\]  

(12)
which simplifies to

$$\tau_n \bigg|_{f_0, x = 1} = \frac{1 + c}{1 - c}$$  \hspace{1cm} (13)

In the case of the three stage cascaded RRAPF as in Fig. 1b, it can be shown that the normalized group delay $\tau_n$ which is the sum of the individual normalized group delay $\tau_{ni}$ is induced by each single stage ring resonator. For a lossless waveguide $\tau_n$ is given by

$$\tau_n(\omega) = \sum_{i=1}^{3} \tau_{ni}(\omega) = \sum_{i=1}^{3} \frac{1 - c_i^2}{1 + c_i^2 - 2c_i \cos(\omega)}$$  \hspace{1cm} (14)

The extra tunable phase shifters of each single stage can be added to yield a broadband group delay. Therefore, in this case, the result in (14) in term of $\cos(\omega)$ which is replaced by $\cos(\omega + \phi_i)$, where $\phi_i$ is an additional phase shift of each ring.

Similarly, by using Eqs. (7) and (8), we obtain from (9) the normalized group delay $\tau_n$ of the lattice RRAPF 2×2 array (as Fig. 2b) is expressed by

$$\tau_n(\omega) = \sum_{i=1}^{2} \tau_{ni}(\omega) = \sum_{i=1}^{2} \left[ \frac{1 - c_{1i}^2}{1 + c_{1i}^2 - 2c_{1i} \cos(\omega - \phi_i)} \left( 1 + \frac{1 - c_{2i}^2}{1 + c_{2i}^2 - 2c_{2i} \cos(\omega)} \right) \right]$$  \hspace{1cm} (15)

where $c_{1i}$, $c_{2i}$ are the coupling intensity coefficients for the throughput path of the $i$-th column for lattice RRAPF 2×2 array. The resonance of Fig. 2a occurs at frequency where $\omega = 2\pi$ and $\phi = \pi$, due to the fact that the light from the upper ring must pass through the coupled arm of the upper coupler twice. The value of (15) at resonance for Fig. 2a is then given by

$$\tau_n \bigg|_{\omega_r} = \frac{2(1 - c_1)}{(1 + c_1)(1 - c_2)}$$  \hspace{1cm} (16)

Using identical symmetrical couplers $\kappa_1 = \kappa_2$, the normalized group delay in (16) at resonance simplifies to small value of $2/(1 + c_1)$.

4. Simulation results

The normalized group delay response of single RRAPF in Fig. 1a is shown in Fig. 3. The parameters of the circuit used for this simulation were the design frequency (wavelength) $f_\omega = 193.1$ THz ($\lambda_\omega = c/f_\omega = 1552.52$ nm), $M_r = 1931$, FSR = 100 GHz and $n_g = 3.46$ (for the III–V semiconductor materials waveguide), which determines
the circumference of the ring as \( L = M_r \lambda_0 / n_g = 0.86 \) mm. The internal ring losses are assumed to be fully compensated \((\alpha = 0)\). The normalized group delay response is periodic functions of the frequency of 100 GHz, which is the same as the FSR of the ring resonator and it has been found that as \( \kappa \) is decreased it became sharper and steeper at the resonant point.

Figure 4 is a plot of the normalized group delay of Fig. 1a by varying six values of round trip losses coefficient \( x \) based on Eq. (10). The coupling coefficient is fixed to be \( \kappa = 0.2 \) and the other parameters are the same as those used for Fig. 3. The critical value of the round trip losses coefficient is calculated to be \( x_c = (1 - \kappa)^{1/2} = 0.894 \). For \( x > x_c \), the normalized group delay has a positive peak at resonance indicating that the signal is trapped and spends a relatively long time circulating in the ring. After
decreasing $x$, while keeping $\kappa$ fixed, $\tau_n$ becomes sharper and large positive as $x$ approaches its critical value, then flips to a large negative value and sharper as $x$ is in the region $x < x_c$ and finally decreases in magnitude (remaining negative and broader) as $x$ is further decreased. As we see, the result in Fig. 4 is following: the parameter is $x$ when $\kappa$ is fixed. Similarly, the same group delay response is realized for a fixed $x$ under a variable $\kappa$ as shown in Fig. 5. Here, the round trip losses coefficient is set to be $x = 0.894$, which results in critical value of $\kappa_c = (1 - a)^{1/2} = 0.2$.

A difficulty with the single-stage RRAPF is that the group delay response and the bandwidth over which a desired response can be approximated are limited. By using multi-stage RRAPF, a desired response can be more closely approximated, and it can be achieved over a broader portion of the period compared to single-stage RRAPF. Figure 6 shows the group delay response using three stage cascaded RRAPF as in Fig. 1b for compensation filter dispersion. The parameters of each ring resonator

![Fig. 5. Normalized group delay response of single RRAPF as in Fig. 1a by keeping $x$ fixed and varying coupling coefficient as: $\kappa = 0.35 (a)$, $\kappa = 0.30 (b)$, $\kappa = 0.25 (c)$, $\kappa = 0.15 (d)$, $\kappa = 0.10 (e)$, and $\kappa = 0.05 (f)$.](image1)

![Fig. 6. Normalized group delay response of three stage cascaded RRAPF with lossless as in Fig. 1b for various identical coupling coefficients in each stage of $\kappa = 0.1, 0.4,$ and $0.7$.](image2)
are identical and the circumference of each ring is 0.86 mm. As a result (shown in Fig. 6), a larger bandwidth or FWHM over which the dispersion can be obtained due to increasing the number of filter stages.

By appropriately adjusting the coupling coefficient $\kappa_i$ and the phase shift $\phi_i$ of each single stage, a broadband group delay response can be achieved as shown in Fig. 7. The delay curve of a ring resonator always has a constant surface independent of the coupling coefficient $\kappa_i$. As a consequence, there is a trade-off between the maximum delay and bandwidth for a certain bandwidth ripple. The filter response result from the sum of each single stage response shows a maximum delay of $\tau_{\text{max}} = 0.16$ ns for bandwidth of $\Delta f_{\text{BW}} = 8.6$ GHz and a ripple ($\Delta \tau$) of 5 ps.

The plot of the normalized group delay at resonance for the two stage lattice RRAPF in Fig. 2a is shown in Fig. 8. After decreasing $\kappa_2$, while $\kappa_1$ is fixed at 0.9,

---

Fig. 7. Normalized group delay response of three stage cascaded RRAPF with lossless as in Fig. 1b, where the coupling coefficients and phase shifters of each single stage have been adjusted to yield a broadband group delay.

Fig. 8. Normalized group delay response of two stage lattice RRAPF with lossless as in Fig. 2a, for various coupling coefficients of $\kappa_2 = 0.1, 0.4$ and 0.7, while $\kappa_1 = 0.9$. 
Multi-stage ring resonator all-pass filters for dispersion compensation

becomes a larger positive value and sharper compared with those with the result in Fig. 3 at the resonant point. The use of such configuration as a dispersion compensation filter is limited by unwanted additional sidemode peaks around the resonant point as shown in Fig. 9. A possible solution for realizing only a single group delay peak is obtained for the coupling coefficient $\kappa_1$, higher than 0.5.

5. Conclusions

Multi-stage ring resonator all-pass filters can be used as dispersion compensation. As shown above, the bandwidth utilization can be increased by increasing the number of filter stages. By appropriately adjusting the coupling coefficient $\kappa_i$ and the phase shift $\varphi_i$ of each single stage for cascaded RRAPF, a broadband group delay response can be achieved. Using lattice RRAPF, a normalized group delay at resonance with larger positive value and sharper is achieved compared with those of the cascaded RRAPF. For lattice RRAPF, a possible solution for realizing only a single group delay peak is obtained for coupling coefficient $\kappa_1$, higher than 0.5.

References


Received March 3, 2008
in revised form August 7, 2008
Low cost incoherent pump solution for Raman fiber amplifier

Ana M. Rocha¹, ², Berta Neto¹, ², Margarida Facão², Paulo S. André¹, ²*

¹Instituto de Telecomunicações, Campus de Santiago, 3810-193, Aveiro, Portugal
²Departamento de Física da Universidade de Aveiro, Campus de Santiago, 3810-193, Aveiro, Portugal
*Corresponding author: pandre@av.it.pt

In this work, we demonstrate the implementation of a low cost broadband pump source for Raman amplification. The advantages of this incoherent pumping scheme were demonstrated by comparing, experimentally and numerically, this solution with the one based on coherent pumping. We observed a decrease by 1.4 dB in the Raman gain ripple and a 16% increase in the bandwidth, for the latter solution.

Keywords: Raman fiber amplifier, incoherent pump.

1. Introduction

One of the most important challenges to the design of Raman fiber amplifiers (RFA’s) is that of achieving a flat gain profile over a broad wavelength band, at a reasonable cost of implementation [1]. Typically, multiple pumps at different wavelengths are used to induce a flat gain in an ultra-wide bandwidth. In this type of amplifiers, the expected gain peak is shifted by 13.2 THz (≈ 100 nm in this spectral region) with respect to the pump signal frequency.

Another option to enlarge the bandwidth of Raman amplifiers makes use of incoherent pumps. This approach has recently been proposed and the first experimental performance results show a 7 dB on/off gain with a gain ripple of approximately 0.8 dB over all the C spectral band [2].

Later numerical studies corroborate the above experimental results. They have also shown that the number of pumps needed to obtain the same gain flatness can be significantly reduced using incoherent pumping [3–5].

In addition, incoherent pumping has also the advantage of reducing nonlinear effects such as stimulated Brillouin scattering or four wave mixing induced by the pump-to-pump, pump-to-signal and pump-to-noise interaction [6].
The last reported studies on Raman incoherent amplification have focused on numerical simulations and they have mainly concerned the design of optimum pump spectra for the purpose of minimizing the gain ripple. The best results achieved were 20 dB on/off Raman gain over a 70 nm bandwidth with a gain ripple better than 0.1 dB [7, 8]. Earlier results also indicate that incoherent pumping decreases the spectral ripple of the Raman gain [9–11].

In this paper, we report the implementation of an RFA solution based on commercially available pumps adapted for incoherent emission. Also, we implemented a computational model in order to estimate the performance for this amplification technology. The gain and noise figure profiles of the counter pumping architecture were analyzed experimentally and numerically.

2. Raman amplification model

The theoretical model used to describe the implemented experimental system was based on the following set of coupled differential equations, which describe a multi-pump Raman amplifier, considering the pump-to-pump, pump-to-signal, and signal-to-signal interactions, the attenuation and the temperature dependent amplified spontaneous emission (ASE) [12, 13]:

\[
\pm \frac{dP_k^\pm}{dz} = - \alpha_k P_k^\pm + \frac{P_k^\pm}{A_{efc}} \sum_{j=1}^{k-1} \frac{g_r\left(\Delta v_{j,k}\right)}{A_{efc} k_p} (P_j^\pm + P_{ASE,j}^\pm) + \\
- \frac{P_k^\pm}{A_{efc}} \sum_{j=k+1}^{N} \frac{v_k}{v_j} \frac{g_r\left(\Delta v_{k,j}\right)}{A_{efc} k_p} (P_j^\pm + P_{ASE,j}^\pm) + \\
- \frac{P_k^\pm}{A_{efc}} \sum_{j=k+1}^{N} \frac{v_k}{v_j} \frac{g_r\left(\Delta v_{k,j}\right)}{A_{efc} k_p} 4h v_k \Delta \nu \eta_{j,k}(T) \tag{1}
\]

\[
\frac{dP_{ASE,k}^\pm}{dz} = - \alpha_{k} P_{ASE,k}^\pm + \frac{P_{ASE,k}^\pm}{A_{efc}} \sum_{j=1}^{k-1} \frac{g_r\left(\Delta v_{j,k}\right)}{A_{efc} k_p} (P_j^\pm + P_{ASE,j}^\pm) + \\
- \frac{P_{ASE,k}^\pm}{A_{efc}} \sum_{j=k+1}^{N} \frac{v_k}{v_j} \frac{g_r\left(\Delta v_{k,j}\right)}{A_{efc} k_p} (P_j^\pm + P_{ASE,j}^\pm) + \\
- \frac{P_{ASE,k}^\pm}{A_{efc}} \sum_{j=k+1}^{N} \frac{v_k}{v_j} \frac{g_r\left(\Delta v_{k,j}\right)}{A_{efc} k_p} 4h v_k \Delta \nu \eta_{j,k}(T) + \\
+ \frac{2h v_k \Delta \nu (P_j^\pm + P_{ASE,j}^\pm) \eta_{j,k}(T)}{A_{efc} k_p} \sum_{j=1}^{k-1} \frac{g_r\left(\Delta v_{j,k}\right)}{A_{efc} k_p} \eta_{j,k}(T) \tag{2}
\]
where $P_k, P_{\text{ASE}, k}$ denotes the power of the $k$-th pump or signal, ASE noise at frequency $\nu_k$ propagating in the $\pm \hat{z}$ direction. $h$ represents the Planck constant, $\alpha$ the fiber attenuation, $A_{\text{efc}}$ the optical fiber effective area, $k_p$ is a constant that takes into account the contribution for the signal polarizations and $\eta(T)$ is a phonon occupancy factor. $\Delta \nu$ is the frequency detuning between the pump and signal. In this model, the effect of single Rayleigh backscattering and anti-Stokes generation are neglected. When considering an incoherent pump source, which spectrum could be approximated by a large number of coherent pumps, spread over the spectral bandwidth of the incoherent signal, the effect induced by this scheme is equivalent to the utilization of multipumps.

The experimental values used for the fiber Raman gain coefficient $g_r$ were taken from earlier experimental results reported in references [14, 15]. The remaining system parameters are also based on data previously used in simulation scenarios: $\alpha = 0.2$ dB/km, $k_p = 2$, $\eta(300 \text{ K}) = 0.1226$ and $A_{\text{efc}} = 80 \mu \text{m}^2$ [16].

Another simplification relies on neglecting the signals and pump spectral shape. The incoherent pump was modeled as a multi-pump Raman amplifier composed of a large number of pumps (100 is considered sufficient) with infinitesimal spectral width.

Our analysis of the RFA performance focuses on the on/off gain $G_{\text{on/off}}$ and effective noise figure $N_{\text{Feff}}$, which are respectively defined as:

\[
G_{\text{on/off}} = 10 \log \left( \frac{P_{s, \text{pump on}}}{P_{s, \text{pump off}}} \right) 
\]

\[
N_{\text{Feff}} = 10 \log \left( \frac{1}{G_{\text{on/off}}} \left( \frac{2P_{\text{ASE}}}{h \nu \Delta \nu} + 1 \right) \right) 
\]

where $P_s$ is the signal output power, $P_{\text{ASE}}$ the forward ASE noise output power, measured in a bandwidth of $\Delta \nu$, for a signal with frequency $\nu$.

3. Description of the experiment

The RFA performance was accessed measuring the Raman on/off gain and the effective noise figure for the counter pumping configuration with coherent and incoherent pumps.

The coherent pumping source used was a high power fiber Bragg grating (FBG) stabilized laser, peaking at 1489 nm (FITEL model-FOL1425RUX-617). The incoherent pump was obtained also from a high power FBG pump laser, where the stabilization grating was removed, causing an emission spectrum with a full width at half maximum (FWHM) of 10 nm, centred at 1498 nm. The pump power was set to achieve the same Raman gain average for both pumps, which implies a pumping power of 160 mW for the coherent pump and 155 mW for the incoherent one. Figure 1 shows optical spectra for the two pumps (coherent and incoherent).
The transmission was analyzed for 21 signals 2 nm spaced, over the 1580–1620 nm spectral region, with an optical power of −5 dBm per signal. For amplification/propagation medium a 40 km standard single mode fiber was used.

The implemented experimental setup was based on the typical schemes of counter pumping Raman amplification [13]. The signals and pumps were joined by a backward pump isolator (DICON), and an optical isolator was placed just before the fiber to protect the signal source. The signal analyses were carried out with an optical spectrum analyser (OSA) (Advantest Q8384).

4. Results and discussion

The implemented amplification scheme achieved a gain average of 3.55 dB for both the incoherent and coherent pumping over the spectral range considered. Figure 2 displays the experimental and simulated gain spectra for the two pumping configuration schemes showing that the simulation results are in good agreement with experimental ones. The average values for the reduced chi-square between the experimental and simulated results are 0.09 and 0.18 for the coherent and incoherent pumps, respectively.

Moreover, Fig. 2 clearly shows a flatter gain spectrum for incoherent pumping over the spectral range considered. The gain spectrum maximum ripples are equal to 1.97 dB and 0.82 dB for coherent and incoherent pumps, respectively. The ripple was calculated in relation to the average gain value and measured over a −0.5 dB bandwidth. The numerical simulation also corroborates these results and shows a 1.4 dB decrease in the ripple value for the incoherent pumping scheme, compared to the coherent one.

The experimental effective noise figures for both pumping schemes are depicted in Fig. 3 in addition to the corresponding simulation results. The simulation results are in good agreement with the experimental ones.
Considering good performance of the computational model in describing the experimental data, we attain the total Raman gain bandwidth making use of simulated results, as presented in Fig. 4.

The gain bandwidth at –1 dB is 34 nm for the coherent pump and 39 nm for the incoherent one. At –3 dB the gain bandwidths are 102 nm and 118 nm for the coherent and incoherent pumps, respectively. It is evident that incoherent pumping makes the Raman amplification bandwidth increase by 16%, as predicted theoretically.
5. Conclusions

We have tested a low cost Raman amplification scheme with an incoherent pump source. We have also implemented a numerical simulation algorithm for coherent and incoherent pumping.

Our results confirmed the advantages of incoherent pumping, such as a decrease of the spectral gain ripple and an increase in the amplification bandwidth.

Furthermore, a good agreement between the experimental and simulation results confirmed the accuracy of the computational model.

Acknowledgements – This work was supported by the POSC program, financed by the European Union FEDER fund and the Portuguese Scientific Program. The authors greatly acknowledge the project ARPA (POSI/EEA-CPS/55781/2004) and the project TECLAR (POCI/A072/2005).

References


Received June 25, 2008
in revised form November 11, 2008
An investigation of the effect of intermediate layer in three-component planar photonic crystal waveguides

HONGJUN SHEN\textsuperscript{1, 2*}, HUIPING TIAN\textsuperscript{1}, YUEFENG JI\textsuperscript{1}

\textsuperscript{1}Key Laboratory of Optical Communication and Lightwave Technologies, Ministry of Education, School of Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, P.R. China

\textsuperscript{2}School of Physics and Electrical Information Science of Ningxia University, Yinchuan, 750021, P.R. China

*Corresponding author: shenhongj2008@hotmail.com

The introduction of a third component into planar photonic crystal waveguides definitely influences the properties of linear defect modes, such as the band diagram, intrinsic loss, group velocity and group velocity dispersion. With the increase of the dielectric constant of the interlayer, the guided modes shift to lower frequencies and the radiative losses decrease in the frequency region of high group velocity of defect mode. The analysis of the sensitivity of a band diagram to the introduction of an interlayer reveals that the wider the planar photonic crystal waveguide and the thicker the slab, the more tolerant the overall structure. When one designs the real planar photonic crystal waveguides, the effect of unintentional intermediate layer on the optical properties of planar photonic crystal waveguides has to be taken into consideration. At the same time, the introduction of an intentional interlayer into macroporous planar photonic crystal waveguides can be utilized to optimize the design.

Keywords: three-component planar photonic crystal waveguide, intermediate layer, group velocity, group velocity dispersion.

1. Introduction

As a potential basic building block for integrated optics, planar photonic crystal waveguide (PPCW) is extensively investigated [1–5]. In these systems, there exist two kinds of modes: truly- and quasi-guided eigenmodes. Truly-guided modes lie below the light line and are theoretically lossless, whereas the quasi-guided modes lie above the light line and are subject to intrinsic radiation losses due to out-of-plane diffraction, so the problem of light line restricts the usage of the PPCW. In addition to the problem of light line, a lot of extrinsic factors, such as insufficient etch depth [6], disorder-induced [4, 7, 8], roughness-induced [9], etc., may contribute to radiative losses.
The sensitivity of the PPCW to the fabrication imperfection is a disadvantage in application. Understanding the effects of the imperfection is thus a topic of both fundamental and practical importance and the effects of the imperfection have to be investigated quantitatively.

In this work, a theoretical treatment of the effect of another kind of fabrication imperfection in PPCW is presented. In the fabrication process of macroporous PPCW by the method of electrochemical etching [10], a third medium with dielectric constant $\varepsilon_i$, which is different from the bulk ones, is introduced like a ring-shaped intermediate layer around the hole’s surface. As a matter of fact, the introduction of the interlayer not only happens in the fabrication process of the PPCW, but also in the practical application it is inevitable. If PPCW does not work in a vacuum, the photonic crystal’s “atoms” may slightly change their properties near the interfaces due to some chemical processes (for example oxidization). The inner pore’s surface of a macroporous slab structure inevitably contains the interlayer.

The unintentional intermediate layer sensitively affects the properties of PPCW. But up to date, little work has been done about this problem. To our knowledge, only the work of Glushko and Karachevtseva [10] focuses on the problem in two-dimensional ideal photonic crystals, which are actually unsuitable for practical application. In order to better understand the optical properties of PPCW, it is necessary to investigate the dependence of the properties of PPCW on various structural parameters, which include those of the intermediate layer.

On the other hand, intentional oxidization of the macroporous structure is widely used [11]. For PPCW, the introduction of the extra component enlarges our freedom for fine-tuning of the dispersion properties of PPCW. As a matter of fact, the fabrication of PPCW which allows to tune the photonic band gap edges by an external influence has been one of important directions of investigations [11, 12].

2. Theoretical model

For a periodic dielectric constant $\varepsilon(r)$, a Bloch state for the magnetic field at the $n$-th band and wave vector $k$ satisfies the Maxwell equation

$$\mathbf{c}^2 \nabla \times \left[ \varepsilon^{-1}(\mathbf{r}) \nabla \times H_{nk}(\mathbf{r}) \right] = \omega_{nk}^2 H_{nk}(\mathbf{r})$$

(1)

$\omega_{nk}$ and $H_{nk}(\mathbf{r})$ are the eigenfrequencies and eigenvectors, respectively. After insertion of a third component, the new dielectric function, eigenfrequencies and eigenvectors are denoted as $\tilde{\varepsilon}(\mathbf{r})$, $\tilde{\omega}_{nk}$ and $\tilde{H}_{nk}(\mathbf{r})$, respectively. Ignoring the higher order terms, from perturbation theory, it is easy to obtain from Eq. (1) that

$$\nabla \times \varepsilon^{-1}(\mathbf{r}) \nabla \times \delta H_{nk} + \nabla \times \left[ \tilde{\varepsilon}^{-1}(\mathbf{r}) - \varepsilon^{-1}(\mathbf{r}) \right] \nabla \times \tilde{H}_{nk}(\mathbf{r}) \approx$$

$$\approx \frac{\omega_{nk}^2 - \tilde{\omega}_{nk}^2}{c^2} H_{nk} + \frac{\omega_{nk}^2}{c^2} \tilde{\delta} H_{nk}$$

(2)
where $\tilde{H}_{nk} = H_{nk} + \delta H_{nk}$. By some mathematic strategies [13], we finally obtain

$$
\left( \frac{\omega_{nk}}{\omega_{nk}} \right)^2 - 1 \approx \int \frac{\left[ \tilde{\varepsilon}^{-1}(r) - \varepsilon^{-1}(r) \right] |D_{nk}(r)|^2 dr}{\int \varepsilon^{-1}(r)|D_{nk}(r)|^2 dr}
$$

(3)

where the integration is over a unit cell. $D_{nk}(r)$ is displacement field, $[\tilde{\varepsilon}^{-1}(r) - \varepsilon^{-1}(r)]$ is nonzero, say, $\delta$, only at the insertion position. In our case $\delta < 0$, so from Eq. (3), it follows that $\omega_{nk} < \omega_{nk}$, which indicates the bands shift downwards. What is more, according to the electromagnetic variational theorem [14], the low frequency modes concentrate their energy in high $\varepsilon$ regions and the high frequency modes concentrate their energy in low $\varepsilon$ regions. The insertion of third medium ($\varepsilon > 1$) affects higher frequency modes more strongly, which implies that the higher frequency modes shift downwards more than the lower frequency modes.

The perturbative approach above explains a physical origin of the shift of bands in photonic crystal (PhC) slab with interlayer. However, ignoring the higher order terms, one can only obtain approximate results by Eq. (3) [13]. In this work, we use the guided-mode expansion (GME) method to calculate more accurately the electromagnetic wave propagation through the 2D dielectric lossless PPCW. GME was described in detail in a lot of papers [15, 16]. The basis of Fourier expansion, which we apply, contains 259 vectors of reciprocal lattice and 4 guided modes of the effective waveguide, which is sufficient for convergence in the calculation of photonic eigenmodes [15, 16]. So photonic eigenmodes obtained by the present approach can be considered reliable and accurate.

For simplicity, we start with the most commonly studied structure, that is W1 waveguide in an air-bridged photonic crystal slab, patterned with a triangular lattice of holes. Figure 1a shows schematic diagrams of the structure. As opposed to the general PPCW, there are additional interlayers which come from fabrication.

![Figure 1](image_url)

Fig. 1. Schematic picture of three-component W1 PPCW in triangular lattice of circular holes with slab thickness $h$, holes radius $R$, thickness of intermediate layer $d$ and waveguide width ($w = w_0 = \sqrt{3} a$) (a); The top view of the unit cell; $\varepsilon_a$, $\varepsilon_b$, $\varepsilon_i$ are the dielectric constant of the air holes, background and interlayer, respectively (b); The supercell for band diagram calculation using GEM method (c).
process on the inner pore’s surface. One supercell with a waveguide for calculation is mapped in Fig. 1e. We define the thickness of slab to be $h$, the width of waveguide to be $w$, the thickness of interlayer to be $d$ and denote by $\varepsilon_b$, $\varepsilon_a$ and $\varepsilon_i$ the dielectric constant of background, air inside holes and intermediate layer, respectively. Infinite periods of the supercells are spatially repeated to form the waveguide in the $\Gamma$–$M$ direction. The waveguide is symmetric. Additional layer on the pore’s surface is characterized by two parameters: dielectric constant $\varepsilon_i$ and thickness $d$. We now restrict ourselves to the consideration of air pores $\varepsilon_a = 1$ in the dielectric with $\varepsilon_b = 12$. This PPCW possesses a very large slab-mode gap for TE-like modes. The larger gap requires a larger filling factor of air holes. However, the larger filling factor of air holes will make the structure more fragile because the mechanical strength of the structure is greatly reduced. Therefore, we choose the radii of air holes to be $0.36a$, where $a$ is the lattice constant. The thickness of the slab is chosen to be $0.3a$. The symmetry of the system is not broken by the insertion of an intermediate layer. The dielectric function of the unit crystal cell is given by [10]

$$
\varepsilon(G) = \begin{cases} 
\frac{2\left(\frac{1}{\varepsilon_a} - \frac{1}{\varepsilon_i}\right)f_1 J_1(GR)}{GR} + \frac{2\left(\frac{1}{\varepsilon_i} - \frac{1}{\varepsilon_b}\right)f_3 J_1(G(R + d))}{G(R + d)} & G \neq 0 \\
\frac{f_1}{\varepsilon_a} + \frac{f_2}{\varepsilon_b} + \frac{f_i}{\varepsilon_i} & G = 0
\end{cases}
$$

where $J_1(x)$ is the Bessel function and $f_i = \frac{\pi}{\sqrt{3}} \frac{(R + d)^2 - a^2}{a^2}$, $f_1 = \frac{\pi}{\sqrt{3}} \frac{R^2}{a^2}$, $f_2 = (1 - f_3)$, $f_3 = \frac{\pi}{\sqrt{3}} \frac{(R + d)^2}{a^2}$.

3. Band diagram of guided modes

Figure 2a shows the calculated band diagram of the PPCW without interlayer for TE-like modes, the photonic mode dispersion is represented in dimensionless units $\omega a/2\pi c$ for the frequency and $ka/\pi$ for the wave vector. A photonic band gap (PBG) exists in the frequency range from 0.33 to 0.47($\omega a/2\pi c$). A series of waveguide modes is formed in the PGB. The dispersion relation of the modes is understood as a mixture of two modes having different natures. One is a PBG-guided mode, which is confined by distributed Bragg reflection due to the periodic structure, and the other is an index-guided mode confined by the total internal reflection. An index-guided mode has its energy concentrated inside the defect and interacts only with the first row of holes adjacent to the defect. A gap-guided mode interacts with several rows of holes, thus it is dependent on the symmetry of the PC and its PBG. An intrinsic interaction between one gap-guided mode and one index-guided mode forms a supermode, which is
An investigation of the effect of intermediate layer ...

represented by the sum of gap-guided and index-guided mode in phase or in antiphase. In our case, four supermodes exist in the PBG region with different lateral symmetry, but only two supermodes which have even and odd symmetry with respect to reflection in the horizontal mid-plane of the channel are shown [17, 18]. We define them as the upper guided mode and the lower guided mode. The index-guided mode and gap-guided mode which interact with each other to form these two supermodes are marked as $\alpha$ and $\beta$, respectively.

For the sake of clarity, the calculated band diagrams in Fig. 2 are shown only when the channel waveguide is monomode for specified parity. The thick solid line denotes the light dispersion in air cladding and every mode above this light line is leaky. When the modes cross light line of air cladding, they become true guided modes which are not theoretically subject to the intrinsic losses. What is more, as shown in Fig. 2d, the dispersion curves of the true guided modes are relatively flat, which results in a small group velocity. This small group velocity leads to a small transmission window and also means that these modes might prove to be susceptible to structural disorder, which must be present in the real samples [19]. This small group velocity can be attributed to the fact that a strong anticrossing occurs between an index-guided mode.

Fig. 2. Band diagram for TE-like mode of W1 air-bridge waveguide without interlayer in photonic crystal slab pattered with a triangular lattice of holes, the thick solid line is the light line of air cladding (a). Dispersion curves of the lower guided mode in PPCW with and without interlayer (b). Dispersion curves of the upper guided mode in PPCW with and without interlayer (c). The tailors of the lower guided modes in part b, the frequency windows corresponding to truly guided modes are highlighted (d). The dielectric constant and thickness of interlayer are 20 and 0.02a, respectively.
and a gap-guided mode close to the Brillouin zone edge of W1 waveguide [1]. In the case of \(d \neq 0\) (with interlayer), due to the introduction of the extra interlayer, the guided modes shift downwards, as shown in the Figs. 1b and 1c, in which, for comparison, the corresponding bands in the case of \(d = 0\) (no interlayer) also be included. In addition, according to the electromagnetic variational theorem mentioned above, the higher the band, the greater will be its downwards shift. This is immediate obvious from comparison between two guided modes in Figs. 1b and 1c. As the guided modes shift downwards, an apparent slope in the region under the light line can be seen. The part of mode below the light line increases significantly (for example, in Fig. 2d about 5 times at \(\varepsilon_i = 20\)). We attribute this to the fact that the introduction of the interlayer significantly shifts the anticrossing between the index-confined mode and gap-confined mode [1]. It is noticed that in a PPCW, the rearrangement of the holes in the row adjacent to the waveguide core is the usual way to achieve large transmission bandwidth for the guided modes [1, 2]. In such case, the magnification of transmission bandwidth is approximately 8 times, which is in the same order of magnification as that achieved in the structure with the interlayer discussed here (5 times). This fact demonstrates that although this additional interlayer is very thin, its effect on the PPCW is not so small as usually expected.

It is well known that the interlayer is usually very thin. According to our systematic stimulations, within a small range of interlayer thickness (< 0.02a), the effect of interlayer depends weakly on its thickness. Therefore in this work, we only discuss structures with an interlayer thickness of \(d = 0.02a\) and with different dielectric constants.

### 4. Group velocity and group velocity dispersion

The slow group velocity and highly dispersive nature are the most distinctive features of a linear defect waveguide. Slow light and strong dispersion will open up new possibilities for functional waveguides, so the properties such as group velocity and group velocity dispersion (GVD) of PPCW have been intensively studied since photonic crystal was found [20–22]. Group velocity of the guided modes is simply the slope of the band curve of guided modes and can be expressed as:

\[
V_g = \frac{\partial \omega}{\partial k} = \frac{c}{n_g}
\]  \hspace{1cm} (5)

where \(k\) is the wave vector along the waveguide direction, \(c\) is the velocity of light in the vacuum, \(n_g\) is the group index of the dispersion material. The GVD parameter \(D\) strongly affects the propagation of optical pulse in the PPCW. Parameter \(D\) can be expressed as the second-order derivative of the dispersion relation

\[
D = \frac{\partial(1 / V_g)}{\partial \omega} = \frac{1}{c} \frac{\partial n_g}{\partial \omega}
\]  \hspace{1cm} (6)
Figure 3 shows the group velocities and the GVD of the waveguide modes with respect to the normalized frequency. In the central region, the linear defect waveguide exhibits relatively high group velocity, and the dispersion value $D$ changes almost linearly. Around the band edges the group velocity is very low and as a result the GVD changes sharply there, toward positive values at one edge and negative values at the other.

Generally speaking, the origin of low group velocity in PPCW is related to three different physical effects [23]:

- The conventional distributed feedback (DFB) effect, which is a coupling between a forward and a backward mode with the same lateral field distribution.
- The DFB effect between a PBG-guided mode and an index-guided mode, which occurs at the anticrossing point.
- Omnidirectional reflection which is a unique feature offered by the photonic crystal environment.

In PPCW, a photonic band gap is presented, light propagating at any angle is reflected in the lateral direction to the waveguide and thus has large lateral wave vector components along the direction perpendicular to the waveguide and in the plane of the slab. Therefore, wave vector components parallel to the waveguides are
negligible near $k \approx 0 \times 2\pi/a$. This indicates that the two supermodes in Fig. 3a have low group velocity near $k \approx 0 \times 2\pi/a$ [23].

In the W1 waveguide, an anti-crossing between a PBG-guided mode and an index-guided mode occurs close to the Brillouin zone edge ($k \approx 0.5 \times 2\pi/a$) [24], which results in slow light near the point. The introduction of the interlayer increases the effective dielectric constants in the core part of PPCW locally, the index-guided and gap-guided modes move downwards but the gap-guided mode moves faster and thus, anticrossing near the Brillouin zone edge is canceled. Only DFB plays a major role in slowing light, the group velocity becomes higher. This is confirmed by what happens in lower frequency region of the lower guided mode in Fig. 3a. However, near the point $k \approx 0 \times 2\pi/a$, a photonic-band-guided mode, which forms the left part of the upper edge of PBG [17], comes closer to the upper guided mode (an index-guided mode), and the coupling between them becomes stronger, which results in lower group velocity in the higher frequency region of the upper guided mode in Fig. 3a.

The introduction of interlayer changes the dielectric constant of PPCW, the eigenfrequency shifts to lower frequency, the change in the eigenfrequency causes a shift of both the group velocity curve and the dispersion curve to lower frequency region. The central gentle inclination region is suitable for device application due to quasi-linear dispersion. In the PPCW without interlayer, the range of bandwidth within which the dispersion is below 1 (ps/a/mm) is $0.0038(\omega a/2\pi c)$ (from 0.3778 to 0.3816 $\omega a/2\pi c$) and increases to 0.03 ($\omega a/2\pi c$) (from 0.3604 to 0.3903 $\omega a/2\pi c$) for less strict constraint of the dispersion 10 (ps/a/mm). While in the corresponding PPCW with interlayer ($\varepsilon_i = 20$) they are $0.005(\omega a/2\pi c)$ (from 0.3686 to 0.3737 $\omega a/2\pi c$) and 0.0314 ($\omega a/2\pi c$) (from 0.3536 to 0.3849 $\omega a/2\pi c$), respectively. It is clear that the introduction of an interlayer results in a broadening of the bandwidth corresponding to certain dispersion region and a downward shift of its frequency region. The rate of frequency shift to bandwidth for $D < 1$ and $D < 10$ is 1.3421 and 0.2267, respectively.

In a high-bit-rate WDM optical communication system, signal pulses, after having been de-multiplexed, are individually reshaped in each channel where a tunable dispersion compensator is required precisely to compensate for dispersion. The group velocity dispersion performances of a PPCW strongly depends upon precise wavelength matching. The introduction of an interlayer causes a downward shift of dispersion curve to lower frequencies, thus lowering operation frequency’s dispersion value from its original position factored into the device’s design [20]. This results in a mismatch between the manufacturer’s compensation value and device’s actual dispersion wave frequency, owing to excessive or insufficient compensation. Naturally, the consequences of this mismatch would be more pronounced at frequencies along the band edges where as our results show, dispersion values increase or decrease exponentially (see Figs. 3b and 3c). Clearly, then, the influence of added interlayer on the dispersive nature of a given PPCW must be carefully taken into consideration.
5. Intrinsic loss of guided modes

The introduction of an interlayer results in the decrease in the hole radius, so the part of electric field scattering out into air decreases and the intrinsic losses decrease [25, 26]. In order to see clearly the influence of interlayer on the intrinsic losses, for different dielectric constants of the interlayer, the intrinsic losses are computed and plotted in Fig. 4. From Fig. 4, it can be observed that as dielectric constant of the interlayer increases, the intrinsic loss of the guided modes reduces in the frequency region of high group velocity of the guided mode. For example, at $\omega = 0.36$, the introduction of the interlayer with $\epsilon_i$ (equal to 6, 12, 20) reduces the intrinsic loss by 3.65, 9.15, 16 dB/mm. Compared to the propagation loss (usually < 10 dB/mm) in many kinds of PPCW which have been optimized by various methods [27, 28] the change in loss which comes from the introduction of an interlayer is not insignificant. In some real PPCWs (which have been optimized to some extent), in order to accurately evaluate the impact of optimized methods on propagation losses, the positive impact on losses which the addition of an interlayer brings about has to be take into account [27, 28].

6. Sensitivity of the effect of an intermediate layer to waveguide width and slab thickness

From the discussion above, one can see that the introduction of an interlayer definitely influences the properties of PPCW. However, for prospective application,
it is necessary to analyze the sensitivity of properties of PPCW to interlayer to obtain a more fabrication tolerant PPCW. Figure 5 shows the amount frequency shift for the lower guided mode \( \Delta \omega = \omega - \omega_0 \) (\( \omega \) is the eigenfrequency of the lower guided mode in PPCW with interlayer, \( \omega_0 \) is the corresponding eigenfrequency of the lower guided mode in PPCW without interlayer), which is brought about by the introduction of the interlayer with different dielectric constants. For simplicity of analysis, we choose right parameters which ensure the two guided modes being in the PBG entirely, we find that the narrower is a waveguide in a thinner slab, the greater is the impact. So, PPCW with wider waveguide width and thicker slab thickness is more tolerant to interlayer. It is noticed that for different frequency range of guided mode, the amount of frequency shift is different. The lower is the frequency, the higher is the shift. This is because in the lower frequency region, the guided modes have a lower group velocity, and a slow mode experiences a larger change in the effective index than a fast mode, despite the fact that the change in the material index \( \Delta n \) is the same [23].

7. Conclusions

We present a comprehensive study about the role of interlayer in the three-component PPCW. We first calculate the band diagram, intrinsic loss, group velocity and group velocity dispersion of PPCW with and without interlayer using GME method. Due to the introduction of the interlayer, the single-mode transmission window of the lower guided mode is substantially widened. The intrinsic losses of guided modes decrease in frequency region of higher group velocity. The group velocity and the group velocity dispersion curve definitely shift to lower frequencies. In prospective application of
An investigation of the effect of intermediate layer...

PPWCs, the performances of which strongly depend upon precise wavelength matching, the effect of the interlayer has to be taken into consideration.

Acknowledgements – This research was supported by the NSFC (No. 60707001, No. 6071140087), National 973 Program (No. 2007CB10705), National 863 Program (No. 2007AA01ZZ47), NCET (07-0110), PCSIRT (No. IRT0609), ISTCP (No. 2006DFA11040), 111 Project (No. B07005) P.R. China.

References


Received August 4, 2008
in revised form October 28, 2008
A 60° photonic crystal waveguide bend with improved transmission characteristics

MALIHE Khatibi Moghaddam¹*, Mir Mojtaba Mirsalehi², Amir Reza Attari²

¹Communications and Computer Research Center, Ferdowsi University of Mashhad, Mashhad, 91779, Iran
²Electrical Engineering Department, Ferdowsi University of Mashhad, Mashhad, 91779, Iran

*Corresponding author: ma_kh303-mail@stu.um.ac.ir

In this paper, a 60° waveguide bend is designed in a two-dimensional (2D) photonic crystal (PC) slab to provide high transmission over a large bandwidth. We apply geometrical modification in the bend region to improve the transmission characteristics. This modification results in increasing the relative bandwidth from 6.5% to 25.7% of photonic band gap width. Using the effective refractive index, the structures designed are simulated by the 2D finite-difference time-domain (FDTD) method. As a specific application, we use the improved bend structure in a PC waveguide directional coupler and show that the drop output increases significantly.

Keywords: photonic crystal, waveguide, bend, bandwidth, slab structures.

1. Introduction

Photonic crystals (PCs) are periodic dielectric structures that exhibit the unique potential of photonic band gap (PBG), i.e., a frequency region in which the propagation of light is not permitted. This property can be utilized to control the light propagation. Unlike 3D structures, 2D PCs can be fabricated easily by integrated circuit technology. PC slabs have vertical confinement using index guiding and confine light in horizontal plane using PBG [1–4]. It was demonstrated that the existence of lossless guided modes can be permitted if their wave vectors are larger than the corresponding values in the cladding medium. This relation is shown with a line which is called light line in the PC band diagram. The modes below the light line are guided modes and those above are leaky [3]. A commonly-used PC structure is manufactured by etching a triangular lattice of air holes in a dielectric slab. This structure provides a suitable band gap for TE-like modes [3].

A linear defect is usually introduced in the PC structure by removing one row of holes and it creates one or more guided modes inside the band gap. Transmission of a PC bend is mainly affected by two characteristics, guided modes of straight
waveguides and coupling efficiency between them [3]. The ability to guide light waves in sharp bends of PC integrated circuits is one of the most important properties of PC waveguides. Reducing the reflection from the bend region in a large bandwidth has been investigated widely in the literature [5–26].

The first results on PC waveguide bends have been reported by Mekis et al. [5], where the transmission through a 90° bend in a 2D photonic crystal with a square lattice structure was investigated. Several techniques are used to improve the transmission of PC bends. In one technique, by introducing a resonant cavity in the bend region and placing some point defects, the transmission of the PC bend has been improved [6–8]. Other reported techniques are changing the width of line defect [9], moving the holes [10–13], varying the hole sizes [7, 12–14], replacing the holes at the outer edge of bend by another design using topology optimization [15–19], placing point defect at the bend edge region [10, 20], removing one hole [21], and tapering the defect size [22]. Also, it has been shown that a curvilinear lattice can be used in the bend part to obtain a large coupling efficiency over a broad bandwidth [19, 23–25].

In this paper, we propose a new bend structure and optimize its geometrical parameters in order to increase the transmission bandwidth.

2. A 60° PC waveguide bend

We report the results of our investigation on improving the transmission characteristics of a 60° bend in a GaAs PC slab with a refractive index of 3.4. Among different geometries of PC structures, we consider the structure of air holes in a triangular lattice which provides a suitable band gap for TE-like modes and therefore can be efficiently coupled to the commonly used optical waveguides [2]. We choose a hole radius of $r = 0.3a$, which is proposed for practical applications [3], where $a$ is the lattice constant. By applying the plane wave expansion (PWE) method, the photonic band structure of PC slab and the guided modes of the PC waveguides can be computed.

![Fig. 1. TE-like band structure of the PC slab shown by solid lines and the corresponding 2D band structure with $n_{\text{eff}} = 2.76$, shown with circles. The dotted lines show the air light lines for air cladding.](image)
A 60° photonic crystal waveguide bend with improved transmission characteristics

For the PC structure considered above, the band gap occurs in the frequency range of 0.256–0.32 \( \frac{c}{a} \), where \( c \) is the velocity of light.

There are two main categories of PC slabs, namely low-index-contrast and high-index-contrast PC slabs, in which the difference between the refractive index of PC and cladding is low and high, respectively. For a low-index-contrast slab, the effective index of slab modes changes slightly for a wide frequency range. Therefore, an average effective index could be used for a wide frequency region in 2D approximation [26]. In a high-index-contrast structure, e.g., the air/Si/air, the effective index method is invalid for such a wide frequency region due to the large changes in index. However, for a narrow frequency range of interest, the 2D approximation using the effective index method may remain valid. As shown in Fig. 1, the 2D PC with \( n_{\text{eff}} = 2.76 \) has the same PBG as the PC slab considered in this paper. As a result, to investigate the transmission and reflection spectra of the 60° PC bend, we utilized 2D approximation of the PC slab, i.e., we applied the 2D finite-difference time-domain (FDTD) method\(^1\) to the structure using a refractive index equal to the effective index of the fundamental mode, which is 2.76 [7, 26]. This approximation reduces the required computational resources (CPU time and memory) significantly.

Figure 2 shows a simple waveguide which is created by introducing a line defect along \( \Gamma K \) direction and illustrates the dispersion curves of corresponding guided modes. As shown in this figure, the frequency range of the even mode below the light line extends from 0.265 to 0.288 \( \frac{c}{a} \). However, the odd mode limits the single mode region of the even mode to the frequency band of 0.265–0.287 \( \frac{c}{a} \). As shown in Fig. 2, the group velocity of even mode is reduced near the frequency of 0.265. This low group velocity decreases the accuracy of FDTD results near the above frequency.

\( ^1 \)The FDTD code is written in MATLAB for a \( 20a \times 20a \) PC lattice. Each unit cell has a dimension of \( 138 \delta x \times 156 \delta y \), where \( \delta x \) and \( \delta y \) represent the mesh size in the FDTD method.
A simple 60° PC bend can be introduced in a triangular structure as shown in Fig. 3. Due to the existence of PBG, light is confined to the bend region and no power will be radiated out of the waveguide. In this simple bend structure, the discontinuity along the guidance path causes large reflection and hence it provides low transmission. The main reason for the low transmission is that the discontinuity at the bend region stimulates higher order modes that are evanescent in the PC waveguide. It should be mentioned that in the frequency range of 0.265–0.287 \( \frac{c}{a} \) the fundamental waveguide mode is below the light line and it cannot be coupled to the cladding mode. Therefore, the transmission is not affected by leakage of the fundamental mode to the cladding.

The transmission and reflection spectra are obtained by integrating the Poynting vector \( \mathbf{S} \), over a surface \( A \), normal to the waveguide path:

\[
\mathbf{S}(\mathbf{r}, \omega) = \frac{1}{2} \mathbf{E}(\mathbf{r}, \omega) \times \mathbf{H}^*(\mathbf{r}, \omega)
\]

(1)

\[
\mathbf{P} = \text{Re} \left( \int_A \mathbf{S}(\mathbf{r}, \omega) \, dA \right)
\]

(2)

where \( \mathbf{E} \) and \( \mathbf{H} \) are the Fourier transforms of the electric and magnetic field components, respectively. Also, as pointed out before, we use 2D FDTD method to extract all field components (\( E_x, E_y, \) and \( H_z \)) at the output ports of the bend. In Figure 3, the transmission and reflection spectra of the simple waveguide bend are shown by solid line and dashed line, respectively. We define the bandwidth of a bend as the frequency interval in which its transmission is more than 0.9. Using this definition, the simple bend structure has a bandwidth of 6.5% (normalized to the width of PC band gap). The PC bend consists of three parts: two waveguides in the \( \Gamma K \) direction; each having two guided modes and one waveguide in the \( \Gamma M \) direction, having one
guided mode. In a simple bend structure, the coupling of energy between these parts is not efficient and, therefore, the transmission of the bend is low.

3. Improving the bend characteristics

The low transmission problem can be solved by placing small holes in the bend so as to decrease the effective index of the waveguide [7]. This modification results in shifting the waveguide modes at the $\Gamma M$ direction upward in the diagram and hence brings about a better coupling of energy between the $\Gamma M$ and $\Gamma K$ direction waveguides. As a result, the bandwidth is increased and the spectrum is made flatter. Figure 4 shows this new structure of the waveguide bend which is realized by placing three holes in the bend region. The transmission spectra for various radii of the introduced holes $r_d$ are shown in this figure. Table 1a presents the computed bandwidth for different values of $r_d$. As can be seen, the maximum bandwidth which is equal to 22.9% is obtained for $r_d = 0.14a$. Introducing more than three holes in the bend region may produce different results. Assuming $r_d = 0.14a$, Tab. 1b presents the bandwidths computed for bend structures with one, three and five holes in the bend region. As shown in this table, the best structure is obtained for three holes.

---

**Table 1a.** Bandwidth of the bend structure shown in Fig. 3 versus radius of middle holes $r_d$. (The bandwidth is normalized to the width of PC band gap.)

<table>
<thead>
<tr>
<th>$r_d$</th>
<th>Bandwidth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1a</td>
<td>14.2</td>
</tr>
<tr>
<td>0.12a</td>
<td>21.7</td>
</tr>
<tr>
<td>0.14a</td>
<td>22.9</td>
</tr>
<tr>
<td>0.16a</td>
<td>12.3</td>
</tr>
</tbody>
</table>

Fig. 4. The transmission spectra a 60° waveguide bend with three additional holes (of radius $r_d$) at the bend region.
The continuity of field in the bend interface, that is necessary for high transmission, depends on the bend length [5]. In addition, if we consider the bend region as a cavity structure, a longer bend which means a bigger cavity may cause more resonance frequencies in the frequency region of interest. By shifting one defect in the bend a longer bend can be achieved. Taking advantage of both the defect shifting and the index reduction, the structure shown in Fig. 5 is obtained. This structure includes five holes in the middle of the bend. Holes adjacent to the waveguide in the bend have an important role on the distribution of fields. Our investigation shows that increasing the radius of holes can result in increasing the transmission bandwidth. We consider three parameters to modify the structure. These parameters are the number of middle holes \( N \), radius of middle holes \( r_d \), and radius of the holes adjacent to the bend \( r' \).

First, we analyze the transmission spectrum of the bend by changing the number of middle holes, where their radius is 0.14\( a \). Table 2a shows the results of bandwidth computation for this case. It can be seen that \( N = 5 \) is the best choice for which a normalized bandwidth of 22% is achieved.

![Fig. 5. The bend structure which is to modify the transmission characteristic. The air holes with variable radii are shown with a circle around them.](image)

Table 2a. Bandwidth of the bend structure shown in Fig. 5 versus number of middle holes \( N \), with \( r_d = 0.14a \). (Bandwidth is normalized to the width of PC band gap.)

<table>
<thead>
<tr>
<th>( N )</th>
<th>Bandwidth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9.3</td>
</tr>
<tr>
<td>1</td>
<td>13.7</td>
</tr>
<tr>
<td>3</td>
<td>21.4</td>
</tr>
<tr>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>7</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Table 2b. Bandwidth of the bend structure shown in Fig. 3 versus number of holes \( N \), with \( r_d = 0.14a \). (The bandwidth is normalized to the width of PC band gap.)

<table>
<thead>
<tr>
<th>( N )</th>
<th>Bandwidth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.3</td>
</tr>
<tr>
<td>3</td>
<td>22.9</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

The continuity of field in the bend interface, that is necessary for high transmission, depends on the bend length [5]. In addition, if we consider the bend region as a cavity structure, a longer bend which means a bigger cavity may cause more resonance frequencies in the frequency region of interest. By shifting one defect in the bend a longer bend can be achieved. Taking advantage of both the defect shifting and the index reduction, the structure shown in Fig. 5 is obtained. This structure includes five holes in the middle of the bend. Holes adjacent to the waveguide in the bend have an important role on the distribution of fields. Our investigation shows that increasing the radius of holes can result in increasing the transmission bandwidth. We consider three parameters to modify the structure. These parameters are the number of middle holes \( N \), radius of middle holes \( r_d \), and radius of the holes adjacent to the bend \( r' \).
A 60° photonic crystal waveguide bend with improved transmission characteristics

Next, we increase the radius of the holes adjacent to the bend $r'$, which changes the effective index along the $\Gamma M$ direction. The bandwidth of the structure for each $r'$ is computed and shown in Tab. 2b. In this case, we assume that $N = 5$ and $r_d = 0.14a$. It can be seen that a normalized bandwidth of 25.7% is obtained for $r' = 0.32a$.

Finally, we investigate the effect of the radius of middle holes on the frequency characteristic of the proposed PC waveguide bend which has five holes in the middle.

<table>
<thead>
<tr>
<th>$r'$ (a)</th>
<th>Bandwidth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3a</td>
<td>22</td>
</tr>
<tr>
<td>0.31a</td>
<td>24.8</td>
</tr>
<tr>
<td>0.32a</td>
<td>25.7</td>
</tr>
<tr>
<td>0.33a</td>
<td>23.2</td>
</tr>
<tr>
<td>0.34a</td>
<td>22.8</td>
</tr>
<tr>
<td>0.35a</td>
<td>17.5</td>
</tr>
</tbody>
</table>

Table 2b. Bandwidth of the bend structure shown in Fig. 5 versus radius of adjacent holes $r'$; $r_d = 0.14a$, $N = 5$. (Bandwidth is normalized to the width of PC band gap.)

<table>
<thead>
<tr>
<th>$r_d$ (a)</th>
<th>Bandwidth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.11a</td>
<td>12.0</td>
</tr>
<tr>
<td>0.12a</td>
<td>13.1</td>
</tr>
<tr>
<td>0.13a</td>
<td>25.3</td>
</tr>
<tr>
<td>0.14a</td>
<td>25.7</td>
</tr>
<tr>
<td>0.15a</td>
<td>22.3</td>
</tr>
<tr>
<td>0.16a</td>
<td>10.7</td>
</tr>
<tr>
<td>0.2a</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2c. Bandwidth of the bend structure shown in Fig. 5 versus radius of middle holes $r_d$; $N = 5$, $r' = 0.32a$. (Bandwidth is normalized to the width of PC band gap.)

Fig. 6. The transmission and reflection spectra of the PC waveguide bend proposed.
of the bend and the radius of its adjacent holes is increased to 0.32\(a\). As can be seen in Tab. 2c, a good performance (25.7\% bandwidth) is obtained for \(r_d = 0.14a\). Figure 6 illustrates the transmission and reflection spectra for the bend structure proposed, in which five middle holes are included and radii of adjacent holes and middle holes are 0.32\(a\) and 0.14\(a\), respectively.

4. Efficient directional coupler using the bend structure proposed

Directional couplers (DCs) are one of the essential parts of optical integrated circuits [27–29]. They can be used as multiplexers, optical switches or channel drop filters in WDM systems. Photonic crystals can be used to design many optical devices including DCs. Directional couplers are created with two parallel PC waveguides. In this structure, each mode of the separated waveguides splits into two modes. Considering the even mode, it splits into two modes that are known as super-modes, as can be seen in Fig. 7. As shown in this figure, in specific frequency region, super-modes have different wave numbers.

![Fig. 7. Guided modes of two parallel waveguides. Super-modes are shown with dark lines, the solid line represents the even mode and the dashed line represents the odd mode. The gray region is the continuum of extended modes above the air light line and the dash-dotted lines show upper and lower frequencies of the PBG.](image)

The coupling in such a structure depends on the overlap of the field patterns of even and odd super-modes. If the phase difference of these modes is equal to an odd multiple of \(\pi\), power is transferred from one waveguide to the other. This requires that both the even and odd super-modes propagate a specific length which is known as the coupling length [26–28]. Considering a length twice this value, the power is transferred back to the first waveguide. In order to prevent that, we must introduce a waveguide bend after one coupling length to direct light to another path.

We investigate the effect of using the modified 60° waveguide bend, proposed in Section 3, on the DC efficiency. For this structure, in order to have power coupling,
A 60° photonic crystal waveguide bend with improved transmission characteristics

the length should be more than $6a$. In this structure, the phase difference is equal to $\pi$ when the normalized frequency is $0.269 \frac{c}{a}$. Hence, in order to have a channel drop filter for wavelengths about 1550 nm, we choose $a = 417$ nm. The modified and unmodified directional couplers and their drop output spectra are shown in Fig. 8. As depicted in this figure, the proposed bend structure improves the transmission efficiency of directional coupler significantly.

5. Conclusions

In this paper, a 60° waveguide bend in a photonic crystal slab waveguide was designed for TE modes. By applying geometrical modifications in the bend region, the frequency characteristic of the bend was improved in a frequency band where the waveguide is single mode. It was shown that the structure proposed has a significant improvement in its bandwidth characteristics. We used the bend structure proposed to design an efficient directional coupler and analyzed the effect of bend transmission on the coupler efficiency.

Acknowledgement – The authors thank the reviewers for their helpful suggestions.

References


A 60° photonic crystal waveguide bend with improved transmission characteristics


Received August 14, 2008
in revised form November 17, 2008
Microwave properties of the generalized Fibonacci quasi-periodic multilayered photonic band gap structure

YOUSEF TRABELSI, MOUNIR KANZARI*, BAHRI REZIG
Photovoltaic and Semiconductor Materials Laboratory, El-Manar University, Tunis, Tunisia

*Corresponding author: mounir.kanzari@ipeit.rnu.tn

The transmission properties in microwave domains (10 GHz to 40 GHz) of generalized dielectric Fibonacci multilayer generated by the rule \( S_{l+1} = S_m S_{l-1} \) with a pair of positive integers \( m \) and \( n \) were studied. The initial generations of generalized Fibonacci sequence are taken as follows: \( S_0 = L \) and \( S_1 = H \), where \( H \) and \( L \) are two elementary layers with refractive indices \( n_L = 1 \) (air) and \( n_H = 3 \) (ceramic). The so-called “trace map method” was used to simulate the transmission spectra of the multilayer structures at normal incidence. Based on the representation of the transmittance spectra in the microwave range an analysis depending on the pair \((n, m)\) is presented. It has been shown that the reflection bands of the proposed quasi-periodic structure could cover the whole spectral range. By comparison, it is impossible to reach this result by using the periodical multilayer structure.

Keywords: generalized Fibonacci multilayer, photonic crystal, Fibonacci dielectric multilayer, microwave band gap structure.

1. Introduction

A great deal of attention has been devoted to photonic crystals (PC), because of their attractive properties and technological application variety, precisely in the microwave domains. They are an artificial material made from periodic arrays of dielectric or metallic building blocks.

Recently, some new periodic structures, such as photonic and electromagnetic band gap structures, have been applied widely to microwave devices [1, 2]. The existence of photonic band gaps (PBG) has brought about an unprecedented power to control and manipulate the propagation of electromagnetic waves [3–5].

This situation is similar to that in semiconductor crystals where the propagation of electron is forbidden in certain energy regions (band gap). The appearance of the band gap can be explained by the concepts of interference and dielectric potential. PBG structures can be one-, two- or three-dimensional periodic structures. The simplest form of a photonic crystal is the one-dimensional periodic structure such as the Bragg mirror [6].
On the other hand, great interest has been observed as regards the properties and applications of one-dimensional spatially periodic, quasi-periodic and random PBG structures [7]. Quasi-periodic systems can be considered as suitable models for describing the transition from the perfect periodic structure to the random structure [8, 9].

Various studies have been based on multilayer systems built recursively according to the Fibonacci sequence, such as those by MACIA [10], who presents an analysis of wave transmission through Fibonacci dielectric multilayer (FDM) structures and demonstrates that they can be used as reflectors. Some of these works were focused on studying the localization of light waves within Fibonacci quasi-periodic multilayer structures in order to create photonic band gaps similar to those existing in periodic structures and developed the omnidirectional band gap [11–14].

The aim of this work is to study, at normal incidence, the transmission properties in microwave domains [10 GHz, 40 GHz] of the one-dimensional multilayer system built according to the generalized Fibonacci sequence. We calculate transmission spectra through these structures using the trace map method. From the numerical results, it has been found that the transmission bands of the quasi-periodic sequence structures can cover the full spectral range by increasing the parameter \( n \) and fixing \( m \) to 1 or vice versa. In addition, extra multi-narrow bands can be obtained and controlled by adjusting the parameter \( m = 2n \) or \( n = 2m \) from the 3rd generalized Fibonacci sequence. Using the proposed analysis, multi-stop band filters in the microwave spectral domains can be easily designed.

2. Fibonacci model

The generalized Fibonacci sequences are a class of quasi-periodic lattices generated by the substitution rules: \( L \rightarrow H^m \) and \( H \rightarrow H^m L^n \), where \( m \) and \( n \) are all positive integers. They can be generated by a recursive relation [15]:

\[
S_{i+1}^{\text{sr}} = S_{i}^{\text{sr}} S_{i-1}^{\text{sn}}
\]  

(1)

Based on the characteristics of the construction of generalized Fibonacci sequences, we consider the matrices of light propagating through the GF\((m, n)\) multilayer of

![Fig. 1. Fibonacci-class quasi-periodic multilayer stack (\(l = 3, m = 2, n = 2\)).](image)
Microwave properties ... 321

the l-st generation \( S_l \) which is sandwiched by two material media types \( L \) and \( H \). Figure 1 shows the 3-rd generation of one-dimensional generalized Fibonacci class quasi-periodic multilayer stacks for \( m = 2 \) and \( n = 2 \). According to Fibonacci rule, the structure contains 10 layers, as shown in Fig. 1.

3. Transmittance spectra through generalized Fibonacci multilayer

The transmission spectra of electromagnetic radiation through the multilayer periodic and aperiodic systems were widely studied by various methods such as the transfer matrix method [16]. The interesting and representative models of Fibonacci-class (FC(\( n \))) and generalized Fibonacci (GF(\( m, n \))) have been extensively reported by KLAUZER-KRUSZYNA et al. [17, 18], who studied the polarized light propagation through optical generalized Fibonacci superlattices.

We use the trace map method to investigate the transmission spectra through the generalized Fibonacci multilayer. The trace-map technique [19] has proven to be a powerful tool to investigate the properties of various aperiodic systems.

The transfer matrices \( A_l \) used in the trace-map technique are written as [11]:

\[
A_1 = P_{ab} P_b P_{ba} \\
A_2 = P_a \\
A_{l+1} = A_l^m A_{l-1}^n
\]

where \( P_{ab} (P_{ba}) \) stands for the propagation matrix from layer \( a(b) \) to \( b(a) \) and \( P_a (P_b) \) is the propagation matrix through a single layer \( a(b) \). They are given by [19]:

\[
P_{ab} = P_{ba}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_a}{n_b} \end{pmatrix}
\]

\[
P_{a(b)} = \begin{pmatrix} \cos \delta_{a(b)} & -\sin \delta_{a(b)} \\ \sin \delta_{a(b)} & \cos \delta_{a(b)} \end{pmatrix}
\]

where \( \delta_{a(b)} = k n_{a(b)} d_{a(b)} \), \( n_{a(b)} \) is the media refraction index \( a(b) \), \( d_{a(b)} \) are the layer thicknesses and \( k \) the wavenumber in a vacuum. The transmission coefficient is expressed as follows:

\[
T_l = \frac{4}{A_l^2 + 2}
\]

where \( A_l^2 \) is the sum of four element squares of the \( A_l \). Since the transfer matrix is unimodular, we can express the transmission coefficient as:
where \(x_l\) and \(y_l\) denote respectively the trace and anti-trace of the transfer matrix \(A_l\). The transmission coefficient is completely determined by the trace and anti-trace. Thus, a complete description of the transmission through general aperiodic multilayer requires both trace and anti-trace map.

Given a matrix \(A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\), the anti-trace of \(A\) is defined as 
\[
y_A = A_{12} - A_{21}.
\]

In order to study anti-trace maps we need the following identity for two unimodular transfer matrices \(A\) and \(B\) [20]: 
\[
y_{AB} = x_B y_A + x_A y_B - y_{BA}.
\]

In this case, we need to know the \(n\)-th power of a unimodular 2×2 matrix \(A\), which can be written as [20, 21]:
\[
A^n = U_n(x_A) A - U_{n-1}(x_A) I
\]
where \(I\) is the unit matrix, and
\[
U_n(x_A) = \frac{\lambda^+_n - \lambda^-_n}{\lambda_+ - \lambda_-}
\]
\[
\lambda_{\pm} = \frac{x_A \pm \sqrt{x_A^2 - 4}}{2}
\]

Here, \(x_A\) and \(\lambda_{\pm}\) denote the trace and the two eigenvalues of \(A\), respectively. Using Eqs. (2) and (7), we can write the recursion relation of the transfer matrix as:
\[
A_{l+1} = \begin{bmatrix} U_n^{(l)} A_l - U_{n-1}^{(l)} I \\ U_{n-1}^{(l)} A_{l+1} - U_n^{(l)} I \end{bmatrix}
\]

From Equation (6) the trace and anti-trace maps are obtained as:
\[
x_{l+1} = U_{l+1}^{(l)} U_n^{(l+1)} v_l - U_{l-1}^{(l)} U_{n+1}^{(l-1)} - U_{l+1}^{(l)} U_{n-1}^{(l-1)}
\]
\[
v_{l+1} = U_{l+1}^{(l)} U_n^{(l+1)} v_l - U_{l-1}^{(l)} U_{n+1}^{(l-1)} - U_{l+1}^{(l)} U_{n-1}^{(l-1)}
\]
\[
y_{l+1} = U_{l+1}^{(l)} U_n^{(l+1)} w_l - U_{l-1}^{(l)} U_{n+1}^{(l-1)} y_{l-1} - U_{l+1}^{(l)} U_{n-1}^{(l-1)} y_{l-1}
\]
\[
w_{l+1} = U_{l+1}^{(l)} U_n^{(l+1)} w_l - U_{l-1}^{(l)} U_{n+1}^{(l-1)} y_{l-1} + \left[ x_{l+1} - U_{l+1}^{(l)} U_{n-1}^{(l-1)} \right]
\]

where \(v_l = x_{A_l A_{l-1}}\) and \(w_l = y_{A_l A_{l-1}}\).

The roles of \(v_l\) and \(w_l\) are subsidiary. Equations (10) and (11) represent the trace map whereas Eqs. (12) and (13) give the corresponding anti-trace map. We choose
appropriate layer thicknesses \( d_a \) and \( d_b \) to make \( n_a d_a = n_b d_b \). Then, we have \( \delta_a = \delta_b = \delta = (k + 1/2)\pi \), with \( \delta_a \) and \( \delta_b \) being the incident angles of light in layers \( A \) and \( B \), respectively, where \( k \) is a positive integer. The propagation matrices become:

\[
P_{a(b)} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}
\]

Finally, the trace and anti-trace maps are completely determined by Eqs. (10)–(13). So, if we know the initial conditions, the transmission coefficients through general aperiodic multilayer can be determined from the trace and anti-trace maps [22].

4. Results and discussion

4.1. The effect of the \( n \) variation with \( m \) set to 1

In the following numerical investigation, we have chosen air (\( L \)) and ceramic (\( H \)) as two elementary layers, with refractive indices \( n_L = 1 \) and \( n_H = 3 \), respectively. The thicknesses \( d_L, d_H \) of the two materials has been chosen to satisfy the Bragg conditions: \( d_L n_L = d_H n_H = \lambda_0/4 \), where \( \lambda_0 = 12 \text{ mm} \) is the central wavelength. According to these conditions, \( d_H = 1 \text{ mm} \) and \( d_L = 3 \text{ mm} \). We use the trace map method to extract the transmission coefficients in the spectral range from 10 GHz to 40 GHz. We show that the corresponding transmission coefficients display interesting properties. As a result, the reflection bands of the multilayer structures cover the entire spectral range by increasing the parameter \( n \) and setting \( m \) to 1. It is interesting to note that this result is impossible to reach by using the periodical multilayer systems.

The Table gives the width \( \Delta f \) of the pseudo-forbidden gaps and their corresponding central frequencies \( f_i \) for the case of \( l = 3 \), \( m = 1 \), \( n = 20 \). It is clear that the width \( \Delta f_i \) increases with an increase of the corresponding central frequency \( f_i \).

<table>
<thead>
<tr>
<th>( f_i ) [GHz]</th>
<th>10.39</th>
<th>11.16</th>
<th>11.95</th>
<th>13.85</th>
<th>15.19</th>
<th>16.9</th>
<th>18.89</th>
<th>21.27</th>
<th>24.46</th>
<th>28.46</th>
<th>34.97</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta f_i ) [GHz]</td>
<td>0.48</td>
<td>0.5</td>
<td>0.86</td>
<td>0.95</td>
<td>1.14</td>
<td>1.15</td>
<td>1.71</td>
<td>1.9</td>
<td>2.38</td>
<td>3.61</td>
<td>4.94</td>
</tr>
</tbody>
</table>

Figure 2 shows typical transmission coefficients of the generalized Fibonacci-class multilayer stack. Many pseudo-band gaps appear in the spectral domain [10 GHz, 40 GHz] and the width of these pseudo-band gaps can cover 75% of the whole spectral domain. By varying the parameter \( m \), the total width of the forbidden gaps increases and can reach the value of 20 GHz for \( m = 10 \) (Fig. 3). Thus, we note a stacking of pseudo-forbidden gap whose sizes increase gradually with the frequency located between 10 GHz and 40 GHz. When we establish the width of each pseudo-band gap and the corresponding central frequency, multi-stop band filter can be easily considered.
The number of photonic band gaps increases linearly with an increase of the parameter $n$. We can modulate the number of photonic band gaps, $N$, according to the parameter $n$ by the following linear variation: $N = 0.6082n - 0.2868$, where $n$ denotes an integer part of real number $N$. From this approximation we can deduce
Microwave properties ...

the number of photonic band gaps for any given \( n > 3 \) (see Fig. 4). Indeed, we verify that for \( n = 100 \), for example, the number of photonic band gaps is equal to 60 as shown in Fig. 5. This allows us to predict the number of peaks without making calculation which becomes complex for large values of the number \( n \).

4.2. The effect of parameter \( m \) variation with \( n \) set to 1

In the case where the parameter \( m \) varies and \( n \) is set to 1, we have found a large zone with 20.54 GHz for \( m \) equal 10. Comparing this result with that of the corresponding case with \( n \) varying and \( m \) set to 1, we have found that \( n \) must be taken equal to 20 in order to reach the same result. Figure 6 shows the transmission spectra for many \( m \) values for the same iteration \( l = 3 \). In this case, we can cover more than 75% of the spectral domain.

We note that the behaviour of \( N \) relative to the parameter \( m \) is not linear as compared with the case where \( n \) varies and \( m \) is set to 1 (Fig. 7). A good approximation of the \( N \)

---

Fig. 5. Transmittance coefficient for \( m = 1, n = 100 \) for third generation of generalized Fibonacci multilayer.

Fig. 6. Transmission coefficient of generalized Fibonacci multilayer, where \( l = 3, m = 15, n = 1 \) (a), and \( l = 3, m = 20, n = 1 \) (b).
variation versus the parameter \( m \) can be made valuable as follows: \( N = 0.38685m^2 + 1.11625m - 3.23171 \) for \( m > 3 \) with a coincidence of 99.99\%, where \( m \) denotes an integer part of the real number \( N \).

### 4.3. Case of \( m \) and \( n \) variations

To study the transmission properties, two cases were taken: i) \( n = 2m \) with \( m = 5, 6, 8, 10 \) and ii) \( m = 2n \) with \( n = 5, 6, 8, 10 \). Hence, the layer numbers of the whole structure increase by varying \( m \) and \( n \) simultaneously. In all the cases, we note an increase of the forbidden band gaps by increasing the parameters \( m \) or \( n \) and the transmission spectra show a multitude of bands which increase by increasing the parameters \( m \) or \( n \). In addition, we can note that for the case where \( m = 2n \) the alternation of higher transmission values (with lower transmission values) is about 40\%.

Thus, reducing the lower transmission values to zero can lead to a good mirror for the 10–40 GHz spectral range. The opportunity of this work is to consider the case of \( m = 2n \) and trying to reduce the lower transmission values by probably introducing a partial or global defect of the whole system as described elsewhere [11].

With generalized Fibonacci quasi-periodic multilayer structure, we show the existence of several forbidden gaps, all of which increase gradually with the system parameters (Fig. 8). Each forbidden gap represents a multi-narrow stop band. This rejection band is localized around the central frequency. The stacking of the forbidden gap leads to the design of the multi-stop band filters in high frequency. In this case, we can order each filter by knowing the frequency centre and the width of corresponding stop band. We deduce that extra narrow-band filters can be obtained using the multilayer structure studied by the trace map method.

However, this method is applicable only to normal incidence and in the case of not normal incidence we must use, for example, the matrix method (MM) [23]. Indeed in our previous works [23, 24] we show that an omnidirectional high reflector with wide bandwidth was obtained for both \( S \) and \( P \) polarizations for the all incident angles in
5. Conclusions

This work is focused firstly on the transmission properties at normal incidence of the multilayer structures built according to the generalized Fibonacci quasi-periodic multilayer GF(m, n) in the microwave spectral domain (10–40 GHz). According to the proposed method (trace and anti-trace), the transmission spectra through the generalized Fibonacci multilayer structure show a stacking multi-narrow stop band. The number of multi-narrow bands can be controlled by varying the parameters m or n. Based on the analysis proposed, multi-stop band filters can be easily designed. In all the cases, by increasing n or m with a fixed Fibonacci iteration, the number of the photonic band gaps increases. By increasing these photonic band gaps, we can obtain high reflector components working in microwave domains, very interesting for the range 0–90°. So, as an alternative work we expect to use the MM and by varying the incident angle we hope to obtain an omnidirectional mirror which covers the spectral range 10–40 GHz.

Fig. 8. Transmission spectra of generalized Fibonacci multilayer, where: $l = 3; m = 5, 10; n = 2m$ (a); $l = 3; m = 2n; n = 5, 10$ (b).
As an alternative work we expect to study the effect of the incident angle on the transmission properties both for S and P polarizations in order to obtain an omnidirectional mirror in the microwave spectral domain.

References


Received September 3, 2008
in revised form December 19, 2008
Real time multiple planar volume clipping based on the programmable graphics process unit

FEINIU YUAN

School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330013, Jiangxi, China; e-mail: yfn@ustc.edu

We propose a real time volume clipping method which is capable of using several analytical planes for virtual clipping, in order to display internal anatomical structures within volumetric data sets. A single proxy plane is used for computation of the direction of a ray that is cast from the viewpoint. Intersections between the rays and the planes are computed on graphics process unit (GPU). The start and end points for each ray are determined by analyzing relationships with the ray direction, intersections and the normal of planes. Then the volume integral is computed along the ray from the start point to the end point. To obtain immediate visual feedback of clipping effects, we implement translation and rotation of planes on GPU to interactively change the shape of clip object. At last, several experiments were performed on a standard PC with a GeForce FX8600 graphics card. Experimental results show that the method can freely clip and clearly visualize volumetric data sets at real time frame rates.

Keywords: planar volume clipping, graphics process unit (GPU) ray casting, single proxy plane.

1. Introduction

Visualization techniques include surface rendering and volume rendering. In surface rendering techniques, intermediate triangles of iso-surfaces must be extracted from 3D volumetric data sets and then the triangles are rendered using traditional computer graphics hardware. It can achieve interactive frame rates. However, the quality of the rendered images is not high, due to the loss of details during the extraction process. While in volume rendering techniques, extraction of polygons is not required and the volumetric data set is directly rendered according to a transfer function specified by users, so it can produce high quality images. But it is memory and time consuming. Ray casting algorithm is one of the image space volume rendering techniques. It can generate high quality images and often be used in many applications. It can achieve high rendering frame rates on high end workstations and specialized volume rendering hardware (such as VolumePro [1], etc.). However, it is unable to obtain interactive frame rates on the popular PC platform without graphics hardware.

Many acceleration techniques were proposed to speed up the brute force ray casting algorithm. One kind of these techniques is the acceleration technique based on space
leaping, which can avoid many empty resamplings without loss of image quality, such as cylindrical approximation of tubular organs presented by Vilanova et al. [2], spherical approximation of tubular organs proposed by Sharghi and Ricketts [3]. But these methods require a complicated or time-consuming preprocessing. Another kind of these techniques is based on a tradeoff between image quality and rendering speed, such as two-phase perspective ray casting for interactive volume navigation presented by Brady et al. [4], screen and object adaptive sampling, etc. The third kind of acceleration techniques is based on graphics hardware.

With the rapid development of computer game industry, consumer level graphics cards have huge computation performance. There are two typical approaches based on graphics cards, including texture based volume rendering and graphics process unit (GPU) based volume rendering. The texture based volume rendering was originally presented by Cullip and Neumann [5] and further developed by Cabral et al. [6]. The algorithm can directly utilize the texture mapping capabilities of graphics hardware by proxy resampling planes, which can be either axis-aligned [7] with three sets of 2D texture stacks or view-aligned [8] with one 3D texture, as shown in Fig. 1. It can achieve interactive frame rates, but it produces relatively low image quality, especially in the cases of close views. As shown in Fig. 2, we can clearly observe circular artifacts when the viewpoint is located within the human trachea. The quality of rendered image depends mainly on the number of proxy surfaces. If the number of
proxy surfaces increases, the rendering speed becomes reduced. Up to now, most of the modern graphics cards have flexible capabilities of programmability on GPU. These capabilities lead to the rapid development of GPU based volume rendering. And graphics hardware has become preferred technique for the standard implementation of ray casting. Krüger and Westermann [9] proposed a GPU based ray casting. Bentoumi et al. [10] proposed a GPU based shear-warp algorithm. GPU based volume rendering algorithm can generate high quality images at interactive frame rates. GPU can also be used in traditional graphics rendering tasks. Reis et al. [11] presented high-quality rendering of quartic spline surfaces on the GPU. Kim et al. [12] presented vertex transformation streams based on GPU. Flexibility of GPU improves parallel computation performance in many time-critical applications [13, 14].

Transfer function is very important for volume rendering. However, rapid specification of an appropriate transfer function is usually difficult in practice. Therefore, volume clipping becomes an important compensatory tool for difficulty in designing transfer function. Clip planes are frequently used in texture based volume rendering. For instance, Van Gelder and Kim [15] used clip planes to specify the boundaries of the data set in 3D texture-based volume rendering, thus planar volume clipping was implemented for texture based volume rendering. Westermann and Ertl [8] presented a volume clipping method using a stencil buffer. The clip object has to be rendered for each slice to set the stencil buffer correctly. Weiskopf et al. [16] presented clipping techniques based on a volumetric description of clip objects. Clip objects must be first voxelized and represented by a 3D volumetric texture. Diepstraten et al. [17] proposed another depth-based clipping method for depth sorting semi-transparent surfaces. The method is related to virtual pixel maps [18] and dual depth buffers [19]. Tiede et al. [20] used a similar method to visualize attributed volumes by ray casting. Williams et al. [21] presented a volumetric curved planar reformation for virtual endoscopy.

In this paper, we propose a real time visualization method based on GPU ray casting, which is capable of using multiple planes for convex volume clipping. This paper is organized as follows. Section 2 introduces traditional GPU based ray casting. In Section 3, the single proxy plane based GPU ray casting is presented. Section 4 discusses geometrical transformation of clip planes. In Section 5, some experiments are described. At last, conclusions are given.

2. Traditional GPU ray casting

The algorithm presented by Stegmaier et al. [22] is the classical GPU ray casting, within which the data set is stored as a 3D texture to take advantage of the built-in tri-linear interpolation in graphics hardware. In their algorithm, a bounding box for the data set is created, and coordinates of start and end points for each ray are encoded in the color channel of the rendered surfaces of the box, as shown in Figs. 3a and 3b,
respectively. The color of the front surfaces is regarded as the start point for the ray casting process. And the color of the back surfaces is regarded as the end point. The ray direction at any given pixel can be computed by subtracting the color of front surfaces at the pixel from the color of back surfaces in the same position. Two-pass rendering is performed [23], that is, one pass for the front surface and another for the back.

3. Single proxy plane based GPU ray casting

In the GPU ray casting algorithm proposed by Chu et al. [24], which is different from the classic GPU ray casting, only a single proxy plane is used to compute ray equations and ray-plane intersections are computed by analytical geometry on GPU, instead of the two-pass rendering. In their algorithm, six planes of the bounding box are used. We extend six planes of the bounding box to an arbitrary number of planes, and these planes are not required to be parallel to one of $xy$, $yz$ and $xz$ coordinate planes. Therefore, our method is more flexible and more useful for volume clipping.

3.1. Single proxy plane

In the classic GPU based ray casting algorithm, six faces of the bounding box are rendered first to get the end point for ray termination while front face culling is enabled, and six faces of the same box are rendered again to generate the start point for ray casting. In other words, it has six proxy planes for encoding information of each ray, as shown in Fig. 4. Computation of start and end points is automatically completed by rendering six proxy faces twice.

Ray-plane intersections can also be calculated by mathematical analytic geometry on GPU. A ray direction must be given before computing intersections of rays and planes. As shown in Fig. 5, a plane specified by OpenGL’s GL_QUADS function is first rendered to generate a ray direction. But in the Chu’s algorithm [24], the six planes must be parallel to one of $xy$, $yz$ or $xz$ coordinate surfaces. Because the orientation and the number of planes are fixed, it has no clipping function and that limits the technique to be widely applied. Our algorithm allows users to specify arbitrary orientation
and number of planes by adding some actual clip planes, as shown in Fig. 5. And all the original planes of the bounding box and actual clip planes can be regarded as ordinary clip planes.

3.2. Computation of ray directions

Suppose that $M_{mv}$ is the model view matrix consisting of a translation matrix $M_T$ and a rotation matrix $M_R$. $P_w$ and $P_e$ are the same vertex in the world and eye coordinate systems, respectively. We have

$$P_e = M_{mv}P_w = M_TM_RP_w$$

(1)

As shown in Fig. 6, the left-handed camera coordinate system has the same origin, $y$ and $z$ axes as the right-handed eye coordinate system, but its $x$ axis is just contrary to $x$ axis in the eye coordinate system. A fragment vertex $P_e$ in the proxy plane can be
used to generate the normalized direction of the ray in the eye coordinate system for that pixel. The ray direction $D_e$ can be computed by

$$D_e = \text{normalize}(P_{e \cdot xyz})$$

(2)

In Equation (2), we adopt the symbol denotation used by OpenGL shading language. $P_{e \cdot xyz}$ stands for a new vector consisting of the first three components of the 4D homogeneous coordinates.

3.3. Computation of ray-plane intersections

Since the ray direction is computed in the eye coordinate system and its bounding box is defined in the world coordinates, so we must transform the ray direction back to the world system and compute the intersections in the world coordinate system. So, we have

$$D_w = \text{normalize} \left[ (M_k^t P_e) \cdot xyz \right]$$

(3)

According to the analytic geometry, the parameterized equation of a ray is defined as

$$P_{w \cdot xyz} = P_{\text{eye} \cdot xyz} + D_w \cdot t$$

(4)

where $P_w$ is the 4D homogeneous coordinates $(x, y, z, 1)$, $P_{\text{eye}}$ is the eye position (viewpoint), $t$ is the distance from the eye position. According to the analytic

Fig. 7. Determination of start and end points. The viewpoint outside (a) and inside (b) the clipped volume.
geometry, positive \( t \) stands for points having the distance \( t \) along the ray direction \( \mathbf{D}_w \). And negative \( t \) denotes points with the distance \( t \) along the contrary ray direction \(-\mathbf{D}_w\).

A plane is usually defined as

\[
Ax + By + Cz + D = S \cdot P = 0
\]  

(5)

where \((A, B, C)\) is the normalized normal of the plane, \( S = (A, B, C, D) \) and \( P_w = (x, y, z, 1) \) are homogeneous coordinates. Therefore, a plane equation can be defined as the dot product of two vectors \( S \) and \( P_w \). Combining Eqs. (4) and (5), we obtain the intersection between a ray and a plane by

\[
t = \frac{-Ax_{\text{eye}} + By_{\text{eye}} + Cz_{\text{eye}} + D}{Ax_d + By_d + Cz_d} = \frac{-S \cdot P_{\text{eye}}}{S \cdot \mathbf{n} \cdot \mathbf{D}_w}
\]

(6)

where \( P_{\text{eye}} \) is the homogeneous coordinates of the eye \((x_{\text{eye}}, y_{\text{eye}}, z_{\text{eye}}, 1)\) and \( \mathbf{D}_w \) is the 3D vector of the ray direction \((x_d, y_d, z_d)\).

In Equation (6), when the denominator \( d = S \cdot \mathbf{n} \cdot \mathbf{D}_w \) is equal to zero, \( t \) becomes infinite, and it means that there is no intersection between the ray and the plane. In other words, the ray is parallel to the plane.

3.4. Determination of start and end points

In our method, a plane splits the 3D space into two parts which are called saved part and discarded part. We keep the saved part that the normal of the plane points to, whilst the discarded part is deleted. Several planes are often used for convex volume clipping and each ray may have several intersections with clip planes. How to select two points as start and end points from these intersections is a key problem to correctly implement convex clipping.

3.4.1. A ray is parallel to a plane

If a ray is parallel to a plane and the eye is located in the discarded part, current ray should be discarded as there is no intersection. In the case of Fig. 7a, Ray3 is parallel to Plane1 and the eye is in the discarded part of Plane1, so we can discard Ray3 immediately. If a ray is parallel to a plane and the eye is in the saved part, we need to compute possible intersections with other planes. In the case of Fig 7b, because Ray2 is parallel to Plane5 and the eye is in the saved part of Plane5, Ray2 cannot be discarded and continue to be processed to compute intersections with other clip planes.

The criterion which can determine if the eye \( P_{\text{eye}} \) is in the saved part of plane \( S \) is defined as

\[
Ax_{\text{eye}} + By_{\text{eye}} + Cz_{\text{eye}} + D = S \cdot P_{\text{eye}} > 0
\]

(7)

If Equation (7) is satisfied, we think that the eye \( P_{\text{eye}} \) is located in the saved part of plane \( S \). Otherwise, it is in the discarded part. Equation (7) is just the numerator in Equation (6).
3.4.2. Sorting classified intersections

If a ray is not parallel to a plane, there must be an intersection. Ray-plane intersections are classified into two categories. One category is composed of such intersections with an angle less than 90 degrees which the ray and the plane normal form, and these intersections are denoted by parameters $t_n(1), t_n(2), \ldots, t_n(M)$. Another category consists of intersections with an angle greater than 90 degrees, represented as $t_f(1), t_f(2), \ldots, t_f(N)$. The criterion for judging the angle less than 90 degrees is defined as

$$Ax_d + By_d + Cz_d = S_{xyz} \cdot D_w > 0$$  \hspace{1cm} (8)

Equation (8) is just the denominator $d = S_{xyz} D_w$ in Eq. (6). Then we compute the maximum value of parameter $t_n$ in the first category of intersections and the minimum value of parameter $t$ in the second one. The process can be formulated as

$$t_n = \max \left\{ t_n(1), t_n(2), \ldots, t_n(M) \right\}$$  \hspace{1cm} (9)

$$t_f = \min \left\{ t_f(1), t_f(2), \ldots, t_f(N) \right\}$$  \hspace{1cm} (10)

At last, we must decide whether or not the two parameters are valid. If $t_f < t_n$ or $t_f < 0$, the ray has no intersection with the convex clipped geometry. In other words, the two parameters are invalid. When the eye is inside the clipped volume, the parameter $t_n$ for the start point is usually less than 0. In this case, the parameter $t_n$ is set to 0, in order to ensure that the start point is not behind the eye along the ray direction. After the aforementioned processing, $t_n$ and $t_f$ are the parameters for the valid start and end points, respectively. Now, we can trace each ray to compute volume integral from the start point to the end point on the GPU. As shown in Fig. 7a, as for Ray1, the start and end points are determined by parameters $t_2$ and $t_3$, respectively. Figure 7b illustrates the eye inside the clipped volume. We obtain the parameters for start and end points which are $t_2$ and $t_3$, respectively. But $t_3$ is less than 0, so the start point parameter is modified to 0. So, the parameters for start and end points are $t_6$ and $t_3$, respectively.

Figure 8 lists GPU codes for computing start and end points. In our implementation, vector operations are widely applied to take full advantage of intrinsic parallel hardware computation contained on GPU. We pass all the parameters of clip planes from CPU to GPU using uniform variables defined by OpenGL Shading Language. The uniform qualifier is used to declare global variables whose values are the same across the entire primitive being processed. All uniform variables are read-only and are initialized externally either at link time or through the API. The uniform variable
vec3 Dw=normalize(FragVertex.xyz);
int i=0;
tn=-1000.0;
tf=1000.0;
while(i<ClipPlaneNum)
{/*Compute intersections between rays and planes S[i]*/
   n=dot(S[i], Peye);
   d=dot(S[i].xyz, Dw);
   if(abs(d)<=-0.00001)
   {/*A ray is parallel to a plane*/
      if(n<=0.0)
         { /*The eye is located in the discard part of plane S[i]*/
            NoIter=1;//Discard this ray
            break;//Stop searching
         }
   }
   else
     {/*A ray is not parallel to a plane
        t=-n/d;
        if(d>0.0)
           { //an angle less than 90 degrees
              //calculate maximum values for first category
              tn=max(tn,t);
           }
        else
           { //Calculate minimum values for second category
              tf=min(tf,t);
           }
    }
    i++;
}
if(tf<tn||tf<0.0)
NoIter=1;// No intersection
else
  if(tn<0.0)
     { /*Set the start point to the eye when the eye inside the clipped volume.*/
        tn=0.0;
     }
Fig. 8. GPU codes for computing start and end points.

S specifies the parameters of clip planes and the uniform variable ClipPlaneNum stores the number of clip planes. The detailed declarations are listed as follows: uniform vec4 S[64]; uniform int ClipPlaneNum.

In our implementation, the number of clip planes varies from 6 to 64. We can see that the flexibility of our method is obvious.

4. Translation and rotation of clip plane

Translation and rotation of clip planes can change the shape of convex clip objects. According to the definition of a plane [25], the parameter $D$ is just a negative
distance to the origin from the plane. As shown in Fig. 9, the translation vector is \( \mathbf{T} \), so the equation of moved plane becomes

\[
Ax + By + Cz + D' = 0
\]  

(11)

where

\[
D' = D - A \cdot \mathbf{T} \cdot x - B \cdot \mathbf{T} \cdot y - C \cdot \mathbf{T} \cdot z
\]  

(12)

Rotation of a plane is more complex than translation. To implement plane rotation, the standard equation of a plane in 3D space (Eq. (5)) is converted to the point-normal equation. The point-normal equation of a plane is defined as

\[
A(x - x_0) + B(y - y_0) + C(z - z_0) = 0
\]  

(13)

where \((A, B, C)\) is the normalized normal of the plane, \((x_0, y_0, z_0)\) is a point \( \mathbf{P}_0 \) on the plane. For convenience, \( \mathbf{P}_0 \) is defined as an intersection between the plane and a ray which is cast from the origin and has the normalized direction \( \mathbf{N} = (A, B, C, 1) \). So, the ray parameterized equation is

\[
\mathbf{P}_{\mathbf{x}yz} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} A \\ B \\ C \end{bmatrix} \cdot t
\]  

(14)

Combining Eqs. (5) and (14), we can obtain the parameter of intersection \( t = -D \). So, the point \( \mathbf{P}_0 \) is equal to

\[
\mathbf{P}_{0, \mathbf{x}yz} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} A \cdot D \\ B \cdot D \\ C \cdot D \end{bmatrix}
\]  

(15)

Now, we rotate the plane normal around the point \( \mathbf{P}_0 \). As shown in Fig. 10, given a rotation matrix \( \mathbf{M}_R \), the new normal \( \mathbf{N}' = (A', B', C', 1) \) is computed by

\[
\mathbf{N}' = \mathbf{M}_R \mathbf{N}
\]  

(16)

Because the standard equation is used for computing ray-plane intersections on GPU, the point-normal equation should be converted back to the standard equation. Combining Eqs. (13), (14) and (16), we have

\[
D' = -\mathbf{P}_{0, \mathbf{x}yz} \cdot \mathbf{N}', \mathbf{x}yz
\]  

(17)

Thus, we obtain the standard equation of the rotated plane as follows

\[
A'x + B'y + C'z + D' = 0
\]  

(18)

where \( A' = \mathbf{N}', x, B' = \mathbf{N}', y, C' = \mathbf{N}', z \).
Real time multiple planar volume clipping...

Figure 11 gives experimental results of our convex volume clipping by interactively translating and rotating two clip planes.

5. Experiments

We implemented the real time GPU ray casting with our volume clipping techniques using Visual C++ and OpenGL Shading Language, and then several experiments were performed on a standard PC with an Intel E2160 dual core 1.8 GHz processor with a 2 Gbyte RAM. And a GeForce FX8600 graphics card was installed.

First, we test our methods on a volumetric data set with the viewpoint outside the clipped volume. The data set was acquired by a CT scanner from a human in
a hospital. The data set is composed of 112 DICOM images with 2D cross-section image size 512×512, so the 3D data set size is 512×512×112. An iso-surface ray casting is used for rendering. As shown in Fig. 12, we can find the volume to be correctly clipped and the rendered images to have high quality. We can interactively alter the transfer function by simply re-generating the 2D lookup texture or directly specify different ambient, diffuse and specular colors for the illumination model. So, our algorithm is very convenient in cases where users need to frequently modify material optical property.

Second, we navigate into the human trachea of the volumetric data set, in order to make the viewpoint located inside the volume. As shown in Fig. 13, we can find that the volume is also correctly clipped and the rendered images have high quality. Our method can also process RGB color volumetric data sets by modifying the ray
casting algorithm and replacing the color transfer function with direct resamplings of RGB volumes. Figure 14 shows the resulting images with our method for the visible human color data set using three clip planes.

The size of data sets does not influence the rendering speed with our method, as long as the data set can be fully loaded into video memory of graphics card. But the size of viewport obviously influences rendering speed. The number of clip planes has a little influence on the rendering speed. The Table illustrates rendering speed under different viewport sizes and different number of clip planes. Experimental results show that our method can render volumetric data sets with an iso-surface ray casting in real time. The method provides users with immediate visual feedback.

6. Conclusions

We propose a real time clipping method based on GPU ray casting, which is capable of using multiple planes for convex volume clipping. The traditional GPU based ray casting methods often use two-pass rendering of front and back proxy surfaces. Different from the classic GPU based ray casting, one proxy plane is used and ray-plane intersections are computed on GPU. Fragment coordinates encode the direction of a ray that is cast from the viewpoint. And we presented an approach to compute intersections between the ray and clip planes, and then the start and end points of the ray are thus obtained by analyzing relationships between the ray direction, intersections and the normal of each plane. Then the volume integral is computed along the ray from the start point to the end point. To obtain immediate visual feedback of volume clipping effects, we implement translation and rotation of planes on GPU to interactively change the shape of clip object. This implementation of GPU based ray casting has good rendering performance and clipping capability. At last, several experiments were performed on a standard PC with an Intel dual core 1.8 GHz processor and a GeForce FX8600 graphics card. Experimental results show that the method can clip and visualize volumetric data sets clearly at real time frame rates.

Acknowledgements–This project was supported by the Natural Science Foundation of Jiangxi Province (2007GQ0076), Foundation of Education Department of Jiangxi Province (GJJ09297) and Key Project of Jiangxi University of Finance and Economics.

<table>
<thead>
<tr>
<th>Number of clip planes</th>
<th>Viewport size</th>
<th>Frame rates [fps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>512×512</td>
<td>59.4</td>
</tr>
<tr>
<td>9</td>
<td>512×512</td>
<td>58.6</td>
</tr>
<tr>
<td>12</td>
<td>512×512</td>
<td>57.9</td>
</tr>
<tr>
<td>12</td>
<td>600×600</td>
<td>39.7</td>
</tr>
<tr>
<td>12</td>
<td>800×800</td>
<td>22.3</td>
</tr>
</tbody>
</table>

Real time multiple planar volume clipping ...
References


Received October 5, 2008
in revised form December 11, 2008
An image fusion algorithm based on polyharmonic local sine transform (PHLST)

LIU SHANGZHENG¹, HAN JIUQIANG¹, BOWEN LIU², ZHANG XINMAN¹*

¹School of Electronic and Information Engineering, Xi’an Jiaotong University, Xi’an, 710049, China
²Business Intelligence Lab, University of Science and Technology of China (USTC), Hefei, 230026, China

*Corresponding author: ccp9999@sina.com

In this paper, we propose a novel image fusion algorithm based on polyharmonic local sine transform (PHLST). First, we apply PHLST to source image to decompose it into two components: polynomial $p$ and residual $r$. Using the Laplace/Poisson equation solver, we obtain polynomial $p$. Subtracting $p$ from original image, we acquire $r$. In order to reduce noise, $r$ is filtered in frequency domain. Next, we fuse $p$ and $r$ separately. Then we add the composite $p$ and composite $r$ directly to obtain the fused image. Experiments demonstrate outstanding performance of the method proposed.

Keywords: image fusion, performance metric, polyharmonic local sine transform (PHLST).

1. Introduction

The objective of image fusion is to integrate complementary information from multiple sources of the same scene so that the composite image is more suitable for human visual and machine perception or further image-processing tasks. With the availability of multiple image sources in many fields such as remote sensing, medical imaging, machine vision, and military applications, image fusion has emerged as a new and promising research area. Much effort has been devoted to image fusion techniques in recent years.

A set of two or more source images is obtained of a given scene viewed with different sensors or under different imaging conditions. These source images may have complementary information. How to merge this complementary information into a single image without losing vital information and introducing false information is the core of image fusion. A number of algorithms based on multiscale analysis have been proposed [1–5]. However, these methods were unsatisfactory. The brute-force periodization will cause mismatch when we deal with non-periodic signals. This will lead to the Gibbs phenomenon [6]. Wavelets and wavelet packets essentially use brute-
-force periodization so that it again creates the discontinuity at the end points, and this produces large wavelet/wavelet packet coefficients. The large coefficients will introduce false information in the composite image. And the wavelets are not efficient for textured images [6]. In this study, we suggest a fusion algorithm based on polyharmonic local sine transform (PHLST). PHLST decomposes each source image into two components: the “base” \( p \) and the “texture” \( r \) of original image. Next, we merge \( p \)'s and \( r \)'s, respectively, and obtain composite \( p \) and composite \( r \). Then, we add \( p \) and \( r \) directly to obtain fused image. The algorithm proposed does not introduce the brute-force periodization so that it avoids the disadvantages of algorithms based on wavelet/wavelet packets.

The paper is organized as follows. In Section 2, we give a brief discussion of PHLST and a brief review of Laplace/Poisson equation solver proposed by Averbuch, Braverman, and Vozovoi. Section 3 presents the image fusion scheme. In our approach, we apply different fusion rules for \( p \) and \( r \), respectively. Preliminary results of the fusion method proposed are presented and discussed in Section 4, and the paper is concluded in Section 5.

2. Polyharmonic local sine transform

The PHLST was first introduced as a model for image compression. We will briefly explain polyharmonic local sine transform. A more detailed description of polyharmonic local transform may be found in [6].

Assume \( I(x, y) \) is a spatial-domain image. The main idea of PHLST is that an image \( I(x, y) \) can be divided into two parts: \( p \) which we call the polyharmonic component of \( I(x, y) \) and \( r \) which we call the residual of \( I(x, y) \). \( P \) is a polynomial. \( R \) is a geometric series. \( P \) represents “base” or “trend” or “predictable” part of the original image, whereas \( r \) stands for “texture” or “fluctuation” or “unpredictable” part of the original image. This method coincides with the characteristic of human visual system. Human beings first focus on the noticeable parts of an image. The noticeable parts are the fluctuation of an image. So, we extract texture which is in favor for subsequent manipulation.

\( I(x, y) \) is an rectangular image. Let \( I_{\text{inter}} \) be the interior of \( I(x, y) \), \( I_{\text{bou}} \) be the boundary of \( I(x, y) \). For simplicity, \( 0 \leq x \leq 1, 0 \leq y \leq 1 \). By solving polyharmonic equation (1) with given boundary conditions (2), we can obtain the polyharmonic component

\[
\Delta^n p = 0 \quad \text{in} \ I_{\text{inter}}, \quad n = 1, 2, \ldots \tag{1}
\]

\[
\frac{\partial^k I}{\partial n^{k_l}} = \frac{\partial^k p}{\partial n^{k_l}} \quad \text{on} \ I_{\text{bou}}, \quad l = 0, \ldots, m-1 \tag{2}
\]

where \( k_l = 2l \), the even order normal derivatives. We need not to consider the odd order normal derivatives because this is automatically guaranteed [6]. The \( k_0 = 0 \), which
means that \( p = f(x, y) \) on the boundary. These boundary values and normal derivatives ensure the function values and the normal derivatives of orders \( k_1, \ldots, k_{n-1} \) of \( p \) along the boundary to match those of the original image \( I(x, y) \) over there.

For \( n = 1 \), we obtain the following Laplace equation with the Dirichlet boundary condition:

\[
\begin{align*}
\Delta p &= 0 \quad \text{in } \Omega_{\text{inter}} \\
\partial_{\mathbf{n}}^2 p &= I(x, y) \quad \text{on } \Omega_{\text{bou}}
\end{align*}
\]  
\hspace{1cm} (3)

For \( n = 2 \), Eq. (1) becomes biharmonic equation with the mixed boundary condition:

\[
\begin{align*}
\Delta^2 p &= 0 \quad \text{in } \Omega_{\text{inter}} \\
\partial_{\mathbf{n}}^2 p &= \frac{\partial^2 I}{\partial n^2} \quad \text{on } \Omega_{\text{bou}}
\end{align*}
\]  
\hspace{1cm} (4)

We use the Laplace/Poisson equation solver proposed by Averbuch et al. [7, 8] to solve Eqs. (3) and (4). The ABIV method provides more accurate solutions than those based on the finite difference (FD).

There are several versions of the ABIV method. We choose the simplest and most practical one to solve (3) that does not need to estimate any derivative. It follows the recipe

\[
p(x, y) = p_1(x, y) + \sum_{k \geq 1} \left\{ p^{(1)}_{2k} g_k(x, 1 - y) + p^{(2)}_{2k} g_k(y, 1 - x) + p^{(3)}_{2k} g_k(x, y) + p^{(4)}_{2k} g_k(y, x) \right\}
\]  
\hspace{1cm} (5)

where \( p_1(x, y) \) is a harmonic polynomial that matches \( I(x, y) \) at the four corner points of the image. And its simplest form is:

\[
p_1(x, y) = a_3 xy + a_2 x + a_1 y + a_0
\]  
\hspace{1cm} (6)

Let \( p_1(0, 0) = I(0, 0), p_1(0, 1) = I(0, 1), p_1(1, 0) = I(1, 0), p_1(1, 1) = I(1, 1) \), we have

\[
\begin{align*}
I(0, 0) &= a_0 \\
I(0, 1) &= a_1 + a_0 \\
I(1, 0) &= a_2 + a_0 \\
I(1, 1) &= a_3 + a_2 + a_1 + a_0
\end{align*}
\]  
\hspace{1cm} (7)
By solving (7), we can easily obtain the parameters $a_i$. The function $g_k(x, y)$ is defined as follows:

$$g_k(x, y) = \frac{\sin(\pi k x) \sinh(\pi k y)}{\sinh(\pi k)}$$

and $p_{2k}(i), i = 1, 2, 3, 4,$ are the $k$-th 1D Fourier sine coefficients of boundary functions $I(x, 0) - p_1(x, 0), I(0, y) - p_1(0, y), I(x, 1) - p_1(x, 1),$ and $I(1, y) - p_1(1, y)$, respectively, where $0 \leq x \leq 1, 0 \leq y \leq 1$. Subtracting $p(x, y)$ from $I(x, y)$, we obtain $r(x, y)$. It can be written as:

$$r(x, y) = \sum_{i \geq 1} \sum_{j \geq 1} s_{ij} \sin(i \pi x) \sin(j \pi y)$$

where $s_{ij}$ is the 2D Fourier sine coefficients of $r(x, y)$.

For a more precise approximation of $I(x, y)$, we can segment an image $I(x, y)$ into a set of rectangular blocks (of different sizes possible) using the characteristic function. There is no overlap between adjacent patches, but adjacent patches may share the boundaries. Then, we decompose each patch into two components: the polyharmonic component $p$ and the residual $r$, according to the foregoing method.

3. The image fusion scheme

Figure 1 shows a schematic diagram of the basic structure of the image fusion scheme proposed. For simplicity, we make an assumption that there are just two source images, $I_1$ and $I_2$, and the fused image is $F$. We note that all the methods described in this paper can also be extended to cases with more than two source images. An important preprocessing step in image fusion is image registration. Image registration ensures
that the information from each sensor is referred to the same physical structures in
the environment. In this study, we assume that the images to be combined are already
perfectly registered.

3.1. Fusion rules

The objective of image fusion is to combine multiple source images of the same scene
and obtain a better quality image. The straightforward approach to image fusion is to
compute the pixel by pixel average of the input images. Although image averaging is
a simple method, a major drawback is that it can cause a decreased image contrast. To
avoid a loss of detail, the basic strategy here is to fuse \( p \) and \( r \) separately to construct
a fused PHLST representation from the PHLST representations of the original data.

\( P \) represents “base” of the original image. We use the simplest method to compute
\( p \) averaging.

\( R \) represents the “detail” or “texture” of the source image. The larger values
in \( r \) correspond to the sharper brightness changes and thus to the salient features in
the image, such as edges, lines, and region boundaries. Therefore, a good integration
rule is to conserve \( r \) of the two source images at each point. So, we compute the composite
\( r \) by the following equation

\[
  r_F = (r_{I_1} + r_{I_2}) \omega
\]  

where \( r_{I_1} \) and \( r_{I_2} \) represent \( r \)'s, from \( I_1 \) and \( I_2 \), respectively, \( r_F \) is the composite \( r \).

Subsequently, a composite image is constructed by performing an inverse PHLST.
Since the PHLST provides spatial localization, the effect of the direct summing fusion
rule can be illustrated in the following two aspects. If the same object appears more
distinctly (in other words, with better contrast) in image \( I_1 \) than in image \( I_2 \), after fusion
the object in image \( I_1 \) and in image \( I_2 \) will be preserved with better contrast than in \( I_2 \);
in a different scenario, suppose an object appears in the image \( I_1 \), while being absent
in image \( I_2 \), after fusion the object in image \( I_1 \) will be preserved and the contrast of
the composite image will be enhanced.

3.2. Performance measures

Performance measures are essential to determine the possible benefits of fusion as well
as to compare results. Evaluation is usually performed through robust yet impractical
subjective trials [9, 10], which is time consuming. Computational objective fusion
metrics are an efficient alternative as they need no display equipment or complex
organization of an audience. Recent proliferation of image fusion algorithms has
prompted the development of reliable and objective ways of evaluating and comparing
their performance for any given application [11–15]. Four metrics are considered in
this study, which do not require ground truth images. These are: i) entropy to evaluate
the information contained in the fused image; ii) \( Q_p \) as defined by Petrović and
XYDEAS [12], we use the same parameter as well; iii) mutual information (MI) proposed by Qu et al. [14, 16]; iv) $Q_M$, $Q_D$ proposed by Hao Chen and Varshney [17].

4. Experimental results

In this section, we verify the significant performance of the image fusion method proposed by comparing it with three different image fusion methods using four image fusion metrics. The first algorithm is a Laplace pyramid fusion algorithm [18, 19], where the input images are decomposed using a Laplace pyramid decomposition and the fused image is reconstructed by averaging the low resolution components and selecting the coefficients with the largest amplitude for the high resolution coefficients. The second fusion algorithm used in this study is a shift invariant wavelet fusion algorithm [2], where the source images are decomposed using Harr wavelet filter, the coefficients of the integrated image are computed by choosing the corresponding

Table 1. The fused image measurements of medical images.

<table>
<thead>
<tr>
<th>Method</th>
<th>$Q_D$</th>
<th>$Q_M$</th>
<th>$E$</th>
<th>MI</th>
<th>$Q_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laplace</td>
<td>60.718</td>
<td>40.059</td>
<td>4.286</td>
<td>1.926</td>
<td>0.5551</td>
</tr>
<tr>
<td>Wavelet</td>
<td>60.605</td>
<td>39.847</td>
<td>4.242</td>
<td>2.021</td>
<td>0.5371</td>
</tr>
<tr>
<td>NSCT</td>
<td>62.73</td>
<td>41.127</td>
<td>4.305</td>
<td>1.824</td>
<td>0.5289</td>
</tr>
<tr>
<td>PHLST</td>
<td>56.114</td>
<td>36.87</td>
<td>4.449</td>
<td>3.570</td>
<td>0.5919</td>
</tr>
</tbody>
</table>

Fig. 2. Testing image set 1: medical images.
coefficients of input images with the largest amplitude in high frequency bands and by averaging the coefficients of base band. The third fusion algorithm is a non-subsampled contourlet fusion (NSCT) algorithm [20, 21], the fusion rule is the same as the one in wavelet fusion algorithm. The number of decomposition levels is three in all the methods.

4.1. Experimental results of medical imagery CT and MRI

In Figures 2–4: a and b are the input images, c is the fused image using Laplace method, d is the fused image using shift invariant wavelet method, e is the fused image using NSCT method, f is the fused image using PHLST method. The quantitative assessments of fused images are listed in Tab. 1. From this table, we can see that the performance of the algorithm proposed is best according to all metrics. The fused images are illustrated in Figs. 2c–2f.

4.2. Experimental results of visible and IR images

The image quality evaluation results of the fused images by the four methods are given in Tab. 2. The fused images are shown in Figs. 3c–3f. The performance of

<table>
<thead>
<tr>
<th>Method</th>
<th>(Q_D)</th>
<th>(Q_M)</th>
<th>(E)</th>
<th>MI</th>
<th>(Q_p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laplace</td>
<td>33.7695</td>
<td>56.6074</td>
<td>4.5548</td>
<td>1.0301</td>
<td>0.6211</td>
</tr>
<tr>
<td>Wavelet</td>
<td>31.8273</td>
<td>56.6074</td>
<td>4.4736</td>
<td>1.037</td>
<td>0.6162</td>
</tr>
<tr>
<td>NSCT</td>
<td>30.9515</td>
<td>51.8811</td>
<td>4.4189</td>
<td>1.04</td>
<td>0.6110</td>
</tr>
<tr>
<td>PHLST</td>
<td>31.1143</td>
<td>52.1593</td>
<td>4.7058</td>
<td>1.1759</td>
<td>0.6561</td>
</tr>
</tbody>
</table>

Fig. 3. Testing image set 2: visible and IR image.
the algorithm proposed is best according to metric $E$, MI, and $Q_p$. Judging by $Q_D$ and $Q_M$, the best method is NSCT. PHLST ranks as the second, but the differences of the two methods are very small, with only $(31.1143 - 30.9515)/30.9515 = 0.53\%$, $(52.1593 - 51.8811)/51.8811 = 0.54\%$.

4.3. Experimental results of SAR images

The quantitative assessments of fused images are listed in Tab. 3. The fused images are given by Figs. 4c–4f. The performance of the algorithm proposed is best according to metrics $E$, MI, and $Q_D$, $Q_M$. Judging by $Q_p$, the performance of the algorithm proposed ranks as the third, but the difference of the first and the third is very small, with only $(0.8806 - 0.8723)/1=0.83\%$. Notice that through visual inspection, it is fairly difficult to discriminate between the four fused images.

<table>
<thead>
<tr>
<th>Method</th>
<th>$Q_D$</th>
<th>$Q_M$</th>
<th>$E$</th>
<th>MI</th>
<th>$Q_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laplace</td>
<td>112.22</td>
<td>74.213</td>
<td>5.0881</td>
<td>1.8127</td>
<td>0.8704</td>
</tr>
<tr>
<td>Wavelet</td>
<td>104.27</td>
<td>68.912</td>
<td>5.0193</td>
<td>1.8818</td>
<td>0.8784</td>
</tr>
<tr>
<td>NSCT</td>
<td>99.199</td>
<td>65.668</td>
<td>5.0171</td>
<td>1.8780</td>
<td>0.8806</td>
</tr>
<tr>
<td>PHLST</td>
<td>95.915</td>
<td>63.457</td>
<td>5.1377</td>
<td>1.9006</td>
<td>0.8723</td>
</tr>
</tbody>
</table>

Fig. 4. Testing image set 3: SAR images.
5. Conclusions

We have presented in this paper a new approach to multisensor image pixel data fusion using PHLST. We demonstrate image fusion using images with different scene information from the different types of sensor. These examples show that the image fusion method is capable of extracting the important information from the source images and placing it in the fused image. Experimental results demonstrate the significant performance of the algorithm proposed.

Acknowledgements – This work has been supported by the National Natural Science Foundation of China under Project 60602025. The original CT and MRI images were kindly supplied by Dr. Oliver Rockinger, the visible and IR images by Alexander Toet of the TNO Human Factors Research Institute, the remote sensing images by Dr. Vladimir Petrovic of Manchester University. The wavelet and Laplace fusion codes are from The Image Fusion Toolbox for Matlab developed by Oliver Rockinger, Metapix, the nonsubsampled contourlet fusion code is from Nonsubsampled Contourlet Toolbox provided by Jianping Zhou et al. These images and codes are available online at www.imagefusion.org.

References


Received April 10, 2008
The feedback stability research of HDRI system

ZHÉBO CHÉN¹, ², CHÉN SHÉN³, JÍN XÚ¹, ², XÚXIÀNG NÍ¹, ², ZÜKĀNG LÜ¹, ²

¹State Key Lab of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027, China
²NERC for Optical Instrument, Zhejiang University, Hangzhou 310027, China
³West Branch of Zhejiang University of Technology, Quzhou 324000, China
*Corresponding author: chenzb@zju.edu.cn

In this paper, one kind of high dynamic range imaging (HDRI) system is analyzed and the feedback stability is optimized. In this system, space light modulator (SLM) is used to modulate the input illuminance with the feedback signals. Because of the illuminance uncertainty of the scene, the feedback may take too long or turn into oscillations. To acquire the optimized feedback configuration, PID theory is used to analyze the feedback process. After PID parameter is obtained, simulations are applied to study the parameters. The optimized value range and principle of choice for the feedback control are discussed. Lastly, imaging experiments are conducted to obtain high dynamic range images, and the results prove the validity of PID parameters.

Keywords: high dynamic range imaging (HDRI), PID, feedback, imaging.

1. Introduction

Many kinds of HDRI methods have been proposed by now, such as multi-exposure method [1, 2], spatially varying exposure method [3], logarithmic compression sensor [4], and spatial light modulating method [5, 6]. In this paper, the image stability problem of the SLM method is discussed.

This kind of HDRI system contains three parts [7]: optical lenses, SLM, and image sensor. The structure of this system is shown in Fig. 1.

In this HDRI system, a liquid crystal on silicon (LCoS) is used as spatial light modulator (SLM, Figs. 1–5). A polarized beam splitter (PBS) prism is used to split the beam. And a CCD is used as the image sensor (Figs. 1–3). There are two sets of lenses. The LCoS is located at the image plane of the first set of lenses and the object plane of the second set of lenses. The CCD is located on the image plane of the second lenses so that the LCoS and the CCD are on two conjugate planes.

The LCoS acts as a negative feedback device to extend the dynamic range of the image. It is an approximate one-to-one modulator. With fine adjusting processes,
the image on the LCoS can be projected onto the CCD plane with the same size of the effective imaging area on CCD. The LCoS and the CCD are synchronized with the same refreshing rate. The whole system is a negative feedback system.

Due to the negative feedback, LCoS will be less reflective for the high illuminance area and be more reflective for the low illuminance area. Hence, a wide range of illuminance can be imaged and high dynamic imaging range is achieved. But how to realize this negative feedback is a very important problem.

2. System analysis

For the multiformity of the natural scene, changes of an object cannot be predicted properly during the dynamic imaging process. PID control is suitable for dealing with such a problem with unknown system parameters. PID algorithm has many advantages, such as simplicity, stable reliability, and robustness [8].

2.1. Feedback process in system

Figure 2 shows a schematic diagram of the feedback process. Firstly, the output signal is transferred from the image sensor to the digital image process unit. Then, the signal is sent to the SLM unit for feedback. SLM unit modulates the input illuminance

Fig. 1. Sketch of HDRI imaging system.

Fig. 2. Schematic diagram of feedback process.
according to those feedback signals. Hence, a closed-loop feedback system is constructed.

Feedback image can be calculated with the formula:

\[ u(n) = F\left[ g(n), g(n-1), g(n-2) \ldots \right] \quad (1) \]

Here, \( g(n) \) is the output signal of the image sensor, \( n \) is the frame number of the images. The key issue is to find the feedback function \( F \).

### 2.2. PID control

Since the whole HDRI system is a closed-loop control system, PID control theory can be used to find the feedback function \( F \).

With PID theory, \( u(t) \) is supposed to be composed of proportional, integral, and derivative parts. Since all the units mentioned in this system are discrete devices, so \( u(t) \) can be expressed as:

\[ u(t) = K_p \sum_{i=0}^{n} e_i T + T_d \frac{e_n - e_{n-1}}{T} \quad (2) \]

Take the efficiency and cost into account, only the current frame and previous frame images from the image sensor are used to calculate \( u(t) \), so formula (2) can be simplified:

\[ u(t) = k_p g(n) + k_i g(n) + k_d g(n) \quad (3) \]

Assume \( k_a = k_p + k_i + k_d, k_b = k_i - k_d \), then:

\[ u(t) = k_a g(n) + k_b g(n-1) \quad (4) \]

After simplification, there are only two uncertain parameters left: \( k_a \) and \( k_b \).

### 2.3. Determining the parameters

Because the system oscillations are unavoidable, what we can do is to choose the best feedback parameter to balance the imaging speed and image quality. So, our target is find a suitable parameter which leads to the minimum feedback times.

Since output image should be stable, and the input data \( g(n) \) and output \( f(n) \) are all 8-bit data, \( f(n) \in [0, 255] \), the following formula should be satisfied:

\[ k_a + k_b = 1 \quad (6) \]

By now, only one parameter \( k_a \) needs to be determined. Data simulations are taken to study this parameter.
Data simulations are carried out with Matlab. For the convenience of calculation, simplified models for SLM and CCD sensor are used.

The simplified model for SLM is:

$$S(l, t) = \frac{l}{u(t)}$$

(7)

$u(t)$ is the control signal at time $t$, which is just the feedback image, $u(t) \in [0, 255]$. $l$ is the input relative illuminance, $l \in [1000, 62000]$. Relative illuminance here means the relative quantity of real-world illuminance; it belongs to an integer type without any unit.

The simplified model for CCD sensor is:

$$g(l, t) = \begin{cases} 
0 & S \leq 0 \\
S(l, t) & 0 < S < 255 \\
255 & S \geq 255 
\end{cases}$$

(8)

According to the simplified PID formula discussed before, the control signal of SLM can be calculated as follows:

$$u(t) = k_a \cdot g(l, t - 1) + (1 - k_a) \cdot g(l, t)$$

(9)

When $|g(l, t - 1) - g(l, t)| \leq 1$, we think the output image is stable, and record the CCD sensor imaging times, that is, the feedback times for stable output.

Changing $k_a$ from 0.1 to 1, with each $k_a$ value, a feedback times value can be calculated. Plot the $k_a$ values and the corresponding feedback times values, one curve is obtained. Change the relative illuminance $l$ from 1000 to 62000; at last 62 curves are plotted in Fig. 3.

It seems that the feedback times before reaching stable status have a U shape relationship with $k_a$. The minimum feedback time appears when $k_a \approx 0.5$. When $k_a$ equals some other value, the feedback times rise very quickly, almost as much as twenty times the minimum value.

![Fig. 3. Simulation results used to determine the parameter.](image-url)
To find out more precise influence of \( k_a \) to the feedback process, make \( k_a \) consecutively equal to 0.4, 0.5 and 0.6. A step signal showed in Fig. 4 is used as the input signal to carry out the feedback simulation. The feedback process for each \( k_a \) is simulated. At first, the input signal rises from 3000 to 60000, and then drops to 6000. The results are shown in Fig. 5, from which the conclusion can be drawn that the bigger the \( k_a \), the shorter the feedback time and the bigger the variation amplitude.

So, the \( k_a \) choice principle is that when the object scene changes slowly, make \( k_a = 0.4 \), the feedback image will be smoother and more stable, whereas with the object changing very fast, a bigger \( k_a \) must be chosen to meet the faster change of signals, e.g., \( k_a = 0.6 \).

### 3. Experimental results

Use the HDRI system introduced in part one, make \( k_a = 0.5 \) and execute imaging experiments.

In the experiments, one LED lamp is used as a high illuminance object. At the beginning, the lamp is off. At some moment, it is turned on and reaches the highest illuminance in a very short time. After the HDRI output stabilizes, the LED is turned to half of the highest brightness. Form the step signal with three states of the lamp:
off state, on state and half state. Serials images for this whole process are obtained. Typical images are shown in Fig. 6.

In Figure 6, part a is the off lamp, b–e is the feedback process with highest brightness lamp, while f–h is the feedback process with half-high lamp.

Consider the output image data in the rectangle area marked in each image of Fig. 6. The average value of 16 bit CCD data in this area is plotted in Fig. 7.

In the first three frames, the output value is zero, because at these moments, the lens is covered. At about the 6th frame, the first stable status is reached with the shut down lamp. At about the 15th frame, the lamp turns on with the highest brightness, the output value rises immediately. This feedback process takes about 10 frames to reach the stable status for the second step. At about the 27th frame, the lamp turns to about half-brightness, and the output value also decreases, within a few frames, it reaches the third stable status.

With this kind of method, there are several ways to reduce feedback times. One is improving the SLM contrast, which means that the proportion band of the $P$ control

![Fig. 6. Serials images obtained in the experiments.](image)

![Fig. 7. Experimental results for the step signal.](image)
The feedback stability research of HDRI system

will be reduced, the system oscillations will decrease faster. Another way to improve the imaging efficiency is choosing SLM and CCD with higher refreshing rate.

4. Conclusions

PID theory is used to analyze the feedback process in one kind of HDRI system. PID parameter is expressed with a simple formula. Data simulations are taken to study this parameter, the optimized value range and the principle of choice are determined. Imaging experiments are conducted to get high dynamic range image with input step signals. Conclusions can be drawn from the experimental results that HDRI system with PID controller has good effect during the imaging process. The feedback process is fast with few feedback loops. The output images are stable with high dynamic range. And the result proves that the PID parameter assures good effect.

Reference


Received July 18, 2008
in revised form November 19, 2008
Efficient iris segmentation method with support vector domain description

RONGNIAN TANG¹, ²*, JIUQIANG HAN², XINMAN ZHANG²

¹Hainan University, School of Mechanical and Electronic Engineering, Baodao Xincun, Danzhou 571737, P.R. China
²Xi’an Jiaotong University, School of Electronics and Information Engineering, 28 Xianning West Rd, Xi’an 710049, P.R. China

*Corresponding author: Rongnian.Tang@gmail.com

With the aim to improve the performance of iris segmentation method to process images with heterogeneous characteristics, the authors introduce a new method inspired by the support vector domain description (SVDD). A local geometric moment function is used to extract shape features of the iris borders. Then, these features are fed into the trained SVDD classifier for borders recognition followed by the application of Hough transform to solve circumference parameters of iris. The performance of the proposed method and the most cited methods, Daugman’s method and Wilders’ method, had been tested on the UBIRIS database. Compared with the two existing methods, our proposal is not only comparable to them when the iris image has good quality, but has better segmentation performance in the case of poor quality images. The experimental results show that the method proposed does have a higher robustness and is less dependent on the quality of raw iris image.

Keywords: iris segmentation, local geometric moment, support vector domain description (SVDD), border recognition.

1. Introduction

With the increasing emphasis on security, the technology of biometrics has been extensively used to identify an individual by both government and private entities [1]. Biometrics aims to accurately identify each individual using various physiological or behavioral characteristics [2], and will replace traditional security systems in the future. Recently, iris recognition comes into focus in this area.

The human iris, an annular part between the pupil (the black central area in the eye) and sclera (the white area in the eye) as shown in Fig. 1, provides many interlacing minute characteristics such as furrows, freckles, crypts and coronas which
are used to identify an individual. The suitability of iris as an exceptionally accurate biometric derives from the following: i) the texture of iris is highly stable over a person’s lifetime; ii) iris is an internal organ that is externally visible as well; iii) the irises of two eyes of an individual or identical twins are completely independent and uncorrelated; iv) iris has rich physical structure and can provide lots of data. So, the iris recognition has some advantages, including high reliability, uniqueness and noninvasiveness. However, despite the safer and quicker access, some critical problems persist and significant work needs to be done before mass-scale deployment on national and international levels can be achieved [3]. Many issues, including system robustness, speed of enrolment and recognition, the accuracy and robustness of iris image segmentation in various environments, etc., remain to be addressed.

In practice, the condition of the environment and that of the lab is not always satisfactory. So, the captured images always have heterogeneous characteristics, being unclear, unfocused or obscured by eyelashes. All this not only disturbs the performance of the existing segmentation methods, but has direct negative effect on the iris recognition system. To improve the robustness of iris segmentation is a challenging task now.

The aim of this paper is to develop a new iris segmentation method which has higher robustness and is computationally efficient and less time-consuming. We first present our intuitive observations about the characteristics of the iris, and then introduce a new segmentation method inspired by the support vector data description (SVDD). Finally, a series of experiments are performed in order to evaluate the method proposed.

The remainder of this paper is organized as follows. Section 2 reviews related work in the area. Section 3 provides detailed descriptions of the theory and steps of the method proposed. The experimental results and discussion are presented in
Section 4. Finally, in Section 5, we draw conclusions and give suggestion for future work.

2. Related work

Much valuable work on iris segmentation was done in the past. In 1987, Flom and Safir discovered that iris morphology would keep stabilization throughout the whole human life, and developed the first relevant method [4]. In 1993, Daugman [5] proposed one of the most famous iris recognition methods, constructed on the basis of modern iris recognition, and proved scientifically the feasibility of iris recognition. For segmentation, Daugman [5–7] introduced an integrodifferential operator to find the inner and outer borders of the iris. The operator searches over the image domain for the maximum in the partial derivative with respect to increasing radius, of the normalized contour integral along a circular arc of radius and centre coordinates.

In 1997, Wildes [8] proposed the iris segmentation method relying on the binary edge-map operation and the Hough transform. The raw image of the iris is converted into a binary edge-map via gradient-based edge detection, then the iris is segmented by Hough transform. In a similar manner, Ma et al. [9] roughly determine the iris region in advance, and then use the Canny edge detection operator and Hough transform to exactly segment iris.

Like in Daugman’s method, Camus and Wildes [10] proposed another integrodifferential operator, a component-goodness-of-fit operator, in the polar coordinate system that searched over a three-dimensional space, and the parameters of iris borders were obtained as the equation maximized.

Assuming prior segmentation of iris inner border, Du et al. [11] proposed a method to detect the iris outer border. They found the parameters of iris outer border in polar coordinates with the largest horizontal edge resultant from Sobel filtering.

Mira and Mayer [12] solved this problem with morphological operators. They applied threshold, closing and opening operators to detect the borders of the iris.

Based on the assumption that iris is not a perfect circle, Miyazawa et al. [13] developed elliptic model to simulate iris shape, and determined iris parameters throughout maximizing the absolute difference of contour summation, of pixel values along the ellipse.

Proença and Alexandre [14] proposed a new iris segmentation method that consists in selecting three discrete features followed by the application of a fuzzy-clustering algorithm.

From the existed literature, we can conclude two major strategies for iris segmentation: one construct the iris edge-map, and the other maximize the specific equations. Both of them are dependent on the specific image characteristics, brightness and contrast, as well as the existence of noise factors (reflections, eyelids and eyelashes, and so on). So, the methods in both classes show perfect performance for
high quality iris image, but their performance is seriously degraded when the iris image includes some heterogeneous characteristics.

3. Proposed method

We can see from the iris images that the iris borders are always obscured by eyelashes and eyelids, and it is very difficult to segment the iris and to remove the noise factors only depending on the intensity or brightness of image. Obviously, it is not enough to separate the pixels belonging to the iris from the obscuring element. For a more accurate and robust iris segmentation method, it is inevitable to restrain the noise factors from a new viewpoint as well as find the borders of iris.

In fact, iris borders have regular shape as well as the gradient of the borders. Both them are totally different from that of the obscuring element such as eyelashes and eyelid. Intuitively, if we can extract these features and segment the iris by recognizing corresponding borders, then we will obtain a new method with characteristics and noise factors being less dependent on particular image. Based on this idea, our segmentation method is given in the block diagram of Fig. 2 (with outer border segmentation as an example). The essential point is to make the edge-map less dependent on the specific image characteristics. A SVDD classifier is introduced to construct it.

The proposed method begins with the feature extraction where feature vectors are extracted for each image pixel. Then SVDD classifier is applied to recognize the edge of iris borders. Finally, the Hough transform is used to complete the circumference detection task.

3.1. Feature extraction

Image moment theory was used to extract the features in our method. It has been widely used in various areas of computer vision and image processing. Calculating moment
values, geometric information of an object can be captured easily. Local geometric moment function measures the image characteristics of a pixel neighborhood, and adapts to capture the shape property of iris outer border.

\[
wpq(x, y) = \sum_{i = -N}^{N} \sum_{j = -N}^{N} i^p j^q f(x + i, y + j)
\]  

(1)

where, \(wpq\) is the local geometric moment, \(N\) is the width of windows, \(f(x, y)\) is the gray-level value of image.

Besides moment values, the gradient of a pixel is also to be considered. And our feature vector is composed of thirteen discrete components, that is,

\[
v_i = \{g_1, \ldots, g_9, w_{10}, w_{01}, w_{20}, w_{02}\}
\]

(2)

where \(g_i\) is the gradient of a 3×3 neighborhood centered as an image pixel, \(w_{10}, w_{01}, w_{20}\) and \(w_{02}\) are the first order and second order local geometric moments.

In applications, the moment values are obtained using existing masks derived from Eq. (1), and this can efficiently reduce the execution time of the moment function.

### 3.2. SVDD classifier

The identification of the iris inner and outer borders can be seen as a data domain description problem, with the only need being to obtain a description of iris borders. This description should cover the class of inner or outer borders, and ideally should reject all other possible objects in iris space.

TAX and DUIN [15, 16] proposed a method for data domain description called support vector domain description (SVDD) inspired by support vector machine (SVM) [17], and it was used for novelty outlier detection. The basic concept of SVDD is to describe a class of data by finding a sphere with minimum volume which contains this class of the data. The radius and center are obtained from the support vector which is the results of SVDD training. The data is accepted when its distance from the sphere center is smaller than the radius. Just like SVM, SVDD also becomes more flexible by using different kernels, which in turn induces more accurate description.

The SVDD algorithm makes us realize that we can only use sample of iris borders to train SVDD classifier for inner border or outer border. The SVDD classifier can recognize the iris borders according to the description obtained. We can extend SVDD to our SVDD classifier as below.

Consider a set of instance-label pairs \((x_i, y_i)\), \(i = 1, 2, \ldots, N\), the samples of which form the same category, where \(x_i \in \mathbb{R}^n\) and

\[
y_i = +1, \quad i = 1, \ldots, N
\]

(3)

We construct a hypersphere for these samples, then we can get the following quadratic optimization problem:
subject to the constraints

\[ y_i \left( \Phi(x_i) - a \right)^2 \leq R^2 + \xi_i, \quad \xi_i \geq 0, \quad i = 1, \ldots, N \]  

(5)

where \( C \) determines the tradeoff between the volume of the hypersphere and the number of outliers. We introduce the Lagrange multipliers \( \alpha_i > 0 \) and \( \lambda_i > 0 \), and get the Lagrangian

\[ L = R^2 + C \sum_{i=1}^{N} \xi_i - \sum_{i=1}^{N} \alpha_i \left( R^2 + \xi_i - y_i \left( \Phi(x_i) - a \right)^2 \right) - \sum_{i=1}^{N} \lambda_i \xi_i \]  

(6)

Using the Karush–Kuhn–Tucker (KKT) condition, the optimal solution of this optimization problem can be obtained. The kernel trick is used to avoid treating the high-dimensional feature space explicitly. However, as noted in [15, 16], polynomial kernels do not yield tight representations of the clusters, while Gaussian kernels work very well. Then we can design the iris border SVDD classifier

\[ f(x) = \text{sgn} \left[ R^2 - 1 + 2 \sum_{i} \alpha_i y_i K(x, x_i) - \sum \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right] \]  

(7)

\[ R^2 = \frac{1}{k} \sum_i \left[ 1 - 2 \sum_{i} \alpha_i K(x_i, x_i) + \sum \alpha_i \alpha_j K(x_i, x_j) \right] \]  

(8)

where \( x_k \) represents support vector and \( k \) is the number of support vector.

To test a feature vector, the distance to the center of the sphere has to be calculated. With this SVDD classifier, a pixel \( i \) is accepted as the iris inner or outer border if its feature vector \( v_i \) makes \( f(v) > 0 \). Figure 3 is an example of outer border segmentation, and the outer border recognized through our SVDD classifier is shown in Fig. 3b. Obviously, the majority of iris border can be recognized by the SVDD classifier. We can also see that the SVDD classifier has excellent capacity to suppress the obscuring element. From our edge-map, we can also see that our SVDD classifier has excellent capacity for suppressing the obscuring element.

### 3.3. Hough transform

As with Wlides’ method, Hough transform is used to solve the performance of iris border. However, the Hough transform is more computationally efficient in our method because if hardly has redundancy pixels in the edge-map. From Fig. 3c, we can see
that the result of Hough transform is accurate and reasonable, and the outer border is segmented properly.

4. Experiments

A series of experiments were designed to test the performance of our method by comparing it to the two most cited methods of Daguman and Wildes. The experiments were conducted on a PC with Celeron 2.6GHz processor, 192MB RAM, and Matlab 7.0 platform.

4.1. Experimental database

UBIRIS database is used in our experiment because it contains lots of images with several types of noise and can simulate the noncooperative environment [18]. The UBIRIS database contains 1877 images which belong to two distinct sessions: 1214 images in the first and 663 in the second one. Table 1 contains the statistical information [14].

4.2. Experiment results

In our experiment, segmentation results are seen as correct when the circumference parameters corresponding to the inner and outer borders almost fall exactly into

<table>
<thead>
<tr>
<th>Quality</th>
<th>Session 1 (%: good, average, bad)</th>
<th>Session 2 (%: good, average, bad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>(82.94, 13.67, 3.78)</td>
<td>(69.68, 19.45, 10.85)</td>
</tr>
<tr>
<td>Reflections area</td>
<td>(94.56, 2.80, 2.63)</td>
<td>(24.13, 38.61, 37.25)</td>
</tr>
<tr>
<td>Visible iris area</td>
<td>(89.29, 7.16, 3.45)</td>
<td>(22.32, 69.07, 8.59)</td>
</tr>
</tbody>
</table>

Fig. 3. Example of outer border segmentation.

Table 1. UBIRIS database statistics.
the respective borders, as we can see in Fig. 4. Or else, the result is considered as unsuccessful segmentation.

In the experiment, the segmentation performance is mainly analyzed by means of two main parameters: correct ratio and average computation time. Here, the existing methods are implemented according to the published papers [5–8]. All of the experimental data are presented in Tab. 2. The first column identifies the method, the second and third specify the correct ratio and averaged computation time of images from the first UBIRIS session, and the fourth and fifth contain the correct ratio and averaged computation time of images from the second UBIRIS session.

From Table 2, we can observe that the proposed method is clearly less dependent on image conditions and correct ratio degradation of the first and second session images was just about 0.63%, which presents the smallest degradation in the presence of noise factors. Although the correct ratio of our method is slightly less than that of Wildes’ method of the first session, it should be noticed that our method presented the best results, 94.88%, of the second session. Otherwise, we can see that our method requires the shortest time spending whether the quality of images is good or not. All of the data clearly show that the method proposed is satisfactory and can well deal with images in various conditions.

The results from Daugman’s method show that its performance is reasonable of the first image sessions. It can reach 94.87% at 3.86 s. However, the relevant results become clearly degraded as the image quality changes. The correct ratio degradation

<table>
<thead>
<tr>
<th>Method</th>
<th>Session 1 CR [%]</th>
<th>Session 1 T [s]</th>
<th>Session 2 CR [%]</th>
<th>Session 2 T [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daugman</td>
<td>94.87</td>
<td>3.86</td>
<td>90.92</td>
<td>5.02</td>
</tr>
<tr>
<td>Wildes</td>
<td>95.66</td>
<td>5.12</td>
<td>90.12</td>
<td>6.84</td>
</tr>
<tr>
<td>Proposed</td>
<td>95.51</td>
<td>3.59</td>
<td>94.88</td>
<td>3.97</td>
</tr>
</tbody>
</table>

Fig. 4. Segmentation result obtained by our method.
Efficient iris segmentation method with support vector domain description

is 3.95% and the degradation time is 1.16 s. Compared with our method, Daugman’s method presents weaker robustness for image quality changing.

Wildes’ method gave the best results in absolute terms, having 95.66% correct ratio of the first session images. However, the performance of Wildes’ method is degraded seriously as the image quality decreases. Its correct ratio degraded by more than 5% and the time expenditure increased to 1.72 s. These data are totally worse than those of Daugman’s method and the method proposed.

From Table 2, we can see that all the methods experience degradation when the image quality changes. But, it is obvious that our method has stronger robustness and satisfactory performance.

5. Conclusions

We analyzed some of the most cited traditional methods in the iris segmentation literature. Performing the experiments on the UBIRIS database, we proved their weak robustness which was induced by their dependence on the specific image-capture conditions.

We proposed a new iris segmentation method which considers the idea of border recognition based on our SVDD classifier. Experimental results showed the encouraging performance of our method in correct ratio, execution time and robustness. Such performance evaluation and comparison not only verify the validity of our observation and understanding of the characteristics of the iris but also provides help for further research.

Acknowledgements – The authors would like to thank Dr. Hugo Pedro Proença (Universidade da Beira Interior, Portugal) and Dr. Luís A. Alexandre (Universidade da Beira Interior, Portugal) for their hard work on the UBIRIS iris image database. The work was supported by a grant from the National Natural Science Foundation of China (No. 60602025), the Ph.D. Programs Foundation of Ministry of Education of China (No. 20050698025).

References


Received May 3, 2008
Red-emitting LiEuW$_2$O$_8$ phosphor for white emitting diodes prepared by sol–gel process

KYU-SEOG HWANG$^1$, YOUNG-SUN JEON$^1$, SEUNG HWANGBO$^2$, JIN-TAE KIM$^3$*

$^1$Department of Biomedical Engineering, Nambu University, 864-1 Wolgye-dong, Gwangsan-gu, Gwangju 506-824, Korea

$^2$Major in Photonic Engineering, Division of Electronic and Photonic Engineering, Honam University, 59-1 Seobong-dong, Gwangsan-gu, Gwangju 506-714, Korea

$^3$Department of Photonic Engineering, College of Engineering, Chosun University, 375 Seosuk-dong, Dong-gu, Gwangju 501-759, Korea

*Corresponding author: kimjt@mail.chosun.ac.kr

Scheelite-type LiEuW$_2$O$_8$ (LEW) phosphors with the optical function of color conversion from near-UV to red were prepared by sol–gel method using a lithium acetate dehydrate, europium(III) nitrate pentahydrate, and tungsten(VI) chloride as starting materials. Viscous mixing sol was prefired at 300 °C for 120 min in air and then white powder precursor was finally annealed at 500–800 °C for 240 min in Ar. From an X-ray diffraction analysis, all the samples are isostructural and belong to the tetragonal system of scheelite-type. The main emission peak is $^5D_0 \rightarrow ^7F_2$ transitions of Eu$^{3+}$ at 615 nm, other transitions from the $^5D_0 \rightarrow ^7F_1$, $^5D_0 \rightarrow ^7F_3$, and $^5D_0 \rightarrow ^7F_4$ located at 570–700 nm range are weak. The characteristic emission of WO$_4^{2-}$ in LEW is quenched absolutely and only red-light emission of Eu$^{3+}$ appears. Crystallinity, surface properties and red-emission by near-ultraviolet (395 nm) were improved by high-temperature annealing.

Keywords: phosphors, sol–gel, near-UV, red-light emission.

1. Introduction

Recently, great progress has been made as regards the solid-state lighting (SSL) device based on GaN light-emitting diode (LED), since the LEDs with high efficiency and various emission wavelengths are available [1, 2]. White LED is a remarkable resource because it is environment-friendly and energy-saving in the 21st century. White LED can be obtained by combining an InGaN blue LED emitting at 465 nm with a broad-band yellow phosphor, e.g., Y$_3$Al$_5$O$_{12}$:Ce$^{3+}$ (YAG:Ce) [3–5]. However,
such a combination has some drawbacks, i.e., the overall efficiency is decreased rapidly when the correlated color temperature of the device is low, and the output light is deficient in the red region of the visible light spectrum. In addition to the blue LED/yellow phosphor approach, white light can be produced by other methods, such as assembling red, green and blue LEDs in one device; a mixture of the tri-color lights generates white light, and a combination of a near-ultra violet (UV) LED (370–410 nm) with red, green and blue phosphors; by this approach, the tri-color phosphors are excited by near-UV emitted by the LED. The use of red, green and blue light-emitting phosphors coated on the near-UV LED chip, is the focus in material and luminescence research nowadays. Three phosphors-converted white LEDs maintain a very high color-rendering index ($R_a > 90$) and are believed to offer the greatest potential for high efficiency SSL [6, 7]. However, the efficiency of the commercially available red phosphors, $Y_2O_2S:Eu^{3+}$ for blue and near-UV GaN based LED, is about eight times less than that of the blue and green phosphors, and this sulfide-based phosphor is chemically unstable. For achieving the three phosphors-converted white LED, it is imperative to develop new effective red phosphors suitable for near-UV LED chips. Therefore, the lack of proper red phosphors becomes the bottleneck for the SSL devices.

Recently, much attention has been paid to making superior red phosphor for white LED. Previously reported scheelite-type compound, $LiEuW_2O_8$ (LEW), written as $ABM_2O_8$ (A: monovalent metal, B: trivalent rare earth, M: Mo$^{6+}$ or W$^{6+}$), exhibits high red emission efficiency under excitation of blue light [8]. LEW has attractive photoluminescent properties; i) the red emission due to $4f^5–4f$ transition of Eu$^{3+}$ when excited by UV, blue or green light, ii) concentration quenching hardly occurs in scheelite-type LEW [9–11]. This weak concentration quenching property is known to be due to the W–O covalent bond [9, 10].

LEW was mostly produced from metal oxides by conventional solid-state reaction at above 1000 °C [11]. As far as we know, there is little information on a successful sol–gel synthesis of LEW red phosphor.

In this paper, we report the preparation of LEW phosphors with the optical function of color conversion from near-UV to red by sol–gel method. In particular, we focused on the preparation parameter of annealing temperature.

2. Experimental

For the synthesis of LEW, stoichiometric mixtures of lithium acetate dehydrate ($C_2H_3LiO_2·2H_2O$), europium(III) nitrate pentahydrate [$Eu(NO_3)_3·5H_2O$], and tungsten(VI) chloride ($WCl_6$) were dissolved with $H_2O_2$ and iso-propanol, respectively. A weighed quantity of urea ($NH_2CONH_2$) was added to the solution at 100 °C. The molar ratio of urea to the total concentration of metal ions was adjusted to 1.5. After mixing, a homogeneous colorless solution was obtained. The solvent was slowly evaporated off from the solution at 100 °C and a colorless precursor with high viscosity was obtained. Subsequently, the precursor was prefired in a dry oven at
Red-emitting LiEuW$_2$O$_8$ phosphor for white emitting diodes ...

Stirring for 30 min

Mixing

Addition of Urea

Preheating

Final annealing

Fig. 1. Experimental flow chart.

300 °C for 120 min in air to obtain white powder. Finally, the dried powder was annealed in a tube furnace at 500, 600, 700, and 800 °C for 240 min in Ar (heating rate: 3 °C/min). Figure 1 shows the experimental flow chart.

X-ray diffraction profiles were measured using an X-ray diffractometer (XRD, D-Max-1200, Rigaku, Japan), equipped with a CuKα radiation source and a graphite monochromator. Particle size and shapes were observed by field emission–scanning electron microscope (FE-SEM, S-4700, Hitachi, Japan). The excitation and emission spectra were recorded on a fluorescent spectrophotometer (F4500, Hitachi, Japan).

3. Results and discussion

The structural analysis of isomorphic LiY(MoO$_4$)$_2$ [8] shows that the scheelite-type LEW has the following crystallographic properties: i) the crystal structure belongs to a tetragonal system with its space group of I4$_1$/a, ii) randomly distributed dodecahedra of LiO$_8$ and EuO$_8$ are surrounded by WO$_4$ tetrahedral units. Figure 2 gives the XRD patterns of LEW as a function of annealing temperature. All the samples are isostructural and belong to the tetragonal system of scheelite-type. In this work, isomorphic NaYW$_2$O$_8$ (JCPDS No. 48-886) is used for the purpose of comparision, since the crystal structure analysis of LEW has not been reported yet and NaYW$_2$O$_8$ exhibits an identical diffraction pattern to the bulk LEW [12]. In this structure, W$^{6+}$ occupies the tetrahedral sites constructed with O$^{2-}$ composing WO$_4$$^{2-}$ anion complex. Li$^{2+}$ is eight-coordinated with O$^{2-}$, forming a distorted cube. The dopant Eu$^{3+}$ occupies the Li$^{2+}$ site. Since there is no significant shift in the peak location, we can consider that the doped Eu$^{3+}$ has little influence on the host structure.

When the prefired white gel was annealed at 500 °C and above, only the scheelite-type LEW phase appeared, and no peaks from impurities were present. This result clearly indicated that sol–gel method effectively decreases the annealing temperature for obtaining a single phase of scheelite-type LEW.
The major diffraction peak at $2\theta \approx 29^\circ$ (112) reflection and other diffraction peaks occurred at $2\theta \approx 18.8^\circ$ (101), 31.7$^\circ$ (004), 34.4$^\circ$ (200) and 47.5$^\circ$ (204), corresponding to the LEW structure, as shown in Fig. 2. With an increase of annealing temperature, the full widths at half-maximum (FWHM) decrease from 0.2362$^\circ$ at 500 °C to 0.1378$^\circ$ at 800 °C, while the peak intensity shows an increase with annealing temperature, as shown in Table 1. These results indicate that the LEW prepared by sol–gel and heat treated at a high temperature can be expected to have high crystallinity.

Based on the XRD data, the lattice parameter has been estimated to be $a = 0.5201 \sim 0.5207$ Å and $c = 1.1253 \sim 1.1309$ Å, as shown in Tab. 2. These values are similar to the reference value of $a = 0.5208$ Å and $c = 1.1282$ Å for bulk LEW [13].

Figure 3 shows the FE-SEM image of LEW powders according to the annealing temperature. With an increase of annealing temperature, the particle size increases. At 800 °C, the mean size of the particles is about 1 μm, which is suitable for
fabrication of the SSL devices. For the powders annealed at a temperature below 700 °C, the surface of the sample is coarser than that of the sample annealed at 800 °C and the shape of the samples is irregular. Therefore, the emission intensity of the samples annealed at below 700 °C may be lower than that of the sample annealed at 800 °C due to the greater surface scattering.

The excitation and emission spectra of LEW excited by near-UV (395 nm) at ambient temperature as a function of annealing temperature are illustrated in Fig. 4. The luminescence intensity of LEW increases with an increase of annealing temperature, which is because the crystallinity of LEW phosphor increases with annealing temperature, as shown in Fig. 2 and Tab. 1.

Fig. 3. FE-SEM images of LEW powders as a function of annealing temperature.

Fig. 4. Excitation and emission spectra of LEW prepared at 500–800 °C.
The broad excitation curve near 250 nm is assigned as the charge-transfer band (CTB) originated from oxygen to tungsten within the WO$_4^2-$ groups, as discussed by Sivakumar and Varadaraju [14]. The Eu$^{3+}$ excitation spectra of the LEW cover the ranges from long-wavelength UV to visible green-light region (300–500 nm). In the range from 300 to 500 nm, all samples show characteristic intraconfigurational 4f–4f emissive transitions of Eu$^{3+}$: $^7F_0 \rightarrow ^5D_4$ transition for 364 nm, $^7F_0 \rightarrow ^5L_7$ transition for 384 nm, $^7F_0 \rightarrow ^5L_6$ transition for 396 nm, $^7F_0 \rightarrow ^5D_3$ transition for 418 nm, and the $^7F_0 \rightarrow ^5D_2$ transition for 466 nm.

Between the luminescent center and the crystal lattice, two couplings are performed in luminescent materials. One is the strong coupling (WO$_4^2-$ group) with high Huang–Rhys factor. The other belongs to the weak coupling (Eu$^{3+}$ ions) with low Huang–Rhys factor [15]. Generally, the strong coupling of CTB (W$^{6+}$) is predominant, whereas the weak coupling of CTB (Eu$^{3+}$) is subordinate. When CTB (Eu$^{3+}$) is excited, the energy absorbed from charge-transfer state is efficiently transferred to Eu$^{3+}$ ion by a non-radiative mechanism, and generates red-light emission of the $^5D_0 \rightarrow ^7F_j$ transition of Eu$^{3+}$. However, in this work, CTB in the range 200–300 nm for WO$_4^2-$ group is remarkably weak, compared to that of LiEuMo$_2$O$_8$ [11], and the f–f transitions of Eu$^{3+}$ dominate the excitation process. Comparing the LEW phosphors prepared in this work, the Eu$^{3+}$ transitions in LEW excitation spectra at a high annealing temperature show more effective absorption at near-UV (396 nm) and blue (466 nm), and these wavelengths coincide with those of commercial GaN-based LED.

As regards the annealing temperature the emission spectra are similar, corresponding to typical 4f level specific transitions of Eu$^{3+}$. The main emission peak is $^5D_0 \rightarrow ^7F_2$ transitions of Eu$^{3+}$ at 615 nm, other transitions from $^5D_0 \rightarrow ^7F_1$, $^5D_0 \rightarrow ^7F_3$ and $^5D_0 \rightarrow ^7F_4$ located at the range 570–700 nm are weak. The characteristic emission of WO$_4^2-$ in LEW is quenched absolutely and only red-light emission of Eu$^{3+}$ appears. The strong emission peak around 615 nm and relatively weak peak around 699 nm is due to the electric dipole energy transition of $^5D_0 \rightarrow ^7F_{2,4}$. The weak emissions of 592 nm and 651 nm are ascribed to the magnetic dipole transition of $^5D_0 \rightarrow ^7F_{1,3}$. The electric-dipole allowed transition would be dominant if Eu$^{3+}$ occupied the lattice site of noncentrosymmetric environment in the scheelite phases [16]. For this reason, the intensity of $^5D_0 \rightarrow ^7F_{2,4}$ was found to be much stronger than that of $^5D_0 \rightarrow ^7F_{1,3}$. When the annealing temperature is increased, the LEW shows stronger red emission at 615 nm by exciting at near-UV wavelength of 395 nm. The reason for this may be that the emission intensity of LEW phosphor depends strongly on the particle shape and size distribution. A narrow size distribution and spherical-like morphology are necessary for the phosphor with good luminescent properties [17]. Compared with the LEW annealed at below 700 °C, the sample annealed at 800 °C has a homogeneous surface and particle shape (see Fig. 4), which are favorable as regards the luminescent properties because of the lesser contamination or fewer dead layers on the phosphor surface. Furthermore, the doping of Eu$^{3+}$ is easier and more effective in sol–gel process than traditional solid-state reaction since all of the starting materials are mixed at the molecular level.
4. Conclusions

Crystalline scheelite-type LEW phosphors for white-LED were prepared at 500–800 °C by using a sol–gel method. The crystal structure, particle shape and luminescent properties under near-UV (395 nm) excitation of the phosphor have been investigated according to the annealing temperature. In the emission spectra, the strongest emission is the electric dipole transition red emission $^5D_0 \rightarrow ^7F_2$ (615 nm), while the magnetic dipole transition orange emission $^5D_0 \rightarrow ^7F_{1,3}$ (592 and 651 nm) is subordinate. Compared with the LEW annealed at below 700 °C, the sample annealed at 800 °C has a high crystallinity, a homogeneous surface and particle shape, which are favorable as regards to luminescent properties because of the lesser contamination or fewer dead layers on the phosphor surface.

References


Received October 17, 2008
Integrated design for large-scale opto-mechanical structure

ANHU LI1*, YONGMING BIAN1, YONGCHENG LIANG2, YONGJIAN ZHU3

1School of Mechanical Engineering, Tongji University, Shanghai 200092, China
2College of Engineering Science and Technology, Shanghai Ocean University, Shanghai 200090, China
3National Engineering Research Center for High Efficiency Grinding, Hunan University, Changsha 410082, China

*Corresponding author: yfliah@yahoo.com.cn

An integrated design method is discussed which thoroughly considers related parameters of the various subsystems in order to optimize the overall system that mainly consists of opto-mechanical structure CAD, CAE and the integrated information platform PDM. Based on the parameter drive of the virtual main model, the method focuses on the model transformation and data share among different design and analysis steps, and so the concurrent simulation and design optimization are carried out. As an example of application, the integrated design for a large-scale opto-mechanical structure is introduced, including optical design, structure design and analysis, which further validates the advantages of the method. Due to comprehensive consideration of the design and analysis process by CAD and CAE based on PDM, the integrated design well attains the structure optimization with high efficiency.

Keywords: integrated design, CAD/CAE, large-scale structure, optical instrument.

1. Introduction

Large-scale optical instruments, such as large telescopes and laser communication terminals, belong to the optical-mechanical-electric integrated systems, which are characterized by a large-scale clear aperture and high precision [1, 2]. Development of such instruments involves many key techniques, for instance, large mirror mounting design, high-accuracy special base frame, large precision shafting, high-accuracy driving technique, ultrathin optical component support, active control for thin mirror surface deformation, new material and machine technique, and so on.

In the early design of large-scale optical instruments, due to the lack of design experience, sufficient tolerances are usually scheduled for initial design parameters [3].
After the engineering prototype is completed, its actual test results are contrasted to the design indexes to evaluate the design quality, and some structures and parameters may in turn be modified till the instrument design is in agreement with the design requirements. At present, this method is not encouraged for heavy task, high expense, long cycle, and especially unattainable optimization of the design results.

With the wide application of computer techniques in various engineering fields, the techniques of CAD and CAE have rapidly developed and led to the innovation in design methods of modern optical instrument structure. In 1980, Jacob M. Miller, American researcher at Honeywell Electro-Optical Systems Center, firstly proposed the concepts and steps of optical-mechanical-electric integrated design method, and enumerated the software used [4]. Meanwhile he successfully analyzed the optical-electric sensor by using the method. Based on CAD/CAE techniques, the optical-mechanical-electric integrated method is used to analyze and simulate the geometry model and finite element model corresponding to the virtual prototype of the instrument, and overall, considers mutual actions and constraints of various subsystems so that the structure parameters are systemically, consistently and dynamically balanced to finally optimize the whole system parameters.

In the paper, we further discuss an integrated design method, which fully considers the model transformation and optimization of the parameters during the entire design process, mainly including opto-mechanical structure design and analysis by CAD, CAE and especially the integrated information share through product data management (PDM). As an example, a large-scale optical instrument structure is developed by this method. The integrated design and simulation are carried out for the overall system.

2. Opto-mechanical structure CAD model

The mechanical structure as the mounting supports the optics system and ensures the optics performance and the system reliability. The constraints of mechanical structure are generally divided into two types according to their effectiveness in optics system and application environment, the auxiliary parts of optics system and the mounting mechanisms of optics components. The opto-mechanical structures include two aspects of static structure and dynamical one, and both of them collaboratively realize the optics performance under the different application conditions. Therefore, the structure design needs to correspondingly consider static stiffness and motion reliability [5]. As regards a large mechanical structure, especially used in special environment such as space conditions, the weight, volume and power consumption become the main factors to be considered, and some special structure, material and technology are to be adopted to optimize the system design.

Figure 1 shows the opto-mechanical structure ACD flow. Firstly, according to the optics system requirements, a concept of structure design is put forward and preliminary calculation is carried out. Then, a virtual prototype of the whole structure is built, including all parts, all components and overall assembly. Finally, after the geometry dimensions and materials are set, the character parameters and structure
rationality can be tested and modified in turn. Together with the dynamical performance simulation based on virtual motion model, we can fully check the feasibility of the design project and decide whether or not to change the design details or even the project. Moreover, we can either import the CAD model into FEA software by format transformation such as IGES, STEP, DFX, etc., or transfer the CAD model to FEA software through the interface processing program processor.

3. Opto-mechanical structure CAE analysis

Due to the large structure and high precision in a large-scale optical instrument, it is necessary to evaluate the design by finite element analysis method, including structure analysis, thermal analysis and optical analysis [6]. In order to realize the integrated design and system optimization, collaborative simulation and analysis must cover the whole process including project selection, structure design, motion simulation, thermal design, assembly analysis and machining process.

Generally, the steps of FEA method consist of solid modeling, generating meshes, setting conditions, solving and post-processing. The structure statics analysis is intended for research into the structure response shown by strain and stress. For a large-scale structure, the gravity effect must be considered, which usually induces the elastic deformation, especially serious in space environment. The dynamics analysis mainly resolves the vibration mode and gains the dynamic rigidity. In other words, the structure weakness as well as the resisting fracture capability can be found through the vibration mode analysis. Conduction, convection and radiation as three heat transfer modes widely exist in large-scale instruments, including steady and transient temperature field [7, 8]. Through the thermal analysis to get thermal performance, the thermal control project of optical instrument can be implemented, which will guide the structure design to meet the requirements of optics performance.

However, no matter which project of the mechanical structure is used, it must center the optics system, and the final analysis is to improve the optical performance. Figure 2 gives relations of the various analysis processes, which can be realized by the data interchange and integration among different software. The Zernike fitting method is usually employed to evaluate the optics performance such as wavefront
analysis, transformation function, and so on. In Figure 3, an example of our prior FEA on a large-scale opto-mechanical structure is shown [9].

4. PDM information integration

PDM as the integrated platform bridges the design and analysis process, and shares data based on the virtual main model, which approaches the parameter drive and real time modification during the whole design [10].

PDM generally includes CAD model data, technology and file data, FEA analysis and simulation data, etc. Different CAD and CAE subsystems all can share data information by PDM. For instance, in CIMS (computer integrated manufacturing system) based on concurrent engineering, PDM plays a key role for different subsystems with high efficiency. Moreover, with the development of network technology, PDM relies on the CAN (controller area network) and client-server system structure will efficiently build the collaborative and integrated work environment, including project design, system simulation and analysis, product management over the whole lifecycle, and even all the production and service process in the market. Figure 4 shows PDM application in optical instrument design and analysis.

5. A design example of a large-scale optical structure

We develop a Fizeau interferometer with a 360 mm clear aperture by the integrated design method. Figure 5 shows the design and optimization flow of the reference mirror and mounting structure in the interferometer. We use Windchill PDMLink to solve the distributed product data management problem, including file management, model data management, information saving and opening. CAD and CAE software include optics design software of CODE V, 2D and 3D design software of AutoCAD and Pro/Engineer, FEA software of Ansys, calculation software of Matlab, etc.
In Figure 5 the arrows show the data flow between different design modules. Firstly, the virtual model of the optical instrument is conceptually designed by Pro/Engineering and CODE V. Then the finite element method is taken for the static and dynamical analysis, as well as the temperature field analysis, by CAE software of Ansys1.0. Thirdly, the structure and thermal control project is further analyzed and
optimized through the CAE result again and again, and various optical aberrations are solved and corrected. Finally, the whole process is based on PDM, and related data between CAD and CAE model share each other till realizing the overall design optimization. The whole design process is under the integrated framework, and the integrated design method through dynamical data interaction highly improves the design efficiency and quality.

6. Conclusions

In order to overcome the disadvantages of the traditional design of large-scale optical instruments, the integrated design method is introduced for complete consideration of the design process to attain the overall system optimization, which mainly consists of three aspects, including opto-mechanical structure CAD, opto-mechanical structure CAE and PDM information integration. With the developments in network and software techniques, the future large-scale optical instrument design will mainly tend towards two aspects. On the one hand, the integrated design based on Web PDM will play an important role in complicated and dynamical data treatment. On the other hand, the Opto-CAD software emerging will better facilitate the visualization and
parameterization design in the optical instrument field, which will well improve the design quality and efficiency.

Acknowledgments – This work was supported by the national “863” Project of China under Grant No. 2007AA03Z105 and the National Natural Science Foundation of China (50805107). We also acknowledge the support from AC21 Special Project Fund.

References


Received July 18, 2008
Measurements of solution fluorescence
– a new concept

STEFAN GĘBALA

Gdynia Maritime University, Department of Physics, ul. Morska 83, 81-225 Gdynia, Poland;
e-mail: stgebala@am.gdynia.pl

The presented studies are based on the predictions of the behavior of certain types of fluorescence
pigments in water solutions. Their asymmetric structure causes spontaneous surface concentration
of the investigated sample. As a result, a densely packed, organized surface layer is created. This
layer facilitates much better conditions for fluorescence than the rest of a sample. A set-up for
solution surface fluorescence measurements has been assembled. This paper presents the results
of the studies which confirm the occurrence of this phenomenon. The studies involved testing
several biological samples. The described concept of fluorescence measurement has significant
practical aspect.

Keywords: fluorescence, amphophility, surface self concentration, liquid crystal.

1. Introduction

Investigations into solution fluorescence are difficult [1]. Single molecules with
fluorescent properties are surrounded by a great number of solution molecules which
are insensitive to light. This is a reason for their relatively low efficiency. Besides,
pigment molecules are subject to chaotic thermal movements. Components of these
oscillations towards exciting light and towards emission cause the Doppler broadening
of the emitted fluorescence lines. The instantaneous changes in these molecule
concentrations in the measurement area also cause fluctuations of the registered light.
Additionally, the part of inducing light can penetrate deep in the solution and excite
the pigment energetic levels there. Fluorescence light, which is generated in the depth
of a sample may be absorbed and then emitted several times on its way to the surface.
The results of solution fluorescence investigation are difficult to interpret due to
reemission and reabsorption phenomena. They require complicated recalculations
based on not always reliable and clear assumptions. In comparison with these
phenomena, the bathochromic effect of fluorescence, resulting from the pigment
molecules shift at the moment of light absorption and their recoil at the moment of
fluorescence quant emission, seems easy to interpret. Summarizing, the objective
difficulties during the solution fluorescence investigation are so significant that their spectra are not too specific and their analyses provide very limited possibilities of practical applications [2].

2. Amphophilic molecules’ behavior in solutions

Molecules of fluorescent pigments are not symmetric [3]. Some of them have amphiphilic structures, i.e. just like molecules of surface active compounds, they have the endings with different water affinity. One end is more hydrophilic and the other is more hydrophobic (lypophylic). A good example comes with flavons, specific substances of plant origin, which are common in nature. Figure 1 presents a molecule of chrysine, typical flavon. At the upper end of this molecule hydrophobic phenyl ring appears while at the bottom hydrophilic hydroxyl groups are observed. Molecules with such structures should spontaneously concentrate at a free surface of a water solution, i.e. at the border between water and air. The hydrophilic endings are pulled towards the water solution, while the hydrophobic ones are pushed above. When amphophylic molecules occur in the sample, they have a minimum potential on its free surface. They migrate from the depth to the surface. Finally, the pigment molecules are concentrated and generate a compact layer on the free surface. Amphophylic properties cause their uniform orientation in a perpendicular to surface layer direction. Small asymmetries

Fig. 1. Amphophyle molecule of chrysine.

Fig. 2. Amphophyle molecules on the surface of water solution.
repeated in the molecular structure cause smaller interactions which put these molecules in order in the layer.

Therefore, it is possible that concentration of amphophylic pigments on the free surface of water solution creates an ordered fluorescent structure. This structure is similar to a singular layer of liquid, smectic crystal [4]. Fluorescent pigments may create a dense layer of stabilized and uniformly oriented molecules (Fig. 2). Concentration of fluorescent pigments in such structured monolayer should improve the properties of the registered fluorescence so that it should be similar to solid state fluorescence properties. Long-range forces stabilize pigment molecules in this structure. So a decrease in fluctuations of measured fluorescence intensity is observed. There are no any conditions for the Doppler disarray of fluorescent spectra to appear in such layer. Additionally, the bathochromic effect should not be registered. Like in crystals, the momentum of light quanta interacts with the mass of the whole of the pigment surface structure, not with the mass of a single molecule. Moreover, dense packing of fluorescent molecules on the sample surface (almost beyond the solution) should increase fluorescence efficiency and eliminate its disturbances with secondary effects, such as reabsorption and reemission.

3. Apparatus

A traditional measurement of solution fluorescence does not employ the concentration of pigments near the surface (Fig. 3). The illuminating light does not reach there. The part of the sample, where pigment molecules may not occur, is illuminated and all molecules move towards the surface. The traditional instruments for liquid sample fluorescence measurements are based on similar procedures. Therefore, there was a need for instrument modification. The geometry of the measurement had to be changed so that the excitement of the horizontal layer of pigments was possible as well as the registration of the fluorescence light created in this layer. Both basic elements

Fig. 3. Fluorescence measurement in traditional way.
4. Materials and methods

The concept must be verified. Natural polar solutions with lots of asymmetric pigment molecules were looked for. Honey, wine and potato mash were chosen for the studies. One of the criterions for classification of honey available on the market is its botanical origin, i.e., its type. Different floral types of honey contain different kinds of fluorescent pigments from nectar of different types of plants. In favorable conditions, bees make clean types of honey. Besides, the market price structure is such that beekeepers tend to produce clean honey types. Thus samples of various types of honey are easily available, they are diverse and classified by the beekeepers. Moreover, there is an economically justified need for the preparation of an instrumental method for honey floral types identification. Thus the beekeepers were ready to cooperate in this matter.

The choice of wine was also based on common availability of samples. The need for an instrumental method, which would confirm wine authenticity was an additional reason.

Flavon chrysin (Fig. 1) is common in potatoes. Thus their consistence was changed in order to be able to seek the fluorescence from the surface structure of these model amphophilic pigment molecules. Also easily available samples supported the choice of potatoes too.

The studies were carried out using a set-up based on the Fluorat-02-Panorama spectrofluorimeter. A special adapter was built for it in order to change its traditional...
Measurements of solution fluorescence – a new concept

Both the excitation and fluorescent light were transmitted using fiber-optic technology. Methodology was adjusted having in mind the tasks and the study potential. Traditionally, in order to determine the study area, light absorption is measured. Then in order to determine the energy levels, excitation and emission spectra are measured. A method for the measurement of light absorption in a pigment monolayer on a surface has not been fixed yet. The necessary range of measurements was impossible to determine on this way. Preliminary fluorescence measurements in the entire range of the set-up possibilities were made. In this way the available areas of surface fluorescence measurements have been determined. In order to study the fluorescence from surfaces of honey, wine or mashed potatoes, the set-up proved to have been appropriate. The complete spectrum is a 3-D surface and the excitation, emission and synchronous spectra are the transects: $ww$, $ee$ and $ss$ (Fig. 5). The synchronous spectra were placed along the direction with the highest fluorescence diversity, and those were analyzed.

5. Experimental verification

Solutions of natural origin may contain different fluorescent pigments. The presented concept is based on an assumption that the asymmetric structure of certain molecules contents in solution may cause their spontaneous concentration on its free surface. These pigments cover the surface of a sample with a thick fluorescent layer, while the other pigments remain in the solution. In such case, the excitation and registration of fluorescence from the top involves several processes (Fig. 6). Monochromatic light flux (1) reaches the layer from the top, penetrates the layer and is partly absorbed (2). The rest of light is absorbed in the solution (3), until it completely

![Fig. 5. Total fluorescent spectrum of buckwheat honey.](image-url)
disappears (4). The excited energetic levels in the solution become the sources of fluorescence \( F_g \), and on its way outside the sample it is partly absorbed in the surface layer (6). The part that is not absorbed (9), adds to fluorescence from the surface layer: primary (7) and secondary (8). This creates emitted light flux over the sample surface \( F_{p+g} \) (10). With properly weak excitation light flux, secondary processes (6 and 8) may be disregarded. Thus the surface fluorescence will be a difference between fluorescence registered above the sample surface \( F_{p+g} \) and below it \( F_g \). This is very clear for different wine types (Fig. 7) [5]. Similar results were also obtained for samples of different honey types (Fig. 8) [6, 7] and potato mash (Fig. 9) [8]. It can be observed that the surface fluorescence \( F_p \) has got much greater impact on total fluorescence than it can be deduced from the contribution of the volume of this layer in the entire sample volume. This means that on the surface of these solutions there are layers which produce significant part of fluorescence registered above.

6. Conclusions

Although the presented plots of solution fluorescence spectra have not been smoothed, they seem to be exceptionally appropriate. The liquid sample fluorescence measured
Measurements of solution fluorescence – a new concept

Fig. 7. Fluorescence spectra of several types of wine registered above the sample surface \( F_{p\rightarrow g}(\lambda) \), under the surface \( F_g(\lambda) \) and the calculated surface fluorescence spectrum \( F_p(\lambda) \).

Fig. 8. Different fluorescence spectra of different types of honey registered above the sample surface \( F_{p\rightarrow g}(\lambda) \), under the surface \( F_g(\lambda) \) and the calculated surface fluorescence spectrum \( F_p(\lambda) \).
in new geometry appears to be more efficient, stable and repeatable than the traditional one [2, 7].

1. It has been proved that fluorescence layer may appear on the surface of certain water solutions.
2. The phenomenon may occur in other water solutions which contain pigments with amphophylic molecules.
3. The properties of the registered fluorescence indicate potential ordering of liquid crystal in the pigment monolayer.
4. Spontaneous surface concentration and stabilization of pigment molecules dissolved in the solution facilitates easier fluorescence application in natural sample investigations.

References
Measurements of solution fluorescence – a new concept


Received December 28, 2007
in revised form June 2, 2008
Biomechanical model of human eyeball and its applications

WIESLAW ŠRÓDKA

Wrocław University of Technology, Deformable Body Mechanics Faculty Unit, Smoluchowskiego 25, 50-372 Wrocław, Poland; e-mail: wieslaw.srodka@pwr.wroc.pl

Attempts at the mechanical identification of the human eyeball are often not very effective for two reasons: the material parameters determined by tension tests on corneal and scleral tissue specimens are not sufficiently accurate while numerical models of the eye, integrating material and geometric parameters, are often based on unrealistic assumptions. The examples presented here cover refractive surgery, Goldmann applanation tonometry and the optical self-adjustment of the eye. The discussed problems are illustrated with calculations showing that it is possible to effectively use a biomechanical model of the eye to identify its material parameters. Also the handicaps, the Imbert–Fick law among them (numerical calculations do not corroborate this law), lying at the basis of applanation tonometry are demonstrated. The conclusions can help to create a realistic numerical model of the eyeball.

Keywords: eyeball, biomechanical model, tonometry, intraocular pressure (IOP), optical self-adjustment.

1. Introduction

There has been a growing interest in the numerical modelling of the human eyeball in view of its practical applications in tonometry and refractive surgery. The mechanics of the outer shells of the eye is related to optical system since their most flexible part – the cornea – is at the same time the strongest lens in this system. The intraocular pressure – usually the principal load acting on this shell structure – is subject to variation in a relatively wide range, noticeably affecting the displacements of the cornea and the sclera. The deformations translate into changes in optical power and the consequent shifts of the optical focus relative to the fundus of the eye, affecting the sharpness of the image on the retina. Thus the eyeball’s structural (geometrical and material) parameters have a major influence on its optical functions. In recent decades this influence has been exploited to correct the optical power of the eye by surgically altering the profile of the cornea’s outer surface. Complete knowledge about the geometry of the eyeball and the material parameters of the tissues forming it would make it possible to precisely plan the effects of such surgical procedures.

Despite the fact that attempts at the mechanical identification of the eye have been made for nearly a century, only the physicochemical structures of the cornea and
the sclera, and recently also their geometries, have been identified. The mechanical parameters of primary importance for solutions concerning displacements, \textit{i.e.}, moduli of elasticity, are still the subject of controversy.

2. Structure of eyeball model

Ophthalmologists distinguish many layers in the cornea, but only the stroma plays a principle role in eyeball mechanics. The stroma takes up over 90\% of the corneal thickness. For this reason the material of the cornea is equated with this single kind of tissue. Because of its specific structure, the shell is globally isotropic in the directions tangent to the middle layer. Moreover, because of its outer shape and material characteristic the cornea can be treated as a membrane. Therefore in calculations this tissue is usually treated as an isotropic material [1]. Similar simplifications are applied to the sclera.

In the corneal-scleral shell the transitional zone, called the limbus, plays a special role. The material of the limbus shows noticeable anisotropy [2], but the area which it occupies between the cornea and the sclera is so small that the anisotropy of this zone is usually neglected. But the ciliary body together with the muscle which controls the lens as well as the choroid cannot be neglected. These tissues clearly increase the overall limbus and sclera stiffness and so this effect should be taken into account in investigations of eyeball deformations caused by intraocular pressure.

3. Material constants

3.1. Exponential characteristic

Although the cornea material curve is still sometimes approximated by a linear function, today both cornea material nonlinearity and anisotropy (and even rheology) are commonly taken into account. However, not always the above effects and not all of them at the same time must be taken into account.

The elastic nonlinearity of the stroma and that of the sclera, under uniaxial stress, is usually described by the exponential relation proposed by Woo \textit{et al.} [3]:

\[
\sigma = A \left[ \exp (\alpha \varepsilon) - 1 \right]
\]

where: \(\sigma\) – stress, \(\varepsilon\) – strain, \(A\) and \(\alpha\) – material constants. For a complex state of stress one should additionally reduce the main stress tensor components to uniaxial stress [4] according to the formula:

\[
\sigma^* = \left\{ \frac{1}{2} \left[ (\sigma_1 - \sigma_2)^2 + (\sigma_1 - \sigma_3)^2 + (\sigma_2 - \sigma_3)^2 \right] \right\}^{0.5}
\]
The main strain tensor components are converted in a similar way

\[ \varepsilon^* = \left\{ \frac{2}{9} \left[ (\varepsilon_1 - \varepsilon_2)^2 + (\varepsilon_1 - \varepsilon_3)^2 + (\varepsilon_2 - \varepsilon_3)^2 \right] \right\}^{0.5} \]

3.2. Longitudinal modulus of elasticity

The material of the eye shells is often described by a secant elasticity modulus, i.e., a ratio of total stress to total strain in a uniaxial state of stress. This parameter can be compared for different materials at a fixed level of stress. In the case of the cornea material, the (two-axial) steady-stress component in the apex, amounting to about 20 kPa at the nominal intraocular pressure, can be considered to be such a level. The elasticity modulus \( E \) (measured \textit{in vitro} on samples excised from the cornea) found in the literature, ranges widely from 0.026 MPa [5] to 57 MPa [6] and to as much as 115 MPa reported by \textit{Uchio et al.} [7].

4. Biomechanical models of eyeball

The eyeball as a structure seems to be uncomplicated and relatively amenable to mechanical analysis. This is really so in some special applications, e.g. when modelling the effects of dynamic loads caused by quickly moving glass slivers or an air bag impact. In other applications, e.g. in refractive surgery, the most important aspect of the eye model is often its optics and then the analysis does not end with the determination of the displacement field and the stress field as in the case of, let us say, the thigh bone. The configuration of the loaded structure requires further calculations to determine the change in the position of the optical focus relative to the fundus caused by deformation. The results of such calculations are highly sensitive to displacement solution accuracy and so to the preliminary geometry and material assumptions, the simplifications made and the adopted boundary conditions. It is extremely difficult to obtain correct results and, in the author’s opinion, they are rare exceptions in the literature on the subject.

The models found in the literature fall into two groups: analytical models [8, 9] and numerical models [1, 10–14, 17]. An analytical model would have this advantage over a numerical model that by providing a closed solution it would make it possible to investigate the influence of individual parameters on its optical functioning. Thus any effects of changes in the parameters (e.g., an increase in intraocular pressure resulting in a change in optical power) would have a physical justification. Unfortunately, the current analytical capacities in this regard are insufficient and models which can be solved in this way are limited to the cornea (with a constant thickness and made of a linear material) alone, which is too large simplification. The numerical model has no such limitations. It can cover the whole eyeball (the cornea
together with the sclera) with any geometry and it can be equipped with nearly any material: anisotropic, nonlinear or inelastic. Such boundary conditions (the way in which the model is fixed) do not encounter any computing barriers. Thus the model’s potential seems to be limitless.

But the numerical model has one major drawback – the solution it provides is in the form of numerical tables. Each relation investigated in this model requires a series of separate solutions and the obtained functions come from approximations. The latter, however, are performed arbitrarily and so have no physical justification.

Despite the above inconveniences, the numerical model of the eyeball has become a powerful investigative tool. Thanks to computer tomography, the geometry of all the structural details of the eye has been precisely determined. But little is known about the mechanical properties of the tissues forming the eyeball. The longitudinal moduli of elasticity measured by different researchers vary by as much as four orders (see Sec. 3.2) and it is difficult to distinguish between correct and worthless results. A similar scatter characterizes the other measured mechanical parameters. We have found ourselves in a rather uncomfortable situation when the possibilities offered by numerical techniques have got much ahead of our laboratory potential as regards the investigation of the mechanical properties of the eye’s tissues and its optical functions. The results yielded by the tensile test turn out to be so uncertain that researchers have turned to the numerical model of the cornea and the sclera to identify the tissues in the mechanical respect. The model material or geometry parameters are matched to make the model behave in the same way as the real eye. By imposing
constraints on the parameters, one can determine their proper ranges. This approach has turned out to be more effective. An axial symmetric finite element eyeball model, on which this kind of material identification was made, is described in [13]. The model is built from 2D solid quadrilateral 8-node body of revolution elements shown in Figs. 1 and 3.

5. Numerical identification of corneal material

The eyeball models found in the literature were designed mostly to simulate the flattening of the corneal apex in Goldmann applanation tonometry [8, 10, 11] or the change in the eyeball’s optical power after surgical correction of its geometry [1, 12, 14]. In both cases, the model is also used to identify the cornea material. The results are often far from ideal because of the questionable assumptions or outright errors made in the creation of the model. Despite this, the obtained parameter values are within a much narrower range than the ones determined experimentally. Sometimes they are quite reliable, although obtained on the basis of flawed assumptions or after gross simplifications. Numerical models seem to be much more researcher-friendly than biological preparations. It often happens that a model which was not previously verified is used to identify the cornea material or to investigate the influence of its thickness on tonometrically measured intraocular pressure. The trust placed on the model by its creator sometimes seems to be boundless. The few problems of model’s applications in ophthalmology are described below.

5.1. Refractive surgery

The model is verified by introducing the same changes as the ones made by surgery into the geometry of the cornea and then calculating its radius of curvature in the apex under intraocular pressure. The change in optical power calculated for the model is compared with the change observed after the surgery. If the model is correct, the respective results, evaluated by the quality of the image on the retina, should be similar.

Geometry modifications made by photorefractive keratectomy (PRK) are the easiest to introduce into the model. The surgery performed with a laser consists in changing the curvature radius of the corneal apex (over a diameter of about 7 mm) through ablation (vaporization) of its outer layers. During the operation the intraocular pressure does not change. If no astigmatism is corrected, but only optical power, the cornea after the new profile is introduced, still remains (approximately) axially symmetric. It seems quite easy to create a model and numerically solve the problem. Unfortunately, “it seems” is the most certain element in this thesis.

An attempt at a numerical solution encounters a formidable difficulty for quite an inconspicuous reason. An eyeball model always comes into existence through a design, i.e., a configuration of the structure prior to loading. This applies to the geometry before and after the operation. The problem derives from the fact that in clinical conditions the geometry in both cases is unavailable. Only the final configuration
is known, i.e., the dimensions of the cornea under pressure before surgery and
the postoperative dimensions of the cornea deformed by both the surgery and the in-tra-
ocular pressure. Only after the solution one can find out whether the cornea model,
both the one before the surgery and the one after the surgery (they are two different
models), acquired correct dimensions, but the configuration of the model under load
depends on both the initial geometry and the assumed material elasticity. The injustice
which the analyst suffers here consists in the fact that the surgeon is completely
unaware of the problem – the measurements before the surgery, the surgery itself and
the postoperative checkup are conducted at a (roughly) invariable intraocular pressure.
The geometry of the cornea not subjected to load does not occur here at all.

One can easily guess what the basic error in the numerical PRK simulation is –
the change in the eye’s optical power is calculated for a “frozen” eyeball. The model
(not subjected to load) acquires eyeball dimensions from clinical measurements, i.e.,
carried out on the eyeball subjected to load. Then it is being solved (changes its
dimensions, and so also its optical power) and the geometry correction caused by PRK
is introduced into the obtained model which is now considered to be stiff. For this final
model configuration the ultimate optical power of the cornea is calculated. Is such
an algorithm admissible?

The answer depends on the cornea’s elasticity modulus. If it is close to 8 MPa,
as indicated by many reliable experiments, carried out mainly by HJORTDAL [2],
the answer is yes since the shell characterized by this elasticity modulus is so stiff
that a change in pressure from 0 to 2.135 kPa (16 mmHg) has little influence on its
configuration. If, however, the elasticity modulus is close to 0.3 MPa, as indicated by
other equally reliable experiments [7, 8], including ours [13], then the answer is no.
One could get the impression that this question is decided by a vote if it were not for
the fact that the former figure comes from measurements while the latter in most cases
is the result of numerical simulations. The existing experience suggests that the latter
figure is correct. The role which the cornea’s elasticity modulus plays in predicting
changes in the optical power of the model after PRK and the significance of
the simplifications made in such calculations were investigated by the authors in [14].

Radial keratotomy is an example of another difficulty, this time associated with
the cornea material itself. The surgery is performed using a special scalpel and consists
in making several deep incisions arranged radially on the peripheries of the cornea, as
is seen in Fig. 2. The apex flattening caused by intraocular pressure corrects myopia.
To the above difficulties, the numerical simulation of the surgical procedure adds
another one: this kind of change in the geometry of the cornea results in high stress
gradients on the bottom of the incisions. Then the assumption that the material is
anisotropic seems untenable. But one can easily find attempts at such solutions [1, 12]
and it is by no means certain that they are basically flawed.

5.2. Goldmann applanation tonometry
The numerical simulation of intraocular pressure (IOP) measurement by means of
an applanation tonometer is readily used to verify the cornea material adopted
in the model. This measuring method [15] is based on the assumption of equality of pressures on both sides of the flattened corneal apex, i.e., the external pressure (denoted as $IOP_G$) caused by the (flat) measuring tip of the instrument and the internal pressure (denoted as $IOP$):

$$IOP_G = IOP$$  \hspace{1cm} (1)

The former pressure is associated with the name of Goldmann. After the pressure force and the diameter of the zone of contact between the tonometer tip and the cornea are measured, the average $IOP_G$ pressure is calculated. Eq. (1), called the Imbert–Fick law, holds good, according to Goldmann, only for the so-called calibration dimensions:
- applanation zone diameter $D = 3.06$ mm,
- apex cornea thickness $CCT = 0.52$ mm,
- apex cornea curvature radius $R = 7.80$ mm.
The measurement of pressure on a cornea with dimensions other than the calibration ones requires a correction for CCT and $R$. But the dependence of $IOP_G$ on the pressure level and the cornea material parameters is not taken into account.

Applanation tonometry has an over one hundred years old tradition and it would seem that the theoretical basis of the mechanics of this measurement is well understood. However, some arguments and beliefs expressed in the literature, although they do not arouse controversy there, are at odds with our numerical solutions. An example here is the argument for calibration dimensions put forward by Goldmann and Schmidt [15] and repeated by others [16]: justifying diameter $D = 3.06$ mm it is assumed (Goldmann did not do any such calculations) that externally applied pressure $IOP_G$ acting over diameter $D$ is counteracted from the inside by pressure $IOP$ and by the bending resistance of the shell not exposed to load. Thus $IOP_G$ is always higher than $IOP$ and condition (1) is satisfied only when also the force of attraction between the tonometer tip and the cornea is taken into account. The force originates from the surface tension in the lacrimal film connecting the two surfaces and its magnitude counterbalances the forces bending the shell exactly at $D = 3.06$ mm. As is apparent, the result is understood here as a superposition of two solutions: 1) for a shell (membrane) devoid of bending rigidity – then the external pressure and the internal pressure, acting in the applanation zone, are exactly equal, regardless of its dimensions and the value of $IOP$ and 2) for a shell with low bending rigidity, loaded by only this (constant) part of pressure $IOP_G$ which is needed to bend the shell at $IOP = 0$.

Our studies of the eyeball model indicate that the above assumption is incorrect and leads to not only large errors, but also to a paradox. The problem is best illustrated by the solution obtained for a linearly elastic model. But the most serious consequences follow from the solution for a realistic nonlinear model. The two results are presented as graphs in Fig. 4. Each of the models (the linear model and the nonlinear one) has such a cornea material that at nominal intraocular pressure $IOP = 16$ mmHg the calculated average pressure exerted from the outside by the tonometer tip, $IOP_G$, also amounts to 16 mmHg. According to current applanation tonometry, the measured pressure $IOP_G$ is a linear function of $IOP$, represented by grey dashed line in the figure, i.e., written as Eq. (1). If the force originating from surface tension were neglected (as it is done in numerical models), then the function graph would be a grey dashed line vertically shifted by the initial value of $IOP_G$ (at $IOP = 0$).

The solution for the linear model, shown in Fig. 4, is, of course, far from the reality, but it clearly puts in doubt the applanation tonometry’s assumption that the difference between $IOP_G$ and $IOP$ does not depend on $IOP$. As the diagram for the linear model shows, the influence of the intraocular pressure is so strong that even the functional trend between the variables has been reversed. One can say that in the linear model by increasing $IOP$ one helps $IOP_G$ to flatten the apex.

According to Fig. 4, the nonlinear model behaves quite differently in this respect. But the functional dependence for this model does not coincide with the applanation tonometry predictions. At a low IOP (below the nominal value) applanation pressure $IOP_G$ is, as assumed by Goldmann, actually higher than $IOP$. But as the pressure
increases, a trend similar to the one observed for the linear model emerges: the higher the IOP, the easier it is to flatten the corneal apex. Thus above the nominal pressure, IOP_G becomes lower than IOP. This is the paradox: the pressure on the outside of the flattened disc is lower than the pressure acting from the inside. If Goldmann’s assumption about the influence of the shell’s bending resistance and the surface tension were correct, the cornea’s bending resistance in this case would have to be negative! The rest of the cornea then would attract the flattened disc, instead of repelling it.

This result stands in contradiction to all the theses and the experimental results found in the literature on the subject. All the opinions and measurement data support the above applanation tonometry assumptions and so they are at variance with the result presented here. This raises questions about the quality of the model and the sense of the obtained solutions. Can the bending resistance of a spherical shell being flattened and simultaneously loaded with an external pressure be negative? However strange the answer will sound – this is precisely the case.

The applanation tonometry’s cornea deformation model based on intuition has never been verified before. To the author’s knowledge, the calculation results shown in Fig. 4 are the only attempt to check Goldmann’s assumptions. The numerical model, described in [13, 14], used for this purpose, was carefully prepared with regard to both the selection of materials for the cornea, the sclera and the corneal limbus and its agreement with the known experimental results. According to Fig. 4, applanation pressure IOP_G for the model not loaded with internal pressure is slight, amounting to about 1 mmHg. Thus there are no reasons to question the IOP_G value obtained at IOP = 32 mmHg. The calculated IOP_G = 27.5 mmHg is by about 5 mmHg lower than

![Fig. 4. Calculated IOP_G for model made of: linear-elastic material and nonlinear-elastic material, depending on IOP. Grey lines represent Goldmann idealization: broken line – pressure measured on dry (not wetted with lacrimal fluid) cornea, solid line expresses Imbert–Fick law (Eq. (1)) and so takes into account the influence of surface tension in lacrimal film. All bold lines satisfy condition IOP_G = IOP = 16 mmHg (nominal).](image-url)
IOP, which means that this deviation cannot be linked to the shell’s bending rigidity. Inequality $IOP_G < IOP$ above 16 mmHg, in the nonlinear model is as contrary to our intuition as in the linear model. But in the latter model the numerical solution does not raise doubts as to its quality since the fact that the graph is so radically different from that of function (1) cannot be due to only calculation errors. Intuition in mechanical problems is a poor adviser.

5.3. Corneal profile

An interesting application of the numerical model to structural identification is corneal profile geometry. The corneal profile is often approximated with a circle whereby the cornea’s outer surface is then treated as a segment of a sphere. But the spherical aberration caused by this shape of the lens and the results of topographic examination of the cornea, indicating that the curvature radius decreases as the distance from the eye’s optical axis increases, speak against the sphere. Easier to use and better fitting the geometry of the cornea is the ellipse [17]

$$z(x) = \frac{1}{e^2 - 1} \left[ \sqrt{R^2 + x^2(e^2 - 1)} - R \right]$$

changing its shape depending on eccentricity $e$ from $e = 0$ (a circle), through $e = 1$ (a parabola) to $e > 1$ (a hyperbola). The parabola has a particularly desirable analytical form since being a 2nd degree polynomial it is easy to differentiate and integrate. Is it acceptable to approximate the corneal profile with the parabola in studies of cornea model optics?

One of the few studies devoted to this problem is [18] in which an attempt is made to determine the optimum eccentricity of the ellipse, used to describe the cornea’s outer and inner outline in the linear-elastic model. The optimization was performed with regard to a peculiar aspect of the model, called optical self-adjustment. One should note that the type of function approximating the outer profile of the cornea has a strong influence on the model’s optical system and its dynamics as the model parameters (e.g., IOP or the radius of corneal apex curvature) are changed. The model’s optical focus not only should be located near the fundus of the eye at the nominal value of IOP but also changes in its location, dependent on the fluctuations in IOP, are governed by strict rules. Not every function can meet their requirements.

The optical self-adjustment of the eye is a hypothesis advanced by KASPRZAK [19]. As applied to the real eye, it reads as follows: the quality of the image on the retina of an unaccommodating eye does not depend on IOP. Physiological deviations of IOP from the mean value vary depending on the time of the day, the blood pressure, the body position and many other circumstances, including the health ones. The deviations do not usually exceed 5 mmHg. Of an eyeball model the Kasprzak hypothesis requires that its performance should be relatively easily verified: intraocular pressure
fluctuations around the nominal value, with an amplitude of at least 30% of this value, should not have a noticeable effect on the location of the optical focus relative to the fundus while the lens should retain constant focal power. This model performance is hardly likely if the model is not structurally adapted for this purpose. When the pressure is increased, the eyeball expands – the cornea displaces forward while the corneal apex curvature radius increases. As a result of the stiff displacement of the cornea, the focus shifts forwards whereas the increase in the apex curvature radius shifts the focus backwards. When the numerical model’s structural (geometrical and material) parameters are properly matched, the absolute values of the focus shifts are equal to each other and ultimately the location of the focus does not change. Such a model is optically self-adjusting.

In the discussed paper [18] the self-adjustment of the eyeball model was achieved through a properly matched limbus ring stiffness. The calculation results showed that limbus stiffness depends monotonically on the corneal profile ellipse eccentricity in the self-adjusting model. Initially, at an eccentricity close to zero, the model shows almost spontaneous self-adjustment. In order to increase eccentricity, it was necessary to increase the tension stiffness of the limbus ring. Initially, increments in limbus ring stiffness were small but once $e = 0.5$ was exceeded, they rapidly grew and at $e$ close to 0.65 the limbus stiffness required to retain self-adjustment approached infinity. This result does not depend on the cornea’s Young modulus, provided the ratio of the sclera modulus to the cornea modulus amounts to 5 (this value is justified by both experimental results [3] and physical predictions [13]).

Experiments of this kind show the strategy of identifying eye structures, based on the numerical model, to be highly effective. The obtained result imposes clear limits on the function used to describe the cornea’s topography. In particular, it provides an answer to the question asked above – the parabola (the more so the hyperbola) is unsuitable for corneal profile approximation in the linear model.

6. Conclusions

At present, experimental results which could be used to create a numerical model of the human eyeball are far from satisfactory. The results of strength tests carried out on specimens excised from eye tissues and on whole eyeballs differ too much to be a reliable source of data. The technique of identifying the material of the eye’s shells, consisting in the numerical simulation of clinical tests (such as applanation tonometry) turns out to be more effective.

Numerical eye models, even though they prove to be suitable for the purpose, are still far from perfect, particularly when applied to predict the results of refractive surgery. The cause is not so much the lack of a method, but rather faulty calculation programs. The resulting solution errors are due to, at least partially, the controversies around material constants and the associated simplifications.
The identified experimental and numerical shortcomings contribute to the persistence of conflicting beliefs about solutions achievable today. The results obtained by the author suggest that the principal equation of applanation tonometry (the Imbert–Fick law) cannot be satisfied by the real eyeball, even when the latter has calibration dimensions, since the law is based on false assumptions.

The shortcomings also affect the other aspects of the eyeball model’s structure and behaviour — rheological material parameters, accommodation and fixing in the eye socket, i.e. the boundary conditions. All have an effect on the model’s optical functions. Therefore one can conclude that the problem lies in the too little weight attached to the correctness of assumptions and solutions. The investigation of the optical system of the eyeball by means of a numerical model constitutes a new quality in mechanics and requires a new approach to eyeball design. The eye is not a mechanical structure in the classical sense — its function is not to carry loads, but to see.

References


Received October 12, 2008
in revised form December 9, 2008
What characteristics a clinical CSF system has to have?

ELISA COLOMBO1, 2, LUIS ISSOLIO1, 2*, JAVIER SANTILLÁN1, ROLANDO AGUIRRE1, 2

1Departamento de Luminotecnia, Luz y Visión, Facultad de Ciencias Exactas y Tecnología, Universidad Nacional de Tucumán. Avda. Independencia 1800, San Miguel de Tucumán (4000), Tucumán, Argentina

2Consejo Nacional de Investigaciones Científicas y Técnicas – CONICET, Argentina

*Corresponding author: lissolio@herrera.unt.edu.ar

We discuss the characteristics a system to measure the contrast sensitivity function (CSF) in the ophthalmologic clinic has to have. We propose that this system should be computer based in order to assure flexibility and precision. Besides the original calibration, this equipment needs an auxiliary system to keep the working conditions allowing periodical gamma corrections. We try a calibration method based on visual comparisons and show that it is valid and simple. We propose to use an adaptive psychophysical method to obtain contrast thresholds that assure a good compromise between precision and duration of the whole test. Finally, we propose that the system has to have its own normality curves for the different age ranges allowing the practitioner to perform clinical evaluations. Summing up, we can say that taking into account the above issues, the fidelity of the stimuli will be guaranteed and the challenges entailed in its transference to the clinic will be overcome.

Keywords: contrast sensitivity, ophthalmologic clinic, standards.

1. Introduction

Human spatial vision has the ability to detect and process stimuli of varying size and contrast; this ability can be evaluated through contrast sensitivity function (CSF). This function is obtained from the threshold contrasts that the visual system requires to detect or to discriminate a sinusoidal stimulus of any spatial frequency (the inverse of a grating period). The inverse of the threshold contrast is the contrast sensitivity (CS). A complete CSF curve results from the threshold contrast measurement for the entire range of spatial frequencies to which the system is sensitive.

The clinic utility of CSF is determined because this function gives information about the optic part as well as the neuronal part of the visual system, thus allowing the ophthalmologists to attain an evaluation of functional vision [1]. From the clinical standpoint, the CSF test can detect vision loss and can give way to get a better diagnosis
of early-stage diseases [2], such as early retina and optic nerve problems, in which visual acuity does not present alterations in its first stages [3]. In this way, CSF provides information not only about the ability to detect objects in a vast range of situations [4], but also to detect anomalies [5, 6].

There has been a renewed interest in CSF for clinical use [7, 8]. Studies show that it can be a helpful tool to be part of a functional visual assessment, especially when a more thorough evaluation is necessary other than the one given by Visual Acuity [9–11]. It is being used successfully as an aid in making decisions about cataract surgery [12]. It has likewise shown to be effective in the diagnosis of glaucoma [13–16], optic neuritis and multiple sclerosis [17, 18], and in the following up of visual performance changes after refractive surgeries [19–23]. It can also help assess visual performance in the elderly [5, 24–27] or in people with low vision [28, 29]. Studies on patients with congenital nystagmus have also shown the usefulness of CSF [30].

Moreover, it is considered advisable to incorporate these kinds of visual tests to determine a person’s eligibility to drive [31]. This is even more so considering that elderly people – whose population percentage is on the rise – undergo an ageing process of their visual capacities [32, 33] and are more sensitive to factors such as glare, all of which are aggravated by the use of glasses when compared to the use of contact lenses [34].

There are several commercial tests, some of them are based on printed or translucent cards and they use letters [35–37] or sinusoidal gratings [2, 38, 39]. The printed cards specify the illuminance levels that must be measured on them in order to achieve measurement (FACT, for instance), or they are loaded onto a system that ensures the fulfillment of these conditions (CST Digital). The translucent ones incorporate the lighting system from behind (CSV-1000). Research shows that the contrast sensitivity curve provided by sine-wave grating tests is more sensitive and informative than the results obtained from low-contrast letter acuity systems [2]. Tests based on printed cards have a limited ability to discriminate the subject’s answer since the stimuli have a fixed number of contrast values [40]; the result obtained is only one because the process of determination is carried out only once; the control of illumination is not always easy to get; and reference ranges are not discriminated by age.

As a counterpart to the possibilities mentioned, there is computer-based equipment that has sophisticated and expensive graphic cards as the VSG (Cambridge Research Systems). There are also other systems that need libraries for visual stimuli generation, such as the Video Toolbox and the Psychophysics Toolbox, both of which require expensive software (a C compiler or MatLab, respectively). These alternatives have neither the facilities needed to be adapted to clinical measurement nor the elements to carry out diagnostic tests because they are oriented towards the research on visual sciences. However, even though the usefulness of this CS function in clinic is well known and there exist a variety of commercial systems, its use has not extended so much in the ophthalmology.

Our goal was to evaluate different methods and technologies available at present to propose the main characteristics a CSF system has to have to be useful in the clinical
practice, in order to achieve reliability, accuracy and flexibility. For this purpose we were checking each one with our own system developed. First of all the system has to be computer based and to have an easy procedure to periodically calibrate it. Besides, it is necessary to select a psychophysical method to determine the CSF and to incorporate ranges by age to compare the measurements from the patients.

2. Hardware components and calibration aspects

The development of a CSF system implies having equipment capable of presenting low contrast sinusoidal gratings in the entire range that the visual system is sensitive to. The use of a system based on a PC is the more appropriate alternative to get the grey scale resolution guaranteeing the fidelity of the stimuli. The minimum requirements of the PC are a Pentium processor, 512 MB of RAM and a graphic card that allows two video cards (an AGP port in our case) to be controlled.

2.1. Display selection

The display should be able to reproduce the attributes of an image with such a precision that the separation of the responses from the normal conditions could be attributed to visual problems. Nevertheless, issues such as size cost and power consumption are factors that should not go unwatched once the main objective has been achieved. It was necessary to choose a given technology from among the three existing ones on the market. At present, plasma displays are not considered a valid choice since they do not have the versatility required in ophthalmology and they are highly expensive. In order to decide between LCD (liquid crystal display) or CRT (cathode rays tube) displays we compared the most important parameters for reproducing sinusoidal stimuli of very small and highly controlled contrast:

1. Dynamic range (DR) of luminance: generally, LCD screens have greater maximum luminance than CRTs, however, minimum luminance is also higher, which leads to greater contrast proportions for CRTs.

2. Gray level resolution: rather than the absolute value of this number, what is important is the relationship between luminance and DAC voltage. While CRT presents an area of greater saturation towards high values of DAC voltage, LCD screens tend to show a flat area not only at high digital levels but also at low levels. In this sense, the advantage of working with CRT is remarkable.

This means that nowadays only CRT monitors have DR characteristics and gray level resolution which make it possible to show low contrast sinusoidal gratings. However, in terms of size, LCD is at an advantage of CRT monitors, but also this advantage is compensated by the price of CRTs. Another disadvantage of LCD screens is that emerging light is not homogeneous in all space directions, as is the case of CRT, although this deficiency is not very important because the task is foveal. We know that in a near future the CRT screens will replace the selected CRT ones.

Two monitors of the same brand were compared, one LCD and the other CRT. They bore similar dimensions, pixel size, and response time and rest consumption.
Although maximum luminance in LCD was slightly more than double, contrast proportion for CRT is 6 times higher; horizontal and vertical vision angles are slightly higher in CRTs, as well as maximum resolution and colours borne. In the end, a 19” CRT display was chosen, with a screen size of 35.2 cm × 26.4 cm (Samsung SyncMaster 955 DF). The video card resolution was set at 1600 × 1200 pixels with 32 bits for colour representation. A monitor of 15” or 17” can be used if no spatial 2AFC paradigm is implemented.

2.2. Gray level resolution of the system

The gray scale resolution of the standard VGA video card is insufficient to reproduce the lowest contrast and by this reason it is necessary to consider auxiliary tools. In order to expand it we incline toward the option of a video attenuator [41] and a module for the adaptation of the monitor to the video signal. Both auxiliary systems enable the color resolution of the monitor jumps from the 8-bit to more than 13 bits.

The video attenuator integrates the output signals $r$, $g$ and $b$ of the graphic card into one unique output by means of a passive network that assigns each input signal with a different attenuation level. This implementation was originally designed to be used with achromatic monitors due to the only output attenuator; meanwhile, the authors proposed the use of two graphic cards for colour monitors with an attenuator that integrates channels $r$, $g$ and $b$ of both [41]. Instead of this solution, in our system, we connect the monochromatic attenuator to the three joint inputs of the colour monitor and modify their input impedance so as to adapt it to the output impedance (75 ohms) of the monochromatic attenuator.

The beat-stealing algorithm [42] instead of a video attenuator could be considered but it allows around 10.8 bits, an insufficient number of gray levels to achieve a precise reproduction of a sinusoidal stimulus of as lower a contrast as 0.2% used in the measurement of the greatest sensitivities. Besides, the beat-stealing technique requires accurate information about the effect of the radiant energy on the human eye, which could vary substantially for individual observers reducing the accuracy of the process.

2.3. Calibration

The system has to have an original (or initial) careful process of calibration to correct the different nonlinearities in the video-monitor card system and so asserting the fidelity of the stimuli (for a revision, see [43]).

The gamma correction parameters should be determined by means of a photometric procedure measuring the absolute luminance with a luminance meter and computing the parameters $k$ and $\gamma$ of the relation between the luminance and the DAC voltage:

$$L = k D^\gamma$$

(1)

However, the user needs update this calibration by a periodical procedure. One alternative is by means of a psychophysical method based on visual comparison [44]
What characteristics a clinical CSF system has to have?

which does not require a luminance meter. In this work, we show that this method could be successfully implemented on a standard PC. We checked it as follows: a 400-pixel diameter circle is presented on the screen, in one of its halves (A), the maximum DAC signal is applied (a value of 255) to a previously established pixel fraction; and in the other semi-circle (B), the signal applied has a varying value of 0 and 255 on all pixels, thus achieving a given luminance level (Fig. 1). This signal is controlled by the subject that calibrates the equipment, and his task consists in varying the luminance in B until it matches the brightness of half A (brightness matching task). These matching values give the levels of voltage in B associated to the luminance value in A, that is, in the proportion of lit pixels. Once the process for pixels lit for ten different rates is over we fit the data to obtain the parameters $\gamma$ and $k$. Due to the fact that the different luminance percentages generated in the upper field present a striped pattern that makes the matching task difficult, a diffuser glass was necessary to be introduced in the system that covers both fields of the stimulus, which facilitates the subject’s task.

In order to evaluate the reliability of the visual calibration we compare it with a photometric one in the equipment proposed. The experiment was carried out with 7 observers, each of whom repeated the test 4 times. Among the observers, some had no experience in this kind of psychophysical determination, while others did. There is high repeatability in each observer and among observers, as shown in Figs. 2a and 2b, respectively. The values obtained were $\gamma = 1.72 \pm 0.04$ and $k = 0.98 \pm 0.01$, for the visual calibration, while photometric calibration had the values of $\gamma = 1.733 \pm 0.004$ and $k = 0.992 \pm 0.001$. The conclusion is that the parameters are the same within experimental errors, thus validating the visual calibration technique. The results indicate reliability of the method, a reduced measuring time of approximately 10 minutes and ease of performance, all the features being sought by the practitioners.

Another aspect of the monitor behaviour to be taken into account by an initial calibration is the MTF of the display. In spite of the MTF being determined only for vertical gratings because of the interaction between neighbouring pixels provoking a contrast decreasing [45], the MTF of the display in the direction of the luminance

Fig. 1. The bipartite field used to perform the brightness match task in the visual calibration process: without diffuser (a), with diffuser (b).
variation of our quasi horizontal gratings has to be determined [46] since we also found in this case contrast reduction. We measured the maximal and minimal luminance of gratings and then computed the effective contrast in the screen. This measurement was done for all the spatial frequencies used in the system (characterized in pixels per cycle) and for seven low contrasts in the same range used for the determination of contrast sensitivity. To this end, we used an LMT 1009 luminance meter with approximation lenses that allows measuring a single pixel. The results are plotted in Fig. 3 showing the slopes of the lines to be fitted to the data decreasing as the spatial frequency grows. These slopes are plotted as a function of spatial frequency in Fig. 4 and represent

Fig. 2. Luminance vs. voltage: data obtained for subject AS (a), data for the seven intervenient subjects in the experiment, each point is the mean of four measurements (b).

Fig. 3. Contrast measured in grating counterclockwise direction (+7°) as a function of nominal contrast for seven spatial frequencies characterized by the number of pixels per cycle. The lines represent linear fits that show the different slopes obtained for each spatial frequency.
What characteristics a clinical CSF system has to have?

the MTF of the display. From the operative point of view, when sinusoidal gratings are generated, compensation has to be introduced incrementing the nominal contrast by a coefficient in inverse proportion to the modulation factor for the corresponding spatial frequency.

Both gamma correction and the monitor’s MTF measurement have to be carried out taking into account the fact that the stimuli appear in the central area of the monitor, thus avoiding misconvergence in the peripheral area that translates onto the image as an unwanted blur. On the other hand, calibration stimuli have to be small enough to assure that the mean luminance drives the value of the display current supply approximately constant; this mean luminance has to be the same when taking measurements as when performing some calibration. In this way, the calibration parameters are guaranteed to be valid in the measuring process.

3. CSF measurement: software, methodology and experimental set-up

Special purpose software for measuring CSF has to be developed. The choice of methodology is crucial for providing a transferable system for the clinic. Once the stimulus fidelity is assured and good precision in visual function determination is achieved, the measurement should be done in relatively short periods of time.

3.1. Stimuli

We propose static sinusoidal gratings to be used as stimuli [2]; they were slanted 7° in relation to the horizontal one, both clockwise and counterclockwise (Fig. 1), maintaining the same response property, since both orientations belong to a unique orientation column in the striate cortex [47]. The presentation time for each stimulus in its nominal contrast could be around 500 ms. Times below this value produce
an insufficient detection of the pattern while longer times unnecessarily extend the measurement [48], the visual size has to be slightly greater than the foveal size – around 6° or 7° is good enough. The spatial frequency values used have to cover the low, medium and high ranges of spatial frequencies [32, 33, 39, 49] – from 1 to 24 c/deg.

The maximum luminance of a CRT monitor is approximately 150 cd/m². The grating mean luminance should be nearly in the middle of the monitor range of luminance as well as the immediate surrounding luminance, which is not a relevant variable [50, 51]. Although the normal room lighting conditions do not change the visual performance [52], a dark surrounding would be better to avoid distractions. For that reason, the subject should be advised to look at the display every time.

3.2. Distance of vision

From an optometric point of view, it is convenient to perform the measurement as far away as possible from the visualization device [53] so as to minimize the demand for accommodation/convergence, taking for granted a person who is normal or corrected ametropic. Considering limitation of space a compromising solution could be a distance of 1.5 m, which considerably lowers the demands.

3.3. The method and the task

We propose an adaptive method because it has the advantage of shortening measurement times significantly, controlling precision. In this way, we have used the QUEST [54] adaptive method whose algorithm makes a Bayesian inference of each response the observer gives, and hence the following contrast to be presented is established. The observer’s task consists in discerning between the two possible

Fig. 5. Contrast sensitivity obtained for 20, 30 and 40 trials with spatial frequencies of 1, 4 and 12 c/deg. Each point is the mean of 5 measurements and the bars represents a standard deviation.
positions, thus being forced to pick one of them – even if he could no longer see
the stimulus – by pressing the appropriate button.

According to a simple experiment we determined the number of trails needed to
obtain a good estimation of the threshold. We have carried out measurements
considering 20, 30 and 40 trials with spatial frequencies of 1, 4 and 12 c/deg in naïve
and experimented subjects. In Fig. 5, the data obtained for one observer (32 years old)
are plotted. For frequencies of 4 and 12 c/deg, there are no differences among
the number of trials, instead for 1 c/deg, there can be seen differences between
measurements with 20 trials and those performed with 30 and 40 trials. These
observations were confirmed by means of an ANOVA with $p < 0.05$ for 1 c/deg and
$p > 0.05$ for 4 and 12 c/deg. Moreover, if we compare the set of data collected for each
number of trials there are no differences between 30 and 40 trials $(p > 0.05)$ although
it is not the case if we consider 20 and 30 trials $(p < 0.05)$. From these results it was
possible to establish that for 20 trials, there can be an overestimation of the threshold,
while with 30 or 40 measurements, equivalent results were obtained. However, because
time is a very important factor – as in clinics – the option to go with was the one with
30 trials.

4. Reference curves by age range

Finally, probably one of the most important characteristics, is that a CSF clinical system
has to incorporate normality ranges according to age. In this way, the system gets
sensitivity for the diagnosis and follow-ups of visual pathologies. The methodology
used to determine the normal curves must be the same as the one suggested for
the measurement in clinics.

We have determined reference curves with 55 subjects, 11 per age range (20–29,
30–39, 40–49, 50–59, 60–69 years old), all of whom agreed to an ophthalmologic
check-up designed according to APA (American Psychological Association) regula-
tions for the work performed on human beings, and to international norms on bioethics
and the ones stated in the Helsinki Declaration. Since the goal is to establish curves
that will help detect vision anomalies, a visual acuity unit $(AV = 1)$ was required.
The latter would be corrected in situ by an ophthalmologist for the distance of the test
during measurement sessions; an exam of the ocular means (cornea and crystalline
biomicroscopy) and the fondus, as well as an adequate control of intraocular pressure
(contact tonometry). The group of people included in the measurement was free from
ocular pathologies, though some needed refractive correction. Each observer did two
series of measurements with each eye: the first was deemed part of the learning process
and only the data from the second round were considered. Measurement time for
the six spatial frequencies in each eye was approximately 10 minutes, four times less
than when using a constant stimuli method, which certifies the efficiency of clinical
measurements based on adaptive methods.

Of the 5 curves obtained, and according to the standard mean error – which
indicates the degree of precision estimated in the “centre” of the distribution of
means – a separation of the curves in the age ranges 20–49 and 50–69 years old is noticed. In clinical practice, it is common to set normality intervals according to the following criterion [55, 56]

\[
\text{Normality interval} = \mu \pm 2\sigma
\]

where \(\mu\) is the mean for the distribution of means and \(\sigma\) is the standard deviation. This implies that, if we measure a new subject, there is a 95.45% probability that his sensitivity curve will drop within that strip. Applying such criterion, we obtain graphs illustrated in Fig. 6a for the 20–49 age range and in Fig. 6b for the 50–69 age range.

Considering our results and those from the quoted papers we think that these two ranges of normality are enough for the system to get a good ability to discriminate a separation from normality. In addition, the equipment should have the alternative to visualize the course of the measurement through another monitor and the possibility of showing the results obtained and to contrast them with the reference strips as well as to print a report with test results.

5. Discussion

The present challenge in the CSF clinical measurement is to be able to count on reliable equipment with an appropriate measuring methodology, taking into consideration two basic requisites. The first one is for the results obtained to have enough accuracy and reliability. The second one is for the time taken in measuring not to extend too much.

Our decision of choosing a computer based system is in good agreement with earlier works. For instance, RUBIN [57] showed more reliability on CS measurements done with CRT-based equipment with respect to printed-card tests through a test-retest

Fig. 6. Normality interval for the age ranges of 20 to 49 (a) and 50 to 69 (b) years. Each point represents the mean of the sample, vertical bars indicate a standard deviation and the shadowed area represents the two standard deviations.
What characteristics a clinical CSF system has to have?

analysis, and other works have shown the sensitivity of electronic equipment-based
tests for discriminating the normal case from the pathological one [13, 58].

In relation to the methodology we propose an adaptive method because its
superiority as regards the times involved is well known. Although some authors [59]
had already noticed the inefficiency of the constant stimuli method for some
applications, it has continued to be widely used due to its simplicity for experimental
implementation. Our results showed that an adaptive method like QUEST allows more
reliable results than those attained with constant stimuli and in notably less time
consuming. Furthermore, while the QUEST method immediately estimated threshold
contrast, the constant stimuli method requires further processing, consisting in fitting
the data into some mathematical model that describes the psychometric function from
which the threshold contrast can be obtained.

Furthermore, the clinical measurements will be performed in less time than those
obtained when all the spatial frequencies are presented, which is about 10 minutes.
For instance, evaluating a smaller number of spatial frequencies since the practitioner
will not always need information on the entire range. CS measurement in each spatial
frequency takes approximately one and a half minute, which means that in order to
determine the threshold contrast for just 2 spatial frequencies the practitioner needs
only 3 minutes.

On the other hand, the adaptive method-based measurement could be quicker if
the estimated CS value were basically within the area of normality. In this case,
the trial sequence can be interrupted thus reducing measurement time to less than one
minute, since less precision is required. The same criterion may be considered when
encountering the system with CS estimates which are much lower than normal. Only
in those cases where CS estimation is near the inferior limit of the area of normality
will the system demand a greater number of trials which will take up more time to
allow a more precise determination in order to adequately discern the normal case from
the pathological one.

6. Conclusions

We have established the characteristics a CSF system has to have to be used in the
ophthalmologic clinic. First of all we discuss how to generate very low contrast
sinusoidal gratings with high accuracy. Then we propose that for the practitioner to be
able to realize a clinical evaluation of CSF measurement, the system has to have
normality curves for different age ranges. At the same time, we demonstrate that a PC
to which components and calibration procedures are added in order to use it as
a measuring tool, allowed the presentation of grating with contrast lower than 0.2% in
establishing threshold contrasts. Apart from using a monitor to present sinusoidal
gratings, the illumination level to which the patient’s visual system adapts itself is
determined by monitor parameters and by the luminance intended by the equipment
operator taking the measurements. Another important advantage is the incorporation
of a module of visual calibration, allowing it to keep the working conditions independent of a photometer. Finally, we suggest that this kind of system has to have at least two ranges by age.

All these characteristics are necessary if the equipment has to be used in clinic, because: i) measurement can be repeated as often as necessary because the stimuli are presented randomly, ii) it can be periodically calibrated, iii) the adaptation luminance is set mainly by monitor luminance, which makes the system very strong for complex environments, iv) measurements can be compared with normal ones according to age of each patient.

In this way, a computerized system for contrast sensitivity tests with characteristics mentioned in this paper can be used in the clinic as well as in research laboratories, in driving permit controls and in screening of population like that of school-age children.

Acknowledgements – The authors wish to acknowledge the support of Consejo Nacional de Investigaciones Científicas y Técnicas (Fellowship to RA and Grant PIP No. 4948), Consejo de Investigaciones de la Universidad Nacional de Tucumán (Grant 26/E112 to EC), and Agencia Nacional de Promoción Científica y Tecnológica (Grant PICT13-15190 to EC).

References

What characteristics a clinical CSF system has to have?

428 E. COLOMBO et al.


Received July 18, 2008
in revised form October 11, 2008
Coherence characterization of partially coherent flat-topped beam propagating through atmospheric turbulence

F.D. KASHANI, M. ALAVINEJAD, B. GHAFARY

Photonics Laboratory, Physics Department, Iran University of Science and Technology, Tehran, Iran

We study the change in the degree of coherence of a partially coherent flat-topped (PCFT) beam propagating through atmospheric turbulence. It is shown analytically that with a fixed set of source parameters and under a particular atmospheric turbulence model, a PCFT beam propagating through atmospheric turbulence reaches its maximum value of coherence after propagating a particular distance, and the effective width of the spectral degree of coherence also has its maximum value. This phenomenon is independent of the turbulence model used. We also study the effects of beam width values, the structure constant of turbulent media and the degree of coherence on effective width of spectral degree of coherence. The results are illustrated by numerically calculated curves.

Keywords: partially coherent flat-topped beams (PCFT), atmospheric turbulence, degree of coherence, effective width of the spectral degree of coherence.

1. Introduction

Optical beams have been a subject of great interest since the advent of the laser in 1960 [1]. They find useful applications in laser radar, optical communication systems and some optical information processing techniques; for example, it is frequently asserted that high directionality of optical beams is a consequence of high spatial coherence of the sources. However, it was predicted theoretically [2–4] in the late 1970’s and confirmed experimentally [5, 6] soon afterwards that high spatial coherence of the source is not necessary to produce very directional beams. In fact, some partially coherent sources may generate beams which have the same far-zone intensity distribution as a fully coherent single-mode laser beam.

It is well known that partially coherent beams are less influenced by turbulent atmosphere than completely coherent beams [7–10]. Therefore, a considerable number of investigations have paid attention to the characterization of partially coherent light propagation through turbulent atmosphere.

On the other hand, the spectral density, the spectral degree of the polarization and the spectral degree of coherence may change on propagation, even in free space [11].
Many interesting results have been presented in numerous publications on the subject of changes in the spectral degree of coherence of partially coherent beam in any transverse cross-section, see Mandel and Wolf [12]. Fewer papers have appeared focusing on the change in degree of coherence of beam on propagation in free space or through some random media [13].

In recent years, much more attention has been paid to the propagation properties of partially coherent beams concerning their application in wireless optical communication involving both free space and atmosphere, which are always carried out in the scalar approximation [13–16]. Therefore, in addition to the research on changes in the spectral density and the degree of polarization, it is necessary for partially coherent beams to study the change in the degree of coherence on propagation, especially through some random medium.

Spectral density of the PCFT beam has been investigated in some papers [8, 9]. But, to the best of our knowledge, there are no investigations on the spectral degree of coherence of PCFT beam.

The aim of this paper is to study the behavior of change in the degree of coherence of PCFT beam with circular symmetry when it propagates through atmospheric turbulence.

Treatments using scalar theory cannot provide any information about the coherence and polarization properties of beam [13]. Based on the extended Huygens–Fresnel principle [17] and according to a unified theory of coherence and polarization [18], we analyze the evolution of the cross-spectral density matrix that describes the second-order coherence and polarization properties of beam on propagation in atmospheric turbulence and then we give expressions for the cross-spectral density matrix and the spectral degree of coherence of PCFT beam in turbulent media.

We employ the effective width of the spectral degree of coherence of beam to characterize the coherence property of beam on propagation. We investigate the effect of correlation length and order of flatness on it. Finally, we study the influence of turbulent atmosphere and different turbulence models on spectral degree of coherence and its effective width.

2. Cross-spectral density matrix of partially coherent flat-topped beams in atmospheric turbulence and its degree of coherence

We consider a field propagation from the plane $z = 0$, i.e., the partially coherent source plane, into the half-space $z > 0$ where the turbulent atmosphere exists. The electric field vector $E(\rho, z; \omega)$ propagates in a linear medium, in which variations of the refractive index are much smaller than the average value of the refractive index. Based on the extended Huygens–Fresnel principle, we give the expression for the field $E_j(\rho, z; \omega)$ at any point in the half-space $z > 0$:
Coherence characterization of partially coherent flat-topped beam ...

\[ E_j(\rho, z; \omega) = \frac{-ik\exp(ikz)}{2\pi z} \int \int E_j^{(0)}(\rho', 0; \omega) \exp\left(\frac{ik(\rho - \rho')^2}{2z}\right) \exp\left[\psi(\rho, \rho', z; \omega)\right] d^2\rho' \]

where \( E_j(\rho, 0; \omega) \) is the electric field vector at the point \((\rho, 0)\) in the source plane; \( \psi \) is a random phase factor that characterizes the effect of atmospheric turbulence on propagating spherical wave. Here, we assume that the incident field propagates close to the z-axis and further neglect the contribution of the z component [15]. The subscript \( j \) denotes the Cartesian components \( x \) and \( y \) across the transverse plane perpendicular to the propagation direction. The subscripts \( l \) and \( j \) appearing in the remainder of this paper have the same sense as the elements of cross-spectral density matrix that describe the second-order coherence and polarization properties of beam on propagation, which can be given as [13]:

\[ W_{lj}(\rho_1, \rho_2, z; \omega) = \left(\frac{k}{2\pi z}\right)^2 \int d^2\rho_1' \times \int d^2\rho_2' W_{lj}^{(0)}(\rho_1', \rho_2', 0; \omega) \times \exp\left\{\frac{-ik[(\rho_1 - \rho_1')^2 - (\rho_2 - \rho_2')^2]}{2z}\right\} \times \langle \exp\left[\psi^*(\rho_1, \rho_1', z; \omega) + \psi(\rho_2, \rho_2', z; \omega)\right]\rangle_m \]

where \( W_{lj}^{(0)}(\rho_1', \rho_2', 0; \omega) = \langle E_l^{(0)*}(\rho_1'; \omega) E_j^{(0)}(\rho_2'; \omega)\rangle \) is the electric cross-spectral density matrix in the source plane \( z = 0 \). The term \( \langle \ldots \rangle_m \) denotes averaging over the ensemble of statistical realizations of the turbulent medium.

In Equation (2), the angular bracket that describes the turbulence effect can be approximated as follows:

\[ \langle \exp\left[\psi^*(\rho_1, \rho_1', z; \omega) + \psi(\rho_2, \rho_2', z; \omega)\right]\rangle_m \cong \exp\left\{-\frac{\pi^2 k^2 z}{3} \left(\rho_1 - \rho_2\right)^2 + \left(\rho_1 - \rho_2\right) \cdot \left(\rho_1' - \rho_2'\right) + \left(\rho_1' - \rho_2'\right) \int_0^\infty \kappa^3 \phi_n(\kappa) d\kappa\right\} \]

where the quantity \( \int_0^\infty \kappa^3 \phi_n(\kappa) d\kappa \) describes the effect of turbulence, \( \phi_n(\kappa) \) being the spectrum of the refractive-index fluctuations that can be characterized by the Tatarskii
model or the Kolmogorov model [20]. As for the PCFT source, \( W^{(0)}_{ij}(\rho_1^i, \rho_2^j; 0; \omega) \) can be written:

\[
E_{iN}(\rho_1^i; \omega) = \sum_{n=1}^{N} A_l \left( \frac{-n \rho_1^i}{N} \right)^{n-1} \frac{N!}{n!} \exp \left( \frac{-n \rho_1^i}{2 \sigma_s^2} \right)
\]

\[
\xi_{ij}^{(0)}(\rho_1^i - \rho_2^j; \omega) = \sum_{n=1}^{N} B_{ij} \exp \left( \frac{-c(\rho_1^i - \rho_2^j)^2}{2 \sigma_{gij}^2} \right)
\]

\[
W^{(0)}_{ij}(\rho_1^i, \rho_2^j; 0; \omega) = \sum_{n=1}^{N} \sum_{m=1}^{N} \sum_{l=1}^{N} A_l A_j B_{ij} \left( \frac{-1}{N^2} \right) \left( \frac{N!}{n!} \right) \left( \frac{N!}{m!} \right) \times
\]

\[
\exp \left( \frac{-(n \rho_1^i - m \rho_2^j)^2}{2 \sigma_s^2} \right) \exp \left( \frac{-(\rho_1^i - \rho_2^j)^2}{2 \sigma_{gij}^2} \right)
\]

where the coefficients \( A_l, B_{ij}, \sigma_s, \) and \( \sigma_{gij} \) are positive quantities, which are independent of position. The coefficients \( B_{ij} \) satisfy the relations

\[
B_{ij} = 1 \quad \text{when} \quad l = j
\]

\[
|B_{ij}| \leq 1 \quad \text{when} \quad l \neq j
\]

and

\[
B_{ji} = B_{ij}^*
\]

The parameters \( \sigma_s \) and \( \sigma_{gij} \) and characterize the effective source size and the effective width of the spectral degree of coherence of source, respectively.

Parameters characterizing an electromagnetic Gaussian source cannot be chosen arbitrarily due to the sufficient conditions they must satisfy [13]. The restrictions on the choice of parameters of an electromagnetic Gaussian source are provided in [20–22]. For isotropic source, i.e., for the source with \( \sigma_{sl} = \sigma_{sj} = \sigma_s (l, j = x, y) \), the following constraints on the parameters of the source have been derived in [20, 21]:

\[
\frac{A_x^2 \sigma_{gxx}^2}{\sigma_{gxx}^2 + 4 \sigma_s^2} - \frac{2 A_x A_y |B_{xy}| \sigma_{gyy}^2}{\sigma_{gyy}^2 + 4 \sigma_s^2} + \frac{A_y^2 \sigma_{gyy}^2}{\sigma_{gyy}^2 + 4 \sigma_s^2} \geq 0
\]

\[
\frac{\sigma_{gxx}^2}{\sigma_{gxx}^2 + 4 \sigma_s^2} - \frac{2 \sigma_{gxx}^2}{\sigma_{gxx}^2 + 4 \sigma_s^2} + \frac{\sigma_{gyy}^2}{\sigma_{gyy}^2 + 4 \sigma_s^2} \leq 0
\]
where the choice of $\sigma_{gxy}^2$ should be in the range:

$$\max\left\{\sigma_{gxx}^2, \sigma_{gyy}^2\right\} \leq \sigma_{gxy}^2 \leq \min\left\{\frac{\sigma_{gxx}^2}{\sqrt{|B_{xy}|}}, \frac{\sigma_{gyy}^2}{\sqrt{|B_{xy}|}}\right\}$$

(6d)

The inequalities (6) are the necessary and sufficient conditions the parameters of the source must satisfy in order to generate a physically realizable electromagnetic Gaussian source [20].

To evaluate the integration in Eq. (2), we define two arguments $u$ and $v$ such that

$$u = \frac{\rho'_1 + \rho'_2}{2}, \quad v = \rho'_1 - \rho'_2$$

(7)

Substituting Equation (4c) into Equation (2) and calculating, the related integral $W_{lj}^{(0)}(\rho'_1, \rho'_2, 0; \omega)$ is obtained as follows:

$$W_{lj}^{(0)}(\rho'_1, \rho'_2, z; \omega) = \sum_{c=1}^{N} \sum_{m=1}^{N} \sum_{n=1}^{N} A_{ij} A_{lj} \eta \frac{(-1)^{n+m}}{N^3} \binom{N}{n} \binom{N}{m} \frac{1}{4 \alpha_1 \alpha_2 - \alpha_5^2} \times$$

$$\times \exp \left[ \frac{\alpha_1 (\beta_2^2 + \alpha_3^2) - \alpha_3 (\beta_1 \beta_2 + \alpha_3 \alpha_4) + \alpha_2 (\beta_1^2 + \alpha_4^2)}{4 \alpha_1 \alpha_2 - \alpha_5^2} \right]$$

(8)

where:

$$\alpha_1 = \frac{n + m}{4 \sigma_s^2}$$

$$\alpha_2 = \frac{n + m}{16 \sigma_s^2} + \frac{c}{2 \sigma_{gij}^2} + M = \frac{\alpha_1}{4} + \frac{c}{2 \sigma_{gij}^2} + M$$

$$\alpha_3 = \frac{-ik}{2z} \left( \rho_{1x} + \rho_{2x} \right) + M \left( \rho_{1x} - \rho_{2x} \right)$$

$$\alpha_4 = \frac{-ik}{z} \left( \rho_{1x} - \rho_{2x} \right)$$
where the quantity $M$ is given as 0.5465 for the Tatarskii spectrum and as 0.49 for the Kolmogorov spectrum [20], with $\mu$ being the refractive index structure parameter and $l_0$ being the inner scale of turbulence.

Equation (8) is the main result of this paper, by which one can study the change in degree of coherence for the partially coherent beam propagating through the atmospheric turbulence. According to the unified theory of coherence and polarization [18], the degree of coherence can be given as:

$$\mu(\rho_1, \rho_2, z; \omega) = \frac{\text{Tr} W(\rho_1, \rho_2, z; \omega)}{\sqrt{\text{Tr} W(\rho_1, \rho_1, z; \omega) \text{Tr} W(\rho_2, \rho_2, z; \omega)}} =$$

$$= \frac{W_{xx}(\rho_1, \rho_2, z; \omega) + W_{yy}(\rho_1, \rho_1, z; \omega) \times W_{xy}(\rho_2, \rho_2, z; \omega) + W_{yx}(\rho_2, \rho_2, z; \omega)}{\sqrt{W_{xx}(\rho_1, \rho_1, z; \omega) + W_{yy}(\rho_1, \rho_1, z; \omega) \times W_{xy}(\rho_2, \rho_2, z; \omega) + W_{yx}(\rho_2, \rho_2, z; \omega)}}$$

(9)

where Tr denotes the trace of matrix. This equation is the base of the results presented in the next section.

3. Discussion

It can be seen from Eqs. (8) and (9) that the spectral degree of coherence for the PCFT beam is determined by both source parameters and turbulence parameters, simultaneously. We may study the change in the spectral degree of the coherence of beam on propagation in turbulence by choosing different condition parameters and further calculating the absolute value of the spectral degree of coherence.

Considering that only $\sigma_{gij}$ ($j = x, y$) may be used in the calculation, we along with it analyze the special case, i.e., $\sigma_{gxx} = \sigma_{gxy} = \sigma_c$. To represent the partial coherence, the spectral degree of correlation $\sigma_{gij}$ should be smaller than the effective source size [13].

At first, we analyze the change in the degree of coherence of PCFT beam only the increasing propagation distance. It is shown in Fig. 1a that when the length of
the beam path increases, the central peak of spectral degree of coherence generally degrades, which is as expected [13]. However, one can find that the degree of coherence improved after the beam initiates from the source, plane $z = 0$, and the spectral degree of coherence reaches its maximum value when the beam travels a distance about 2.7 km, then it begins degrading and keeps decreasing along with the distance. After sufficiently long propagation distance the degree of coherence almost fades away and the beam can be considered as almost incoherent one. This interesting phenomenon is evidently shown in Fig. 1b. In this figure, we define the width of the spectral degree of coherence $\rho_{\mu}(z)$ of the beam in turbulence as the separation $|\rho_1 - \rho_2|$ of points in a transverse cross-section at which $|\mu(\rho_1, \rho_2, z, \omega)|$ drops from its maximum value of unity (for $|\rho_1 - \rho_2| = 0$) to the value $1/e$.

![Fig. 1. a – The change in the central peak of spectral degree of coherence of PCFT beam. The turbulence is described by Tatarskii spectrum. The curves correspond to different distances in turbulence labeled in the figure. The parameters are taken as: $\lambda = 632.8 \text{ nm}, C_n^2 = 10^{-13} \text{ m}^{-2/3}$ and $I_0 = 5 \text{ mm}, A_x^2 = A_y^2 = 1, B_{xy} = 0, \sigma_s = \sigma = 5 \text{ cm}, \sigma_{gxx} = \sigma_{gxy} = 1 \text{ mm}, N = 4$. b – The effective widths of the spectral degree of coherence for beams generated from three partially coherent sources with different coherence properties, which propagate through atmospheric turbulence. Three sets of source parameters are taken as: $A_x^2 = A_y^2 = 1, B_{xy} = 0, \sigma_s = \sigma = 5 \text{ cm}, N = 4$, and $\sigma_{gxx} = \sigma_{gxy} = 2 \text{ mm}$ (case A), $\sigma_{gxx} = \sigma_{gxy} = 1 \text{ mm}$ (case B), $\sigma_{gxx} = \sigma_{gxy} = 0.1 \text{ mm}$ (case C).]
The behavior of $\tilde{\rho}_\mu(z)$ with increasing distance $z$ is shown in Fig. 1b. One can readily see that the effective width of the spectral degree of coherence increases firstly and then decreases after reaching a maximum value. This can be explained as follows [13]. There are two sets of parameters used to determine the behavior of the beam coherence in the beam propagation, i.e., source parameters and parameters of turbulence spectrum model. When the beam travels in turbulence only over a small distance from the source, the strength of the turbulence is negligible. Therefore, it cannot overcome the effect of coherence properties of the PCFT source. The source parameter determines dominantly the evaluation of the coherence area of the beam in a transverse cross-section of the beam, which increases with $z$, as shown in Fig. 1b (rising part of the curve). At the same time, atmospheric turbulence also plays its role with increasing propagation distance, which as is well known, degrades the beam coherence properties. As the beam propagates a particular distance (about 2.7 km in our example), the strength of atmospheric turbulence matches that of source coherence property and then the width of the spectral degree of coherence reaches its maximum value. After that the effect of atmospheric turbulence dominates the behavior of the width. This causes the value of the effective width of the spectral degree of coherence to decrease gradually (see dropped part of the curve).

Using three different sets of the values of $\sigma_s$ and $\sigma_g$, we study the influence of source parameters on the behavior of the central peak of spectral degree of coherence of beam propagating in atmospheric turbulence characterized by Tatarskii spectrum, as shown in Fig. 1b. It can be seen from this figure that decreasing the correlation length ($\sigma_g$) causes a decrease of the effective width of spectral degree of coherence.

Figure 2 shows the width of spectral degree of coherence of PCFT beam propagating in atmospheric turbulence versus propagation distance $z$. It is evident that increasing the effective size of source causes the effective width of spectral degree of coherence to decrease.

Fig. 2. The effective widths of the spectral degree of coherence for beams generated from three PCFT beams sources with different effective size labeled in the figure. The parameters are taken as: $\lambda = 632.8$ nm, $C_n^2 = 10^{-15}$ m$^{-2/3}$ and $l_0 = 5$ mm, $A_x = A_y = 1$, $B_{xy} = 0$, $\sigma_{sxx} = \sigma_{gxx} = \sigma_{gyy} = 1$ mm, $N = 4$. 

$C_n^2$ and $l_0$.
We also study the influence of an order of flatness on the behavior of the degree of coherence as well as its effective spectral width. Figure 3a shows the behaviors of degree of coherence of the PCFT beam propagation through the turbulences characterized by Tatarskii spectrum. It can be seen that increasing the order of flatness causes its effective spectral width to decrease. Figure 3b shows the effective widths of the spectral degree of coherence for beams generated from PCFT sources with different order of flatness, which propagate through atmospheric turbulence. The source parameters are taken as: $A_x^2 = A_y^2 = 1, B_{xy} = 0, \sigma_{sx} = \sigma_{sy} = 5 \text{ cm}, \sigma_{gxx} = \sigma_{gyy} = 1 \text{ mm}, N = 1, 2, 3, 4$.

Fig. 3. a – The change in the central peak of spectral degree of coherence of PCFT beam. The turbulence is described by Tatarskii spectrum. The curves correspond to different order of flatness in turbulence labeled in the figure. The parameters are taken as: $\lambda = 632.8 \text{ nm}, C_n^2 = 10^{-13} \text{ m}^{-2/3}$ and $l_0 = 5 \text{ mm}, A_x^2 = A_y^2 = 1, B_{xy} = 0, \sigma_{sx} = \sigma_{sy} = 5 \text{ cm}, \sigma_{gxx} = \sigma_{gyy} = 1 \text{ mm}, N = 1, 2, 3, 4$. b – The effective widths of the spectral degree of coherence for beams generated from PCFT sources with different order of flatness, which propagate through atmospheric turbulence. The source parameters are taken as: $A_x^2 = A_y^2 = 1, B_{xy} = 0, \sigma_{sx} = \sigma_{sy} = 5 \text{ cm}, \sigma_{gxx} = \sigma_{gyy} = 1 \text{ mm}, N = 1, 2, 3, 4$.

Another important parameter that affects the effective spectral width is turbulence. As is shown in Fig. 4, the effective spectral width decreases as $C_n^2$ increases.

Finally, we study the behaviors of the degree of coherence of the PCFT beam propagating through the turbulences characterized by Kolmogorov spectrum and Tatarskii spectrum, respectively. It can be seen that there is not much difference
Fig. 4. The effective widths of the spectral degree of coherence for beams generated from PCFT source, which propagate through atmospheric turbulence with different structure constants as labeled in the figure. The source parameters may be chosen as: $A_x = A_y = 1$, $Bxy = 0$, $\sigma_{sx} = \sigma_{sy} = 5\,\text{cm}$, $\sigma_{gxx} = \sigma_{gyy} = 1\,\text{mm}$, $N = 4$.

Fig. 5. Behaviors of the degree of coherence (a) as well as the effective width of spectral degree of coherence (b) for a PCFT beam propagating through two different turbulences characterized by Tatarskii spectrum and Kolmogorov spectrum, respectively. The beam source parameters take the same values as those used in Fig. 4. The propagation distance used in (a) is $10\,\text{km}$.
between the behaviors of the degree of coherence when using the different turbulence models. Figure 5b shows the effective widths of the spectral degree of coherence for both cases. The position that maximum value of the width happens under the Kolmogorov spectrum is closing to its counterpart under the Tatarskii spectrum. Both the results also show that choosing different turbulence models does not bring any appreciable difference in either the behavior of the degree of coherence or the effective width of the spectral degree of coherence.

4. Conclusions

In this paper, we have studied the change in the spectral degree of coherence of PCFT beam propagating through atmospheric turbulence. Both source parameters and atmospheric turbulence models influence the behavior of a beam. Here, we employ the effective width of the spectral degree of coherence to characterize the beam coherence properties on propagation in turbulent atmosphere. In free space, the effective width of the spectral degree of coherence keeps increasing along with the increasing distance, which is determined only with the source parameters, but after a special distance, the effective spectral width becomes constant. However, in the turbulent atmosphere, the change in the effective width of the spectral degree of coherence can be affected by two parts; free space part and turbulence part. The former causes the width to increase and the latter leads to the width decreasing. With the presence of the two mechanisms, the effective width of the spectral degree of coherence has its maximum value when the beam propagates a particular distance. We investigate the flat-topped order effect which makes it clear that the higher order the PCFT beams, the smaller the effective width of spectral degree of coherence. We have also analyzed the effect of correlation length; increasing correlation length causes a decrease in the spectral effective width.

The results of our analysis may find applications in problems involving optical imaging as well as atmospheric optical communication system.

References

[1] Shirai T., Wolf E., Spatial coherence properties of the far field of a class of partially coherent beams which have the same directionality as a fully coherent laser beam, Optics Communications 264(1–6), 2002, pp. 25–31.


Received August 4, 2008
Instructions for Authors

1. Papers will be considered for publication in Optica Applicata if they have not been previously published elsewhere. They should deal with original research in applied optics and may contain experimental as well as theoretical results. All papers will be subjected to refereeing.

2. Manuscripts should be typed in double spacing on one side of the page only. Printed manuscripts should be sent in duplicate to the Editorial Office. We also accept electronic submission, preferably in Word, txt or rtf format.

3. The mathematical formulae should be typed and care should be taken that the symbols easily confusing (e.g., superscripts and subscripts) are very clearly distinguished. Formulae should be numbered on the right-hand side in round parentheses.

4. Each paper should be preceded by an abstract of no more than 100 words. It must include the scope of the research and principal findings.

5. All illustrations, drawings and photographs should be submitted on separate sheets. Electronic submission as bmp or tiff format with minimum resolution of 300 dpi is preferred. Arial or Helvetica fonts type should be used (bold and italic are avoided). All illustrations, including photographs, must have the figure number, the author's name and the title of the article. Figure captions should be listed on a separate sheet at the end of the manuscript.

6. Tables should be submitted also on the separate sheets, numbered in Arabic numerals, and referred to the text. The size of a table must not be larger than A3.

7. References should be numbered consecutively as they are cited in the text and listed after text.
   a) author's name and initials,
   b) the title of the journal,
   c) journal volume, year of issue, and the first page,
   d) if a book is cited: publishing house, place, and year of issue.

8. The authors receive one copy of the manuscript (with changes suggested by the Editor) for their acceptance and this copy sent back to the Editor will be the basis for all the proof-reading made by the editorial staff. The Editors reserve themselves the right to charge the author for the alterations in the proof other than the correction of printer's errors.

9. The authors of each article will receive 25 reprints of their paper free of charge.

10. The author(s) of the paper(s) submitted to publication and published in Optica Applicata agree to exclusive transfer of the due copyright to the publisher.
Other Periodicals of the Wroclaw University of Technology

*Operations Research and Decisions* does not only discuss theoretical problems but it also provides solutions to be applied in practice. In addition to original extensive research papers, reviews, communications and reports from conferences sponsored by Polish and foreign research centres are presented.

*Studia Geotechnica et Mechanica* is a unique journal in Poland entirely devoted to theoretical and experimental problems of engineering sciences on soils and rocks.

*Environment Protection Engineering* covers the field of environmental protection engineering and its technical application: purification, wastewater treatment, water reuse, neutralization and utilization of industrial gases, noise and vibrations, ecological problems, solid waste disposal, legislation and forecasting dealing with the environmental problems.

*Materials Science* is an interdisciplinary journal devoted to experimental and theoretical research into the synthesis, structure, properties and applications of materials.

*Systems Science*. The papers published in this periodical are devoted to a general theory of systems, their mathematical models as well as to various applications of the analytical methods to industry, information science, biology, and other disciplines.

Orders should be addressed to:
Oficyna Wydawnicza Politechniki Wrocławskiej
Wybrzeże Wyspińskiego 27, 50-370 Wrocław
CHZ „Ars Polona” S.A., Dział Prasy, ul. Obrónców 25, 00-933 Warszawa
„Ruch” S.A., Oddział Krajowej Dystrybucji Prasy
ul. Jana Kazimierza 31/33, 01-248 Warszawa
or to any „Ruch” agency in Poland

PL ISSN 0078-5466