ISSN: 1947-3931 (Print), 1947-394X (Online) Volume 2, Number 1, January 2010

Engineering

Prof. David L. Carroll Wake Forest University, USA

www.scirp.org/journal/eng/

Scientific Research

Journal Editorial Board

ISSN: 1947-3931 (Print), 1947-394X (Online)

http://www.scirp.org/journal/eng

Engineering, **2010, 2, 1-84**
Published Online January 2010 (http://www.scirp.org/journal/eng).

TABLE OF CONTENTS

Engineering

Journal Information

SUBSCRIPTIONS

Engineering (Online at Scientific Research Publishing, www.SciRP.org) is published quarterly by Scientific Research Publishing, Inc., USA.

E-mail: eng@scirp.org

Subscription rates: Volume 2 2010

Print: $50 per copy.

Electronic: free, available on www.SciRP.org.

To subscribe, please contact Journals Subscriptions Department, E-mail: eng@scirp.org

Sample copies: If you are interested in subscribing, you may obtain a free sample copy by contacting Scientific Research Publishing, Inc at the above address.

SERVICES

Advertisements

Advertisement Sales Department, E-mail: eng@scirp.org

Reprints (minimum quantity 100 copies)

Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.

E-mail: eng@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:

E-mail: eng@scirp.org

Engineering, **2010, 2, 1-11**
doi:10.4236/eng.2010.21001 Published Online January 2010 (http://www.scirp.org/journal/eng/).

Using Microgripper in Development of Automatic Adhesive Glue Transferring and Binding Microassembly System

R. J. CHANG[1], C. C. CHEN[2]
[1]*Department of Mechanical Engineering, National Cheng Kung University, Tainan, Taiwan, China*
[2]*Micro System Technology Center, Industrial Technology Research Institute, Tainan, Taiwan, China*
E-mail: *rjchang@mail.ncku.edu.tw, Daniel_chen@itri.org.tw*
Received August 13, 2009; *revised September*12, 2009; *accepted September* 20, 2009

Abstract

A system using microgripper for gluing and adhesive bonding in automatic microassembly was designed, implemented, and tested. The development of system is guided by axiomatic design principle. With a compliant PU microgripper, regional-edge-statistics (RES) algorithm, and PD controller, a visual-servoing system was implemented for gripping micro object, gluing adhesive, and operating adhesive bonding. The RES algorithm estimated and tracked a gripper's centroid to implement a visual-servoing control in the microassembly operation. The main specifications of the system are: gripping range of 60~80μm, working space of 7mm×5.74mm×15mm, system bandwidth of 15Hz. In the performance test, a copper rod with diameter 60μm was automatically gripped and transported for transferring glue and bonding. The 60μm copper rod was dipped into a glue container and moved, pressed and bonding to a copper rod of 380μm. The amount of binding glue was estimated about 5.7nl.

Keywords: Micro Gripper, Adhesive Bonding, Microassembly, Visual Servo

1. Introduction

In the manufacturing cycle of microsystems, assembly is a crucial operation in production [1]. The assembly tasks are varied according to the areas and methods of microproducts in manufacturing. There are several methods such as anodic bonding, eutectic bonding, welding bonding and adhesive bonding which can be utilized for micro joining [1,2]. Anodic bonding, eutectic bonding, and welding bonding usually are achieved under some critical conditions, such as high temperature and pressure. In comparing with other methods, adhesive bonding has the advantages of free heating, cleanness, speediness and no influence on parts.

In the application of adhesive bonding, a traditional method is to use a glue dispenser [3] for the injection of a micro drop on the binding surface of a component. Then, another component is applied and pressed for bonding. By using a dispenser, it is very difficult to apply small amount of glue on a specific small surface of a micro component [3]. The limitation of injection of dispenser and operational working space makes the technique of micro joining most severe in the production of Microelectromechanical System (MEMS). Instead of using a dispenser, an effective technique for micro adhesive bonding is to transfer glue by holding and dipping a slender tool into a glue container and applying it to the surface for bonding [2]. However, a direct approach by utilizing a manipulator to grasp and hold a component instead of using a tool can be applied for transferring glue in the micro adhesive bonding. By utilizing a manipulator for the adhesive bonding operation, it is most effective to hold and dip the surface of a component into a glue container and to move the component touching, pressing, and bonding to another component.

Recently, the automation of microassembly has become an important technology which attracts many investigators [4–10]. Although there are several different designs in the microassembly system, a systematic engineering approach from conceptual design to realization is still not proposed. Actually, in the development of an automatic microassembly system, the stringent engineering requirements such as small size, high reliability, clean operation, fast response, and high accuracy usually make a conventional engineering scheme ineffective in the process from design to manufacturing. Systematic tools in realization concerning the relationship between design and manufacturing are most useful to assure ef-

fective development of the systems [11,12].

In realizing an automatic microassembly system, a visual control through microscope has been a challenging task. Actually, visual servo has been a viable method in industry for assembly and material handling jobs. Since the early contribution of Shirai and Inoue, considerable efforts have been devoted to develop visual control manipulators in manufacturing [13]. There are many visual estimation algorithms developed for tracking control [14]. These algorithms mainly include optical flow method, sum of square difference (SSD) method, model-based method, motion energy method, and template matching method. Although template matching method is widely employed in automatic microassembly system [6–9], different algorithms usually have their advantages and disadvantages in different applications. The issues of robustness, resolution, and efficiency have been identified for further improvement of algorithms in the visual tracking applications [15].

In the present research, a visual-servoing automatic system with microgripper for adhesive bonding is developed and tested. By employing the precision design axioms on system design, a visual-servoing automatic microassembly system is developed to achieve the requirements of adhesive bonding in micro manipulations. For achieving the operation of precise and accurate gripping and transportation of micro objects in assembly task, a compliant microgripper actuated by piezoelectric actuator is designed and fabricated. In order to provide an efficient, robust and accurate estimation for the automation of microassembly operation, an algorithm of regional edge statistics (RES) is developed and implemented. For implementing a closed-loop control, the system model will be identified and the controller is synthesized. The performance of a prototype microassembly system is tested. Finally, the research on the development of the visual-servoing microassembly system is concluded.

2. Preliminary Design Consideration

The major steps in system development consist of conceptual design, preliminary design, and detail design [11]. Conceptual design forms the fundamental step in the whole development process and a physical realization should capture the essence of the conceptual design. Because of the lack of detailed information in the early development process, a prescriptive model to aid decision making is most useful in system development [16]. For the effective development of a microassembly system under the stringent engineering constraints such as small size, high reliability, clean operation, fast response, and high accuracy, a precision design method guided by axiomatic design principle is employed to aid decision making in the system development [12].

2.1. Design Objective and Constraints

In the area of microassembly, there are numerous functional principles which can be applied for the manipulation of micro objects [1,17]. In the present design, the objective is to develop an automatic microassembly system utilizing mechanical micro gripper to handle bonding of micro parts in the clean room of MEMS industries. For the microassembly operation, a reliable automatic operation of picking, gluing, attaching, and binding is required. No lubrication and no wear are essential constraints for the clean room operations. The objects or tools to be picked up, glued, and transported for assembly are slender hard components with diameter or width around 60~80μm. The bonding of micro parts is to be operated in room temperature around 25 ℃. Under the constraints of limited working space, micro fabrication technology, geometrical configuration of gripper, and the size of micro object, the planar size of gripper mechanism with width about 500μm and length about 700μm is required.

2.2. Design Principle

For the development of a microassembly system, the optimal design of the microassembly system is to achieve two design axioms in the mappings from Functional Requirements (FRs) to Design Parameters (DPs) and from DPs to Process Variables (PVs). The two design axioms are stated as follows [12,18]:

Axiom 1 (Functional Independence). An optimal design of a micro assembly system must maintain the independence of functional requirements of subsystems.

Axiom 2 (Information Minimization). The best design of a micro assembly system is a design of functionally independent subsystems with minimum information content. Here, the information content is a measure of uncertainties in physical realization of the design specifications of a system and subsystems.

The microassembly system is to be realized under the guidance of precision design axioms. By employing the axiom of Functional Independence, a micro assembly system can be designed and realized with the merits of independent module design, independent functional testing, and degrees-of-freedom in controller implementation. With the axiom of Information Minimization in the processes of design, assembly, and manufacturing, a micro assembly system can be effectively realized to satisfy the stringent requirements in assembly operations.

2.3. Conceptual Design

For realizing an adhesive-bonding operation by using a micro gripper, a conceptual design on the microassembly system is described. The configuration of the microm assembly system is designed by following the axiomatic design principle. The highest FRs of the microassembly

system is identified as independent functions: FR1= Gripping and releasing of object, FR2= Carrying gripper and components in assembly operation, and FR3= Acquiring working states during bonding process. The corresponding DPs of the microassembly system is identified as DP1= Micro Gripper System, DP2= Working Stages, and DP3= Visual System. The functional mapping between the FRs and DPs in the first level can be formulated as (1):

$$\begin{bmatrix} \text{FR1} \\ \text{FR2} \\ \text{FR3} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \text{DP1} \\ \text{DP2} \\ \text{DP3} \end{bmatrix} \quad (1)$$

where the matrix by $[a_{ij}]$ is to be characterized in the physical realization of the microassembly system. The mapping between FRs and DPs in (1) satisfies a decoupled module design. In addition, the micro gripper system can be uniquely designed without considering the effects of other modules. By following the axiom of Functional Independence, the functional mapping between the FRs and DPs can be realized by utilizing sensor, actuator, mechanism, and controller. The decomposition of FRs is to be unique and independent and will be used as the basis for formulating DPs which are corresponding to FRs, respectively. In the decomposition and mapping, the branches of FRs are mapped into physical domain to develop lower levels of DPs as:

(1) Micro Gripper System

FR1-1: Gripper opening to fit object size

FR1-2: Gripper closing and releasing

FR1-3: Afford gripping force

DP1-1: Gripper Mechanism

DP1-2: Gripper Controller

DP1-3: Gripper Actuator

(2) Working Stages

FR2-1: Carry object or tool for transferring glue

FR2-2: Moving gripper

FR2-3: Carry glue container and assembly part for bonding

DP2-1: Object/Tool Stage

DP2-2: Gripper Carrier Stage

DP2-3: Glue Container and Assembly Part Stage

DP2-4: Controller

(3) Visual System

FR3-1: Monitor and control gripper working states

FR3-2: Automatic control the motion of gripper stage

FR3-3: Acquire image for tracking

DP3-1: User Interface

DP3-2: Tracking Algorithm

DP3-3: CCD Subsystem

A system structure of the microassembly system to satisfy the FRs by synthesizing the corresponding DPs is depicted as Figure 1. From Figure 1, the subsystem consists of three modules as Micro Gripper System, Working Stages, and Visual System.

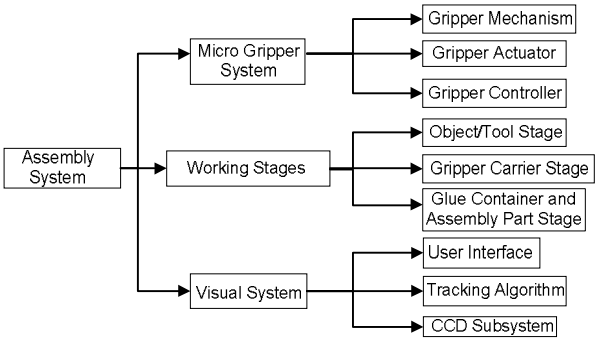


Figure 1. System structure.

2.4. Design Procedure and Constraints on Design Parameters

For the development of a microassembly system, the design procedure in general can be obtained from (1). The mapping between FRs and DPs in (1) can be decoupled and realized starting from a_{11}. From the mapping by a_{11}, the Micro Gripper System will be first realized. With the Micro Gripper System and a_{21}, the a_{22} is obtained by realizing Working Stages to satisfy the functional requirement. With the mapping set up by a_{11}, a_{21}, a_{22}, a_{31}, and a_{32}, the a_{33} is obtained finally by realizing Visual System to satisfy the functional requirement. As a result, the design procedure is given by

$$\text{FR1} \rightarrow \text{DP1} \rightarrow \text{FR2} \rightarrow \text{DP2} \rightarrow \text{FR3} \rightarrow \text{DP3}$$

By following above design procedure, the design constraints in implementing DPs will be analysed. The DP1 as Micro Gripper System is constructed by Gripper Mechanism, Gripper Controller and Gripper Actuator. In the design and realization of a Micro Gripper System, the selections of material, manufacturing processes, system configuration, and components assembly are crossly related. From the requirement of design objective, the constraints of DPs in the design of Micro Gripper System can be identified as clean environment, precision dimension, micron operation, and gripper size. In the process domain, the PVs are stated as material selection, fabrication method, configuration, and assembly works. The functional mapping between the constraints and PVs can be formulated as (2):

$$\begin{bmatrix} \text{Clean environment} \\ \text{Precision dimension} \\ \text{Micron operation} \\ \text{Gripper size} \end{bmatrix} = \begin{bmatrix} B_{ij} \end{bmatrix} \begin{bmatrix} \text{Material selection} \\ \text{Fabrication method} \\ \text{Configuration} \\ \text{Assembly works} \end{bmatrix} \quad (2)$$

The B_{ij} in (2) in general is not independent and constraints in (2) are to be considered in the physical realization of the Micro Gripper System.

2.5. Design Features by Employing Axiomatic Design

The design axioms are employed first for the design of a Micro Gripper System. For the micron scale of a Gripper Mechanism to be operated in a clean room, a one-piece compliant gripper, without assembly works, is selected and designed to provide accurate tip motion and gripping force transmission. The design of the Gripper Mechanism satisfies the axiom of Information Minimization. In the realization of assembling Gripper Mechanism and Gripper Actuator, physical uncertainties are to be minimized by the axiom of Information Minimization. A packed piezoelectric actuator with very rigid steel case is selected since the case can be employed as a structural frame to align and install Gripper Mechanism. The design feature of independent Gripper Mechanism and Gripper Actuator satisfies the axiom of Functional Independence.

For the design consideration of Working Stages, the precision of Working Stages for microassembly is required to be one micron since the working range of gripper belongs to micron scale. The selection of three independent stages in industrial use satisfies the axioms of Functional Independence and Information Minimization.

A Visual System is composed of User Interface, Tracking Algorithm, and CCD Subsystem. The CCD Subsystem consists of CCD, Illumination Light, Microscopic Lenses, and Image Processing Card. The selection of Visual System satisfies the design axioms of Functional Independence and Information Minimization.

In the final consideration of control hardware and software, a friendly software tool is to be selected to implement a User Interface on Display for system monitoring and control. For the consideration of real-time and precise processing of image signal and piezoelectric control signal, a functional independent PC with Digital Signal Processor (DSP), and interface cards are selected.

3. Gluing-Adhesive Microassembly System

Based on the conceptual design, a system structure of image-based automatic gluing-adhesive system to realize the DPs is shown in Figure 1. The detail design of micro gripper system, working stages, and visual system is described in the following sections.

3.1. Micro Gripper System

Micro Gripper System consists of Gripper Mechanism (DP1-1), Gripper Actuator (DP1-3), and Gripper Controller (DP1-2). At first, the design of Micro Gripper system will be considered to satisfy FR1-1 to FR1-3. The Micro Gripper System is expected to grip a slender object with diameter or width around 60~80μm. For the object with slenderness ratio, characteristic length to width ratio, above 2 and in dry condition, the micro sticking force will not be an issue in object releasing operation. Therefore, the design and fabrication of the Micro Gripper System is expected to provide the operating functions of precise and accurate gripping of micro objects.

Compliant polymer micro grippers have been designed to provide accurate tip motion and sufficient gripping force in a clean room operation [18–20]. The selection of a lumped-compliant gripper mechanism is to minimize the effects of creeping, relaxation, and hysteresis loop due to polymer material [18]. A lumped-compliant gripper mechanism is composed of pseudo linkages and compliant joints. By following the axiom of Information Minimization, the topological structure of the Gripper mechanism is designed to be symmetric with minimum number of compliant joints. The micro compliant gripper mechanism and its associated PLM (Pseudo linkages model) [20] is designed as shown in Figure 2. In Figure 2, the contour line shows a geometrical shape of the Gripper Mechanism and its structural frame. The PLM is modeled as a six-linkage mechanism for providing one degree of freedom in input-output motion. When an actuating force F is applied, the actuator will drive link 4 to produce displacement. Due to the constraints of structure frame 1, both link 3 and 5 will rotate and translate to cause the gripping operation by tips C and C'.

The assumption of small deformation is used in the following derivation. From Figure 2, the horizontal and vertical displacement gains can be derived as the ratio between the tip displacement Δx , Δy and input displacement Δi of slider 4, respectively as

$$G_x = \frac{\Delta x}{\Delta i} = \frac{L_2 \cos b}{L_1} \tag{3}$$

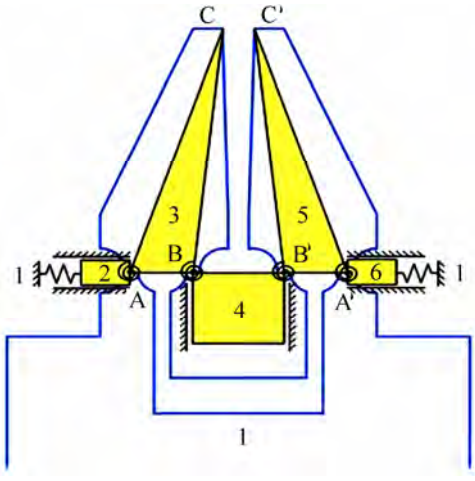


Figure 2. PLM of kinematic relation of microgripper.

and

$$G_y = \frac{\Delta y}{\Delta i} = \frac{L_2 \sin b}{L_1} \tag{4}$$

where $L_1 = \overline{AB}$, $L_2 = \overline{AC}$ and β is the angle between a vertical line and $\overline{AC}$. The ratio between the horizontal and vertical displacements is

$$\frac{\Delta x}{\Delta y} = \cot b \tag{5}$$

From (3) to (5), there are two important parameters G_x and β to be designed. For achieving the function of stable operation, the horizontal displacement is required to be much larger than the vertical displacement in gripping. Therefore, it is expected to have smaller β from (5). For the compliant gripper, the lateral stiffness of gripper arms against pay load needs to be taken into consideration. It is essential when a gripper is fabricated by using elastic polymer material. The lateral stiffness depends on the material and geometry of a gripper. The ratio of L_2 to L_1 in (3) and (4) gives a measure of the lateral stiffness for a fixed thickness in gripper. The ratio is smaller for higher lateral stiffness.

The formulation from input force F to output horizontal displacement can be derived by employing the conservation of energy. In closing operation, ideally, the input work is equal to the strain energy stored by the compliant joints to give

$$F \times \Delta i = 2 \times (\frac{1}{2} k_{1s} \Delta s_1^{\,2} + \frac{1}{2} k_A \Delta\, \theta^2 + \frac{1}{2} k_B \Delta\, \theta^2) \tag{6}$$

Since $\Delta\theta = \frac{\Delta i}{L_1}$ and $\Delta s_1 = L_1 \sin(\Delta\theta) = \Delta i$, then

$$F \times \Delta i = 2 \times (\frac{1}{2} k_{1s} \times (\Delta i)^2 + \frac{1}{2} k_A \times (\frac{\Delta i}{L_1})^2 + \frac{1}{2} k_B \times (\frac{\Delta i}{L_1})^2)$$
$$= (\frac{L_1^{\,2} k_{1s} + k_A + k_B}{L_1^{\,2}}) \Delta i^2 \tag{7}$$

Equation (7) can be expressed to give

$$F = K_{eq} \Delta i \tag{8}$$

where K_{eq} is the equivalent stiffness as

$$K_{eq} = \frac{L_1^{\,2} k_{1s} + k_A + k_B}{L_1^{\,2}} \tag{9}$$

The optimal shape and size of the microgripper is synthesized based on the following considerations. The size of gripping object is of primary consideration. The maximum size of outlines and the minimum size of the compliant joints are constrained by the capability of micro fabrication. The design objective is to trade off the parallel gripping and maximum horizontal displacement gain under the geometrical constraints. For gripping a hard object with size 60μm, the present design is to trade off the lateral stiffness and horizontal gripping to give $G_x = 4$, and $b = 21°$. The selection of thickness of the Gripper Mechanism is constrained by the fabrication material, micro fabrication techniques, and the lateral stiffness of the gripper. After synthesizing the geometrical size of the gripper, the shape and size of the opening of gripping surfaces can be further modified. For helping the gripping object fed into the opening of gripping surfaces and maintaining parallel gripping, an opening with 2° in slope and round tips are designed. With the constraint of the geometrical size 500μm×700μm, the final geometrical shape of the Gripper Mechanism is designed and to be fabricated to give one tenth of a mask as shown in Figure 3.

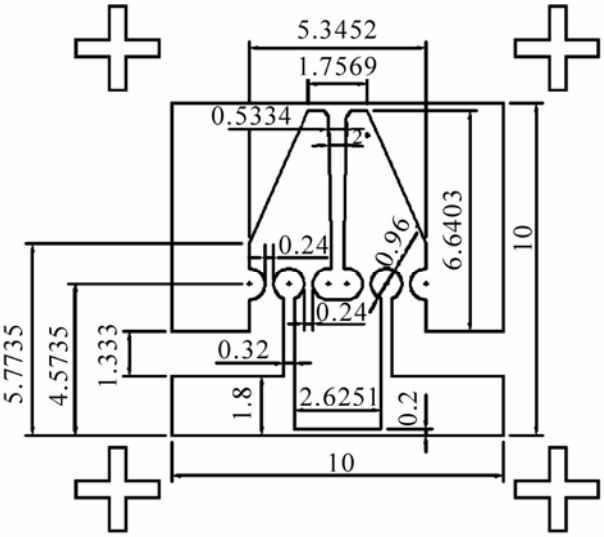


Figure 3. Mask designed for fabricating microgripper (Unit: mm).

The detail design of a Gripper Mechanism based on the axiomatic design principle satisfies the FRs of the conceptual design. By employing axiomatic design as a tool of robust design in considering the relationship between design and manufacturing, the stringent functional requirements of a micro gripper mechanism is to be realized effectively. The design performance of the Gripper Mechanism will be further analysed. The Gripper Mechanism is to be fabricated by using a Polyurethane (PU) film of thickness 0.2mm. The Young's modulus of the PU is E=7.775×10^{7}N/m. By employing finite element analysis through Ansys with planar linear elastic model, the stress-strain behavior of the gripper is analysed. From the results by Ansys and through the equivalent model by PLM, the horizontal gain is obtained as 3.85 and the equivalent stiffness is 62.2N/m. The error of horizontal gain and equivalent stiffness between the Ansys results and PLM design is 3.75% and 5%, respectively. These

results support the validity of utilizing PLM in design.

The fabrication of the micro Gripper System is considered by following the axiom of Information Minimization. The PU micro Gripper Mechanism is fabricated by employing a mask projection method instead of a direct write method. By employing an excimer laser, Exitech 2000, the fabricated micro Gripper Mechanism through mask is reduced by the optical lens of 10x. The fabricated Gripper Mechanism with Gripper Actuator is assembled through a metal coupler by utilizing simple hand tools and with adhesive glue. The Gripper Actuator with Gripper Controller can provide 10μm linear displacement. With much higher stiffness of Gripper Actuator to drive the lower stiffness of Gripper Mechanism, the gripper assembly can provide sufficient gripping force in the gripping range 60μm to 80μm by input 0 to 80 volts from the Gripper Controller.

3.2. Working Stages

Working Stages include Object/Tool Stage (DP2-1), Gripper Carrier Stage (DP2-2), and Glue Container and Assembly Part Stage (DP2-3). By following the guidance of design axioms, these stages are implemented by selecting industrial products to satisfy FR2-1 to FR2-3. Two three-axes stages for Object/Tool Stage and Glue Container and Assembly Part Stage are obtained from Newport. The three-axes stages are driven by stepping motors and can provide 1.3μm resolution with 15mm stroke. Gripper Carrier Stage obtained from Physik Instrumente is implemented to give 0.5μm resolution with 15mm stroke for planar X-Y motion.

3.3. Visual System

Visual System consists of User Interface (DP3-1), Tracking Algorithm (DP3-2), and CCD Subsystem (DP3-3). By following the guidance of Information Minimization in implementing an image acquisition system for visual servo, the installation of Visual System is to adopt Endpoint Closed Loop (ECL) scheme and the structure of visual servo is to adopt the Dynamic Image-Based Look-and-Move (DIBLM) scheme.

For image acquisition to satisfy FR3-3, hardware is set up to include JAI CVS3200 CCD with lens, TI vDB image grabber card, and TI DSP320C6711 DSK board. The imaging model adopts affine projection. In this case, a point with coordinates as ${}^{c}\mathbf{P}$, expressed with respect to the Cartesian coordinate frame of a camera, will project onto an image plane as $[x\ \mathrm{y}]^T$ by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{A}\,{}^{c}\mathbf{P} + \mathbf{c} \tag{10}$$

where **A** and **c** are calibrated 2x3 and 2x1 matrix, respectively. The model is purely linear. The **A** and **c** are easily computed by using linear regression techniques. Since affine projection does not correspond to any specific imaging situation, the issue of camera calibration is greatly simplified. The affine projection can be simplified to a scaled orthographic projection if **A** is reduced to a scalar and **c** is zero.

The DIBLM scheme provides several advantages. First, the system with DIBLM structure doesn't need coordinate transformation and consequently, the control performance is independent of the calibration. Second, it is more efficient in computation. Thus, the computational load on control structure is not stringent in implementation. Third, it is suitable for the present realization of a planar visual-servoing microassembly operation.

The design axioms provide an efficient and effective tool to help logical reasoning and decision making in obtaining an optimal design of the microassembly system. The system framework of hardware installation is shown in Figure 4. The image-based automatic operation of gluing-adhesive bonding is to implement control software on a PC and through a motion control card, NI-3744 and communication interface, RS-232. LabView is selected to implement control windows for DP3-1. The selection of LabView for implementing control windows to satisfy FR3-1 is based on the guidance of Minimum Information. For the microassembly system, all the assembly works are operated under the field of view of a visual system by DP3-1 to DP3-3. The detail designs of DP3-1 and DP3-3 were described in this section. A tracking algorithm to realize DP3-2 will be developed in next section.

4. Tracking Algorithm

The performance of an algorithm which tracks an object from the features of its image plays a key role to implement a visual-servoing system. In general, edges are most essential features of an object employed in image tracking. If the shape of an object has clear edge features in the image, the geometric features such as length, width, area, and centroid, depending on the representative of these edge features, can be fully or partially determined to construct a feature space of the object. An object with symmetrical geometry in image is a typical object that the edge is representative of its centroid. For estimating the geometric centroid of an object from the edge features of the image, it is required that the image is under conditions:

1) The edge of an object is representative of its centroid.

2) There is no high frequency spatial noise in background.

By employing the concepts of interesting region and edge feature, a dynamic image tracking algorithm, Regional Edge Statistics (RES), is developed. In the initial time instant k, the initial center of a tracking region was defined as the centroid of the micro gripper and repre-

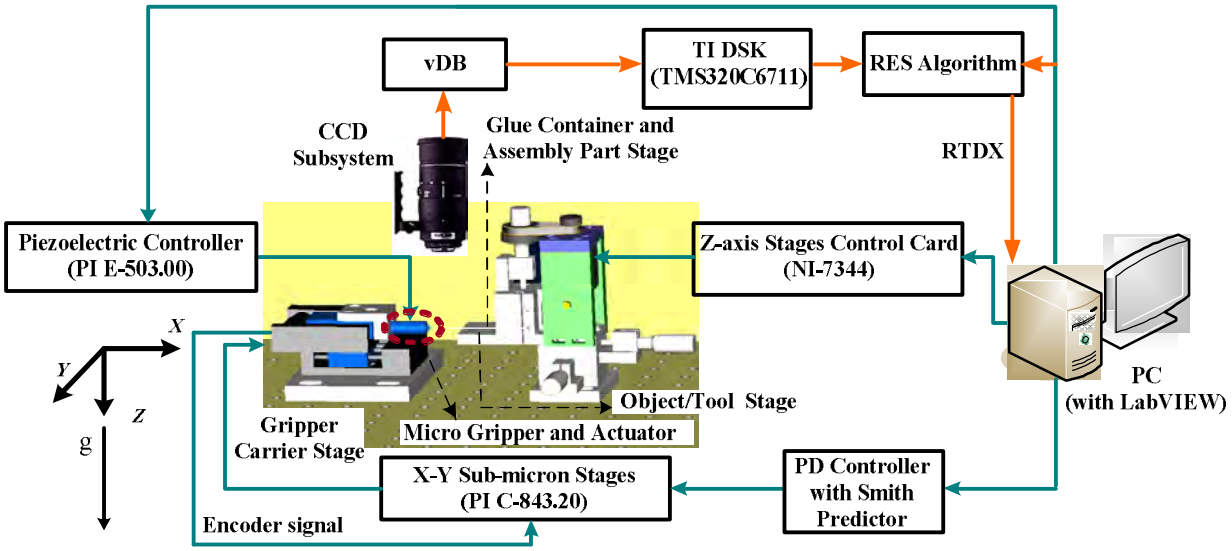


Figure 4. System framework of hardware installation.

sented as $(X_c(k),Y_c(k))$ in RES algorithm. The height and the width of the tracking region are determined by the height L_o and width W_o of the microgripper, the maximum velocities (V_v, V_h) of the microgripper, and the sampling period T in moving one step. The height L and the width W of the tracking region are given as (11) and (12), respectively,

$$L = L_o + V_v T \tag{11}$$

$$W = W_o + V_h T \tag{12}$$

The limits of boundary of the tracking region in height are $X_{\max}(k)$ and $X_{\min}(k)$ as shown in (13) and (14), respectively,

$$X_{\max}(k) = X_c(k) + (L/2) \tag{13}$$

$$X_{\min}(k) = X_c(k) - (L/2) \tag{14}$$

The limits of boundary of the tracking region in width are $Y_{\max}(k)$ and $Y_{\min}(k)$ as shown in (15) and (16), respectively,

$$Y_{\max}(k) = Y_c(k) + (W/2) \tag{15}$$

$$Y_{\min}(k) = Y_c(k) - (W/2) \tag{16}$$

The present algorithm utilizes a 2-D mask to extract the edge features of the target on an image. For a 2-D mask with height $2a$ and width $2b$, the image response $G(x,y,k)$ with a mask $W(s,t)$ on an original image $F(x,y,k)$ can be expressed as

$$G(x,y,k) = \sum_{s=-a}^{a}\sum_{t=-b}^{b} W(s,t)F(x+s,y+t,k) \tag{17}$$

The horizontal and vertical Sobel masks $(W_{Sobel_v}(s,t), W_{Sobel_h}(s,t))$ with given appropriate threshold values $(V_{Sobel_v} V_{Sobel_h})$ are utilized, respectively, to extract the horizontal and vertical edge features of the microgripper. The vertical edge features extracted is given by (18) or (19),

$$E_v(k) = \left\{ \begin{matrix} e_v(k) \mid G(x,y,k) > V_{Sobel_v}, \\ X_{\min}(k) < x < X_{\max}(k), Y_{\min} \\ (k) < y < Y_{\max}(k) \end{matrix} \right\} \tag{18}$$

$$E_v(k) = \left\{ e_v(k) \mid \sum_{s=-a}^{a}\sum_{t=-b}^{b} W_{Sobel_v}(s,t)F(x+s,y+t,k) > V_{Sobel_x}, \right.$$
$$\left. X_{\min}(k) < x < X_{\max}(k), Y_{\min}(k) < y < Y_{\max}(k) \right\} \tag{19}$$

The horizontal edge features extracted is given by (20) as

$$E_h(k) = \left\{ e_h(k) \mid \sum_{s=-a}^{a}\sum_{t=-b}^{b} W_{Sobel_h}(s,t)F(x+s,y+t,k) > V_{Sobel_y}, \right.$$
$$\left. X_{\min}(k) < x < X_{\max}(k), Y_{\min}(k) < y < Y_{\max}(k) \right\} \tag{20}$$

From (19) and (20), the set of edge features of the microgripper is extracted to give $E(k)$ as

$$E(k) = \{ e(k) \mid e(k) \in E_v(k) \cup E_h(k) \} \tag{21}$$

The RES algorithm finally utilizes the set of edge features to estimate the centroid of an object $(X_o(k),Y_o(k))$ as given by the following two equations,

$$X_0(k) = \frac{1}{c_v + c_h}\left(\frac{\sum_{i=1}^{M} x_{v_i}^e}{M} c_v + \frac{\sum_{j=1}^{N} x_{h_j}^e}{N} c_h \right) \tag{22}$$

$$Y_o(k) = \frac{1}{c_v + c_h}\left(\frac{\sum_{i=1}^{M} y_{v_i}^e}{M} c_v + \frac{\sum_{j=1}^{N} y_{h_j}^e}{N} c_h\right) \quad (23)$$

In (22) and (23), the $(x_{v_i}^e, y_{v_i}^e)$, and $(x_{h_j}^e, y_{h_j}^e)$ are the ith and jth image coordinates of elements in the vertical and horizontal edge features, respectively. The C_v, C_h which are between 0 and 1 are given as the statistical weighting coefficients of vertical and horizontal features, respectively. The numbers of elements of vertical and horizontal edge features, respectively, as *M* and *N* are counted in the procedure of edge extraction. When the object is moving one step, the center of tracking region in next time instant *k*+1 is replaced by the centroid which was estimated in time instant *k* as (24) and (25),

$$X_c(k+1) = X_o(k) \quad (24)$$

$$Y_c(k+1) = Y_o(k) \quad (25)$$

The application of RES algorithm for tracking the motion of a microgripper is illustrated. For the present microassembly system, the microgripper for RES algorithm satisfies two conditions described above. The images of a microgripper under RES are shown in Figure 5. Figure 5(a) is an original image of the microgripper. In Figure 5(b), the gray region is shown as a tracking region. The horizontal edge features are extracted with $C_v = 0$ as shown in white. In Figure 5(c), the bright white spot is the centroid extracted. Figure 5 reveals that the centroid of the microgripper as a tracking target can be estimated accurately by employing the RES algorithm.

The RES method is compared with other methods. Since a template matching method has good performance in robustness and resolution, it is selected for the comparisons. The simulated results of RES, Gray-Scale Correlation (GSC), Normalized Gray-Scale Correlation (NGSC), and Normalized Gray-Scale Correlation with Image Pyramid (NGSC-IP) methods are shown in Table. 1. From Table 1, it is observed that the RES method can provide high resolution as those of GSC and NGSC. However, the computational efficiency is much improved compared with GSC related methods.

A template matching method utilized GSC related methods to find out a position which gives maximum correlation between a template and image. The position is then

(a) Original image **(b) Edge features** **(c) Centroid extracted**

Figure 5. Images of microgripper under RES.

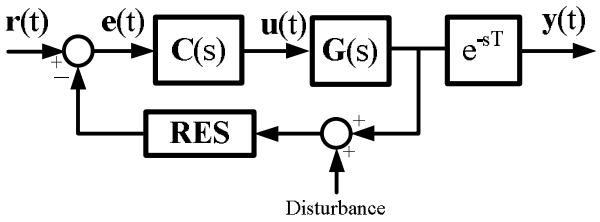


Figure 6. Block diagram with ideal Smith predictor.

Table 1. Comparison of RES and template matching related methods.

	Template size (pixel)	Resolution (pixel)	Computation load (second)	Tracking area (pixel)
RES	51x45	1	0.01	11x11
GSC	51x45	1	0.03	11x11
NGSC	51x45	1	0.06	11x11
NGSC-IP	51x45	2	0.11	31x31

used to track any target through the known template. If the selected template has no apparent feature or template has noisy feature on the image, the GSC related methods may fail to find out the correct position of the selected template. However, the RES algorithm can statistically collocate weighting coefficients to improve the robustness and computational load. The positioning accuracy is improved by the high resolution in image tracking. Actually, if a tracking region always covers the target on an image, the feature-based RES algorithm is expected to track the centroid of the target fast and precisely.

5. System Modeling and Controller Design

A visual-based automatic microassembly system utilizing the architecture of DIBLM is implemented. The present system utilizes the RES algorithm to estimate and track the microgripper for assembly task. The size of the tracking region by the RES algorithm is 51x45 pixels. The modeling and controller design for the system will be described.

5.1. System Modeling

In the DIBLM architecture, the stage controller, PI C-843.20, and the sub-micron stage, M1-111.DG, are considered as a plant. Two axes of the sub-micron stage are orthogonal and can be modeled independently. The image controller is denoted as $C(s)$. In the preliminary step-response tests of both axes with $C(s) = 3$, the responses showed dead-time delays and overshoots. By employing an ideal model of Smith predictor [21], the system architecture is simplified as shown in Figure 6.

For the X-axis, the step-response test with $C_x(s) = 3$

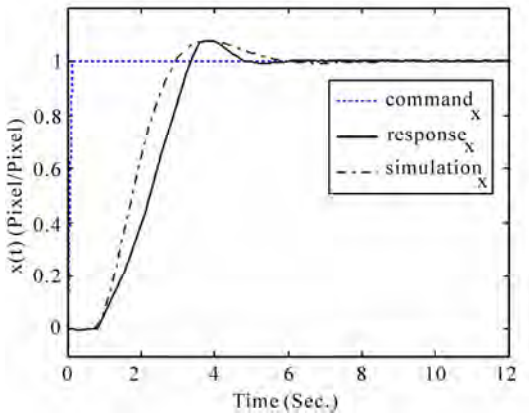


Figure 7. Normalized step response of X-axis servo.

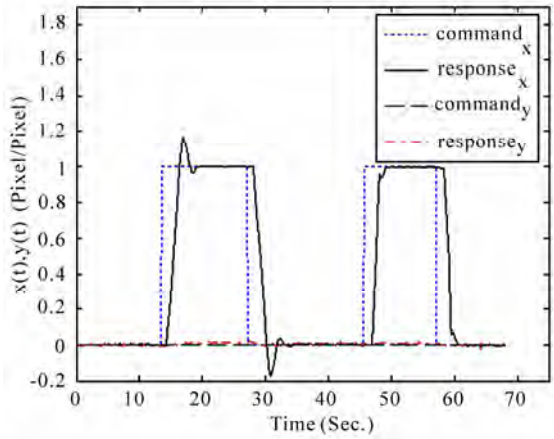


Figure 8. Normalized step response of X-axis servo without employing PD and then with PD controller after 40 seconds.

is shown in Figure 7 by a solid line. The overshoot is about 7.68% and the delay time is 0.695 seconds. The peak time is about 3.125 seconds. By assuming the transfer function of a plant as $G_x(s)e^{-sT_x}$ and utilizing a second-order model to approach the step response, one derives damping ratio $V = 0.633$ and natural frequency $w_n = 1.299$. The closed-loop transfer function of X-axis is presented as $T_x(s)e^{-sT_x}$ by (26),

$$T_x(s)e^{-sT_x} = \frac{1.687}{s^2 + 1.644s + 1.687}e^{-0.692s} \quad (26)$$

The plant is derived to give $G_x(s)e^{-sT_x}$ as (27),

$$G_x(s)e^{-sT_x} = \frac{0.562}{s(s+1.644)}e^{-0.692s} \quad (27)$$

A closed-loop step response of the model is compared with that of experimental test as shown in Figure 7. From Figure 7, the simulated result does not fit very well with the experimental one. A higher order model may be utilized to improve the modelling error. However, the identified model fits experimental result accurately in delay time, peak time, and overshoot. These specifications are actually concerned in the present gluing and assembly operations.

By following the same procedure as the X-axis for modeling and testing of the Y-axis in the sub-micron stage, the closed-loop transfer function is obtained to give

$$T_y(s)e^{-sT_y} = \frac{4.071}{s^2 + 2.381s + 4.071}e^{-0.81s} \quad (28)$$

The plant is derived to give

$$G_y(s)e^{-sT_y} = \frac{1.018}{s(s+2.381)}e^{-0.81s} \quad (29)$$

5.2. Controller Design

In the present microassembly operation, the task is to use a microgripper gripping an object, gluing adhesive, and bonding with another object automatically. Organic glue which can be hardened in room temperature around 25℃ is utilized for the assembly task. For the assembly operation, the desired control performance is described in the following:

1) The overshoot in response is minimized since it will cause collision to make the assembly fail.

2) Response must be fast enough to avoid the glue hardening.

In order to reduce the overshoot, a PD controller is selected. For the servo in the X-axis, the closed-loop transfer function with PD controller becomes

$$T_x(s)e^{-sT_x} = \frac{0.562(D_x s + P_x)}{s^2 + (1.644 + 0.562D_x)s + 0.562P_x}e^{-0.692s} \quad (30)$$

The PD controller is finally designed as $C_x(s) = 4 + 1.21s$ to minimize the overshoot and increase the response speed as shown in Figure 8. For the Y-axis servo, the PD controller is designed as $C_y(s) = 1.9 + 0.367s$ to satisfy the control objective.

The stability robustness due to modelling error ΔG needs to be investigated. Since the plants of both axes are type 1, the gain margins should be infinite if the time delays are zero. The phase margins of the servo system due to delay time in both axes are analyzed. If both axes have no delay time, these X-axis and Y-axis servos can provide phase margins of 78.5 and 80 degrees, respectively. Therefore, for the two axes of a servo system operated with natural frequency of 1.5rad/second, the system is stable with sufficient margin.

By the present servo design on both axes, there is no overshoot in response and the settling time is less than 5 seconds. The PD controller satisfies the control objective and can be implemented for the microassembly task.

6. Gluing and Adhesive Bonding Tests

Automatic gluing and adhesive bonding test by the micro assembly system is to implement an automatic operation for a microgripper gripping a copper rod with diameter 60μm, gluing adhesive, and bonding to another copper rod with diameter 380μm. The copper rod grasped by microgripper can be considered as a glue transfer tool or a cylindrical component. The whole system except PC and controllers are enclosed in an acrylic case to maintain a constant temperature about 25 ℃. In the following performance test, an automatic operation takes seven steps to complete a visual-servoing microassembly task.

First, the system will estimate both centroids of the microgripper and the copper rod with diameter 60μm. Also, the system will define the centroid of the microgripper as the origin of system. Second, the system will drive the microgripper to grip the copper rod and go back to the origin as defined in the first step. Third, the system estimates the centroid of glue in a container and the tip position of another copper rod with diameter 380μm. Fourth, the system estimates the pose and the tip position of the copper rod which has been gripped by the microgripper. The pose of the copper rod, which was held by the microgripper, provides information for the system to avoid collision. A copper rod gripped by the microgripper is shown in the left of Figure 9. Fifth, the system initiates RES algorithm to estimate the centroid of the microgripper with gripping copper rod. The microgripper gripping the copper rod is to glue adhesive for three times. The copper rod gluing adhesive on the tip is shown in the middle of Figure 9. Sixth, the microgripper gripping the 60μm copper rod is bonding to a 380μm copper rod. The microgripper stays and waits for ten minutes for glue hardening. Finally, the microgripper releases the 60μm copper rod as shown in the right of Figure 9 and goes back to the origin as defined before.

In the present system, the steady-state error in visual tracking is less than 1 pixel. The bandwidths of system in X-axis and Y-axis are about 16Hz, and 19Hz, respectively.

The microassembly system takes about 3 minutes to finish one operation and 10 minutes to wait for glue hardening in room temperature. The volume of glue in assembly can be estimated from the image of Figure 9.

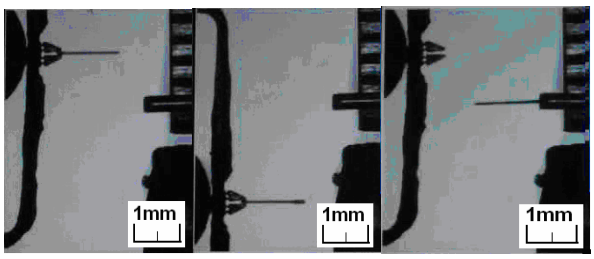


Figure 9. Microassembly procedure from gripping (left), gluing (middle), to bonding and releasing (right).

Since the image resolution is 22μm/ pixel and the thickness of glue in system is about 1 pixel, the thickness of glue is estimated as 22μm. Therefore, the volume of binding glue is about 5.7nl in the microassembly operation.

7. Conclusions

A PC-based visual-servoing automatic microassembly system was developed and tested. By employing the precision design axioms, a visual-servoing automatic microassembly system is realized efficiently and effectively to achieve the requirements of adhesive bonding in micro manipulations. The micro adhesive bonding by a PU microgripper is highly reliable without failure in the compliant joints during more than a hundred of operational tests. The system utilized the RES algorithm to track the microgripper and achieve the task of adhesive bonding. The present RES algorithm can accurately track and estimate the microgripper in a real-time operation. The steady-state error in visual tracking is less than 1 pixel. The system bandwidth is about 15 Hz. The performance was tested for gripping a copper rod with 60μm in diameter, gluing adhesive, and bonding to another copper rod with 380μm in diameter. The volume of binding glue in the microassembly operation is about 5.7nl.The system took about 3 minutes to finish one assembly operation if the waiting time for glue hardening was ignored.

8. Acknowledgments

The authors would like to thank the NSC for the support under contract No. (95)-2221-E-006-157.

9. References

[1] H. V. Brussel, J. Peirs, D. Reynaerts, A. Delchambre, G. Reinhart, N. Roth, M. Weck, and E. Zussman, "Assembly of microsystems," Annals of the CIRP, Vol. 49, pp. 451–472, 2000.

[2] T. Tanikawa, Y. Hashimoto, and T. Arai, "Micro drops for adhesive bonding of micro assemblies and making a 3-D structure `Micro scarecrow'," IEEE International Conference on Intelligent Robots and Systems, Vol. 2, pp. 776–781, 1998.

[3] J. Bergkvist, T. Lilliehorn, J. Nilsson, S. Johansson, and T. Laurell, "Miniaturized flowthrough microdispenser with piezoceramic tripod actuation," Journal of Microelectromechanical Systems, Vol. 14, No. 1, pp. 134–140, 2005.

[4] S. Fatikow, J. Seyfried, A. Buerkle, and F. Schmoeckel, "A Flexible microrobot-based microassembly station," Journal of Intelligent and Robotic Systems, Vol. 27, pp. 135–169, 2000.

[5] P. Korhonen, Q. Zhou, J. Laitinen, and S. Sjovall, "Auto-

matic dextrous handling of micro components using a 6 DOF micrograpper," Proceedings of 2005 IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp. 125–131, 2005.

[6] A. Ferreira, C. Cassier, and S. Hirai, "Automatic microassembly system assisted by vision servoing and virtual reality," IEEE/ASME Transactions on Mechatronics, Vol. 9, No. 2, pp. 321–333, 2004.

[7] K. B. Yesin and B. J. Nelson, "A CAD model based tracking system for visually guided microassembly," Robotica, Vol. 23, pp. 409–418, 2005.

[8] Y. H. Anis, J. K. Mills, and W. L. Cleghorn, "Automated microassembly task execution using vision-based feedback control," IEEE Proceedings of the 2007 International Conference on Information Acquisition, Vol. 9, No. 2, pp. 321–333, 2007.

[9] L. Ren, L. Wang, J. K. Mills, and D. Sun, "Vision-based 2-D automatic micrograsping using coarse-to-fine grasping strategy," IEEE Transactions on Industrial Electronics, Vol. 55, No. 9, pp. 3324–3331, 2008.

[10] G. Yang, J. A. Gaines, and B. J. Nelson, "Optomechatronic design of microassembly systems for manufacturing hybrid microsystems," IEEE Transactions on Industrial Electronics, Vol. 52, No. 4, pp. 1013–1023, 2005.

[11] B. S. Blanchard, "System Engineering Management," New York, Wiley, 1991.

[12] N. P. Suh, "The Principles of Design," New York, Oxford University Press, 1990.

[13] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," IEEE Transactions on Robotics and Automation, Vol. 12, No. 5, pp. 651–670, 1996.

[14] E. Trucco and K. Plakas, "Video tracking: a concise survey," IEEE Journal of Oceanic Engineering, Vol. 31, No. 2, pp. 520–529, 2006.

[15] G. Zhu, Q. Zeng, and C. Wang, "Efficient edge-based object tracking," Pattern Recognition, Vol. 39, pp. 2223–2226, 2006.

[16] D. Braha and O. Maimon, "The design process: properties, paradigms, and structure," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 27, No. 2, pp. 146–166, 1997.

[17] S. Fatikow and U. Rembold, "Microsystem technology and microrobotics," Berlin, Springer, 1997.

[18] R. J. Chang, H. S. Wang, and Y. L. Wang, "Development of mesoscopic polymer gripper system guided by precision design axioms," Precision Engineering, Vol. 27, pp. 362–369, 2003.

[19] R. J. Chang, P. W. Shih, R. Z. Huang, and K. L. Lin, "Application of piezo-driven polymer microgripper in automatic transportation of micro object," IEEE International Conference of Mechatronics, pp. 140–144, 2005.

[20] R. J. Chang and Y. L. Wang, "Integration method for input-output modeling and error analysis of four-bar polymer compliant micromachines," ASME Journal of Mechanical Design, Vol. 121, pp. 220–228, 1999.

[21] O. J. Smith, "A controller to overcome dead time," ISA Journal, Vol. 6, pp. 28–33, 1959.

Engineering, **2010, 2, 12-21**
doi:10.4236/eng.2010.21002 Published Online January 2010 (http://www.scirp.org/journal/eng/).

Scientific Research

Optimum Load Shedding in Power System Strategies with Voltage Stability Indicators

P. AJAY-D-VIMAL RAJ, M. SUDHAKARAN
Department of Electrical and Electronics Engineering, Pondicherry Engineering College, Pondicherry, India
E-mail: *ajayvimal@yahoo.com, karan_mahalingam@yahoo.com*
Received August 20, 2009; *revised September* 23, 2009; *accepted September* 28, 2009

Abstract

An optimal load shedding strategy for power systems with optimum location and quantity of load to be shed is presented in this paper. The problem of load shedding for avoiding the existence of voltage instability in power systems is taken as a remedial action during emergency state in transmission and distribution sector.Optimum location of loads to be shed is found together with their optimum required quantity. L-Indicator index is in used for this purpose with a modified new technique. Applications to be standard 6 bus Ward-Hale test system and IEEE – 14 bus system are presented to validate the applicability of the proposed technique to any system of any size.

Keywords: Load Shedding, Voltage Instability, Power systems, Transmission and Distribution Power system

1. Introduction

The major objective of power systems is to supply electricity to its customers. During emergency state of the power system, it may shed partial loads to ensure the power supply to important loads, as the last resort to maintain system integrity. Load shedding in bulk power system has been studied many years [1–6].

In electric power systems heavy loading may lead to voltage instabilities or collapses or in the extreme to complete blackouts. One technique of avoiding voltage instability is to shed some consumer's loads in Order to draw the operating point for away from the critical voltage value.

Many techniques have been developed [7–13] to minimize the load curtailment without violating the system security constraints. The emergency state in the power System with distributed generations has been formulated the load shedding is solved as an optimization problem [14]. Ying Lu *et al*. [15] has proposed a load shedding scheme, working with various load models, such as single-motor model, two-motor model, and composite model. Armanda *et al*. [16] have adopted a Distributed Interruptible Load Shedding (DILS) program instead of the traditional methodologies based on the separation of some users and/or entire distribution feeders, according to programmed plans of emergency. Andrzej Wiszniewski [17] have formulated a new method for estimating the voltage stability margin, which utilizes local measurements and applied criterion is based on the very definition of the voltage stability. Zhiping Ding *et al*. [18] have developed an expert-system-based load shedding scheme (LSS) for ship power systems. Emmanuel J. Thalassinakis *et al*. [19] have built an efficient computational methodology that can be used for calculating the appropriate strategy for load shedding protection in autonomous power systems.

In this paper a fast method for determining the location and quantity of the load to be shed in order to avoid risk of voltage instability is presented. The method defines clearly the bus where load shedding should make. A relation between voltage stability indicator changes and load power to be shed is developed. Using the relation the amount of load to be shed is determined. An algorithm has been developed for determining the location and quality of load to be shed and tested with tbe standard 6 buses Ward-Hale test system and IEEE – 14 bus system. The proposed method is valid for any system of any size at any loading conditions.

2. Mathematical Calculation for Load Shedding Using Voltage Stability Indicator- METHOD I

From the indicator proposed in [3] voltage stability indi-

cator at bus j can determine by

$$B_j = \left| 1 - \frac{\sum_{i \in \alpha G} C_{ji} V_i}{V_j} \right| \qquad j \in \alpha L \qquad (1)$$

where

α_L : Set of load buses.

α_G : Set of generator buses.

V_i : Voltage of generator bus i.

V_j : Voltage of load bus j.

C_{ji} : Elements of matrix C determined by

$$[C] = - [B_{LL}]^{-1} [B_{LG}] \qquad (2)$$

$[B_{LL}]^{-1}$ is the imaginary part of the matrix $[Y_{LL}]^{-1}$

$[B_{LG}]$ is the imaginary part of the matrix $[Y_{LG}]$

$[B_{LL}]^{-1}$ and $[B_{LG}]$ are susceptance matrices.

$[Y_{LL}]$ and $[Y_{LG}]$ are submatrices for the Y-Bus matrix.

The indicator at bus j, determined by equation (1) can be separated into real and imaginary parts (B_j^R, B_j^I):

$$B_j = 1 - \frac{\sum_{i=1}^{\alpha G} C_{ji} |V_i| \underline{|\delta i}}{|V_j| \underline{|\delta j}} \qquad (3)$$

$$B_j^R + jB_j^I = 1 - \frac{\sum_{i=1}^{\alpha G} C_{ji} |V_i| \underline{|\delta i - \delta j}}{|V_j|} \qquad (4)$$

$$B_j^R = 1 - \frac{\sum_{i=1}^{\alpha G} C_{ji} |V_i| \cos(\delta i - \delta j)}{|V_j|} \qquad (5)$$

$$B_j^I = - \frac{\sum_{i=1}^{\alpha G} C_{ji} |V_i| \sin(\delta i - \delta j)}{|V_j|} \qquad (6)$$

δi , δj : voltage angles at buses i and j

$|V_i|$, $|V_j|$:voltage magnitudes at buses i and j

The voltage stability indicator at each bus is a function of voltage angle and magnitude. The real imaginary part of indicator can be expressed as:

$$B^I = F_1(\delta, |V|)$$
$$B^R = F_2(\delta, |V|) \qquad (7)$$

The partial derivative of Equations (5) and (6) with respect to voltage angle and magnitude changes can be determined as:

$$\begin{bmatrix} \Delta B^I \\ \Delta B^R \end{bmatrix} = \begin{bmatrix} \frac{\partial B^I}{\partial \delta} & \frac{\partial B^I}{\partial |V|} \\ \frac{\partial B^R}{\partial \delta} & \frac{\partial B^R}{\partial |V|} \end{bmatrix} \begin{bmatrix} \Delta \delta \\ \Delta |V| \end{bmatrix} = [T] \begin{bmatrix} \Delta \delta \\ \Delta |V| \end{bmatrix} \qquad (8)$$

Matrix [T] is sensitivity matrix between indicator changes and voltage angle and magnitude changes

Coefficient of matrix [T] can be determined as

$$\frac{\partial B_J^I}{\partial \delta i} = - \frac{C_{ji} |V_i| \cos(\delta i - \delta j)}{|V_j|} \qquad i \in \alpha_G, i \neq s \qquad (9)$$

(S indicates the swing bus)

$$\frac{\partial B_j^I}{\partial \delta j} = \frac{\sum_{i \in \alpha G} C_{ji} |V_i| \cos(\delta i - \delta j)}{|V_j|} = -(B_j^R - 1) f \qquad (10)$$

$$\frac{\partial B_j^R}{\partial \delta i} = \frac{C_{ji} |V_i| \sin(\delta i - \delta j)}{|V_j|} \qquad (11)$$

$$\frac{\partial B_j^R}{\partial \delta j} = - \frac{\sum_{i \in \alpha G} C_{ji} |V_i| \sin(\delta i - \delta j)}{|V_j|} = B_j^I \qquad (12)$$

$$\frac{\partial B_j^I}{\partial |V_j|} = \frac{\sum_{i \in \alpha G} C_{ji} |V_i| \sin(\delta i - \delta j)}{|V_j|^2} = - \frac{B_j^I}{|V_j|} \qquad (13)$$

$$\frac{\partial B_j^R}{\partial |V_j|} = \frac{\sum_{i \in \alpha G} C_{ji} |V_i| \cos(\delta i - \delta j)}{|V_j|^2} = - \frac{B_j^R - 1}{|V_j|} \qquad (14)$$

From conventional Newton-Raphson load flow we obtain a linear relation between changes in voltage phases/magnitudes and active/reactive power injections as:

$$\begin{bmatrix} \Delta \delta \\ \Delta |V| \end{bmatrix} = [j]^{-1} \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} \qquad (15)$$

Sub (15) in (8) we get a relationship between real and imaginary part indicators and injected power as:

$$\begin{bmatrix} \Delta B^I \\ \Delta B^R \end{bmatrix} = [T][j]^{-1} \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} \qquad (16)$$

A relationship between changes in indicators at load bus j and power injections at all load buses can be obtained:

$$\Delta B_j^I = S_{11} \Delta P_j + S_{12} \Delta Q_j \qquad (17)$$

$$\Delta B_j^R = S_{21} \Delta P_j + S_{22} \Delta Q_j \qquad (18)$$

The active and reactive loads are not independent; one cannot shed active load, without curtailing reactive loads. Usually, a relation between active and reactive load at bus j can be obtained as follows. Here the load power factor is assumed to be constant at each load bus.

$$Pf_j = \frac{Q_j}{P_j} \qquad j \in \alpha_L$$

$$Pf_j = \frac{\Delta Q_j}{\Delta P_j}$$

$$\Delta Q_j = Pf_j \Delta P_j \tag{19}$$

Sub (19) in (17and18) we get a relationship between changes of the indicator at bus j and changes in active power injected at the same bus can be obtained as

$$\Delta B_j^I = S_{11}\Delta P_j + S_{12} Pf_j \Delta P_j \tag{20}$$

$$\Delta B_j^R = S_{21}\Delta P_j + S_{22} Pf_j \Delta P_j \tag{21}$$

$$\Delta B_j^I = S_1 \Delta P_j \tag{22}$$

$$\Delta B_j^R = S_2 \Delta P_j \tag{23}$$

where

$$S_1 = S_{11} + S_{12} Pf_j$$

$$S_2 = S_{21} + S_{22} Pf_j$$

We know

$$\Delta B_j = \sqrt{(\Delta B_j^I)^2 + \left(\Delta B_j^R\right)^2} \tag{24}$$

$$\Delta B_j = \sqrt{S_1^{\,2}\Delta P_j^{\,2} + S_2^{\,2}\Delta P_j^{\,2}} \tag{25}$$

$$\Delta B_j = \sqrt{(S_1^{\,2} + S_2^{\,2})\Delta P_j^{\,2}}\, s$$

$$\Delta B_j = \Delta P_j \sqrt{(S_1^{\,2} + S_2^{\,2})} \tag{26}$$

$$\Delta P_j = \frac{\Delta B_j}{\sqrt{(S_1^{\,2} + S_2^{\,2})}} \tag{27}$$

$$\Delta Q_j = Pf_j \Delta P_j \tag{28}$$

Using Equation (28) reactive power to be shed at bus j can be obtained if the active power to be shed at bus j is known.

3. Improved Method for Load Shedding By Voltage Stability Indicator

There are several methods for improving the voltage conditions in a power system as suggested in various articles.

In the improved method, the voltage profile is enhanced by determining the location and quantity of load to be shed, such that voltage instability can be avoided.

This method is based on the indicators of risk of voltage instability suggested by P. Kessel and H. Glavitsch [3]. The improved technique is a modification of the technique for previous section. It defines clearly the bus where the load shedding should be made. A relation between indicator changes and load powers to be shed is developed here. Using this relation, the amount of load to be shed is determined, for any operational situation.

4. Mathematical Calculation for Load Shedding Using Improved Voltage Stability Indicator- METHOD II

Kessel and Glavitsch [3] have developed a voltage stability indicator at load bus j

$$L_j = \left| 1 - \frac{\sum_{i \in \alpha G} F_{ji} V_i}{V_j} \right| \qquad J \in \alpha L \tag{29}$$

where

α_L : Set of load buses.

α_G : Set of generator buses.

V_i : Voltage of generator bus i.

V_j : Voltage of load bus j.

$$[Y] = \begin{bmatrix} Y_{GG} & Y_{GL} \\ Y_{LG} & Y_{LL} \end{bmatrix} \tag{30}$$

$$[F] = -[Y_{LL}]^{-1}[Y_{LG}] \tag{31}$$

Y_{GG} , Y_{LL} , Y_{LG} , Y_{GL} : Elements of system admittance matrix

A global voltage stability indicator of a power system is given by L, $0 \le L \le 1$

0: far away from voltage instability point.

1: at voltage instability point.

The indicator at bus j determined by Equation (29) can be separated into real and imaginary part (L_j^R, L_j^I)

$$L_j^R + jL_j^I = 1 - \frac{\sum_{i \in \alpha G} |F_{ji}||V_i| \underline{\theta_{ji} + \delta_i - \delta_j}}{|V_j|} \tag{32}$$

$$L_j^R = 1 - \frac{\sum_{i \in \alpha G} |F_{ji}||V_i| \cos\underline{\theta_{ji} + \delta_i - \delta_j}}{|V_j|} \tag{33}$$

$$L_j^I = -\frac{\sum_{i \in \alpha G} |F_{ji}||V_i| \sin|\theta_{ji} + \delta_i - \delta_j}{|V_j|} \quad (34)$$

The voltage stability indicator at each bus is a function of voltage angles and magnitudes. The real and imaginary parts of indicators can be expressed as:

$$L^I = F_1 d(\delta, |V|) \quad (35)$$

$$L^R = F_2(\delta, |V|) \quad (36)$$

The partial derivative of Equations (33) & (34) with respect to voltage angle and magnitude changes can be determined as

$$\begin{bmatrix} \Delta L^I \\ \Delta L^R \end{bmatrix} = \begin{bmatrix} \frac{\partial L^I}{\partial \delta} & \frac{\partial L^I}{\partial |V|} \\ \frac{\partial L^R}{\partial \delta} & \frac{\partial L^R}{\partial |V|} \end{bmatrix} \begin{bmatrix} \Delta\delta \\ \Delta|V| \end{bmatrix} = [T] \begin{bmatrix} \Delta\delta \\ \Delta|V| \end{bmatrix} \quad (37)$$

Matrix [T] is the sensitivity matrix between indicator changes and voltage angle and magnitude changes

Coefficient of matrix [T] can be determined as

$$\frac{\partial L_J^I}{\partial \delta i} = -\frac{|F_{ji}||V_i| \cos(\theta_{ji} + \delta_i - \delta_j)}{|V_j|} \quad i \in \alpha G,\ i \neq s \quad (38)$$

(S indicates the swing bus)

$$\frac{\partial L_j^I}{\partial \delta j} = \frac{\sum_{i \in \alpha G} |F_{ji}||V_i| \cos(\theta_{ji} + \delta_i - \delta_j)}{|V_j|} = -(L_j^R - 1) \quad (39)$$

$$\frac{\partial L_j^R}{\partial \delta_i} = \frac{|F_{ji}||V_i| \sin(\theta_{ji} + \delta_i - \delta_j)}{|V_j|} \quad (40)$$

$$\frac{\partial L_j^R}{\partial \delta_i} = -\frac{\sum_{i \in \alpha G} |F_{ji}||V_i| \sin(\theta_{ji} + \delta_i - \delta_j)}{|V_j|} = L_j^I \quad (41)$$

$$\frac{\partial L_j^I}{\partial |V_j|} = \frac{\sum_{i \in \alpha G} |F_{ji}||V_i| \sin(\theta_{ji} + \delta_i - \delta_j)}{|V_j|^2} = \frac{L_j^I}{|V_j|} \quad (42)$$

$$\frac{\partial L_j^R}{\partial |V_j|} = -\frac{\sum_{i \in \alpha G} |F_{ji}||V_i| \cos(\theta_{ji} + \delta_i - \delta_j)}{|V_j|^2} = \frac{L_j^R - 1}{|V_j|} \quad (43)$$

Form conventional Newton-Raphson Load flow we obtain a linear relation between changes in voltage phase/magnitudes and active/reactive power injections as:

$$\begin{bmatrix} \Delta\delta \\ \Delta|V| \end{bmatrix} = [j]^{-1} \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} \quad (44)$$

Sub (44) in (37) we get a relationship between real and imaginary parts of indicators and injected power as:

$$\begin{bmatrix} \Delta L^I \\ \Delta L^R \end{bmatrix} = [T][j]^{-1} \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} \quad (45)$$

A relationship between changes in indicators at load bus j and power injections at all load buses can be obtained:

$$\Delta L_j^I = S_{11}\Delta P_j + S_{12}\Delta Q_j. \quad (46)$$

$$\Delta L_j^R = S_{21}\Delta P_j + S_{22}\Delta Q_j \quad (47)$$

The active and reactive loads are not independent: one cannot shed active loads without reducing reactive loads. Usually a relation between active and reactive load can be obtained as follows. Here the load power factor is assumed to be constant at each load bus.

$$Pf_j = \frac{Q_j}{P_j} \qquad j \in \alpha_L$$

$$Pf_j = \frac{\Delta Q_j}{\Delta P_j}$$

$$\Delta Q_j = Pf_j \Delta P_j \quad (48)$$

Sub (48) in (46and47) we get a relationship between changes of the indicator at bus j and changes in active power injected at the same bus can be obtained as

$$\Delta L_j^I = S_{11}\Delta P_j + S_{12} Pf_j \Delta P_j \quad (49)$$

$$\Delta L_j^R = S_{21}\Delta P_j + S_{22} Pf_j \Delta P_j \quad (50)$$

$$\Delta L_j^I = S_1 \Delta P_j \quad (51)$$

$$\Delta L_j^R = S_2 \Delta P_j \quad (52)$$

where

$$S_1 = S_{11} + S_{12} Pf_j$$

$$S_2 = S_{21} + S_{22} Pf_j$$

$$\Delta L_j = \sqrt{(\Delta L_j^I)^2 + (\Delta L_j^R)^2} \quad (53)$$

$$\Delta L_j = \sqrt{S_1^2 \Delta P_j^2 + S_2^2 \Delta P_j^2} \quad (54)$$

$$\Delta L_j = \sqrt{(S_1^2 + S_2^2)\Delta P_j^2}$$

$$\Delta L_j = \Delta P_j \sqrt{(S_1^2 + S_2^2)} \quad (55)$$

$$\Delta P_j = \frac{\Delta L_j}{\sqrt{(S_1^2 + S_2^2)}} \quad (56)$$

$$\Delta Q_j = Pf_j \Delta P_j \quad (57)$$

Using Equation (57) reactive power to be shed at bus j can be obtained if the active power to be shed at bus j is know.

5. Algorithm for Calculation Load to Be Shed

The step by step procedure of load shedding algorithm is given as follows

Step 1

Carry out load flow by Newton Raphson Method.

Step 2

Calculate voltage stability indicator for all load buses and find Lmax.

Step 3

Check Lmax $\geq$ Lcritical, if exceeds we have to shed a part of load at that bus with

maximum value of L, goto next step. Else there is no need to shed the load stop.

Step 4

Using Equations (56 and 57), calculate the required load to be shed.

Step 5

Remove this load and goto step (1)

i.e. (subtract this ΔP_j and ΔQ_j from the j^{th} load bus).

6. Numerical Examples, Simulation Results and Analysis

The study has been conducted on test cases with standard 6 bus Ward-Hale test system and IEEE-14 bus test system. The voltage stability indicator for all load buses are computed by two methods for various load patterns and results are tabulated.

To verify the feasibility of the improved method, two different power systems were tested, under various base cases. The solutions were compared by their solution quality and computation efficiency. From the experiences of many experiments the optimum load shedding algorithms have been used to solve the above test cases and are results are tabulated. For implementing the above algorithm, the simulation studies were carried out on P-IV, 2.4 GHz, 512 MBDDR RAM systems in MATLAB environment. The load shedding results for the first test case with the corresponding base cases are tabulated in Tables 1-6.

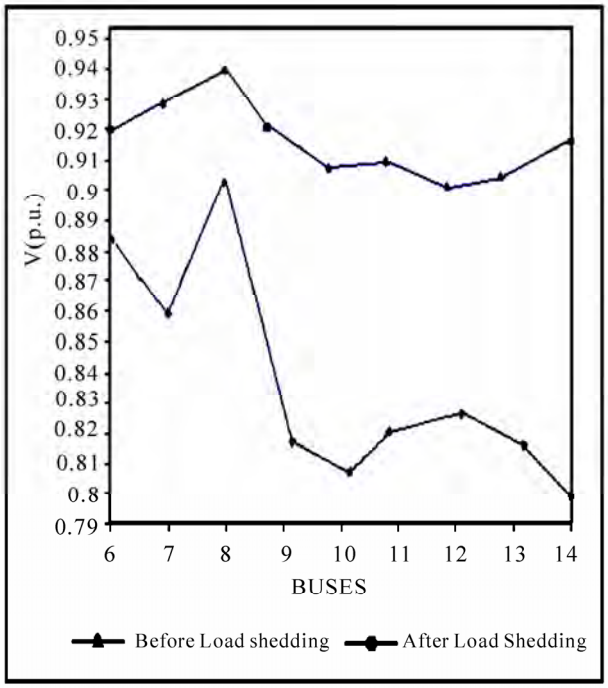


Figure 1. Voltage profile for standard 6 bus ward-hale test system.

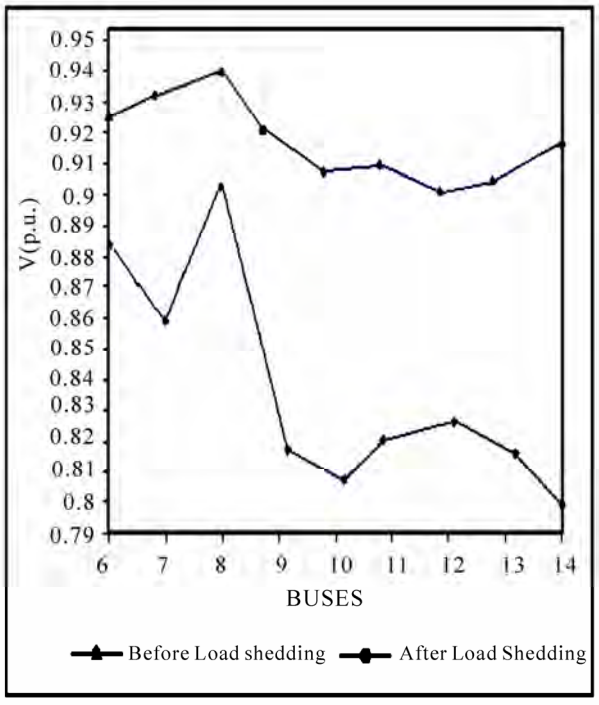


Figure 2. Voltage profile for IEEE-14 bus test system.

Table 1 illustrates the results of load flow solution of 1.5 times the base power; the maximum value of the B indicator is 0.572051 at bus 5. Therefore it is the bus at which load must be shed. Using the proposed method,

the quantity of load to be shed is found and shown in the same table. The B indicator after shedding and the bus voltage magnitude after each load shed is given in the same Table 1.

In case study I using method 1 the maximum value of indicators occur at bus 3, 5 and 6. Using the improved method the impact of maximum value of indicators occurred at bus 3, 5 and 6 got reduced and illustrated in Table 2.

The successive load shedding and improvement in voltage magnitude with the maximum value of B indicator less than the B critical value for load flow solution 1.1 times and 1.0 times the base power is illustrated for both the methods were illustrated in Tables 3,4,5 and 6 for 6 bus Ward-Hale test system.

The results from the improved method is compared with the method proposed by T. Quoc Taun *et al.* [4–5] in Tables 7,8,9,10,11,and 12 for IEEE-14 bus system for different loading conditions respectively. Analysis of these tables shows that shedding of selected loads at selected buses improves the voltage magnitude at all buses. In addition, the stability of the system is improved.

The computation time of these two methods for different loading conditions are tabulated in Tables 13 and 14. From these tables, it is very clear that the computation time of the proposed method is slightly higher than method-I .The various results obtained by the two methods show that both the methods are quite effective. But, in the proposed method both resistance and reactance are taken into account hence this method is more accurate and yields more computation time. Figure 1 and Figure 2 show the improved voltage profile of the proposed algorithm for standard 6 bus ward-hale test system and IEEE-14 bus test system The figure shows that the algorithm is capable of obtaining a faster convergence for the three unit thermal system in a very few generations and the solution is consistent.

Test case 1: 6 bus ward-hale test system
Case study i
Method –i

Table 1. Load flow solution 1.5 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude(after load shedding)
3	0.792526	0.552319	-0.278049-j0.065721	0.299996	0.907653
4	0.821274	0.421552	0.00	0.230428	0.933661
5	0.761915	0.572051	-0.196304-j0.117783	0.299993	0.889731
6	0.788985	0.535638	-0.123482-j0.012348	0.294669	0.917190

Alpha: 1.5 Bcri: 0.3

Method –II

Table 2. Load flow solution 1.5 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude(after load shedding)
3	0.792526	0.565819	-0.272274-j0.064356	0.299995	0.905982
4	0.821274	0.427228	0.00	0.230818	0.931995
5	0.761915	0.589754	-0.190293-j0.114176	0.299906	0.887288
6	0.788985	0.537283	-0.121310-j0.012131	0.296168	0.915063

Alpha: 1.5 Bcri: 0.3

Case Study 2
Method I

Table 3. Load flow solution 1.1 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude (after load shedding)
3	0.910748	0.309107	-0.013415-j0.003171	0.299995	0.915424
4	0.929561	0.239028	0.00	0.231807	0.934368
5	0.891220	0.312879	-0.019480-j0.011688	0.298756	0.898976
6	0.911773	0.296007	0.00	0.286505	0.918155

Alpha: 1.1 Bcri: 0.3

Method II

Table 4. Load flow solution 1.1 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude(after load shedding)
3	0.910748	0.313655	-0.019744-j0.004430	0.299995	0.917204
4	0.929561	0.240298	0.00	0.229818	0.936187
5	0.891220	0.318878	-0.026911-j0.016146	0.298020	0.901871
6	0.911773	0.295298	0.00	0.281989	0.920540

Alpha: 1.1 Bcri: 0.3

Case Study 3
Method I

Table 5. Load flow solution 1.0 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude(after load shedding)
3	0.931663	0.268746	0.00	-	-
4	0.948248	0.208422	0.00	-	-
5	0.914858	0.270784	0.00	-	-
6	0.933040	0.257166	0.00	-	-

Alpha: 1.0 Bcri: 0.3

Method II

Table 6. Load flow solution 1.0 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude(after load shedding)
3	0.931663	0.272399	0.00	-	-
4	0.948248	0.209258	0.00	-	-
5	0.914858	0.275604	0.00	-	-
6	0.933040	0.256247	0.00	-	-

Alpha: 1.0 Bcri: 0.3

Test Case 2: IEEE 14 Bus System
Case Study 1
Method I

Table 7. Load flow solution 1.75 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude (after load shedding)
6	0.885987	0.076577	0.00	0.065605	0.926245
7	0.861489	0.103337	0.00	0.068974	0.935048
8	0.903179	0.052732	0.00	0.045991	0.939599
9	0.826483	0.188466	-0.117981-j0.066389	0.116584	0.915637
10	0.817095	0.181116	0.00	0.116787	0.904157
11	0.835082	0.101553	0.00	0.068160	0.912634
12	0.832003	0.065987	0.00	0.046524	0.909214
13	0.825003	0.090781	0.00	0.056859	0.906149
14	0.797535	0.225521	-0.22901-j0.076913	0.100380	0.910447

Alpha: 1.75 Bcri : 0.12

Method II

Table 8. Load flow solution 1.75 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude (after load shedding)
6	0.885987	0.071513	0.00	0.060637	0.919887
7	0.861489	0.101100	0.00	0.068528	0.923378
8	0.903179	0.048410	0.00	0.041661	0.933292
9	0.826483	0.187316	-0.112015-j0.063032	0.119214	0.901171
10	0.817095	0.18039	-0.001460-j0.000094 1	0.119999	0.888367
11	0.835082	0.101179	0.00	0.070317	0.894098
12	0.832003	0.067257	0.00	0.049879	0.887929
13	0.825003	0.090801	0.00	0.059871	0.885136
14	0.797535	0.225110	-0.220811-j0.074098	0.105331	0.892381

Alpha : 1.75 Lcri : 0.12

Case study II
Method I

Table 9. Load flow solution 1.5 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude (after load shedding)
6	0.965883	0.056862	0.00	0.054954	0.968854
7	0.961972	0.071512	0.00	0.066872	0.965849
8	0.975622	0.039895	0.00	0.038464	0.978367
9	0.939840	0.126675	0.00	0117218	0.946047
10	0.932435	0.121099	0.00	0.113175	0.937631
11	0.944580	0.068419	0.00	0.064546	0.947266
12	0.944571	0.044103	0.00	0.041059	0.946053
13	0.937257	0.060500	0.00	0.054197	0.940925
14	0.918004	0.148539	-0.092214-j0.030944	0.119998	0.932353

Alpha : 1.5 Bcri : 0.12

Method II

Table 10. Load flow solution 1.5 times the base power.

Bus number	Bus voltage magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude (after load shedding)
6	0.965883	0.051551	0.00	0.050210	0.968655
7	0.961972	0.069175	0.00	0.064798	0.965592
8	0.975622	0.035267	0.00	0.034432	0.978183
9	0.939840	0.124902	0.00	0.115731	0.945636
10	0.932435	0.119784	0.00	0.112075	0.937287
11	0.944580	0.067704	0.00	0.063903	0.947088
12	0.944571	0.044896	0.00	0.041999	0.945956
13	0.937257	0.060239	0.00	0.054229	0.940652
14	0.918004	0.147165	-0.085875-j0.028817	0.119997	0.931398

Alpha : 1.5 Lcri : 0.12

Case Study III
Method I

Table 11. Load flow solution 1.25 times the base power.

Bus number	Bus voltage Magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude (after load shedding)
6	0.991775	0.045028	0.00	-	-
7	0.984258	0.056165	0.00	-	-
8	1.000579	0.031853	0.00	-	-
9	0.972796	0.098070	0.00	-	-
10	0.967549	0.093527	0.00	-	-
11	0.978900	0.052892	0.00	-	-
12	0.980368	0.034090	0.00	-	-
13	0.974269	0046519	0.00	-	-
14	0.956955	0.113685	0.00	-	-

Alpha : 1.25 Bcri : 0.1

Method II

Table 12. Load flow solution 1.25 times the base power.

Bus number	Bus voltage Magnitude (before load shedding)	B indicator (before load shedding)	Sheddable load	B indicator (after load shedding)	Bus voltage magnitude (after load shedding)
6	0.991775	0.040850	0.00	-	-
7	0.984258	0.054596	0.00	-	-
8	1.000579	0.027980	0.00	-	-
9	0.972796	0.097214	0.00	-	-
10	0.967549	0.092971	0.00	-	-
11	0.978900	0.052608	0.00	-	-
12	0.980368	0.034715	0.00	-	-
13	0.974269	0.046498	0.00	-	-
14	0.956955	0.113325	0.00	-	-

Alpha : 1.25 Lcri : 0.12

Table 13. Computation time in sec for 6 bus Ward-Hale test system.

Cases	Computation time (secs.)	
	Method I	Method II
Case 1	0.66	0.82
Case 2	0.61	0.66
Case 3	0.22	0.27

Table 14. Computation time in sec for IEEE-14 bus system.

Cases	Computation time (secs.)	
	Method I	Method II
Case 1	1.21	1.42
Case 2	0.88	1.15
Case 3	0.49	0.55

7. Conclusions

A simple new method is developed to determine the optimum location and the optimum quantity of load to be shed in order to prevent the system voltage from going to the unstable. This method is based on indicators of risk of voltage instability. It can be implemented for large power system to estimate voltage instability. Successive load flow runs are required to accomplish this method. The proposed method can be used for real time applications in power systems. The computation speed of these indicators is fast compared to other methods.

8. Acknowledgment

The authors gratefully acknowledge the the Management

of Pondicherry Engineering College, Pondicherry, INDIA for their continued support, encouragement and the facilities provided to carry out this research work.

9. References

[1] C. W. Taylor, "Concept of under voltage load shedding for voltage stability," IEEE transactions on Power Delivery, Vol. 7, No. 2, pp. 480–488, 1992

[2] S. A. Niern berg, D. A. Meinnis, and K. D. Sparks, "Fast acting load shedding," IEEE Transactions on Power Systems, Vol. 7, No. 2, pp. 873–877, 1992.

[3] P. Kessel and H. Glavitsch, "Estimating the voltage stability of a power system," IEEE Transactions on Power Delivery, Vol. 1, No. 3, pp. 346–354, 1986,

[4] T. Quoc Taun, J. Fandino, N. Hadjsaid, and J. C. sabonnadiere, H. Vu, "Emergency load shedding to avoid risks of voltage instability using indicators," IEEE Transactions on Power Systems, Vol. 9, No. 1, pp. 341–351, 1994.

[5] T. Quoc Taun, J. Fandino, N. Hadjsaid, and J. C. sabonnadiere, "Fast calculation of indicators of risk of voltage instability," Proc .of 27th Universities Power Engineering Conference, UPEC–92 , pp. 612–615, 1992.

[6] C. W. Taylor, "Power system voltage stability," Mc.Graw Hill, 1993.

[7] L. P. Hajdu, J. Peschon, W. F. Tinney, and D. S. Piercy, "Optimum load-shedding policy for power system," IEEE Tran, PAS-87, No. 3, pp. 784–794, 1968.

[8] D. K. Subramanian, "Optimum load shedding through programming techniques," IEEE Trans, PAS-90, pp. 89–94, 1971.

[9] S. M. Chan and F. C. Schweppe, "A generation rellocation and load shedding algorithm," IEEE Trans., PAS-98, No. 1, pp. 26–34, 1979.

[10] M. M. Adibi and D. K. Thorne, "Local load shedding," IEEE Trans. On Power System, Vol. 3, No. 3, pp. 1220–1226, 1988.

[11] S. Shah and S. M. Shahldehpour, "A heuristic approach to load shedding scheme," IEEE Trans, on Power System, Vol. 4, No. 4, pp. 1421–1429, 1989.

[12] R. Billinton and J. Satish, "Effect of rotational load shedding on overall power system adequacy indices," IEE Proc. C, Vol. 143, No. 2, pp. 181–187, 1996.

[13] P. Wang and R. Billinton, "Optimum load-shedding technique to reduce the total customer interruption cost in a distribution system," IEE Proc. - Gener. Transm. Distrib., Vol. 147, No. 1, pp. 51–56, 2000.

[14] D. Xu and A. A. Girgis, "Optimal Load Shedding Strategy in Power Systems with Distributed Generation," IEEE, pp. 788–793, 2001.

[15] Y. Lu, W. S. Kao, and Y. T. Chen, "Study of applying load shedding scheme with dynamic d-factor values of various dynamic load models to taiwan power system," IEEE Transactions on Power Systems, Vol. 20, No. 4, pp. 1976–1984, 2005.

[16] R. Faranda, A. Pievatolo, and E. Tironi, "Load shedding: A new proposal," IEEE Transactions on Power Systems, Vol. 22, No. 4, pp. 2086–2093, 2007.

[17] A. Wiszniewski, "New criteria of voltage stability margin for the purpose of load shedding," IEEE Transactions on Power Delivery, Vol. 22, No. 3, pp. 1367–1371, 2007.

[18] Z. P. Ding, S. K. Srivastava, D. A. Cartes, and S. Suryanarayanan, "Dynamic simulation-based analysis of a new load shedding scheme for a notional destroyer-class shipboard power system," IEEE Transactions on Industry Applications, Vol. 45, No. 3, pp. 1163–1174, 2009.

[19] E. J. Thalassinakis, E. N. Dialynas, and D. Agoris, "Method combining anns and monte carlo simulation for the selection of the load shedding protection strategies in autonomous power systems," IEEE Transactions on Power Systems, Vol. 21, No. 4, pp. 1574–1582. 2006.

Engineering, **2010, 2, **-****
doi:10.4236/eng.2010.21003 Published Online January 2010 (http://www.scirp.org/journal/eng/).

Computer Aided Modeling and Deign of a New Magnetic Sealing Mechanism in Engineering Applications

Jeremy (Zheng) LI
School of Engineering, University of Bridgeport, Bridgeport, USA
E-mail: *zhengli@bridgeport.edu*
Received June 6, 2009; *revised August* 3, 2009; *accepted August* 10, 2009

Abstract

This article introduces a new type of magnetic sealing mechanism that reduces the lubrication oil pollution and media gaseous leakage in general reciprocating machinery including air compressors and refrigerators. The feasible function and reliable performance of this new sealing mechanism are introduced and analyzed in this paper. The computer aided design, modeling and analysis are being used to study this new sealing mechanism, and the prototype of this sealing mechanism is being tested. The study indicated the proper function of this sealing mechanism. The major advantages of this sealing mechanism include: improved sealing capacity to prevent the gaseous leakage and oil leakage, simple and compact in structure, lower precision requirement on surfaces of reciprocating pistons and shafts in production and manufacturing, and longer services in sealing life span. Also there is almost no frictional loss during the reciprocating motion of piston or shaft.

Keywords: Magnetic Sealing, Magnetic Flux, Reciprocating Machinery, Self-Lubricated System

1. Introduction

The gaseous leakage and oil pollution in reciprocating machines including compressors and refrigerators are common problems that have not been well resolved and it directly affects the machinery performance [1–4]. The design and development of new sealing mechanism are continued in these years [5–8].

In this research, a new magnetic sealing mechanism using rare-earth magnet as permanent magnet is developed to solve these problems based on theoretic analysis, computational modeling simulation, and prototype tests. The permanent magnet is made from the materials that stay magnetized. Materials that can be magnetized are called ferromagnetic including rare earth magnets. The current research and development of rare earth permanent magnets have brought the renovation in the field of magnetic separation and provided the magnetic products that are an order of magnitude stronger than that of conventional ferrite magnets. This leads to the development of high-intensity magnetic circuits that operated energy free and surpasses the electromagnets in strength and effectiveness. Common applications of rare-earth magnets include: computer hard drives, audio speakers, bicycle dynamos, fishing reel brakes, mag-lev wind turbines, and LED throwies.

The prototype testing of this new magnetic sealing mechanism indicated that this sealing mechanism can significantly reduce the leakage problem in reciprocating machines including compressors and oil pollution in cryogenic regenerator. It also shows that this sealing mechanism can replace the oil separation system in refrigerating compressors. Through the prototype tests, the sealing function of this new mechanism is better than regular rubber seal, diaphragm seal, corrugated pipe seal and magnetic fluid seal.

2. Magnetic Circuit in Sealing Mechanism

The rare-earth magnet steel can be performed as a permanent magnet steel which has the higher density of magnetic flux Br, strong magnetic field H_g, and larger product of magnetism and energy $(BH)_{max}$ as shown in Figure 1. All these good features allow the magnetic particles to be firmly adhered onto the inside wall of magnet steel. The major advantages of this magnetic circuit include higher Br in working gap of the circuit, longer and durable in sealing lifetime, compact in system configuration, light in unit weight, higher in performance efficien-

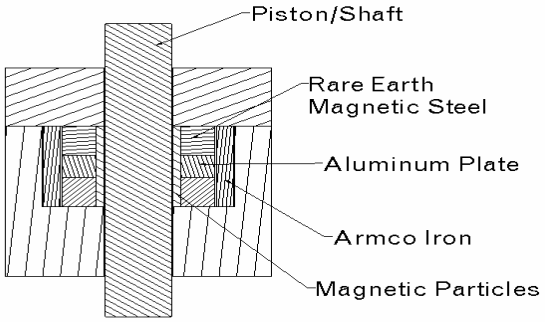


Figure 1. Magnetic sealing mechanism.

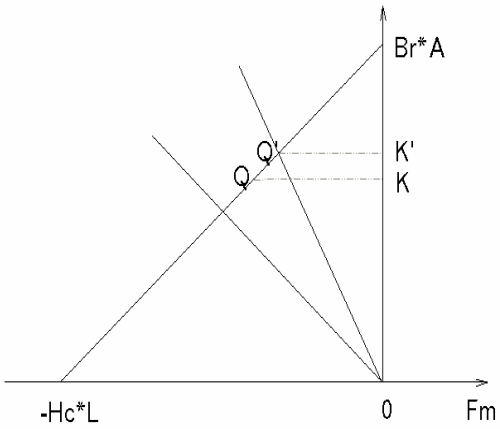


Figure 2. Magnetic curve of circuit.

cy, and stable in sealing functioning.

The concept of this new magnetic sealing mechanism can be briefly described as follows. When piston/shaft reciprocates inside of the cylinder, lubricating oil is sealed by magnetic particles, which are firmly adhered on the inside surface of magnet steel, as oil particles move to the seal. Then the oil droplets drop to the main shaft chamber at the bottom of compressors by its gravity which can prevent the oil in crank chamber from entering the gas cylinder. Also the gaseous leakage can be prevented because the gas could not pass through the strong adhesive layers of magnetic particles. In this new magnetic sealing design, two critical factors that should be considered to keep its well function are density of magnetic flux and magnetic stability of the magnet steel. Thus the magnetic flux in magnetic circuit of this sealing mechanism must be maintained over a long period of time and magnetic field of this magnet steel should be stable enough to withstand the external/disturbed magnetic fields, temperature change, mechanical vibration/shock, and severe environmental fluctuation. The surplus density of magnetic flux Br, surplus intensity of magnetic field H_g, and maximum product of magnetism and energy $(BH)_{max}$ are required to keep their peak values in this magnetic sealing mechanism design.

The magnetic circuit in this sealing mechanism is in static condition which can be analyzed using ampere enclosed circuit and H-B curve of this rare-earth magnet steel. This magnetic circuit can be considered as a series magnetic circuit mainly made up from magnet steel and working gap. Refer the Figure 1, the following equations can be derived:

$$H * L + H_g * L_g = 0 \quad (1)$$

$$H * L = - \frac{Lg * \Phi}{U_0 * Ag} \quad (2)$$

Let $Fm(\Phi) = H * L$, the intersection of $Fm(\Phi)$ and straight line $-[L_g / (U_0 * A_g)] * \Phi$ at ordinate in Figure 2 is the magnetic flux in working gap that required to be determined. This gap decreases from Lg to Lg' after magnetic particles being added into the space in magnetic circuit gap. When the thickness of magnetic particles in the gap between surfaces of magnet steel and cylinder changes from 0 to b, the working point of magnet steel changes along the straight line QK. The corresponding solution of magnetic flux in working gap can be found from line QK. The magnetic field is well distributed / maintained in this sealing mechanism that has been verified from computational modeling simulation. The coefficient of magnetic efficiency f is used to judge if the magnetic field in this sealing mechanism is properly designed. Here,

$$f = \frac{Bg^2 * Vg}{(B * H)_{max} * V} \quad (3)$$

The higher f value indicates more feasible and reasonable design on the magnetic circuit. The f value is normally 40% in standard conditions. The computational modeling solution shows that f value in this sealing magnetic mechanism is 48.5% which verifies the proper magnetic circuit design in this sealing mechanism.

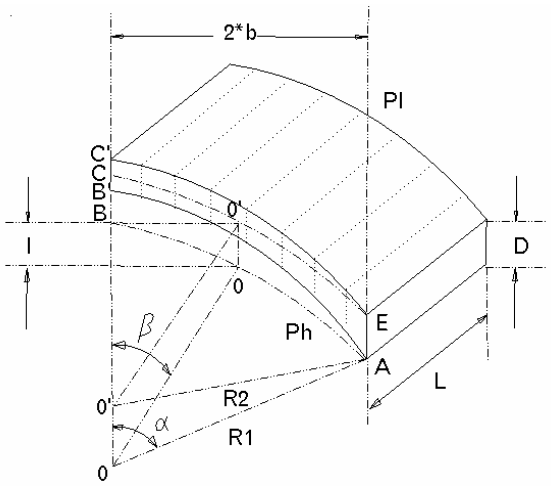


Figure 3. Cross section view of magnetic steel.

3. Analysis of Sealing Capacity

The formula of sealing capacity ΔP can be derived from energy balancing theory as follows.

Referring the cross section of this magnetic seal in Figure 3,

$$R_1 = \frac{2*b}{\sin(\alpha)} \quad (4)$$

$$R_2 = \frac{2*b}{\sin(\beta)} \quad (5)$$

$$S_1 = R_1 * \alpha \quad (6)$$

$$S_2 = R_2 * \beta \quad (7)$$

$$\Delta S = S_2 - S_1 = [\frac{\beta}{\sin(\beta)} - \frac{\alpha}{\sin(\alpha)}] * 2 * b \quad (8)$$

$$|OO'| = 2 * \beta * [\text{ctg}(\alpha) - \text{ctg}(\beta)] \quad (9)$$

Because the work that each magnetic line of force applied is $T * \Delta S$, total work that magnetic lines of force applied in magnetic circuit is:

$$W_1 = B*D*T*2b*[\frac{\beta}{\sin(\beta)} - \frac{\alpha}{\sin(\alpha)}] \quad (10)$$

At the same time, the work that media pressure applied to the body of magnetic particles is:

$$W_2 = 4 * b^2 * \Delta P * [\frac{\sin(\alpha) + \pi*[\cos(\alpha) - 2]^2}{8*\sin(\alpha) - \frac{\sin(\alpha)}{4}}] \quad (11)$$

Based on energy balancing theory, the work that magnetic lines of force applied in magnetic circuit should be equal to the work that media pressure exerted to the body of magnetic particles. So,

$$\Delta P = \frac{(\frac{1}{2})*B*D*T*[\sin(\alpha) - \alpha]}{\{2*b*(\sin\alpha) + \frac{\alpha}{8}*[\cos^2(\alpha) - 4*\cos(\alpha) + 3]\}} \quad (12)$$

This formula can be calculated by computational modeling with numerical solution. The optimized computational simulation indicated that, when α and β are changed to certain values, the ΔP_{max} can be determined as follows:

$$\Delta P_{max} = \frac{C*B*D*T}{2*b} = 28.5 \text{ Kg/cm}^2 \quad (13)$$

This result shows that the seal capacity of this magnetic sealing mechanism can prevent the oil leakage/pollution from crank chamber into the cylinders of reciprocating machinery and refrigerating regenerators. It can also keep the compressors from gaseous leakage. The above mechanism analysis and computational simulation have been verified through the prototype tests. Furthermore, the sealing capacity in this mechanism can be improved by increasing the number of this magnetic seal, improving the magnetic material composite and optimizing the magnetic circuit design.

4. Prototype Testing Results

The prototype of this new magnetic sealing mecha-

Table 1. Estimated air leakage at different piston linear speed.

Piston Linear Speed (Ft/Min)	Estimated Air Leakage (SCFM)	Piston Linear Speed (Ft/Min)	Estimated Air Leakage (SCFM)	Piston Linear Speed (Ft/Min)	Estimated Air Leakage (SCFM)
5	0.001	45	0.036	90	0.083
10	0.003	50	0.043	100	0.088
15	0.006	55	0.048	105	0.092
20	0.010	60	0.054	110	0.095
25	0.014	65	0.061	115	0.099
30	0.019	70	0.066	120	0.104
35	0.025	75	0.071	125	0.111
40	0.031	80	0.077	130	0.119

Table 2. Estimated air leakage at different air pressure.

Air Pressure (PSIG)	Estimated Air Leakage (SCFM)	Air Pressure (PSIG)	Estimated Air Leakage (SCFM)	Air Pressure (PSIG)	Estimated Air Leakage (SCFM)
50	0.002	450	0.030	800	0.063
100	0.005	500	0.033	850	0.067
150	0.007	550	0.037	900	0.070
200	0.011	560	0.042	950	0.074
250	0.015	600	0.045	1000	0.079
300	0.018	650	0.049	1050	0.083
350	0.022	700	0.054	1100	0.088
400	0.026	750	0.058	1150	0.094

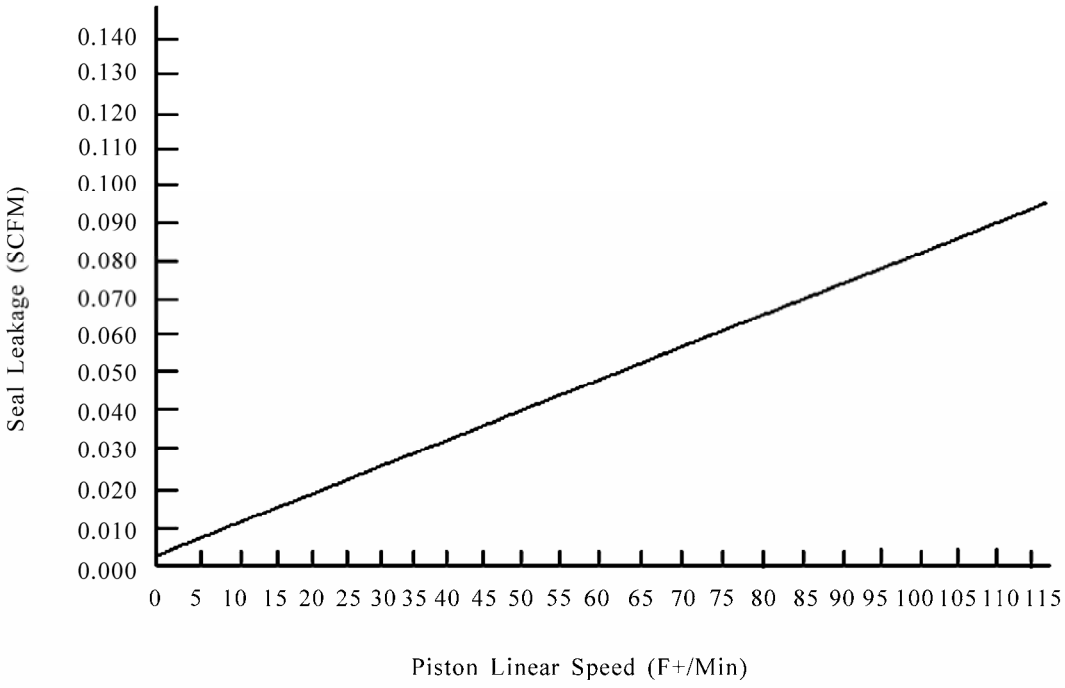


Figure 4. Air leakage vs. piston linear speed.

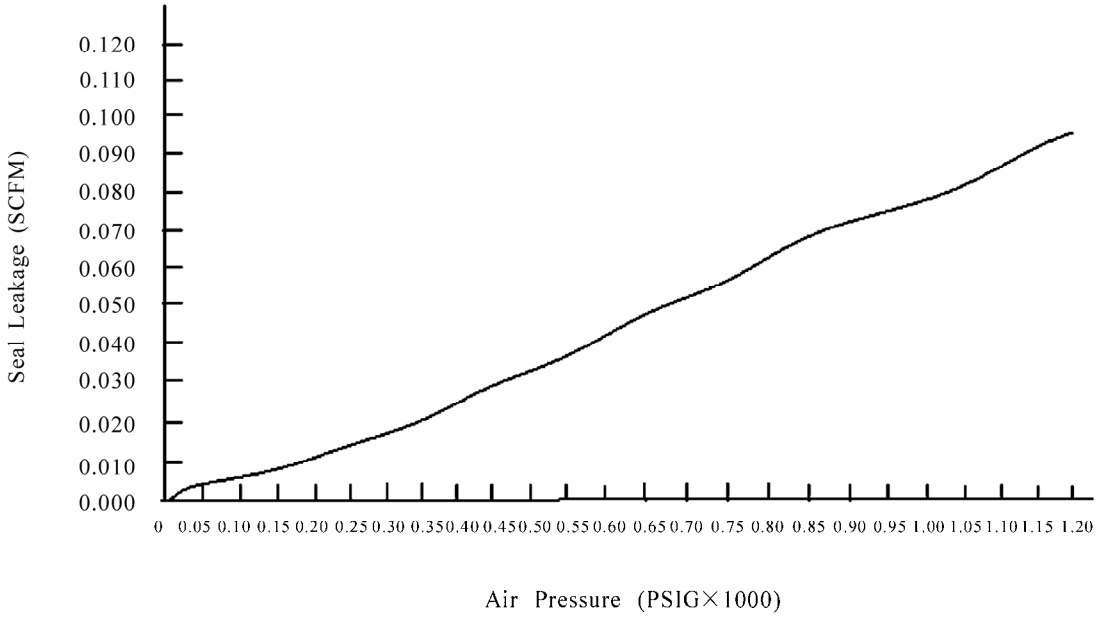


Figure 5. Air leakage vs. air pressure.

nism has been tested and the preliminary results are shown in Tables 1 and 2.

5. Computational Simulation Results

This new magnetic sealing system has also been simulated by computational solution and results are indicated in Figures 4 and 5.

Based on the above, the preliminary results from prototype testing and computational simulation are closed to each other, and this verifies the creditability and feasibility of this new magnetic sealing mechanism.

6. Conclusions

Today the oil and gaseous media leakages are the tough and difficult engineering problems that affect the reciprocating machinery function and performance. This new magnetic sealing mechanism has been developed to reduce the oil and gaseous media leakages in reciprocating machinery. All the theoretical mechanical and magnetic analysis, computational simulation, and prototype tests indicated that this sealing mechanism can significantly decrease the oil and gaseous media leakages in reciprocating machinery. Its sealing performance is reliable due to the firmly adhesive and strong forces between the magnetic particles and reciprocating pistons/shafts. This seal mechanism is also durable if compared with regular seals including rubber seal, diaphragm seal, corrugated pipe seal because of less frictional force between surfaces of seal and pistons/shafts in this new sealing mechanism. Moreover, the development of this magnetic sealing mechanism will further contribute to the exploitation, popularization, and application of the rich rare-earth elements/materials in today's modern industrial world.

7. References

[1] H. Quelle, G. González-Jorge, and L. Domarco, "Magnetic field cancellation on the surface of superconducting rings: Influence on the flux creep," Cryogenics, Vol. 48 pp. 2–5, 2008.

[2] J. E. Hirsch, "The Lorentz force and superconductivity," Physica A, Vol. 315, pp. 474–476, 2003.

[3] H. González-Jorge, B. Liñares, I. Quelle, E. Carballo, and G. Domarco, "Measurement of critical current in superconducting rings," Measurement Vol. 40, pp. 378–380. 2007.

[4] H. González-Jorge, J. Peleteiro, E. Carballo, L. Romaní, and G. Domarco, "Procedure to induce a persistent current in superconducting cylinders or rings," Applied Physica, Vol. 81, pp. 420–421, 2002.

[5] H. González-Jorge, J. Peleteiro, E. Carballo, and G. Domarco, "Inducting technique and trapped field in ring-shaped Superconductors," Physica, Vol. 433, pp. 132–134, 2005.

[6] P. W. Anderson, "Theory of flux creep in hard superconductors," Physica, Vol. 9, pp. 309–311, 1962.

[7] U. Topal, L. Dorosiskii, H. Ozkan, and H. Yabuz, "Effect of thermal neutron irradiation in boron-doped melt- textured YBCO superconductors," Physica, Vol. 388, pp. 401-403, 2003.

[8] L. Shlyk, G. Krabbes, G. Fuchs, K. Nenkov, and B. Schupp, "Flux pinning and magnetic relaxation in melt-processed $YBa_2Cu_3O_{7-d}$ dopped with Li," Journal of Applied Physica, Vol. 96, pp. 371–374, 2004.

Nomenclature

A_g = cross section area of working gap
B_g = density of magnetic flux in working gap
B_r = density of magnetic flux
$(BH)_{max}$ = maximum product of magnetism and energy
C = coefficient
D = Width of magnetic steel
f = coefficient of circuit efficiency
H = intensity of magnetic field of magnet steel
H_g = intensity of magnetic field in working gap
L = length of magnet steel
L_g = length of working gap
T = intensity of magnetization
U_0 = magnetic conductivity of vacuum
V_g = volume of working gap
V = volume of magnet steel
B = half length of working gap
Φ = magnetic flux

Engineering, 2010, 2, **-**
doi:10.4236/eng.2010.21004 Published Online January 2010 (http://www.scirp.org/journal/eng/).

Scientific Research

'Fading is Our Friend!': A Performance Comparison of WiMAX-MIMO/MISO/SISO Communication Systems

Arya ROHAN[1,2], Nicolas DAILLY[1], Palash KUSUMWAL[1,2]
[1]*ESIEE-Amiens, Ecole Supérieure d'Ingénieurs en Electrotechnique et Electronique, France*
[2]*NIT-Trichy, National Institute of Technology, Trichy, India*
E-mail: *aryarohan07@gmail.com, dailly@esiee-amiens.fr, palashkusumwal@gmail.com*
Received August 24, 2009; *revised September* 22, 2009; *accepted September* 28, 2009

Abstract

Research work for some time now has shown that fading wireless channels present enormous advantages if properly exploited through a Multiple-Input Multiple-Output (MIMO) communication model. In this paper, we demonstrate the advantages of implementing the MIMO communication model by investigating three communication techniques, namely, Single-Input Single-Output (SISO), Multiple-Input Single-Output (MISO) and MIMO for WiMAX communication systems. The performances of these communication techniques are analyzed and compared for three scenarios - rural environment, TGV (high-speed train) environment and urban environment by using the models to investigate several communication parameters.

Keywords: MIMO, MISO, SISO, Wimax, Space-Time Coding, Alamouti Space-Frequency Block Coding

1. Introduction

In wireless communication systems, MIMO, pronounced my-moh or mee-moh, refers to a link for which the transmitting end as well as the receiving end is equipped with multiple antenna elements. Like MIMO, MISO is another smart antenna technology, but characterized by multiple antennas only at the transmitting end. To understand smart antenna technology, it is best to consider an example in which, say, you are in a room. Someone in the room is talking to you and, as he speaks, he begins moving around the room. Your ears and brain have the ability to track where the user's speech is originating from as he moves throughout the room. This is similar to how smart antenna systems operate. They locate the users, track them, and provide optimal RF signals to them as they move throughout the base station's coverage area. MIMO is rapidly becoming the face of smart antenna technology. On the other hand, SISO, which has a single antenna at both transmitting and receiving ends, is the simplest and cheapest to implement among the three and has been in use since the birth of radio technology. MIMO promises to resolve the bottlenecks of traffic capacity in the forthcoming high speed wireless broadband wireless internet access networks like Worldwide Interoperability for Microwave Access (WiMAX), 3G-Long Term Evolution (see [3]) and beyond.

In this paper we have limited our analysis to the WiMAX system and/or mobile-WiMAX system, which were based on the IEEE 802.16-2004 standard and IEEE 802.16e-2005 standard respectively. In essence, WiMAX is a 4G technology for a state-of-the-art ''last mile'' telecommunication infrastructure (see [4,9]). WiMAX is poised to replace a number of existing broadband telecommunication infrastructure for wireless local loop, while mobile-WiMAX can replace cellular networks.

There are several ways to implement MIMO systems, such as, BLAST described by G. J. Foschini (see [1,2]), space-time coding (see [5–7,10,11,13]) and more. However, we have stuck to the Alamouti space-time block code proposed by Siavash Alamouti in 1998 (see [12]). This code achieves transmit diversity by correlating the transmit symbols spatially across the two transmit antennas, and temporally across two consecutive time intervals. The only condition is that the channel should remain stationary over two consecutive symbols. Although the Alamouti code achieves the same rate as SISO, it attains maximum diversity for two transmit antennas. The greatest advantage it offers is that the coding and the decoding mechanisms it symbolizes are remarkably simple and equally effective. The code also provides the lowest probability of error and implementation complexity among all MIMO implementation techniques. At the receiver end we use Maximum Likelihood

(ML) detection technique which largely does an exhaustive search among all received signals in order to find the optimum received signal (see [7]). Importantly, the performance of the Alamouti code depends on an accurate estimation of the channel between the transmitter and the receiver. Transmission of training symbols is used to perform channel estimation (discussed in Section 2).

In this paper we have analyzed the performance of a WiMAX-MIMO system (by means of Bit Error Rate (BER)) vis-à-vis WiMAX-MISO and WiMAX-SISO systems using the MATLAB simulation tool. Analyses for the following scenarios have been performed.

- Rural environment: We consider an environment with no obstacles, hence no fading takes place. Also, the transmitter and the receiver are in zero relative motion (stationary).
- Train à Grande Vitesse (TGV)/High-speed train environment: We consider a doppler fading environment. The transmitter is stationary and the receiver is sitting in a TGV moving at its top speed of 574.8 km/hr.
- Urban environment: We consider a static multipath environment with a LOS link between transmitter and receiver. Again, the transmitter and receiver are in zero relative motion.

We will also compare the behavior of the three models to varying SNR and number of input bits in all the three environments. This paper is structured as follows. Section 2 introduces the Alamouti space-frequency block code; it's encoding and decoding scheme and how it differs from the well-known Alamouti space-time block code. In Section 3, we present our WiMAX- MIMO/ MISO/SISO communication models with a detailed description of the complete layout of each model. Section 4 is devoted to the results obtained from computer simulations for different analyses performed to compare the communication performances of the three models and also proposes a hybrid model (part SISO - part MISO - part MIMO) which can be implemented for rural environments. Finally, Section 5 presents the conclusion.

2. Alamouti Space-Time Block Code

In this paper we have focused on the Alamouti coding scheme, precisely the Alamouti space-frequency block code which is a slight variant of the Alamouti space-time block code (see [11]).

For implementing MIMO for WiMAX systems, we have employed the desired diversity differently at the reception and transmission. The reception employs Alamouti space-time block code (STBC) while the transmission employs an Alamouti space-frequency block code (SFBC). The motivation behind such a variation is that STBC requires the channel to be stationary over two consecutive OFDMA symbols (also see [14–17]). However, in a fast-fading radio channel, this is

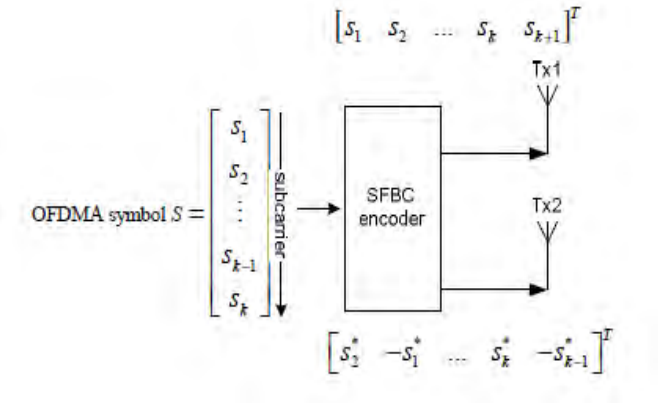


Figure 1. SFBC encoding scheme.

not always true. In SFBC, the coding is implemented across two consecutive sub-carriers in the frequency domain and thus within the OFDMA symbol. This eliminates the aforementioned handicap posed by STBC.

Figure 1 illustrates the encoding scheme for SFBC's. As clearly visible, the mapping scheme is designed in such a way that the first antenna transmits the entire symbol stream without any modification, also facilitating the system to act as a SISO system provided antenna two is switched off. However, it is assumed that two adjacent sub-carriers in the frequency domain experience correlated fading. This assumption holds in channels where the delay spread is low enough for the resulting coherence bandwidth to exceed twice the sub-channel spacing. Also, this is the reason why SFBC cannot be used for reception.

For transmitted symbols X_1 and X_2 the receiver antenna obtains the received symbol r_1 and r_2 for a 2x1 MISO system as

$$r = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2^* \\ x_2 & -x_1^* \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = Xh + z$$

where Z is the Additive White Gaussian Noise (AWGN) and h_1 and h_2 are the channel coefficients.

The optimal estimates for h_1 and h_2 can be obtained by linear processing at the receiver, and are given by

$$\tilde{h} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \frac{1}{\|x\|^2} X^* r = \begin{pmatrix} h_1 + \tilde{z}_1 \\ h_2 + \tilde{z}_2 \end{pmatrix}$$

where

$$\tilde{z}_1 = \frac{x_1 z_1 + x_2^* z_2}{\|x\|^2}; \quad \tilde{z}_2 = \frac{x_2 z_2 + x_1^* z_2}{\|x\|^2}$$

These channel estimates can then be used to detect the next pair of code symbols. After the next code symbols are decoded, the channel estimate can be updated using

those decoded symbols (see [9]). When the channel variation is slow, the receiver improves stability of the decoding algorithm by averaging old and new channel estimates.

The decision for the two transmitted symbols or in other words the estimate of the two transmitted symbols according to ML estimation, $\sim X_1$ and $\sim X_2$, is given by:

$$\tilde{x}_1 = h_1^* r_1 + h_2 r_2^*$$
$$\tilde{x}_2 = h_2^* r_1 + h_1 r_2^*$$

Similarly, the above scheme can be extended to 2 receiver antennas and hence to N_r receiver antennas.

3. Wimax-MIMO/MISO/SISO Communication Model

This section presents the communication model block diagram of all the three models and talks about the nuances of each of them. It is interesting to note that the MIMO and the MISO models have a far more complex implementation than the simple SISO model (Figures 5, 6 & 7).

SISO communication systems are vulnerable to environments characterized by problems caused by multipath effects. Figure 2 illustrates a real-time model of a SISO system. On the other hand, the MISO transmission strategy maximizes the received SNR by adding up the received signal from all transmit antennas in-phase and by allocating more power to the transmit antennas. MISO wireless communication system exhibits transmitter diversity. Some of the transmitter diversity techniques include frequency weighting, antenna hopping, delay diversity and channel coding (see [8]). The real-time model of a MISO system is similar to the one in Figure 3; however, there is only one antenna at the receiver end. The MIMO system exhibits both transmitter diversity and receiver diversity. While the transmitter diversity techniques have already been discussed, some of the receiver diversity techniques include selection diversity, antenna diversity, maximal ratio combining and equal gain combining (see [11]). Figure 3 illustrates a real-time model of a MIMO system.

The advantages of using MIMO systems are increased spectral efficiency, throughput, coverage, capacity, better BER and resistivity to fading effects to name a few. However, the greatest challenge it faces is the necessity of complex DSP circuitry and the fact that its promise of better communication performance hold true most only for scattering-rich environments.

Figures 5, 6 & 7 illustrate the block diagram for urban environment for the WiMAX-SISO, WiMAX-MISO and WiMAX-MIMO communication model discussed in this paper. We have taken a 200x200 black & white image (Figure 4) as the input to the communication system.

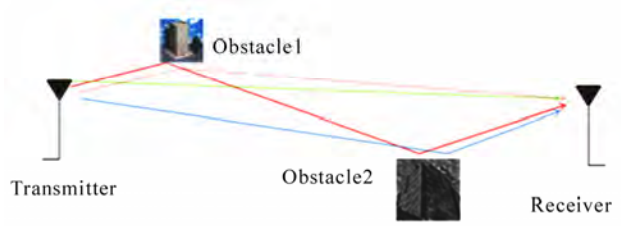


Figure 2. Real Time SISO communication model.

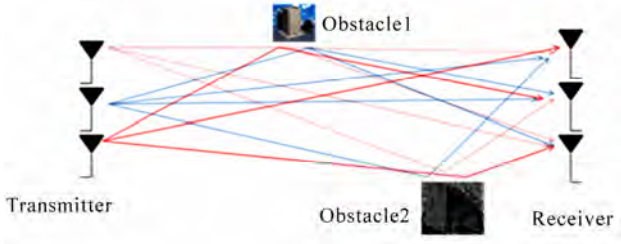


Figure 3. Real Time MIMO communication model.

Figure 4. 200×200 black&white input image.

We have used an in-built MATLAB function 'qammod (data, index) to perform 16-QAM and similarly 'qamdemod (data, index)' to perform 16-QAdeM. The SFBC encoder and decoder are designed in accordance with the equations already talked about in Section 2. The characteristics of the models for the three environments that we have used for our analysis are given below.

- Rural environment: There is no 'Multipath Rayleigh filter 'block in the block diagram and the output of 'peak power clipping' goes to the 'add Gaussian noise' block.
- TGV environment: The 'Multipath Rayleigh filter' block is replaced by the 'Doppler fading filter' block. We have used an in-built MATLAB function ''rayleighchan (sampling freq, Doppler spread, 'path-delays', 'path-gains')'' to realize the doppler fading environment by setting Doppler spread as maximum of 1300 Hz, corresponding to 574.8 km/hr. Path delay is set to zero.
- Urban environment: We again use the same in-built MATLAB function mentioned in TGV environment to realize the multipath fading environment. We have taken a total of 15 multipath delays generated randomly.

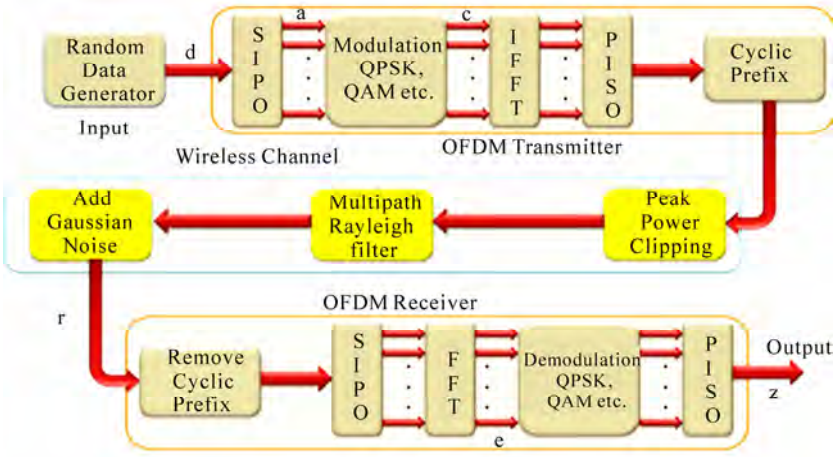


Figure 5. WiMAX-SISO communication model for urban environment.

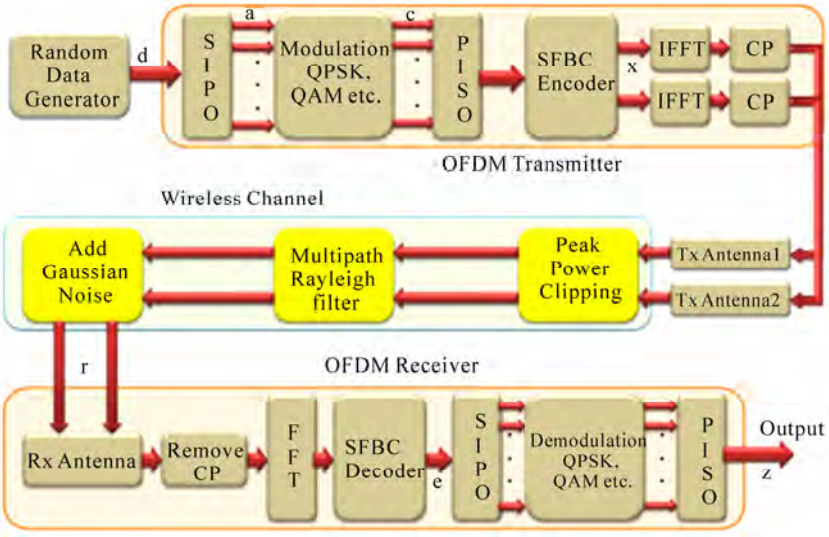


Figure 6. WiMAX-MISO communication model for urban evironment.

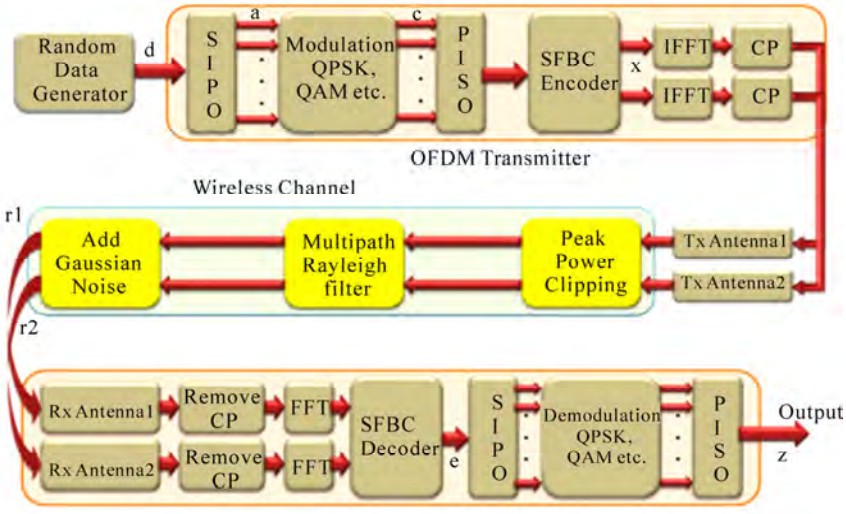


Figure 7. WiMAX-MIMO communication model for urban environment.

Figure 8. Comparison of SISO, MISO & MIMO model for rural, TGV and urban environments.

The output is compared with the input data and the BER is calculated. Also, importantly, it should be noted that we have not used channel state information at the transmitter (CSIT) for our analysis.

4. Simulation Results

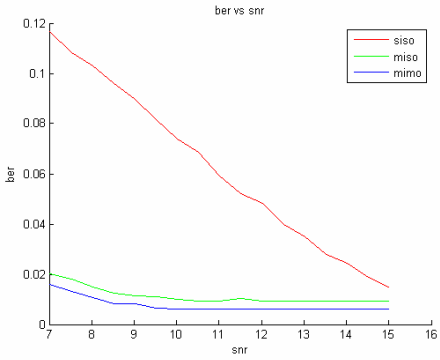
4.1. SISO / MISO / MIMO Comparison in Rural, Urban and High Speed Environments.

This section contains the results of simulations carried out to compare the communication performance (BER) of the three models in the aforementioned three environments. Also, we analyze the performance of the three models by varying the SNR and the number of input bits for all the three environments.

For the first part of our analysis, we have fixed the E_b/N_0 value to be 3dB and compared the behavior of the three communication models in all the three environments. We have used the image in Figure 4 as the input data.

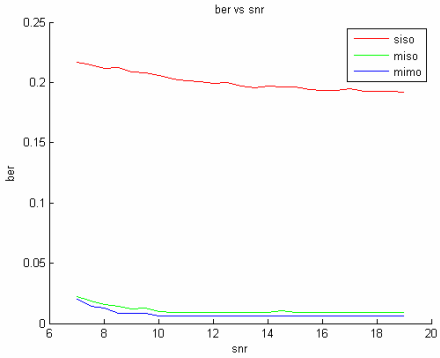
Figure 8 illustrates the results thus obtained. Here 'q'

Rural environment



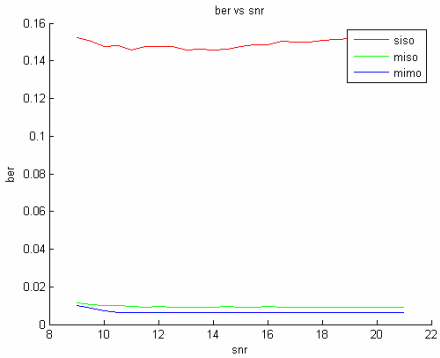
MODEL	AVERAGE	VARIANCE
SISO	0.0622	0.0011
MISO	0.0112	1.0911e-5
MIMO	0.0077	8.3562e-6

TGV environment



MODEL	AVERAGE	VARIANCE
SISO	0.2003	5.8135e-5
MISO	0.0109	1.1977e-5
MIMO	0.0075	1.0903e-5

Urban environment



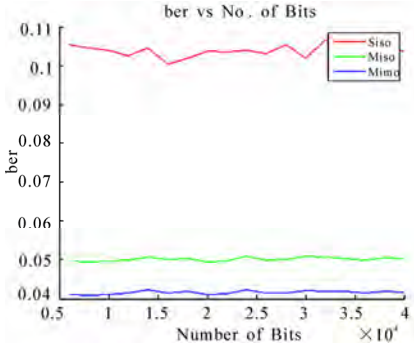
MODEL	AVERAGE	VARIANCE
SISO	0.1491	5.7380e-6
MISO	0.0094	3.0720e-7
MIMO	0.0064	8.2893e-7

Figure 9. Bervs SNR for SISO, MISO & MIMO models for rural, TGV and urban environments.

is the number of wrong bits when compared to the input data. The following conclusions can be drawn from the simulation results.

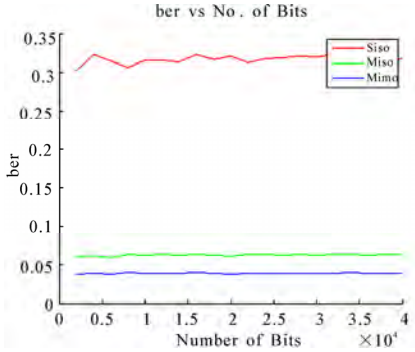
- The image is least distorted for the MIMO model followed by the MISO model and finally the SISO model for all the three environments. Also, there is a great difference between the MISO/MIMO model and the SISO model especially for the TGV and urban environment. However, there is not much difference between the MIMO model and the MISO model for any environment since the only difference between them is the inherent receiver diversity in the MIMO model.
- Another notable point is that the Bit- Error Rate (BER) for MIMO model decreases for urban and TGV environment as compared to rural environment. However, for the other two models, i.e. MISO model and SISO model, the BER increases for urban and TGV environment as compared to rural environment. However, unlike in the SISO model, the increase is a slight one for the MISO model. This result vindicates the point that a fading environment improves the performance of the MIMO model. Hence, fading is our friend!

Rural environment



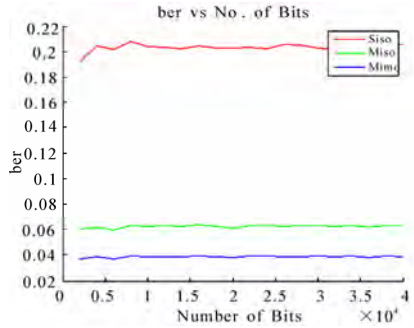
MODEL	AVERAGE	VARIANCE
SISO	0.1038	2.0936e-6
MISO	0.0500	2.1333e-7
MIMO	0.0417	1.6768e-7

TGV environment



MODEL	AVERAGE	VARIANCE
SISO	0.3165	2.6215e-5
MISO	0.0621	1.5024e-6
MIMO	0.0384	3.3734e-7

Urban environment



MODEL	AVERAGE	VARIANCE
SISO	0.2035	1.0377e-5
MISO	0.0623	1.2111e-6
MIMO	0.0384	3.7565e-7

Figure 10. Ber vs. number of bits for SISO, MISO & MIMO models for rural, TGV and urban environments.

For the second part of our analysis, we have varied the E_b/N_0 value, hence the SNR value, and found out the Bit-Error Rate (BER) value for all the three communication models each time. Again, we have used the image in Figure 4 as our input data. We have then plotted the BER vs. SNR curve for all the three models in the same graph. Also, we have performed the analysis for all the three environments. Figure 9 illustrates the results obtained from the aforementioned simulation process. The following conclusions can be drawn.

- The BER is highest for the SISO model and the least for the MIMO model for all the three environments, as can be inferred from the graph and the average BER values from the tables.
- It can also be inferred from the graphs that, as we increase the SNR values, the BER decreases for the SISO model in all the three environments with the exception of urban environment. However, for MISO and MIMO models, as we increase the SNR value, the BER decreases till a certain SNR value and then become steady at a fixed value irrespective of the SNR. The exception of the SISO model curve in urban environment can be attributed to the inability of the model to counter multipath effects.
- Another point to be noted is that, the average BER for MIMO model decreases for urban and TGV environment on comparison to rural environment. But, this is not the scenario for SISO model. Hence, we can infer

that multipath and similar fading environments truly help MIMO in its communication performance.

- Variance is not the greatest parameter to be compared for this analysis.

4.2. Influence of Data Size on Transmission Performances

For the third part of our analysis, we have fixed the E_b/N_0 value to 3 dB and varied the number of input bits and found out the Bit-Error Rate (BER) for all the three communication models each time. Here we have used randomly generated input data for this analysis. We have then plotted the BER vs. number of bits curve for all the three models in the same graph. Again, we performed the analysis for all the three environments. Figure 10 illustrates the results obtained from the aforementioned simulation process. The following conclusions can be drawn.

- The BER is highest for the SISO model and the least for the MIMO model for all the three environments, as can be inferred from the graph and the average BER values from the tables.
- Also, it can be inferred from the graphs that as we increase the number of bits, the BER values remain more or less constant for all the three models with the MIMO model providing the best results in terms of stability of the curve. Hence, we can infer that the communication performance of all the three models is irrespective of the number of bits.
- The variance of MIMO model is the least for all the three environments as compared to the SISO and MISO model. This states that the MIMO model provides us with the maximum stability in the communication performance for the aforementioned analysis.
- Also, it is interesting to note that the average BER for the MIMO model decreases for urban and TGV environments on comparison to rural environment. But, this is not the case for SISO. This again vindicates the point that multipath and other fading environments prove to be favorable for MIMO's communication performance.
- Finally, we can infer that fading environments prove to be a big downfall for the SISO model as the BER increases enormously for TGV and urban environments.

4.3. Hybrid Model for Mobile Power Saving

The purpose of this final part is to propose a hybrid model where the mobile will switch between SIMO, MISO and MIMO communications models, depending of radio transmission performances. The aim is to obtain an optimal radio performance over power consumption ratio. Using multiple antennas at the MS will utilize much power, which is a great source of concern as the MS has limited battery power. Such a hybrid model will activate antennas as a function of radio conditions.

To investigate this hybrid model, we have considered the transmitted power and the received power for various E_b/N_0 values. Then, for each power value thus calculated we have found the distance between the transmitter (base station-BS) and the receiver (mobile terminal-MS) by using the relation given below (also see [18])

$$P_r = P_t G_t G_r \frac{h_t^2 h_t^2}{d^4}$$

The above expression is valid for flat-terrain mobile communication environments and hence can be applied for our rural environment. The values for Pr (received power) and Pt (transmitted power) are found from the model using MATLAB. We assume Gt and Gr to be19dBi and 10dBi respectively, while ht and hr are assumed to be 50m and 1.5m respectively. All the as-

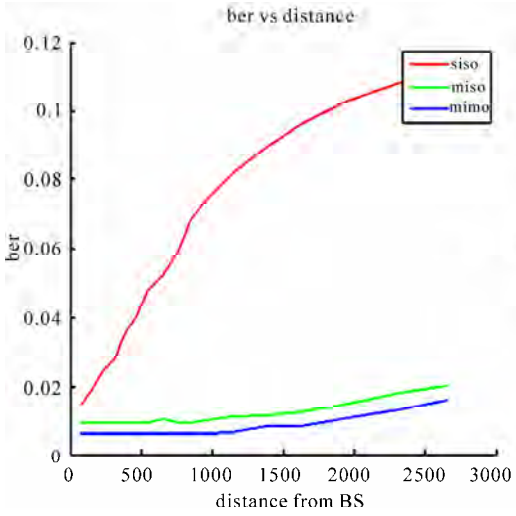


Figure 11. Ber vs distance for SISO, MISO &MIMO.

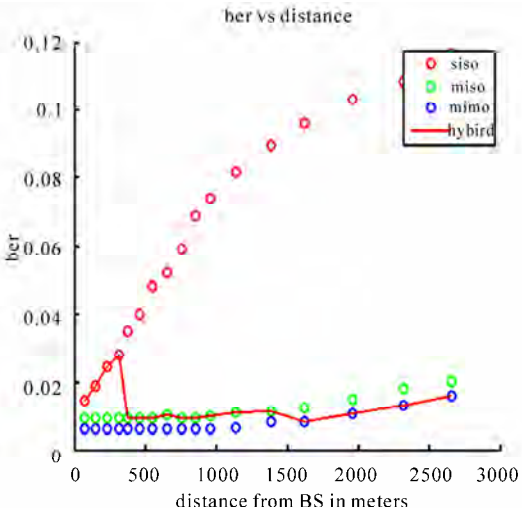


Figure 12. Ber vs distance for hybrid model.

sumed values are in accordance with the WiMAX standards (see [4,9]). We have then plotted the BER vs. distance curve for all the three models in the same graph, as illustrate in Figure 11.

As a result, a hybrid model scenario for rural environment is proposed in Figure 12. Depending of the MS-BS distance, the BER and the power consumption, the mobile choose the optimal system: SISO, MISO or MIMO.

5. Conclusions

This paper shows the use of Alamouti space-frequency block codes, a slight variant of the well-known Alamouti space-time block code, to design MISO and MIMO communication models for WiMAX systems for three environments, namely, rural, TGV/high-speed train and urban environment. The performances of the three models (SISO, MISO & MIMO) are compared for all the three environments with MIMO model clearly surpassing the other two models in every environment. This paper also notifies the improvement in the performance of MIMO systems in fading environments and also how such environments prove to be a downfall for the other models. The simulation results obtained from BER vs. number of bits analysis confirm that the MIMO model offers the maximum stability even if we have large input data bits. While those obtained for the BER vs. SNR analysis emphasize on the growing need for implementing MIMO enhanced communications systems (in this paper WiMAX system) especially for fading environments similar to urban and TGV (high-speed train) environment discussed in this paper. Such a step, if taken, will not only increase the coverage area of the communication system, but also allow for uninterrupted communication service to be possible even at the edges of the hexagonal cell. However, such an implementation would increase the power consumption at the user end. To counter this problem we have proposed our hybrid model, as of now for the rural environment, where-in, the communication system can switch from a SISO to MISO to MIMO depending upon the communication parameters (here BER). Hence this will ensure controlled power consumption as well as good communication performance. However, more research in this direction needs to be done especially at the various network layers (MAC layer). Also, such a system needs to be expanded to the urban environments as well.

6. References

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," Bell Laboratories Technical Journal, Vol. 2, pp. 41–59, October 1996.

[2] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," URSI International Symposium on Signals, Systems, and Electronics, 1998.

[3] E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, and S. Parkvall, "Key features if the LTE radio interface," Ericsson Review, No. 2, 2008.

[4] IEEE 802.16a Standard and Wimax Igniting Broadband Wireless Access, White paper, WiMAX Forum, September 2003.

[5] D. Gesbert, University of Oslo, "Lecture to 5-th year NTNU students," September 2003.

[6] D. Gesbert, M. Shafi, D. Shiu, P. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," IEEE Journal on selected areas of communication (JSAC), April 2003.

[7] M. Jankiraman, "Space-time codes and MIMO systems, "Artech house Publishers, July 2004.

[8] D. Tse, P. Vishwanath, "Fundamentals of wireless communication," Cambridge University Press, June 2005.

[9] S. Ahson and M. Ilyas, WiMAX: Technologies, Performance Analysis, and QoS, CRC, 1 Edition, pp. 41–96, September 2007.

[10] G. Tsoulos, "MIMO system technology for wireless communications," CRC, 1 Edition, March 2006.

[11] B. Vucetic, J. H. Yuan, "Space-time coding," Wiley, pp. 49–148, 245-264, June 2003.

[12] S. M. Alamouti, "A simple transmit diversity scheme for wireless communications, IEEE Journal on selected areas of communication (JSAC)," Vol. 16, pp 1451–1458, October 1998.

[13] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal design," IEEE Transactions on Information theory, Vol. 45, pp. 1456–1467, July 1999.

[14] H. Bolcskei and A. Paulraj, "Space-frequency codes for broadband fading channels," Proceedings IEEE International Symposium on Information Theory, pp. 219, 2001.

[15] B. Muquet, E. Biglieri, A. Goldsmith, and H. Sari, MIMO Techniques for Mobile Wimax Systems, SEQUANS Communications White Paper.

[16] Mai Tran, Angela Doufexi, and Andrew Nix, Mobile WiMAX MIMO Performance Analysis: Downlink and Uplink, Centre for communication research, University of Bristol, 2008, IEEE.

[17] A. A. Hutter, S. Mekrazi, B. N. Getu, and F. Platbrood, "Alamouti-based space-frequency coding for OFDM, Wireless Personal Communication Review, Vol. 35, pp. 173–185, October 2005.

[18] T. S. Rappaport, Wireless Communication Principles and Practices, Prentice Hall PTR, July, 2002.

Engineering, 2010, 2, **-**
doi:10.4236/eng.2010.21005 Published Online January 2010 (http://www.scirp.org/journal/eng/).

Consolidation Solutions of a Saturated Porothermoelastic Hollow Cylinder with Infinite Length

Consolidation Solutions of Porothermoelastic Media

Bing BAI
School of Civil Engineering, Beijing Jiaotong University, Beijing, China
Email: *Baibing*66@263.*net*
Received June 4, 2009; *revised July* 28, 2009; *accepted August* 10, 2009

Abstract

An analytical method is derived for the thermal consolidation of a saturated, porous, hollow cylinder with infinite length. The solutions in Laplace transform space are first obtained and then numerically inverted by Stehfest method. Two cases of boundary conditions are considered. First, variable thermal loadings are applied on the inner and outer pervious lateral surfaces of the hollow cylinder, and a variable mechanical loading with time is applied on the outer surface; while the displacement of the inner surface remains fixed. Secondly, variable thermal and mechanical loading are applied on the outer pervious surface, and the inner surface remains fixed, impervious and insulated. As two special problems, a solid cylinder with infinite length and a cylindrical cavity in a half-space body are also discussed. Finally, the evolutions of temperature, pore pressure and displacement with time along radial direction are analyzed by a numerical example.

Keywords: Porothermoelastic Media, Hollow Cylinder, Variable Thermal Loading, Consolidation Solutions, Stehfest Method

1. Introduction

The studies on the thermo-hydro-mechanical responses of saturated porous materials are widely used in various engineering fields such as the disposal of high-level nuclear waste, extraction of geothermal energy, storage of hot fluids, biomechanics to materials sciences, concrete resistance against fire, reliability of airfield.

There exists a substantial and growing literature to account for non-isothermal consolidation behavior of fluid-saturated porous materials [1–5]. Up to now, some analytical solutions to boundary and initial value problems have been developed under various scenarios. Booker and Savvidou [6] have presented solutions for the temperature, pressure and stress fields arising from a spherical heat source buried in a thermally consolidating material of infinite extent. McTigue [7] presented resolution methods and established exact solutions for a semi-infinite porous medium subjected to a constant surface temperature or heat flux with either drained or undrained boundary conditions. Smith and Booker [8] presented the Green's functions for a system of fully coupled linear equations governing thermal consolidation in a homogeneous isotropic material, and later gave a boundary integral method of numerical analysis. Giraud *et al*. [9] analyzed the case of a heat source that decreases exponentially with time by considering a low-permeability clay for nuclear waste disposal. Wang and Papamichos [10,11] discussed solutions for a cylindrical wellbore and a spherical cavity subjected to a constant temperature change and heat flow rate. Blond *et al.* [12] developed a closed-form solution for a porothermoelastic half-space submitted to a cyclic thermal loading, and a pressure-diffusion equation that governs the fluctuation of the interstitial pressure was established. Bai [13] developed a solution approach for a planar thermal loading with variable intensity with time on the surface of a semi-infinite space. Bai [14] later derived an analytical method for the responses of saturated porous media subjected to cyclic thermal loading by using the Laplace transform and the Gauss-Legendre method of Laplace transform inversion. Abousleiman and Ekbote [15] presented the analytical solutions for an inclined hollow cylinder in a transversely isotropic material subjected to thermal and stress perturbations. Kanj *et al.* [16,17] applied an anisotropic porothermoelastic solution to an unjacketed hollow cylinder in a triaxial set-up. Bai [18] derived an analytical method for the thermal consolidation of layered, saturated porous half-space to variable thermal loading with time.

In previous studies, the most concerned domains are the geometries such as a half-space body, a cylindrical cavity in a semi-infinite space and a cylindrical body, etc. In fact, hollow cylindrical geometries are also widely used in the laboratory for measurements of material properties and the understanding of the subsidence phenomena, formation consolidation and borehole stability. In this paper, analytical solutions for a hollow cylinder of porothermoelastic media with infinite length are derived aiming at experimental studies of the thermal consolidation of saturated porous materials such as soil that are carried out under non-isothermal conditions. The solutions in Laplace transform space are first obtained and then numerically inverted by Stehfest method. As two special problems, a solid cylinder with infinite length and a cylindrical cavity in a half-space body are also discussed. Based on the proposed solutions, numerical analyses are carried out to demonstrate the evolutions of temperature, pore pressure, displacement as well as radial and tangential stresses with time.

2. Governing Equations

For saturated, homogeneous, isotropic porous materials, the equilibrium equation of thermo-hydro-mechanical coupling consolidation may be written as [8,10,12,14]

$$M\,\nabla^2\varepsilon_v-\alpha\nabla^2 p-\beta\nabla^2\theta=0 \tag{1}$$

where $M=\lambda+2G$ is the confined drained isothermal modulus; λ and G are Lamé constants; ∇^2 is the Laplace operator; ε_v is the volumetric strain; p is the excess pore pressure; $\theta=T-T_0$ is the increment of temperature above the ambient temperature, T_0 is the ambient absolute temperature, T is the current absolute temperature; $\alpha=1-C_s/C$ is Biot's coefficient, C and C_s are the coefficients of volumetric compression of the solid skeleton and grains respectively; $\beta=3\alpha_s/C$ is the thermal expansion factor, and α_s is the linear thermal expansion coefficient of solid grains.

According to Darcy's law and the continuity condition of seepage, the equation of mass conservation can be written as [12,14]

$$\int_0^t\frac{k}{\gamma_w}\nabla^2 p\,\mathrm{d}t-\alpha\varepsilon_v+Y\theta-\alpha_p p=0 \tag{2}$$

where k is the hydraulic conductivity; γ_w is the unit weight of pore water; t is time; $\alpha_p=n(C_w-C_s)+\alpha C_s$; $Y=-3n(\alpha_w-\alpha_s)-3\alpha\alpha_s$, C_w is the coefficient of volumetric compression of pore water; α_w is the linear thermal expansion coefficient of pore water; and n is the porosity of the medium.

According to Fourier's law of heat conduction, the equation of energy conservation can be written as [12,14]

$$\int_0^t\frac{K}{T_0}\nabla^2\theta\,\mathrm{d}t-Z\theta-\beta\varepsilon_v+Y p=0 \tag{3}$$

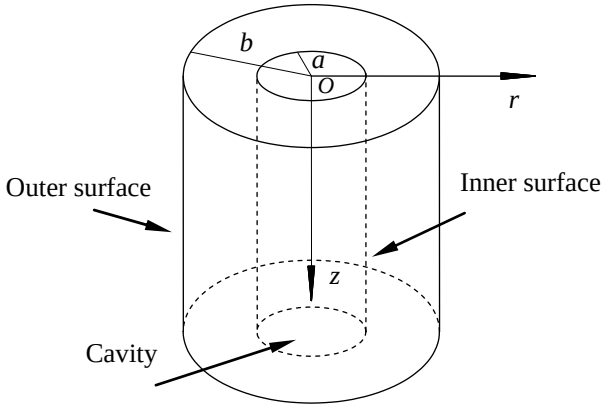


Figure 1. Mathematical model.

where K is the coefficient of heat conductivity, $Z=[(1-n)\rho_s c_s+n\rho_w c_w]/T_0-3\beta\alpha_s$ is a coupling parameter, with ρ_w and ρ_s being the densities of pore water and solid grains, respectively, and c_w and c_s being the heat capacities of pore water and solid grains, respectively.

3. Mathematical Model

Consider the problem of a saturated, porous hollow cylinder with infinite length (see Figure 1). The following two cases of boundary conditions are imposed here.

Case 1: Variable thermal loading $\theta_a(t)$ and $\theta_b(t)$ are respectively applied on the inner and outer pervious lateral surfaces of the hollow cylinder. At the same time, a variable mechanical loading $p_b(t)$ is also applied on the outer surface; while the displacement of the inner surface remains fixed. The origin O of the cylindrical coordinate system is selected at the center of the cylinder and the z-axis is the axis of rotational material symmetry, so that

$$\theta(a,t)=\theta_a(t)\cdot H(t),\quad p(a,t)=0,\quad u_r(a,t)=0\quad (t>0) \tag{4}$$

$$\theta(b,t)=\theta_b(t)\cdot H(t),\quad p(b,t)=0,\quad \sigma_r(b,t)=p_b(t)\cdot H(t)\quad (t>0) \tag{5}$$

where a and b are the inner and outer radius of the hollow cylinder, respectively; $\theta_a(t)=T_{w1}(t)-T_0$, $\theta_b(t)=T_{w2}(t)-T_0$, with $T_{w1}(t)$ and $T_{w2}(t)$ are the current absolute temperatures of the inner and outer surfaces, respectively; $p_b(t)$ is the mechanical loading of the outer surface; u_r is the radial displacement; and $H(t)$ means the Heaviside unit step function.

Case 2: Variable thermal and mechanical loading $\theta_b(t)$ and $p_b(t)$ are respectively applied on the outer pervious surface of the hollow cylinder. The inner surface remains fixed, impervious and insulated. Thus, the boundary conditions of the outer surface can still be expressed by Equation (5); while the boundary conditions of the inner surface are

$$\frac{\partial \theta(a,t)}{\partial r}=0\,,\quad \frac{\partial p(a,t)}{\partial r}=0\,,\quad u_r(a,t)=0 \qquad (6)$$

It is noted that all applied boundary conditions on the cylinder may be time-dependent.

4. Solution Approach

4.1. Solutions of the Governing Equations

It is assumed that the initial conditions (Figure 1) are: $\theta(r, 0)=0$, $p(r, 0)=0$, and $u_r(r, 0)=0$. Then, upon Laplace transformation, Equations (1) to (3) become

$$M\,\nabla^2\bar{\varepsilon}_v-\alpha\nabla^2\bar{p}-\beta\nabla^2\bar{\theta}=0 \qquad (7)$$

$$\frac{k}{\gamma_w}\nabla^2\bar{p}-\alpha s\bar{\varepsilon}_v+Y s\bar{\theta}-\alpha_p s\bar{p}=0 \qquad (8)$$

$$\frac{K}{T_0}\nabla^2\bar{\theta}-Z s\bar{\theta}-\beta s\bar{\varepsilon}_v+Y s\bar{p}=0 \qquad (9)$$

where $\nabla^2=\partial^2/\partial r^2+(\partial/\partial r)/r$, and $\bar{L}=\int_0^\infty e^{-st}L\,dt$ ($L=\theta, p, \varepsilon_v$) and s is the Laplace transform variable.

Equation (7) can be rewritten as

$$\nabla^2[M\,\bar{\varepsilon}_v-\alpha\,\bar{p}-\beta\,\bar{\theta}]=0 \qquad (10)$$

The integration of Equation (10) twice over r yields

$$M\,\bar{\varepsilon}_v-\alpha\,\bar{p}-\beta\,\bar{\theta}=h_1(s)\ln r+h_2(s) \qquad (11)$$

where h_1 and h_2 are arbitrary functions of s to be determined from the boundary conditions.

From Equation (11), one has

$$\bar{\varepsilon}_v=\frac{\alpha\,\bar{p}+\beta\bar{\theta}}{M}+D_1\ln r+D_2 \qquad (12)$$

where $D_1=h_1/M$, $D_2=h_2/M$.

Substituting Equation (12) into Equations (8) and (9) results in, respectively

$$a_2\nabla^2\bar{p}+a_3 s\bar{\theta}+a_4 s\bar{p}=\alpha(D_1\ln r+D_2)s \qquad (13)$$

$$b_1\nabla^2\bar{\theta}+b_3 s\bar{\theta}+b_4 s\bar{p}=\beta(D_1\ln r+D_2)s \qquad (14)$$

where $a_2=k/\gamma_w$, $a_3=Y-\alpha\beta/M$, $a_4=-\alpha_p-\alpha^2/M$, $b_1=K/T_0$, $b_3=-Z-\beta^2/M$, $b_4=Y-\alpha\beta/M$.

Here, $c=-a_2/a_4=k/[\gamma_w(\alpha_p+\alpha^2/M)]$ is defined as the coefficient of thermal consolidation defined in previous work [8,11], $\kappa=-b_1/b_3=(K/T_0-\alpha_w K_w S_w)/(Z+\beta^2/M)$ is defined as the thermal diffusivity. In fact, the ratio c/κ reflects the relative rate of pore pressure dissipation to heat conductivity.

Using Equations (13) and (14), the elimination of term $\nabla^2\bar{\theta}$ leads to

$$\bar{\theta}=\frac{f_1}{s}\nabla^2\bar{p}+f_2\bar{p}+f_3(D_1\ln r+D_2) \qquad (15)$$

where $f_1=-a_2/a_3$, $f_2=-a_4/a_3$, $f_3=\alpha/a_3$.

Substitution of Equation (15) into Equation (14) results in

$$g_1\nabla^4\bar{p}+g_2 s\nabla^2\bar{p}+g_3 s^2\bar{p}+g_4 s^2(D_1\ln r+D_2)=0 \qquad (16)$$

where $g_1=b_1f_1$, $g_2=b_1f_2+b_3f_1$, $g_3=b_3f_2+b_4$, $g_4=b_3f_3-\beta$.

It can be proved that the general solution of Equation (16) is

$$\bar{p}=A_1I_0(\sqrt{s\xi}\,r)+B_1K_0(\sqrt{s\xi}\,r)+A_2I_0(\sqrt{s\eta}\,r)+B_2K_0(\sqrt{s\eta}\,r)-\frac{g_4}{g_3}(D_1\ln r+D_2) \qquad (17)$$

where A_1, B_1, A_2 and B_2 are arbitrary functions of s to be determined from the boundary conditions, I_0 and K_0 are the modified first-kind and second-kind Bessel functions of order zero, respectively,

$$\xi=(-g_2-\sqrt{g_2^2-4g_1g_3})/(2g_1)$$

and

$$\eta=(-g_2+\sqrt{g_2^2-4g_1g_3})/(2g_1)\,.$$

Substitution of Equation (17) into Equation (15) results in

$$\bar{\theta}=(f_1\xi+f_2)[A_1I_0(\sqrt{s\xi}\,r)+B_1K_0(\sqrt{s\xi}\,r)]+(f_1\eta+f_2)[A_2I_0(\sqrt{s\eta}\,r)+B_2K_0(\sqrt{s\eta}\,r)]+(f_3-f_2\frac{g_4}{g_3})(D_1\ln r+D_2) \qquad (18)$$

Substitution of Equations (17) and (18) into Equation (12) results in

$$\bar{\varepsilon}_v=E_1[A_1I_0(\sqrt{s\xi}\,r)+B_1K_0(\sqrt{s\xi}\,r)]+E_2[A_2I_0(\sqrt{s\eta}\,r)+B_2K_0(\sqrt{s\eta}\,r)]+E_3(D_1\ln r+D_2) \qquad (19)$$

where $E_1=[\alpha+\beta(f_1\xi+f_2)]/M$, $E_2=[\alpha+\beta(f_1\eta+f_2)]/M$, $E_3=-\alpha g_4/(Mg_3)+\beta(f_3-f_2g_4/g_3)/M+1$.

4.2. Displacement, Stress and Strain

Introduce a displacement potential function $\psi(r, t)$, then

$$u_r=\frac{\partial\psi}{\partial r} \qquad (20)$$

Using Equation (19) and $\varepsilon_v=\partial^2\psi/\partial r^2+(\partial\psi/\partial r)/r=\nabla^2\psi$, one has

$$\nabla^2\bar{\psi}=E_1[A_1I_0(\sqrt{s\xi}\,r)+B_1K_0(\sqrt{s\xi}\,r)]+E_2[A_2I_0(\sqrt{s\eta}\,r)+B_2K_0(\sqrt{s\eta}\,r)]+E_3(D_1\ln r+D_2) \qquad (21)$$

The general solution of Equation (21) can be ex-

pressed as

$$\overline{\psi} = \frac{E_1}{s\xi}[A_1 I_0(\sqrt{s\xi}\, r) + B_1 K_0(\sqrt{s\xi}\, r)] + \frac{E_2}{s\eta}[A_2 I_0(\sqrt{s\eta}\, r) + B_2 K_0(\sqrt{s\eta}\, r)] + \frac{E_3}{4}[D_1 r^2(\ln r - 1) + D_2 r^2] + (D_3 \ln r + D_4) \quad (22)$$

where D_3 and D_4 are arbitrary functions of s to be determined from the boundary conditions.

Using Equations (20) and (22) results in

$$\overline{u}_r = \frac{E_1}{\sqrt{s\xi}}[A_1 I_1(\sqrt{s\xi}\, r) - B_1 K_1(\sqrt{s\xi}\, r)] + \frac{E_2}{\sqrt{s\eta}}[A_2 I_1(\sqrt{s\eta}\, r) - B_2 K_1(\sqrt{s\eta}\, r)] + \frac{E_3}{4}[D_1 r(2\ln r - 1) + 2D_2 r] + \frac{D_3}{r} \quad (23)$$

where u_r is radial displacement, I_1 and K_1 are the modified first-kind and second-kind Bessel functions of order one, respectively.

Using Equation (23) and generalized Hooke's law, radial stress σ_r and tangential stress σ_φ can be obtained:

$$\overline{\sigma}_r = -2G\{\frac{E_1}{\sqrt{s\xi}\, r}[A_1 I_1(\sqrt{s\xi}\, r) - B_1 K_1(\sqrt{s\xi}\, r)] + \frac{E_2}{\sqrt{s\eta}\, r}[A_2 I_1(\sqrt{s\eta}\, r) - B_2 K_1(\sqrt{s\eta}\, r)] + \frac{E_3}{4}[D_1(2\ln r - 1) + 2D_2] + \frac{D_3}{r^2}\} - MD_1 \ln r - MD_2 \quad (24)$$

$$\overline{\sigma}_\varphi = -2G\{E_1 A_1[I_0(\sqrt{s\xi}\, r) - \frac{1}{\sqrt{s\xi}\, r} I_1(\sqrt{s\xi}\, r)] + E_1 B_1[K_0(\sqrt{s\xi}\, r) + \frac{1}{\sqrt{s\xi}\, r} K_1(\sqrt{s\xi}\, r)] + E_2 A_2[I_0(\sqrt{s\eta}\, r) - \frac{1}{\sqrt{s\eta}\, r} I_1(\sqrt{s\eta}\, r)] + E_2 B_2[K_0(\sqrt{s\eta}\, r) + \frac{1}{\sqrt{s\eta}} K_1(\sqrt{s\eta}\, r)] + \frac{E_3}{4}[D_1(2\ln r + 1) + 2D_2] - \frac{D_3}{r^2}\} - MD_1 \ln r - MD_2 \quad (25)$$

4.3. Inversion of the Laplace Transform

Equations (17) to (19), (23) to (25) constitute the solutions in the Laplace transform domain. In reality, solutions in the real domain can all be obtained by inverting the above solutions. There are many numerical Laplace inversion schemes reported in literatures [10,17]. The Stehfest method has been extensively used, due to its accuracy, efficiency and stability. This method is based on sampling inversion data according to a delta series. The present study uses the Stehfest method.

5. Determination of the Integration Functions

In a cylindrical coordinate system, the following equilibrium is also used to determine the unknown coefficients in addition to the boundary conditions:

$$\frac{\partial \sigma_r}{\partial r} + \frac{\sigma_r - \sigma_\varphi}{r} = 0 \quad (26)$$

Substitution of Equations (24) and (25) into Equation (26) results in D_1=0. The remaining six arbitrary coefficients (A_1, B_1, A_2, B_2, D_2 and D_3) can be determined by solving Equation (A9) (see Appendix A).

For the special case of $a \to 0$ (i.e. a solid cylinder with infinite length), there are three arbitrary coefficients (A_1, A_2 and D_2), which can be determined by solving Equation (A10) (see Appendix B). For the special case of $b \to \infty$ (i.e. a cylindrical cavity in a half-space body), there are also three arbitrary coefficients (B_1, B_2 and D_3), which can be determined by solving Equation (A11) (see Appendix C).

6. Numerical Examples

6.1. Material Properties and Loading

The material properties used in the analysis are given as follows: the elastic modulus E=6.0×10^5 Pa, the Poisson ratio μ=0.3, the bulk modulus of solid grains K_s=2×10^{10} Pa, the bulk modulus of pore water K_w=5×10^9 Pa, the thermal expansion coefficient of solid grains α_s= 1.5×10^{-5}/°C, the thermal expansion coefficient of pore water α_w=2.0×10^{-4}/°C, the porosity n=0.4, the heat capacity of solid grains c_s=800 J/(kg °C), the heat capacity of pore water c_w=4200 J/(kg °C), the density of pore water ρ_w=1.0×10^3 kg/m^3, the density of solid grains ρ_s =2.6×10^3 kg/m^3, the coefficient of heat conductivity K=0.5 W/(m °C) and Biot's coefficient α=1.0.

Geometrically, the hollow cylinder has an inner radius a=0.02m; and an outer radius b=0.08m (see Figure 1). As such, the thickness of the cylinder wall d=0.06m. For convenience, the following problem is discussed (i.e. Case 1): the thermal loading $\theta_b(t)$ and mechanical loading $p_b(t)$ on the outer surface remain both constant; while an exponentially increasing temperature variation $\theta_a(t)$ is applied on the inner surface. It can be written as

$$\theta_a(t) = \theta_0[1 - \exp(-\omega t)] \quad (31)$$

where θ_0=100°C, ω=–0.00384s^{-1}.

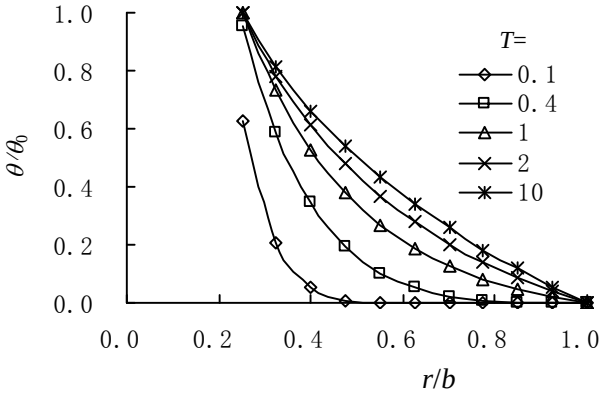


Figure 2. Distributions of temperature along radial distance.

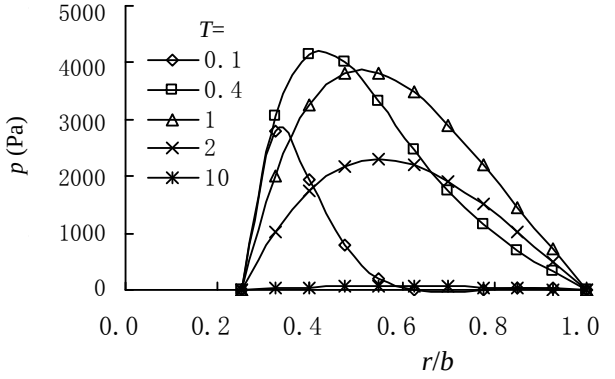


Figure 3. Pore pressure varying with *T* under thermal loading.

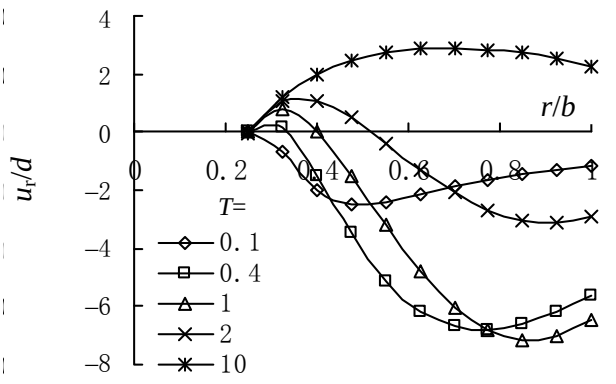


Figure 4. Radial displacement varying with *T* under thermal loading.

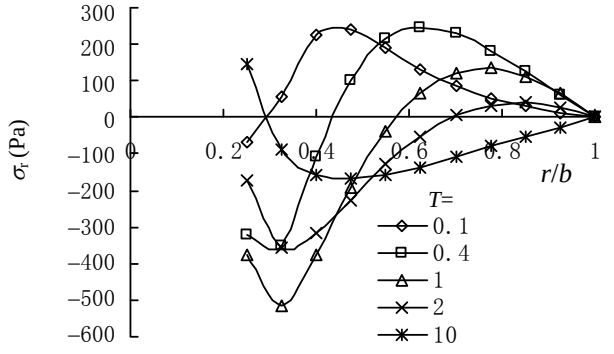


Figure 5. Radial stress varying with *T* under thermal loading.

6.2. Responses under Thermal Loading

In this section, we assume that the thermal loading $\theta_b(t)=0$, and the mechanical loading $p_b(t)=0$; while the thermal loading given by Equation (31) is applied on the inner surface. A dimensionless time is defined as $T=\kappa t/b^2$ (i.e. time factor). Figures 2 to 6 present respectively the temperature, pore pressure, radial displacement, radial stress and tangential stress distributions along radial distance for various time factors (e.g. T=0.1, 0.4, 1, 2, 10, i.e. t=3.9, 15.6, 39.0, 78.1, 390.4min; c/κ=1).

It can be seen from Figure 2 that, with the elapsed time, the temperature is gradually conducted from the inner surface of the hollow cylinder to the points away from the surface. As the time factor T increases continuously (e.g. T=10), the temperature values finally reach a quasi-steady state. At this time, the temperature distribution along the radial distance remains a steady temperature gradient. Certainly, the values of temperature gradient near the inner surface (e.g. r/b=0.25–0.4) are greater than those of the points in the vicinity of the outer surface (e.g. r/b=0.8–1.0).

It can be seen from Figure 3 that at early times (e.g. T=0.1, 0.4) the pore pressure value in the vicinity of the inner surface takes on a rising trend, and then is conducted from the inner surface of hollow cylinder to the outer surface; while the peak value gradually moves to the outer surface. At later times (e.g. T=2, 10), the pore pressure begins to decrease quickly due to the completely pervious lateral surfaces of the hollow cylinder, and finally is dissipated to zero along the radial distance.

Figure 4 shows that, at early times (e.g. T=0.1) the radial displacement takes on an expanding trend (i.e. negative displacement) along the whole radial distance; however with time factor T increasing (e.g. T=0.4, 1, 2), the radial displacements of the points in the vicinity of the inner surface begin to contract (i.e. positive displacement) due to the strong coupling between the expansion of grains and the drainage of pore water. With the successive dissipation of pore pressure (e.g. T=10), the radial displacement eventually takes an increasing trend with the radial distance increasing.

As shown in Figures 5 and 6, in the whole processes of consolidation, the radial and tangential stress distributions are very complicated due to the coupling effects of pore pressure dissipation and thermal stress. In fact, at early times (e.g. T=0.1, 0.4, 1) the radial stress in the vicinity of the inner surface begins to increase (being negative values, i.e. stretching stress), then (e.g. T=2) decreases with time factor increasing, and finally (e.g.

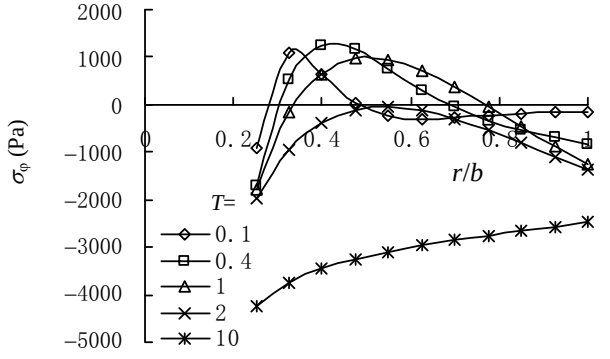


Figure 6. Tangential stress varying with *T* under thermal loading.

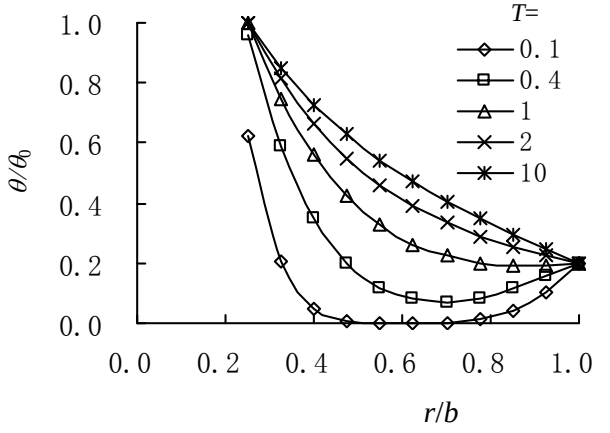


Figure 7. Distributions of temperature along radial distance (*c*/*k*=1).

T=10) takes on a positive value (i.e. compressive stress). On the other hand, the tangential stress in the middle part of the wall of the hollow cylinder initially takes a compressive state (i.e. positive value), and eventually presents a stretching state (i.e. negative value).

6.3. Responses under Thermal and Mechanical Loading

The responses of the porothermoelastic hollow cylinder under thermo-hydro-mechanical coupling are discussed in this section. Here, the thermal loading $\theta_b(t)$=20°C, and the mechanical loading $p_b(t)$=100kPa; while the thermal loading given by Equation (31) is applied on the inner surface. Figures 7 to 11 present respectively the temperature, pore pressure, radial displacement, radial stress and tangential stress distributions as a function of the radial distance for various time factors (e.g. *T*=0.1, 0.4, 1, 2, 10; c/κ=1).

It can be seen from Figure 7 that the developing trend of the temperature along the radial distance is similar to the temperature distributions in Figure 2 except for the values on the outer surface of the hollow cylinder, which

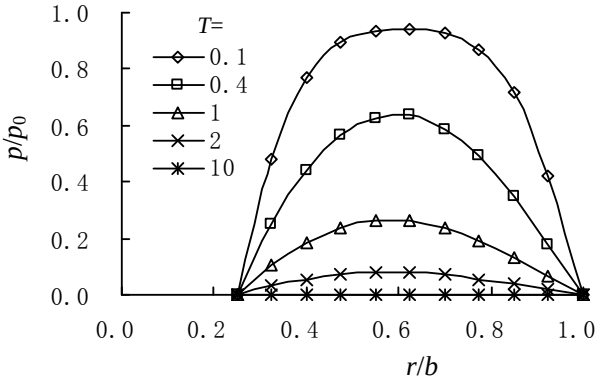


Figure 8. Distributions of pore pressure along radial distance (*c*/*k*=1).

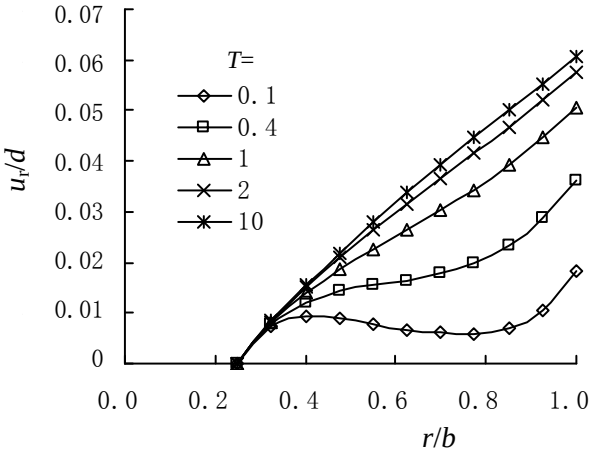


Figure 9. Distributions of radial displacement along radial distance (*c*/*k*=1).

is due to the difference of boundary thermal loading. Calculation results for various c/κ (e.g. c/κ=0.1, 1, 2) show that the coupling effects of displacement and stress fields on temperature field can be generally neglected.

By virtue of the imposed lateral boundary conditions, the pore pressure drops almost instantaneously at the inner and outer boundaries (i.e. *r*=*a* and *r*=*b*) as indicated in Figure 8. As such, there exists a peak value of pore pressure in the inner layers. As time progresses, the pore pressure peak value gradually diffuses and flattens. It should be noted that the pore pressure in the vicinity of the outer boundary seems to dissipate more quickly, which is due to the greater drainage surface of the outer boundary than that of the inner boundary.

Figure 9 shows that the radial displacement anywhere in the cylinder contracts (i.e. being positive value) with the diffusion of the pore water. However, it is noticed that, at early times (e.g. *T*=0.1, 0.4), the radial displacement in the local range of the wall (e.g. 0.4<*r*/*b*<0.9) is even smaller than that in the vicinity of the inner boundary (here, noting $u_r(a, t)$=0). This may be attributed to the

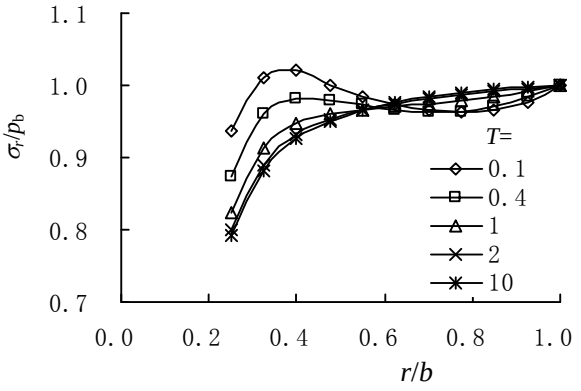


Figure 10. Distributions of radial stresses along radial distance (*c*/*k*=1).

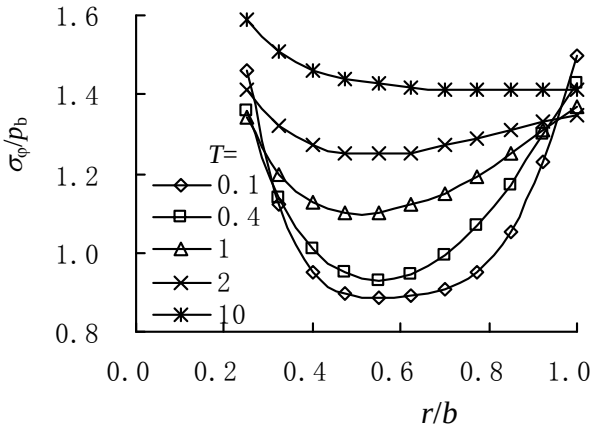


Figure 11. Distributions of tangential stresses along radial distance (*c*/*k*=1).

consolidation deformation (i.e. shrinkage of the cylinder) caused by the rapid dissipation in the vicinity of the pervious inner and outer surfaces. Obviously, this phenomenon vanishes as time progresses (see Figure 9). Furthermore, the radial displacement u_r will take on a linear relation with radial distance r at time factor T tends to infinity.

It can be seen from Figure 10 that the evolution of radial stress is very complicated. At early times (e.g. T=0.1, 0.4), the radial stress in the local range of the wall (e.g. $0.3<r/b<0.4$) is even greater than the applied mechanical loading p_b on the outer surface (i.e. $\sigma_r/p_b>1$). Obviously, as time factor T increases, the radial stress takes on a monotonically increasing trend with radial distance, and finally reach a steady state at time factor T tends to infinity.

Figure 11 presents a tangential stress concentration at each of the boundaries of the cylinder. These stresses are generated as a result of the hoop effects that accompany the inner and outer diffusion fronts. Moreover, this stress concentration is more severe at early times. With diffusion, the tangential stress concentrations at the lateral surfaces weaken and diminish in magnitude while higher compressive tangential stresses ($\sigma_\varphi/p_b>1$) are noted to form inside the cylinder.

7. Conclusions

1) An analytical method is derived for the thermal consolidation of a saturated, porous, hollow cylinder with infinite length. The solutions in Laplace transform space are first obtained and then numerically inverted by Stehfest method. As two special problems, a solid cylinder with infinite length and a cylindrical cavity in a half-space body are also discussed.

2) The responses of a porothermoelastic hollow cylinder subjected to exponentially increasing thermal loading with time on the inner surface are discussed. Calculations show that, as the temperature is gradually conducted from the inner surface to the points away from the surface, the pore pressure value initially takes on a rising trend, and then begins to decrease quickly due to the pervious lateral surfaces. In addition, the radial displacement initially increases with the increase of temperature and then contracts with the diffusion of pore water, which is due to the strong coupling effects of the expansion of grains, the drainage of pore water, and the radial and tangential stress varying with time under thermal loading.

3) The responses of a porothermoelastic hollow cylinder under thermo-hydro-mechanical coupling have also been analyzed. Numerical results indicate that the temperature difference generates both pore pressure and stress distributions in the cylinder. A pore pressure field, as well as a radial displacement, a radial and tangential stress fields, are created as a result of the compaction of the cylinder and the heating of the borehole wall. Correspondingly, this results in the complicated evolution processes of all the variables, and can be explained by the consolidation deformation due to the thermal stress and the rapid dissipation of pore pressure in the vicinity of the pervious inner and outer surfaces.

8. Acknowledgements

Financial support from the National Natural Science Foundation of China (NSFC) under approved grant No. 50879003 is gratefully acknowledged.

9. References

[1] M. Kurashige, "A thermoelastic theory of fluid-filled porous materials," International Journal of Solids and Structures, Vol. 25, No. 9, pp. 1039–1052, 1989.

[2] X. Li, "A generalized theory of thermoelasticity for an anisotropic medium," International Journal of Engineering Science, Vol. 30, No. 5, pp. 571–577, 1992.

[3] H. N. Seneviratne, J. P. Carter, D. W. Airey, and J. R. Booker, "A review of models for predicting the ther-

momechanical behavior of soft clays," International Journal for Numerical and Analytical Methods in Geomechanics, Vol. 17, No. 10, pp. 715–733,1993.

[4] M. Bai and Y. Abousleiman, "Thermoporoelastic coupling with application to consolidation," International Journal for Numerical and Analytical Methods in Geomechanics, Vol. 21, No. 2, pp. 121–132, 1997.

[5] R. W. Zimmerman, "Coupling in poroelasticity and thermoelasticity," International Journal of Rock Mechanics and Mining Sciences, Vol. 37, No. 1, pp. 79–87, 2000.

[6] J. R. Booker and C. Savvidou, "Consolidation around a spherical heat source," International Journal of Solids and Structures, Vol. 20, No. 5, pp. 1079–1090, 1984.

[7] D. F. McTigue, "Flow to a heated borehole in porous, thermoelastic rock: Analysis," Water Resources Research, Vol. 26, No. 8, pp. 1763–1774, 1990.

[8] D. W. Smith and J. R. Booker, "Green's function for a fully coupled thermoporoelastic materials," International Journal for Numerical and Analytical Methods in Geomechanics, Vol. 17, No. 2, pp. 139–163, 1993.

[9] A. Giraud, F. Homand, and G. Rousset, "Thermoelastic and thermoplastic response of a double-layer porous space containing a decaying heat source," International Journal for Numerical and Analytical Methods in Geomechanics, Vol. 22, No. 2, pp. 133–149, 1998.

[10] Y. Wang and E. Papamichos, "An analytical solution for conductive heat flow and the thermally induced fluid flow around a wellbore in a poroelastic medium," Water Resource Research, Vol. 36, No. 5, pp. 3375–3384, 1994.

[11] Y. Wang and E. Papamichos, "Thermal effects on fluid flow and hydraulic fracturing from wellbores and cavities in low-permeability formations," International Journal for Numerical and Analytical Methods in Geomechanics, Vol. 23, No. 15, pp. 1819–1834, 1999.

[12] E. Blond, N. Schmitt, and F. Hild, "Response of saturated porous media to cyclic thermal loading," International Journal for Numerical and Analytical Methods in Geomechanics, Vol. 27, No. 11, pp. 883–904, 2003.

[13] B. Bai, "Response of saturated porous media subjected to local thermal loading on the surface of semi-infinite space," Acta Mechanica Sinica, Vol. 22, No. 1, pp. 54–61, 2006.

[14] B. Bai, "Fluctuation responses of saturated porous media subjected to cyclic thermal loading," Computers and Geotechnics, Vol. 33, No. 4, pp. 396–403, 2006.

[15] Y. Abousleiman and S. Ekbote, "Solutions for the inclined borehole in a porothermoelastic transversely isotropic medium," Journal of applied Mechanics, Vol. 72, No. 2, pp. 102–114, 2005.

[16] M. Kanj, Y. Abousleiman, and R. Ghanem, "Poromechanics of anisotropic hollow cylinders," Journal of Engineering Mechanics, Vol. 129, No. 11, pp. 1277–1287, 2003.

[17] M. Kanj and Y. Abousleiman, "Porothermoelastic analyses of anisotropic hollow cylinders with applications," International Journal for Numerical and Analytical Methods in Geomechanics, Vol. 29, No. 1, pp. 103–126, 2005.

[18] B. Bai, "Thermal consolidation of layered porous halfspace to variable thermal loading," Applied Mathematics and Mechanics, Vol. 27, No. 11, pp. 1531–1539, 2006.

Appendix A

For the first case of boundary conditions (i.e. Case 1), from Equations (4) and (5), when $t>0$, $\bar{\theta}(a,t)=\bar{\theta}_a$, $\bar{p}(a,t)=0$, $\bar{u}_r(a,t)=0$, $\bar{\theta}(b,t)=\bar{\theta}_b$, $\bar{p}(b,t)=0$ and $\bar{\sigma}_r(b,t)=\bar{p}_b$, which yields

$$A_1 r I_0(\sqrt{s\xi}\,a)+B_1 r K_0(\sqrt{s\xi}\,a)+A_2 q I_0(\sqrt{s\eta}\,a)+B_2 q K_0(\sqrt{s\eta}\,a)+D_2 v=\bar{\theta}_a \quad (A1)$$

$$A_1 I_0(\sqrt{s\xi}\,a)+B_1 K_0(\sqrt{s\xi}\,a)+A_2 I_0(\sqrt{s\eta}\,a)+B_2 K_0(\sqrt{s\eta}\,a)+D_2 u=0 \quad (A2)$$

$$A_1\frac{E_1}{\sqrt{s\xi}}I_1(\sqrt{s\xi}\,a)-B_1\frac{E_1}{\sqrt{s\xi}}K_1(\sqrt{s\xi}\,a)+A_2\frac{E_2}{\sqrt{s\eta}}I_1(\sqrt{s\eta}\,a)-B_2\frac{E_2}{\sqrt{s\eta}}K_1(\sqrt{s\eta}\,a)+D_2\frac{E_3 a}{2}+\frac{D_3}{a}=0 \quad (A3)$$

$$A_1 r I_0(\sqrt{s\xi}\,b)+B_1 r K_0(\sqrt{s\xi}\,b)+A_2 q I_0(\sqrt{s\eta}\,b)+B_2 q K_0(\sqrt{s\eta}\,b)+D_2 v=\bar{\theta}_b \quad (A4)$$

$$A_1 I_0(\sqrt{s\xi}\,b)+B_1 K_0(\sqrt{s\xi}\,b)+A_2 I_0(\sqrt{s\eta}\,b)+B_2 K_0(\sqrt{s\eta}\,b)+D_2 u=0 \quad (A5)$$

$$A_1\frac{-2GE_1}{\sqrt{s\xi}\,b}I_1(\sqrt{s\xi}\,b)+B_1\frac{2GE_1}{\sqrt{s\xi}\,b}K_1(\sqrt{s\xi}\,b)+A_2\frac{-2GE_2}{\sqrt{s\eta}\,b}I_1(\sqrt{s\eta}\,b)+B_2\frac{2GE_2}{\sqrt{s\eta}\,b}K_1(\sqrt{s\eta}\,b)+D_2 w-D_3\frac{2G}{b^2}=\bar{p}_b \quad (A6)$$

where $r=f_1\xi+f_2$, $q=f_1\eta+f_2$, $u=-g_4/g_3$, $v=f_3-f_2 g_4/g_3$, $w=-M-GE_3$.

For the second case of boundary conditions (i.e. Case 2), from $\bar{u}_r(a,t)=0$ and Equation (5), the Equations (A3) to (A6) can be obtained easily; while using Equation (6), when $t>0$, $\partial\bar{\theta}(a,t)/\partial r=0$ and $\partial\bar{p}(a,t)/\partial r=0$, which yields

$$A_1 r\sqrt{s\xi}\,I_1(\sqrt{s\xi}\,a)-B_1 r\sqrt{s\xi}\,K_1(\sqrt{s\xi}\,a)+A_2 q\sqrt{s\eta}\,I_1(\sqrt{s\eta}\,a)-B_2 q\sqrt{s\eta}\,K_1(\sqrt{s\eta}\,a)=0 \quad (A7)$$

$$A_1\sqrt{s\xi}\,I_1(\sqrt{s\xi}\,a)-B_1\sqrt{s\xi}\,K_1(\sqrt{s\xi}\,a)+A_2\sqrt{s\eta}\,I_1(\sqrt{s\eta}\,a)-B_2\sqrt{s\eta}\,K_1(\sqrt{s\eta}\,a)=0 \quad (A8)$$

Hence, the following equation is obtained:

$$\begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \\ d_{41} & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} \\ d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} \\ d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} \end{bmatrix}\begin{bmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \\ D_2 \\ D_3 \end{bmatrix}=\begin{bmatrix} \bar{\theta}_a \text{ or } 0 \\ 0 \\ 0 \\ \bar{\theta}_b \\ 0 \\ \bar{p}_b \end{bmatrix} \quad (A9)$$

where the coefficients d_{11}, d_{12}, …, d_{65} and d_{66} can be given correspondingly by Equations (A1) to (A6) (for Case 1) or Equations (A7), (A8) and (A3) to (A6) (for Case 2).

Appendix B

Due to $K_1(\sqrt{s\xi}\,a)=\infty$, $K_1(\sqrt{s\eta}\,a)=\infty$ and $1/a=\infty$ at $a\to 0$ in Equations (A7), (A8) and (A3), one has $B_1=0$, $B_2=0$ and $D_3=0$. This implies that all variables must be finite. Hence, Equations (A7), (A8) and (A3) are satisfied automatically (noting $I_1(0)=0$). The remaining three arbitrary coefficients (A_1, A_2 and D_2) can be determined by solving Eqution (A4) to (A6) simultaneously. At this time, Equation (A9) reduces to the following expression:

$$\begin{bmatrix} d_{41} & d_{43} & d_{45} \\ d_{51} & d_{53} & d_{55} \\ d_{61} & d_{63} & d_{65} \end{bmatrix}\begin{bmatrix} A_1 \\ A_2 \\ D_2 \end{bmatrix}=\begin{bmatrix} \bar{\theta}_b \\ 0 \\ \bar{p}_b \end{bmatrix} \quad (A10)$$

where the coefficients d_{41}, d_{43}, …, d_{63} and d_{65} can be given correspondingly by Equations (A4) to (A6).

Appendix C

Due to $I_0(\sqrt{s\xi}\,b)=\infty$ and $I_0(\sqrt{s\eta}\,b)=\infty$ at $b\to\infty$ in Equations (A4) to (A6), one has $A_1=0$ and $A_2=0$. Noting $K_0(\infty)=0$ and $K_1(\infty)=0$, for the satisfaction of Equations (A4) to (A6), one must let $D_2=0$. The remaining three arbitrary coefficients (B_1, B_2 and D_3) can be determined by solving Equations (A1) to (A3) simultaneously. At this time, Equation (A9) reduces to the following expression:

$$\begin{bmatrix} d_{12} & d_{14} & d_{16} \\ d_{22} & d_{24} & d_{26} \\ d_{32} & d_{34} & d_{36} \end{bmatrix}\begin{bmatrix} B_1 \\ B_2 \\ D_3 \end{bmatrix}=\begin{bmatrix} \bar{\theta}_a \\ 0 \\ 0 \end{bmatrix} \quad (A11)$$

where the coefficients d_{12}, d_{14}, …, d_{34} and d_{36} can be given correspondingly by Equations (A1) to (A3).

Engineering, **2010, 2, 46-54**
doi:10.4236/eng.2010.21006 Published Online January 2010 (http://www.scirp.org/journal/eng/).

Vibration Monitoring of Rotating Systems

K. N. EDE[1], E. A. OGBONNAYA[2], M. T. LILLY[3], S. O. T. OGAJI[4], S. D. PROBERT[4]
[1]*Method Engineer, Netco-Dietsmann Nigeria Limited, Port Harcourt, Nigeria*
[2]*Department of Marine Engineering, Rivers State University of Science and Technology, Port Harcourt, Nigeria*
[3]*Department of Mechanical Engineering, Rivers State University of Science and Technology, Port Harcourt, Nigeria*
[4]*School of Engineering, Cranfield University, Bedfordshire, United Kingdom*
E-mail: *kingston*2004*@yahoo.com, ezenwaogbonnaya@yahoo.com, s.ogaji@cranfield.ac.uk*
Received June 6, 2009; *revised August* 3, 2009; *accepted August* 10, 2009

Abstract

Most energy-conversion machines (e.g. vehicle engines and electric motors) involve rotating components (e.g. roller bearings and gears), which generate vibrations. The behavior of a pump which includes a deliberate fault was chosen to illustrate this assertion. The test bearing at the driven end of the pump's motor was deliberately damaged using a 1.5mm wire-cutting method and an adjustable coupling disk introduced to impose a shaft misalignment of 4^0. The resulting undesirable behavior of the pump was observed. Experimental data were measured at various speeds of the rotor. The sample period at various operating frequencies were 0.9, 0.6 and 0.45s respectively. The ball-passage frequency was observed at 4.4, 8.8, 13.2 and 17.6Hz. A computer-based analytical model was developed, in visual basic, for monitoring the machine failures: this led to an integrated system-process algorithm for diagnosis of faults in rotating components.

Keywords: Vibration Measurement, Rotating Component, Ball-Passage Frequency, Alarm Limit

1. Introduction

The use of vibration analysis as a fundamental tool for condition monitoring of equipment has evolved over the last 35 years. With parallel developments in electronic equipment, transducers and computers, the monitoring of machines is now almost completely automated. Onboard microprocessors provide the ability to capture pertinent measurements and analyze them via suitable algorithms, then store and display the conclusions. Several accelerometers, velocity transducers and displacement pick-ups have been developed and adapted to suit the pertinent requirement of industrial applications [1]. Modification of a rolling-element bearing activity monitor (REBAM) system offers a high signal-to-noise ratio relative to those for a casing-mounted accelerometer or velocity transducer, as shown in Figure 1 [2]. A system's signal-to-noise ratio is defined as the ratio of the amplitude of the desired signal to the amplitude of the noise signal. Increased signal-to-noise ratios for accelerometers have been achieved through the use of electronic filters [3].

REBAM vibration signals can be separated into rotor-vibration and prime-spike regions. This signal separation improves the signal-to-noise ratio for both regions. By measuring directly the vibrations at the rolling-element bearing outer-ring and displacements are relative to the machine casing.

This isolates the signal of interest from extraneous vibrations (e.g., due to structural resonances, steam throttling, pump cavitation, gear noise, etc.) which often mask the bearing-defect signals when casing-mounted transducers are employed. The results obtained indicate that this signal separation technique makes REBAM twice to eight times more sensitive to bearing faults than casing-mounted transducers [4]. A significant feature of this frequency analysis is the effective computation of the fast Fourier-transformation (FFT), which converts digital information from the time domain to the frequency domain and thereby achieves a more rapid spectral-analysis [5].

Long-term data storage is now a well accepted practice. However difficulties occurred when random-vibration

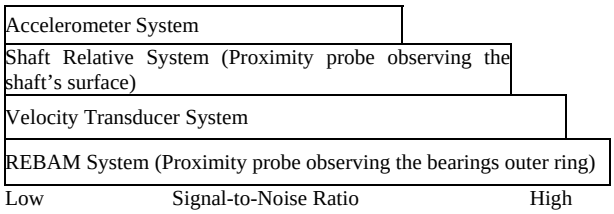


Figure 1. Qualitative indication of relative magnitudes of signal-to-noise ratios for various transducers.

envelope spectrum algorithms for identifying defects in rolling-element bearings were applied to gearing. Extra lines appeared in the envelope spectrum due to the dynamic loads applied to the bearings by the gears. Additionally, spectral lines associated with the rolling-element detector based on the power spectrum [7]. The ensuant effective predictive maintenance can result in an 8% maintenance-cost saving and a further 8% increase in productivity [8], and hence the associated energy-thrift.

1.1. Glossary

An accelerometer, in the present context, is a transducer whose electrical or mechanical output is directly proportional to the acceleration experienced.

An alarm limit is the maximum permitted predicted value of the considered parameter, i.e. it indicates when attained that a dangerous situation is developing and so the alarm is triggered.

The ball-passage frequency is the frequency of the ball in a bearing component that will generate specific frequency dependent upon the bearing's geometry and its rotational speed.

The bearing-passage frequencies are the frequencies of rotation of the ball, outer race, inner race and cage.

A coupling-disc or flexible connecting system is a robust general purpose pin/buffer coupling, that facilitates a reliable fail-safe transmission of the torque between two nominally-coaxial shafts; it possesses the capability of being able to function despite a misalignment of the shafts.

A fast Fourier-transform is a numerical operation commonly used for transposing data rapidly from the time domain to the frequency domain, and usually accomplished via computer.

The gap, in the present context, is the air-filled separation between the rotor and stator.

Harmonics are components of a spectrum which are integral multiples of the fundamental frequency.

In the context of this article, imbalance (or a lack of balance) results from the distribution of mass according to the radial direction of the rotating system.

Outer-race-ball passage frequency is the rate at which a point on the specified rolling element passes a point on the outer bearing-race.

Inner-race-ball passage frequency is the rate at which a point on the specified rolling element passes a point on the inner bearing-race.

Prime spike is a term used by Bently Nevada to describe a vibration frequency range which includes those bearing frequencies that are generated by the rolling elements experiencing either an inner or outer race flaw.

A probe, in the present context, is a sensor capable of detecting the vibration signal.

The raceway is the ball passage frequency relative to the inner and outer ring of the bearing.

Rack-configuration software is a tool for configuring bearings occasionally disappeared from the spectrum. Investigations were conducted in order to overcome these complications [6]. This led to the introduction of the demerit of a bi-spectrum-based quickest-change detector, which was an extension of the analogous quickest integrated control-switching. It is designed to have a 3500/40 proximity monitor and a four-channel monitor that accepts inputs that can trigger alarms.

Rolling-element frequency is the frequency at which either the balls or rollers revolve about their own centerlines in a bearing.

Ringing of the bearing occurs when ever the ball hits the flaw on the outer race of the bearing.

A signal, in the present context, is an electric voltage or current, which is an analogue of the vibration being measured.

In the context of this article, Spike energy is a measure of the energy generated by the repetitive impacts of the rolling elements against the defect in the bearing.

A time base is a horizontal line (representing time) about which the waveform representing the vibration signal occurs, i.e. the representation of a vibration signal in the time domain is a wave form.

A transducer is a device that converts the magnitude of one physical parameter into the value of another parameter, usually an electrical signal. The transducer used in vibration measurements is usually an accelerometer.

2. The Aim

This study investigates the occurrences of common machine-faults (i.e. rolling-element bearing and gear failures, misalignment of the shaft and imbalance of the shaft) that lead to energy wastages. The bearings were tested when undamaged and the resulting data were used as a bench-mark. An accelerometer was employed to obtain the vibration data from on-line measurements. The resulting information made available by the sensor and amplifier, was translated into useful knowledge about the forces on the machine and vibration patterns: the latter was classified according to the defects detected. Alarm limits were setup and programmed for the data collected. Fault-diagnosis software (FDS) for vibration monitoring of pumps and turbines has as a result been developed in this investigation: this is available from the authors.

2.1. Instrumentation

The characteristics of the employed standard vibration-measuring instrument for the applied range of 0.5 to14000Hz, the power supply was a 4mA constant current at 23VDC, and 3 contact probes were available. Bently Nevada eddy-probes were installed on the test rig. A 3500/40 proximity four-channel monitor that accepts inputs to drive the alarms and programs is shown schematically in Figure 2.

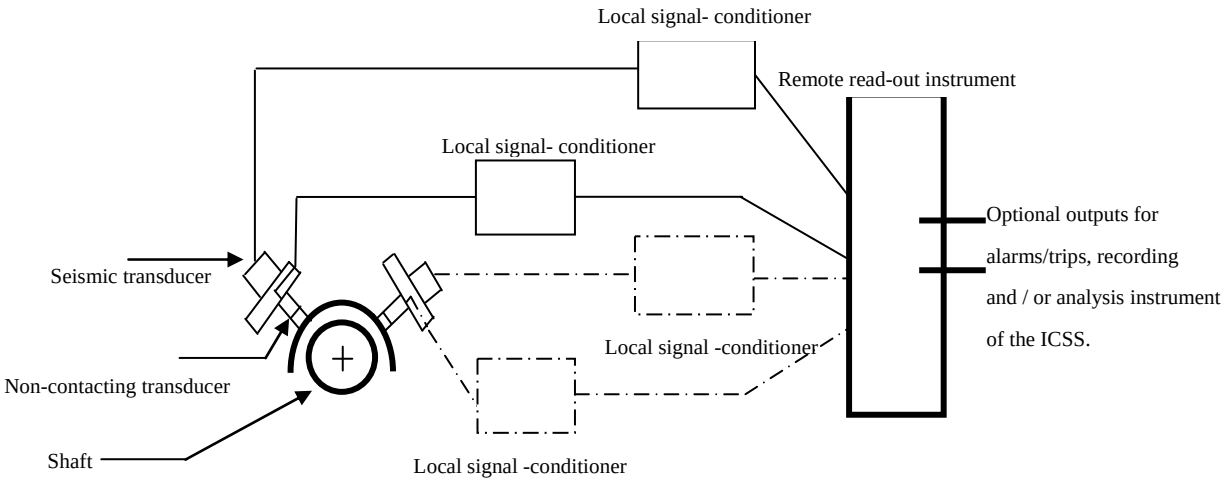


Figure 2. The probe arrangement.

2.2. Time Period Sampled

The relationship between frequency and time is given as:

$$F=I/P \quad (1)$$

For this analysis, the instrument was set up to observe the shaft rotating at 10 or 15 revolutions per second.

$$\textit{Total time (s)}=60(\textit{Number of Revolutions}) / \textit{RPM} \quad (2)$$

With this instrument used, it became necessary to set an equivalent F_{max} setting [9]. The appropriate F_{max} setting can be calculated as;

$$F_{max}\ (CPS) = \textit{Lines of revolution x RPM/60(No. of revolutions desired)} \quad (3)$$

The SKF pattern No. 6226/C3 bearing used has a single row of balls. Because the defect was inflicted deliberately on the outer race of the ball bearing, at various speeds of the rotor, spikes were generated. These spikes corresponded to the outer race-ball passage frequencies [10]. Readings collected while using this damaged bearing is shown in table.

$$ORBP=\frac{RPM}{2*60}\left(1-\frac{B_d}{P_d}\cos\theta\right)z \quad (4)$$

The cage rotating frequency, (i_c) [10].

$$i_c=\frac{RPM}{2*60}\left(1-\frac{B_d}{P_d}\cos\theta\right) \quad (5)$$

Table 1. Standard bearing 6226(SKF) used in this investigation.

Cage's Rotating-Frequency (i_c)	Inner Race-Ball Passage Frequency *(IRBP)*	Outer Race-Ball Passage Frequency *(ORBP)*	Rolling Element (i_R)
0.49	5.29	3.71	5.49

A coupling disc was then designed so that it could impose a shaft misalignment onto the undamaged bearing. The disc was capable of moving relatively by adjusting the pin/buffer coupling. This forces the disc of 70mm diameter on the shaft to move and produce an angular misalignment of 4^0.

2.3. Bench-Mark Alarm

This type of alarm was used to trigger an alarm when the measured value exceeds its bench-mark value times a coefficient, which in this case WAS assumed to be 1.5. That is,

$$\textit{Threshold}=1.5\ (\textit{Bench mark value}) \quad (6)$$

Vibration data were automatically compared with bench-mark values. User-defined alarm limits for the data collected are illustrated in Table 2.

3. Results and Discussions

The accelerometer was set up to record over a long timespan in order to observe the cycle of spike occurrences for probe data when the shaft was rotating at 10 or 15 revolutions per second. The periods necessary to accomplish 10 and 15 revolutions at 17 RPS were 0.6 and 0.9s respectively, at 25 RPS are 0.4 and 0.6s respectively, and

Table 2. User-defined upper and lower threshold alarm-limits.

Type of alarm	Definition of Alarm
"High"	
DG+ 0	DANGER if measured value>DG+
AL+ 0	ALARM if AL+<value<DG+
AL- 0	Otherwise acceptable.
DG- 0	
"Low"	
DG+ 0	DANGER if measured value<DG-
AL+ 0	ALARM if DG-<value<AL-
AL- 0	Otherwise acceptable.
DG- 0	

at 33 RPS are 0.3 and 0.45 seconds respectively.

When a ball was located near the point where the probe is mounted (left side, 45^0 from the top) see Figure 3, the outer race, due to the ball pressing on the defect, interacts and the resulting impacts excited a natural resonant frequency of the machine, caused the bearing to ring, and deform away from the bearing centre and towards the probe. This corresponds to the high peak-points in the time-domain plot shown in Figure 4.

3.1. Damaged Bearing

This was located at the driven-end bearing of the motor assembly and imposed by a flexible coupling to the AC motor. Three rotational speeds were used in the experiment to show the bearing-passage frequencies (see Table 3). The spectrum in Figure 5 indicates a mass imbalance of the shaft (*F0*) at *2x*. The *Outer Race-Ball Passage Frequency* of the bearing at 17 RPS, i.e., the *2xORBP=8.2*.

Harmonic frequencies were generated due to 4^0 misalignment of the shaft which disturbs the natural full motion of the shaft and this is often characterized by rubbing. The friction between the rubbing parts produced a broadband of high frequencies, some of the kinetic energy is released in the form of harmonic vibrations. The *2x* harmonics is indication of failure of the bearing, i.e. the most dominant amplitude. Lower frequency vibration was also transmitted through the coupling and even was amplified on the other end of the machine pump; this is indication of low amplitude on the frequency domain. However, the vibrations observed were often symptoms accentuating other vibration problems which gradually propagated and generated increased noise level and temperature of the equipment.

4. Fault-Diagnostic Software (FDS) Program

This program was written in visual basic. It is an on-line-based data simulation and monitoring process for a pump or turbine defects. The FDS was composed using rack-configuration software. For this reason, the flow chart shown in Figure 6 was developed to describe the adopted monitoring procedure. The process accepted basic process information, such as acquired data or signals from a PLC on the ICSS on-line [11]. The electronic control of the test equipment and the conventional connection of proximity probes for on-line data-acquisition via the PLC unit that was associated with the monitor (4-channel orbit analysis), therefore, was able to capture radial vibration, thrust position, and eccentricity. The module received input from many types of displacement transducers.

4.1. Vibration Measurements

To monitor this equipment on-line, the measured data were compared with their preset threshold values (in order to indicate alarm and danger).

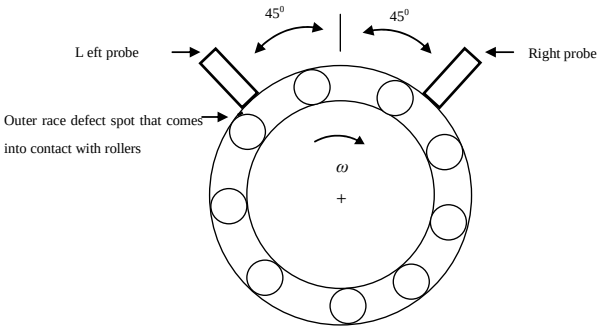


Figure 3. Probe orientations and outer race defect spot causing "ringing" of the bearing.

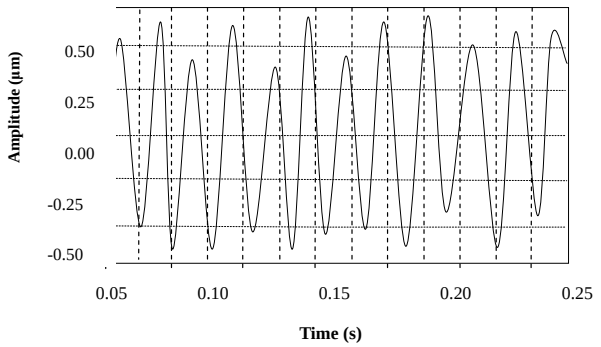


Figure 4. Outer-race deflection of the shaft which is rotating at 17 RPS.

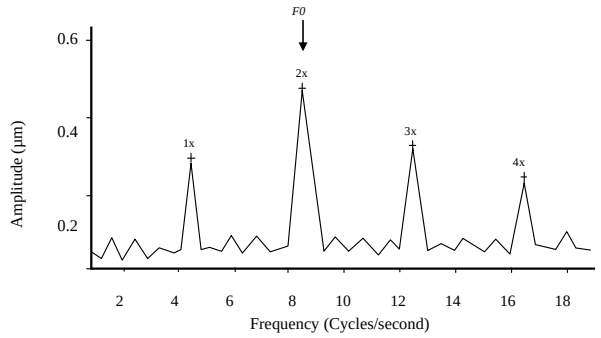


Figure 5. Frequency spectrum of the bearing operating at 17 RPS.

4.2. Simulation of All Alarms

Three categories of alarm severity were observed during the series of tests.

1) Alert 1 (Yellow alarm) indicated that small in behavioral changes are occurring: it provides an early warning of a deteriorating situation, i.e. an on-line displacement defect, as seen in Tables 4(b) to 4(f) and Figure 7(a).

2) Alert 2 (Amber alarm) was the result of the next measurement and indicated that maintenance planning had to be scheduled.

3) Danger (Red alarm). This serious situation requires immediate attention, as the measurements show consistent evolution of the defect(s), which was an indication of the severity observed as shown in Tables 4(a) to (e) for on-line imbalance and misalignment of the shaft, as shown in Tables 4(e) and 4(f) and Figures 7(c) and 7(d).

Table 3. Bearing-passage frequencies of the bearing.

Shaft speed *(RPS)*	Outer Race-Ball Passage Frequency *(cycles/second)*	Theoretically bearing-passage frequency *(cycles/second)* *(ORBP=zi_c)*	Experimentally observed bearing-passage frequency (cycles/second)
17	1 x *ORBP*	4.4	4.1
	2 x *ORBP*	8.8	8.2
	3 x *ORBP*	13.2	12.3
	4 x *ORBP*	17.6	16.4
25	1 x ORBP	8.8	8.2
	2 x ORBP	17.6	16.4
	3 x ORBP	26.4	24.6
	4 x ORBP	35.2	32.8
33	1 x ORBP	13.2	12.3
	2 x ORBP	26.4	24.6
	3 x ORBP	39.6	36.9
	4 x ORBP	52.8	49.2

Table 4(a). On-line bench-mark measurements for undamaged bearing: December 14th, 2005.

	2 (DER) Undamaged bearing	3 (DER) Damaged bearing	4 (DER) Undamaged bearing	5 (NDEA) Undamaged bearing	Alarm limit
Displacement of shaft (μm)	**15.9 (Amber)**	**26.5 (Amber)**	**21.9 (Amber)**	**22.9 (Amber)**	DG+60, AL+40 AL-0, DG-0
Gap (μm)	**1208 (Amber)**	**1254 (Amber)**	**1255 (Amber)**	**1230 (Amber)**	DG+0, AL+0 AL-1280, DG-1180
Imbalance of shaft (μm)	**10.09 (Amber)**	**12.04 (Amber)**	**11.12 (Amber)**	**15.10 (Yellow)**	DG+20, AL+15 AL-0, DG-0
Misalignment of the shaft (μm)	**0.0076 (Amber)**	**0.0011 (Amber)**	**0.0070 (Amber)**	**0.0056 (Amber)**	DG+.1, AL+.05 AL-0, DG-0

Table 4(b). On-line results of undamaged and damaged bearing: April 2nd, 2006.

	2 (DER) Damaged bearing	3 (DER) Damaged bearing	4 (DER) Undamaged bearing	5 (NDEA) Undamaged bearing	Alarm limit	Diagnosis
Displacement of shaft (μm)	**50.9 (Yellow)**	**42.5 (Yellow)**	**41.9 (Yellow)**	**58.2 (Yellow)**	DG+60, AL+40 AL-0, DG-0	Deterioration behavior of displacement of the shaft defects on the motor bearing and pump bearing.
Gap (μm)	**1211 (Amber)**	**1254 (Amber)**	**1255 (Amber)**	**1270 (Amber)**	DG+0, AL+0 AL-1280, DG-1180	There is no gap defect
Imbalance of shaft (μm)	**16.13 (Yellow)**	**14,04 (Amber)**	**16.16 (Yellow)**	**17.91 (Yellow)**	DG+20, AL+15 AL-0, DG-0	Imbalance of the shaft defects on the motor bearing and pump bearing.
Misalignment of the shaft (μm)	**0.0885 (Yellow)**	**0.0918 (Yellow)**	**0.0880 (Yellow)**	**0.0911 (Yellow)**	DG+.1, AL+.05 AL-0, DG-0	Misalignment of the shaft due to misalignment of 4^0 deliberately introduced

Table 4(c). On-line results of undamaged and damaged bearing: April 3rd, 2006.

	2 (DER) Damaged bearing	3 (DER) Damaged bearing	4 (DER) Undamaged bearing	5 (NDEA) Undamaged bearing	Alarm limit	Diagnosis
Displacement of shaft (μm)	**51.9 (Yellow)**	**45.5 (Yellow)**	**44.9 (Yellow)**	**60.9 (Red)**	DG+60, AL+40 AL-0, DG-0	Displacement of the shaft defects on the motor bearing and pump bearing.
Gap (μm)	**1235 (Amber)**	**1255 (Amber)**	**1259 (Amber)**	**1271 (Amber)**	DG+0, AL+0 AL-1280, DG-1180	There is no gap defect
Imbalance of shaft (μm)	**21.84 (Red)**	**18.01 (Red)**	**28.19 (Red)**	**25.11 (Red)**	DG+20, AL+15 AL-0, DG-0	Imbalance of the shaft defects on the motor bearing and pump bearing.
Misalignment of the shaft (μm)	**0.0921 (Yellow)**	**0.0941 (Yellow)**	**0.0841 (Yellow)**	**0.0971 (Yellow)**	DG+.1, AL+.05 AL-0, DG-0	Misalignment of the shaft due to misalignment of 4^0 deliberately introduced

Table 4(d). On-line results of undamaged and damaged bearing: April 4th, 2006.

	2 (DER) Damaged bearing	3 (DER) Damaged bearing	4 (DER) Undamaged bearing	5 (NDEA) Undamaged bearing	Alarm limit	Diagnosis
Displacement of shaft (μm)	**51.9 (Yellow)**	**45.7 (Yellow)**	**50.1 (Yellow)**	**61.9 (Red)**	DG+60, AL+40 AL-0, DG-0	Displacement of the shaft defects on the motor bearing and pump bearing.
Gap (μm)	**1273 (Amber)**	**1254 (Amber)**	**1265 (Amber)**	**1273 (Amber)**	DG+0, AL+0 AL-1280, DG-1180	There is no gap defect
Imbalance of shaft (μm)	**28.09 (Red)**	**25.19 (Red)**	**34.19 (Red)**	**34.17 (Red)**	DG+20, AL+15 AL-0, DG-0	Evolution of Imbalance of the shaft defects on the motor bearing and pump bearing.
Misalignment of the shaft (μm)	**0.1032 (Red)**	**0.1117 (Red)**	**0.1011 (Red)**	**0.1322 (Red)**	DG+.1, AL+.05 AL-0, DG-0	Misalignment of the shaft due to misalignment of 4^0 deliberately introduced

Table 4(e). On-line results of undamaged and damaged bearing: April 5th, 2006.

	2 (DER) Damaged bearing	3 (DER) Damaged bearing	4 (DER) Undamaged bearing	5 (NDEA) Undamaged bearing	Alarm limit	Diagnosis
Displacement of shaft (μm)	**52.5 (Yellow)**	**54.5 (Yellow)**	**54.0 (Yellow)**	**66.5 (Red)**	DG+60, AL+40 AL-0, DG-0	Displacement of the shaft defects on the motor bearing and pump bearing.
Gap (μm)	**1279 (Amber)**	**1264 (Amber)**	**1275 (Amber)**	**1277 (Amber)**	DG+0, AL+0 AL-1280, DG-1180	There is no gap defect
Imbalance of shaft (μm)	**28.94 (Red)**	**29.81 (Red)**	**35.34 (Red)**	**34.31 (Red)**	DG+20, AL+15 AL-0, DG-0	Evolution of Imbalance of the shaft defects on the motor bearing and pump bearing.
Misalignment of the shaft (μm)	**0.1051 (Red)**	**0.1314 (Red)**	**0.1108 (Red)**	**0.1535 (Red)**	DG+.1, AL+.05 AL-0, DG-0	Misalignment of the shaft due to misalignment of 4^0 deliberately introduced

Table 4(f). On-line results of undamaged and damaged bearing: April 6th, 2006.

	2 (DER) Damaged bearing	3 (DER) Damaged bearing	4 (DER) Undamaged bearing	5 (NDEA) Undamaged bearing	Alarm limit	Diagnosis
Displacement of shaft (μm)	**52.3 (Yellow)**	**56.5 (Yellow)**	**54.0 (Yellow)**	**69.2 (Red)**	DG+60, AL+40 AL-0, DG-0	Displacement of the shaft defects on the motor bearing and pump bearing.
Gap (μm)	**1279 (Amber)**	**1266 (Amber)**	**1275 (Amber)**	**1279 (Amber)**	DG+0, AL+0 AL-1280, DG-1180	There is no gap defect
Imbalance of shaft (μm)	**29.02 (Red)**	**32.11 (Red)**	**35.51 (Red)**	**34.52 (Red)**	DG+20, AL+15 AL-0, DG-0	Imbalances of the shaft defect on the motor bearing and pump bearing. Imbalance increases with temperature
Misalignment of the shaft (μm)	**0.1181 (Red)**	**0.1669 (Red)**	**0.1301 (Red)**	**0.1632 (Red)**	DG+.1, AL+.05 AL-0, DG-0	Misalignment of the shaft due to misalignment of 4^0 deliberately introduced.

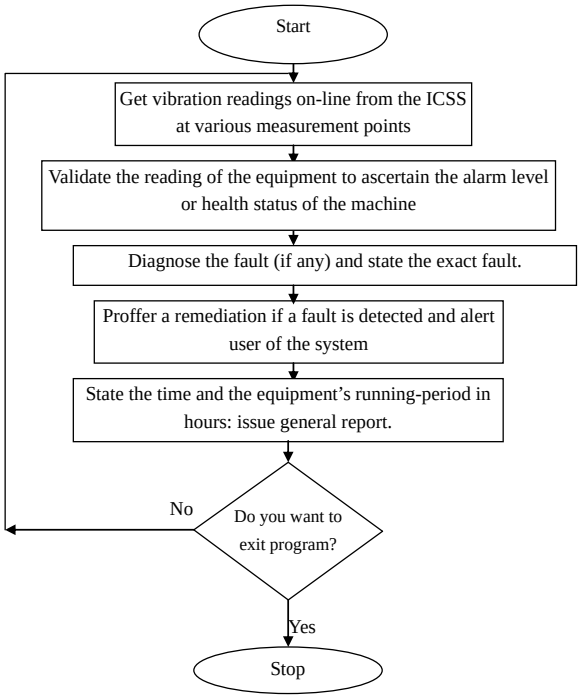


Figure 6. Process algorithm for fault diagnosis.

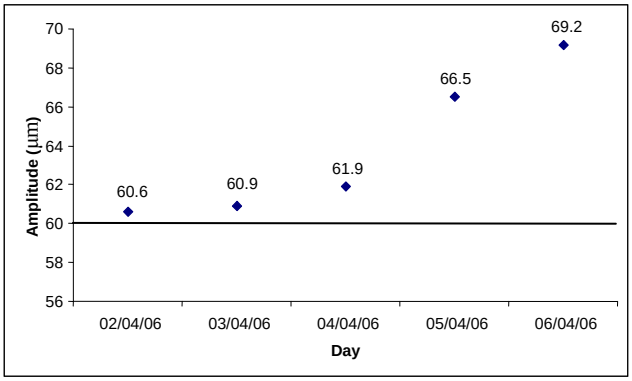


Figure 7(a). On-line displacement of shaft defect for the 5 (NDEA) pump bearing.

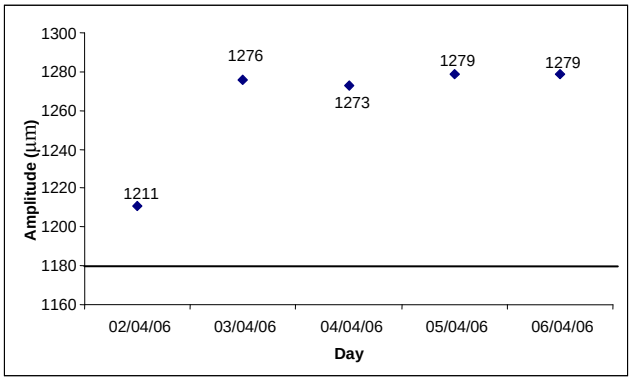


Figure 7(b). On-line gap defect for the 2 (DER) motor bearing.

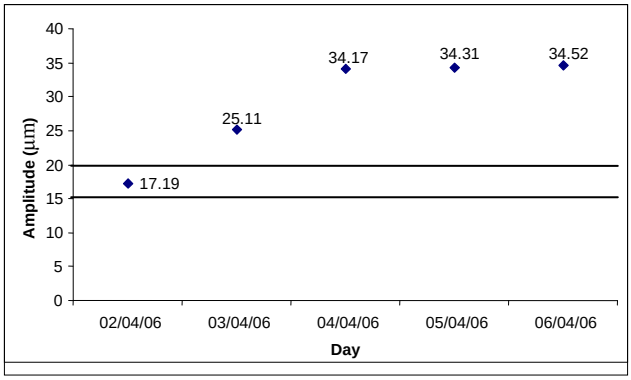


Figure 7(c). On-line imbalance of shaft defect for the 5 (NDEA) undamaged pump bearing.

5. Conclusions

These series of tests were set up to observe shaft rotations for speeds of 10→15 revolutions per second, and all key events were noted. Installed probes, on the driven end of the damaged bearing of the motor that was used for measuring the outer race deflections, also were used to monitor the misalignment and imbalance of the shaft. Defect locations in the raceway were determined from the spike locations in the frequency domain.

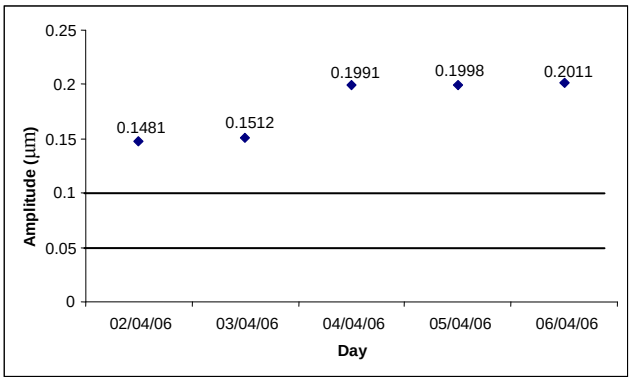


Figure 7(d). On-line misalignment of the shaft defect for the 3 (DEA) motor bearing.

Time-waveform analysis is an excellent analytical tool for fault diagnosis and prognostics: it enhances FFT information for on-line monitoring.

6. Recommendations

Accurately diagnosing and locating faults in machinery involving rotating components (e.g. pumps and turbines) can be achieved effectively by on-line vibration monitoring. This practice will reduce the waste of energy, effort expenditure, time consumption and drudgery experienced with off-line data-acquisition procedures and will facilitate wiser, maintenance decision making.

7. Acknowledgements

The authors wish to thank Netco-Dietsmann Nigeria Limited and Aminam/Kpono Field-Nigeria for logistics support and the provision of vital information that led to the successful execution of this work.

8. References

[1] R. Wilfried, S. Ulrich, P. Oliver, S. Christian, and W. B. Fiedrich, "Basics of vibration monitoring for fault detection and process control," Non-Destructive Testing Department, Institute of Materials Science, University of Hanover, Garbsen, Germany, pp. 23–28, 2003.

[2] J. S. Hansen, Harker, and G. Roger, "A new method for rolling-element bearing monitoring in the petrochemical industry," Proceeding of the Vibration Institute Machinery Vibration Monitoring and Analysis Meeting, New Orleans, pp. 34–45, 1984.

[3] D. E. Bently, "Monitoring rolling-element bearings," Orbit Article, Bently Corporation, Nevada, pp. 12–23, 1982.

[4] D. B. Spencer and J. S. Hensen, "Better way to monitor bearings," Hydrocarbon Processing, pp. 2–7, 1985.

[5] E. A. Ogbonnaya, "Thermodynamics of steam and gas turbines," Oru's Press Ltd, Port Harcourt, Nigeria, pp. 11, 2004.

[6] V. B. Alexej and N. A. Bakorva, "Diagnostics of gearing and geared couplings using envelope spectrum methods," Vibro Acoustical Systems and Technologies (VAST), Inc., Saint-Petersburg, Russia, 1994. Online Available: http://www.vibrotek.com/bio/anb.htm.

[7] B. E. Parker Jr, H. V. Poor, E. C. Larson, T. A. Hamilton, and J. P. Frankel, Noise Conference, Proceedings of the Statistical Change Detection Using Non-Linear Models, USA, pp. 319–324, 1997.

[8] S. David, "Predictive intelligent machining and machine-monitoring sensors," March Consulting Group (Managing Maintenance into the 1990's), 1989. Online Available: http://www.vibanalysis.co.uk/the author/about. html.

[9] 01db Steel Technologies, Metravib, 2001. Online Available: http://www.01db-metravib.com (Assessed 26 December, 2005).

[10] A. V. Barkov and N. A. Barkova, "Automatic diagnostics of rolling-element bearings using enveloping methods," Proceedings of the 18th Annual Meeting, The Vibration Institute, USA, pp. 21–23, 1995.

[11] E. A. Ogbonnaya and K. E. Koumako, "Basic automatic control," King Tovic International, Port Harcourt, Nigeria, pp. 109–116, 2006.

Abbreviations, Nomenclature and Units

AC	Alternating current (Amp)
AL+	High alarm (i.e. when measured value is greater than pre-alarm limit)
AL-	Low alarm (i.e. when measured value is less than pre-alarm limit)
B_d	Diameter of the ball bearing (m)
DE	Driven end
DER	Driven-end radial
DG+	High danger (i.e. when measured value is greater than pre-alarm limit)
DG-	Low danger (i.e. when measured value is less than pre-alarm limit)
F	Is the frequency in Hz
FDS	Fault-Diagnostic Software
FFT	Fast Fourier-transform
F0	Imbalance
g	Acceleration due to Gravity (ms^{-2})
Hz	Cycles per second
ICSS	Integrated control switching system
IRBP	Inner-race-ball passage frequency (s^{-1})
i_c	Cage rotating-frequency of the bearing (s^{-1})
i_R	Rolling element of the bearing (s^{-1})
NDEA	Non drive-end axial (i.e. measurement location on the pump bearing in the same direction as the shaft centerline)
ORBP	Outer-race-ball passage frequency (s^{-1})
P	Is the period in seconds (the amount of time required to complete one cycle)
P_d	Pitch diameter of the bearing (m)
PLC	Programmable logic control
REBAM	Rolling element bearing activity monitor
RPS	Revolutions per second
SKF	Svenska Kullagerfabriken; a Swedish bearing company
VDC	Direct-current voltage (V)
S	Seconds
z	Number of rolling elements of the bearing. z is equal to 9 in the chosen test case
θ	Contact angle between the rolling element and rolling surface (degrees)
ω	Angular speed (Revolutions per second)

Engineering, **2010, 2, 55-59**
doi:10.4236/eng.2010.21007 Published Online January 2010 (http://www.SciRP.org/journal/eng/).

Analysis for Pull-In Voltage of a Multilayered Micro-Bridge Driven by Electrostatic Force

Yu LIU[1,2,3], Guochao WANG[1,2,3], Hongyun YANG[1,2,3]

[1]*Automobile College, Chongqing University of Technology, Chongqing, China*

[2]*Key Laboratory of Automobile Parts & Test Technique in Chongqing, Chongqing, China*

[3]*Chongqing Engineering Research Center for Automobile Power System and Control, Chongqing, China*

Email: *liuyu_cq*@126.*com*

Received August 8, 2009; *revised September* 4, 2009; *accepted September* 12, 2009

Abstract

A trial solution for bending deflection of a multilayered micro-bridge subject to a voltage induced load is presented. The relation between the applied voltage and the displacements of the micro-bridge in the pull-in state is analyzed by energy method. Furthermore, two analytical expressions about normalized displacement and pull-in voltage are carried out. It's proved that the value of normalized displacement is not influenced by residual stress if axial and shear deformation is ignored. Finally, the theoretical results are compared with that of FEM, and they show good agreement.

Keywords: MEMS, Electrostatic Actuation, Multilayered Micro-Bridge, Trial Solution, Energy Method, Pull-In Voltage

1. Introduction

Moving part is the most frequently used one in MEMS structures. There are various principles that can be used to drive the moving part, including electrostatic, piezoelectric, thermal, magnetism, etc. Compared with the others, the method of electrostatic force drive is more attractive [1,2] because of its larger force caused by micro-effect and non-contact, which is beneficial to high precision. Furthermore, the process of MEMS devices driven by electrostatic force is compatible with IC process.

Micro-bridge structure driven by electrostatic force is familiar in MEMS devices, such as changeable capacitors, RF switches, micro-resonators, pressure sensors, and so on. On one hand, it is a common structure being used to obtain the mechanical parameters of film, just like Young's Modulus, residual stress, yield strength and bending strength [3–5]. But on the other hand, MEMS devices with micro-bridge structure are often with low reliability and worse quality, which is induced by pull-in phenomena. So it's worthy to research the pull-in phenomena of micro-bridge structure subject to a voltage induced load, which is important to MEMS design and optimization. For example, drive voltage must be less than pull-in voltage for most structures driven by electrostatic force, while it's opposite for DMD and self-measure unit of micro-accelerometer [6].

Although many people study the pull-in phenomena of the monolayered micro-bridge structure driven by electrostatic force, and get some valuable conclusions, they, however, cannot be applied to the multilayered micro-bridge structure. So some people [7] devote themselves to research on the pull-in phenomena of the multilayered micro-bridge structure. Whereas the analytical expression is complicated, this depends on the introduction of the assumptions and the form of the trial solution.

This paper presents a trial solution for bending deflection of a multilayered micro-bridge subject to a voltage induced load. The relation between the applied voltage and the displacements of the micro-bridge in the pull-in state is analyzed by energy method. Furthermore, two analytical expressions about normalized displacement and pull-in voltage are carried out. It's proved that the value of normalized displacement is not influenced by residual stress if axial and shear deformation is ignored. Finally, the theoretical results are compared with that of FEM, and they show good agreement.

2. Model

The micro-bridge with a rectangular cross-section is made of n layer materials, illustrated in Figure 1. Here we assume the top of the beam is a conductor, while the others are dielectric. The electrode under the beam is

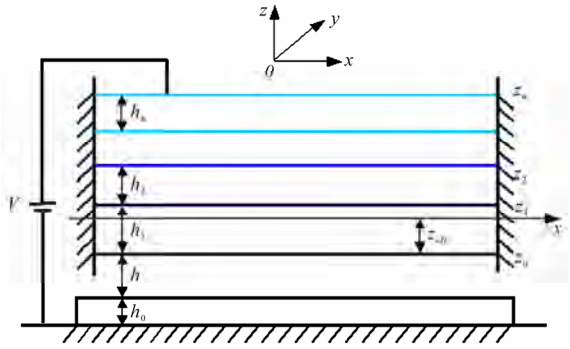


Figure 1. Schematic picture of a multilayered micro-bridge subject to a voltage induced load.

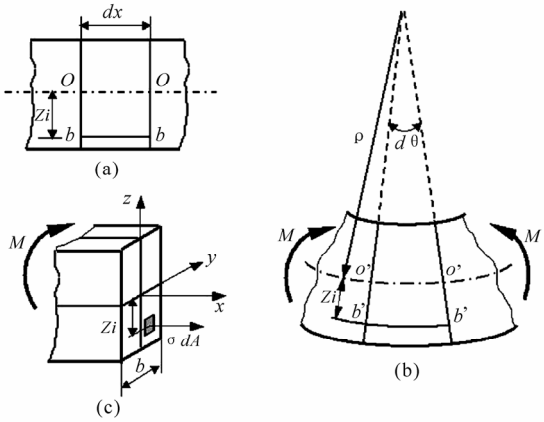


Figure 2. Deformation of an infinitesimal line element *dx*.

covered with the insulated material, which is fixed on the substrate. It's not difficult to imagine that micro-bridge will be bended when the voltage is added to the beam and the electrode.

Let's explain the meanings of some symbols firstly, where l is the length of the beam. h_i, E_i, ε_{ri} are the thickness, Young's Modulus and relative permittivity of the *i*-layer material respectively. h_0, ε_{r0} are the thickness and relative permittivity of the insulated layer above the electrode respectively. V is the voltage added to the beam and the electrode. h is the gap distance between the bottom of the beam and the top of the electrode when $V = 0$.

Set symmetric axis of beam's cross section as z axis, and positive direction is upwards. Set the equivalent neutral axis as y axis, and its position is needed to be confirmed. Then x axis is the perpendicular axis of yz plane and through the origin. The distance from the equivalent neutral axis to the bottom of the beam is z_{eff}. It's clear that $z_{eff} \neq \frac{1}{2}\sum_{i=1}^{n} h_i$ due to the different Young's Modulus of n layer materials.

As shown in Figure 2, according to the plane cross-section assumption on pure bending, we calculate the strain ε_i of the fiber bb while $V \neq 0$, which is in the i-layer and the primary length as well as the gap distance to neutral surface are *dx*, z_i respectively.

$$\varepsilon_i = -\frac{z_i}{\rho}$$

where ρ is curvature radius of the neutral surface.

Based on the assumption on pure bending, that is, there is no normal stress between the longitudinal fibers; we calculate the stress σ_i of the fiber bb according to Hooke's law. Of course, the deformation is at the elastic deformation period.

$$\sigma_i = E_i \varepsilon_i = -E_i \frac{z_i}{\rho}$$

Uncountable micro internal force $\sigma_i dA$ composes a system of parallel force in space, which can be simplified as three internal force components. What we're concerned about is the component parallel to x-axis, namely axial force $N = \int_A \sigma_i dA$. Based on the equilibrium of static equilibrium $\sum X = N = 0$ and the following formula,

$$\begin{aligned} N &= \int_A \sigma_i dA \\ &= \int_{A_1} \sigma_1 dA_1 + \int_{A_2} \sigma_2 dA_2 + \cdots + \int_{A_n} \sigma_n dA_n \\ &= -\frac{1}{\rho}\left[\int_{A_1} E_1 z_1 dA_1 + \int_{A_2} E_2 z_2 dA_2 + \cdots + \int_{A_n} E_n z_n dA_n\right] \\ &= -\frac{1}{\rho}\left[\int_{-z_{eff}}^{h_1 - z_{eff}} E_1 b z_1 dz_1 + \int_{h_1 - z_{eff}}^{h_1 + h_2 - z_{eff}} E_2 b z_2 dz_2 + \cdots + \int_{h_1 + h_2 + \cdots + h_{n-1} - z_{eff}}^{h_1 + h_2 + \cdots + h_n - z_{eff}} E_n b z_n dz_n\right] \end{aligned}$$

We get

$$z_{eff} = \frac{\sum_{j=1}^{n} E_j h_j \left(\frac{h_j}{2} + \sum_{i=1}^{j-1} h_i\right)}{\sum_{i=1}^{n} E_i h_i}$$

Therefore,

$$z_0 = -z_{eff}$$

$$z_i = \left(\sum_{j=1}^{i} h_j\right) - z_{eff} \qquad i = 1, 2, \cdots, n$$

where z_0, z_n are the coordinates of the micro-bridge's top and bottom surface along z-axis, while $z_1, \cdots, z_{n-1}$ are that of layer-layer interface respectively, as shown in Figure 1.

3. Solution

Here we choose a trial solution for bending deflection of

a multilayered micro-bridge subject to a voltage induced load as following [8]:

$$w(x) = A(1-\cos\frac{2\pi x}{l}) = 2A\sin^2\frac{\pi x}{l}$$

where A is the undetermined coefficient and $2A$ is the maximum deflection of the micro-bridge.

It's acceptable to neglect the influence of axial and shear deformation on bending due to the small displacement assumption of the micro-bridge.

3.1. No Consideration of Residual Stress

For this situation, the strain ε_i of the i-layer material only includes flexural strain ε_{bi}. That is,

$$\varepsilon_i = \varepsilon_{bi} = -z\frac{d^2w}{dx^2} = z\frac{4\pi^2 A}{l^2}\cos\frac{2\pi x}{l}$$

So the unit volume deformation energy and the deformation energy of the *i*-layer material are described by the following formula:

$$u_i = \frac{1}{2}E_i\varepsilon_i^2$$

$$U_{si} = \int_{v_i} u_i dv_i = \frac{E_i b}{2}\int_0^l\int_{z_{i-1}}^{z_i}\varepsilon_i^2 dzdx$$

Therefore, the total deformation energy of the micro-bridge U_s is equal to the summation of the deformation energy of n layer materials. That is,

$$U_s = \sum_{i=1}^{n}U_{si} = \sum_{i=1}^{n}\frac{E_i b}{2}\int_0^l\int_{z_{i-1}}^{z_i}\varepsilon_i^2 dzdx$$

$$= \frac{blA^2}{12}(\frac{2\pi}{l})^4\sum_{i=1}^{n}E_i(z_i^3 - z_{i-1}^3)$$

Considering the influence of the insulated layers on the gap distance, we introduce the equivalent gap distance $\tilde{h}$ [9]

$$\tilde{h} = h + \frac{h_0}{\varepsilon_{r0}} + \frac{h_1}{\varepsilon_{r1}} + \cdots + \frac{h_{n-1}}{\varepsilon_{r(n-1)}}$$

where h_i/ε_{ri} is the equivalent air gap distance of the *i*-layer insulated material. Then the model in Figure 1 is simplified as Figure 3, where $\tilde{\delta}$ is the equivalent thickness, which will be discussed in detail later.

So the electrostatic energy can be described as [8]

$$U_e = -\frac{V^2 b\varepsilon_0 l}{2\sqrt{\tilde{h}(\tilde{h}-2A)}}$$

where ε_0 is the permittivity of vacuum.

Therefore, the energy function is described as [10]

$$\prod = U_s + U_e$$

Because the value of the electrostatic force is in inverse proportion to the square of distance, and the distribution of electric field is influenced by structural dimension as well as position, so the issue of bending equilibrium stability should be considered. It exists when the electrostatic force is equal to the elastic force. If the voltage V exceeds the critical value, the equilibrium state is destroyed and the micro-bridge will contact with the fixed plane, which is called pull-in phenomena. The minimum critical voltage is called pull-in voltage, marked as V_{PI}. Here,

$$\frac{\partial\prod}{\partial A} = 0 \quad \text{and} \quad \frac{\partial^2\prod}{\partial A^2} = 0 \tag{1}$$

Mark the maximum deflection of the micro-bridge under V_{PI} as $2A_{cr}$, then

$$\frac{A_{cr}}{6}(\frac{2\pi}{l})^4\sum_{i=1}^{n}E_i(z_i^3 - z_{i-1}^3) = \frac{V^2\varepsilon}{2}\tilde{h}(\tilde{h}^2 - 2\tilde{h}A_{cr})^{-\frac{3}{2}} \tag{2}$$

$$\frac{1}{6}(\frac{2\pi}{l})^4\sum_{i=1}^{n}E_i(z_i^3 - z_{i-1}^3) = \frac{3V^2\varepsilon}{2}\tilde{h}^2(\tilde{h}^2 - 2\tilde{h}A_{cr})^{-\frac{5}{2}} \tag{3}$$

So

$$\frac{A_{cr}}{\tilde{h}} = 0.2 \tag{4}$$

And the normalized displacement

$$\beta = \frac{2A_{cr}}{\tilde{h}} = 0.4$$

3.2. Consideration of Residual Stress

For this situation, the strain ε_i of the *i*-layer material includes not only flexural strain ε_{bi} but also the strain

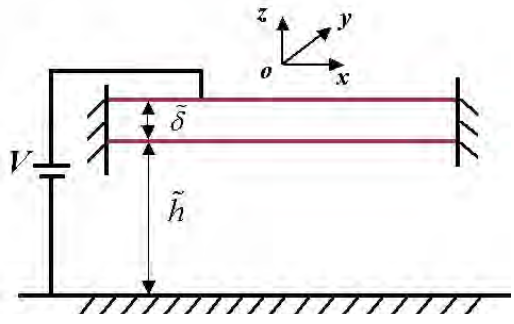


Figure 3. The equivalent model of a multilayered micro-bridge.

induced by the residual stress σ_i of the *i*-layer material. That is,

$$\varepsilon_i = \varepsilon_{bi} + \frac{\sigma_i}{E_i} = -z\frac{d^2w}{dx^2} + \frac{\sigma_i}{E_i} = -z\frac{4\pi^2 A}{l^2}\cos\frac{2\pi x}{l} + \frac{\sigma_i}{E_i}$$

Similarly, the deformation energy of the micro-bridge

$$\begin{aligned} U_s &= \sum_{i=1}^{n}\frac{E_i b}{2}\int_0^l\int_{z_{i-1}}^{z_i}\varepsilon_i^2 dzdx \\ &= \frac{blA^2}{12}\left(\frac{2\pi}{l}\right)^4\sum_{i=1}^{n}E_i(z_i^3 - z_{i-1}^3) + \sum_{i=1}^{n}\frac{blh_i\sigma_i^2}{2E_i} \end{aligned} \quad (5)$$

Because the second part of Formula (5) is not the function of the undetermined coefficient *A*, so the results from Formula (1) are same to that of Subsection 3.1.

It's clear that the value of normalized displacement is not influenced by residual stress if axial and shear deformation is ignored.

Based on Equations (2) and (4), we get

$$V_{PI} = \sqrt{48.290\frac{\tilde{h}^3\sum_{i=1}^{n}E_i(z_i^3 - z_{i-1}^3)}{\varepsilon_0 l^4}} \quad (6)$$

For the multilayered micro-bridge with a rectangular cross-section, the area moment of inertia

$$\tilde{I} = \int_{z_0}^{z_n} bz^2 dz = \frac{b}{3}(z_n^3 - z_0^3)$$

and the flexural stiffness

$$\tilde{E}\tilde{I} = \sum_{i=1}^{n}\int_{z_{i-1}}^{z_i} E_i bz^2 dz = \frac{b}{3}\sum_{i=1}^{n}E_i(z_i^3 - z_{i-1}^3)$$

then the equivalent Young's Modulus

$$\tilde{E} = \frac{\tilde{E}\tilde{I}}{\tilde{I}} = \frac{\sum_{i=1}^{n}E_i(z_i^3 - z_{i-1}^3)}{z_n^3 - z_0^3}$$

and the equivalent thickness

$$\tilde{\delta} = \left(\frac{12\tilde{I}}{b}\right)^{\frac{1}{3}} = \left[4(z_n^3 - z_0^3)\right]^{\frac{1}{3}}$$

Therefore, the formula (6) is simplified as

$$V_{PI} = \sqrt{12.0725\frac{\tilde{E}\tilde{\delta}^3\tilde{h}^3}{\varepsilon_0 l^4}} \quad (7)$$

It's clear that the multilayered micro-bridge can be considered as a monolayered micro-bridge if only the equivalent Young's Modulus $\tilde{E}$ and the equivalent thickness $\tilde{\delta}$ are carried out.

In the above discussion, we assume the top of the beam is a conductor while the others are dielectric. If the conductor is not on the top of the micro-bridge, or more than one layer is conductor, the equivalent gap distance $\tilde{h}$ will differ.

4. Validation

All the above deduction is based on a given trial solution for bending deflection. Are they correct? How about the accuracy? Now we try to prove it by simulation.

Firstly, some examples relating to the monolayered micro-bridges are considered. For the monolayered micro-bridge given in paper [11], the value of V_{PI} is 32.5V, and the simulated values by ANSYS as well as Intellisuite are 37.6V, 39.5V respectively, which are presented in paper [11]. Whereas the analytical value based on the Formula (7) is 34.2V, which is more close to the simulated values. And for that given in paper [12], the simulated value of V_{PI} is 40V, while the analytical value based on the Formula (7) is 40.99V. The error is only 2.48%.

Secondly, a multilayered micro-bridge with a rectangular cross-section is considered. It's made of two layer materials. And the electrode under it is covered with 0.5 μm Si_3N_4.

Model 1: the top material is gold. And the other is Si_3N_4.

Model 2: the top material is Si_3N_4. And the other is gold.

Table 1. Geometry and material parameters of the bi-layered micro-bridge.

Item	Symbol (unit)	Model 1	Model 2
length	l (μm)	400	
width	b (μm)	50	
gap distance (V=0)	h (μm)	2.0	
thickness of the insulated layer	h_0 (μm)	0.5	
thickness	h_1 (μm)	2.0	0.5
	h_2 (μm)	0.5	2.0
relative permittivity	$\varepsilon_{r0} = \varepsilon_{r1}$	8.0	
Young's Modulus	E_1 (GPa)	210	57
	E_2 (GPa)	57	210

Table 2. V_{PI}, β of the bi-layered micro-bridge.

Item	Model	Analytical values	Simulated values	Error
V_{PI}	model 1	37.9671V	40.2187V	-5.60%
	model 2	31.9798V	33.3438V	-4.09%
β	model 1	0.4	0.4069	-1.70%
	model 1#		0.3933	1.70%
	model 2		0.3972	0.70%
	model 2#		0.4132	-3.19%

Note: # Residual stress of the gold layer is 100 MPa.

The geometry and material parameters of the bi-layered micro-bridge are listed in Table 1. The micro-bridge will be bended when subject to a voltage induced load.

For Models 1 and 2, we get the simulated values of the pull-in voltage V_{PI} and the normalized displacement β by FEM. Table 2 lists the simulated performances and the analytical values of V_{PI} and β, as well as the error of them. It shows good correlation. So the trial solution of deflection is acceptable.

5. Conclusions

This paper presents a trial solution for bending deflection of a multilayered micro-bridge subject to a voltage induced load. The relation between the applied voltage and the displacements of the micro-bridge in the pull-in state is analyzed by energy method. Furthermore, two analytical expressions about normalized displacement and pull-in voltage are carried out. The theoretical results are compared with that of FEM, and they show good agreement.

All results in this paper are based on the form of the selected trial solution. The better the selected trial solution is, the higher the accuracy of solution is. Moreover, the axial and shear deformation is ignored, which maybe has influence on the stress and strain. Besides, in order to get an analytical expression of V_{PI}, it's difficult to consider the influence of some secondary effects, such as the marginal effect of electric field. However, this paper provides an analytical model with high accuracy for a multilayered micro-bridge driven by electrostatic force, which is beneficial to design, optimization and application of MEMS devices with the micro-bridge structure.

6. References

[1] H. Y. Ma, Y. H. Wang, M. L Wang, *et al.* "The sequential coupling analysis of a cantilever driving by electrostatic force in an optical switch," Optical Instruments, Vol. 25, No. 3, pp. 17, 2003.

[2] G. Li and N. R. Aluru, "Efficient mixed-domain analysis of electrostatic MEMS," ICCAD, pp. 474, 2002.

[3] T. Y. Zhang, Y. J. Su, C. F. Qian, *et al.* "Micro-bridge testing of silicon nitride thin films deposited on silicon wafers," Acta Mater, Vol. 48, pp. 2843, 2000.

[4] Y. Zhou, C. S. Yang, J. A. Chen, *et al.* "Investigation of Young's modulus and residual stress of copper film micro-bridges by MEMS technology," Modern Scientific Instruments, Vol. 4, pp. 45, 2003.

[5] M. J. Wang, Y. Zhou, J. A. Chen, *et al.* "Measurements of elastic modulus and residual stress of nickel film by micro-bridge testing methods," Electronic Components & Materials, Vol. 23, No. 12, pp. 13, 2004.

[6] H. Yang, M. H. Bao, S. Q. Shen, *et al.* "The displacement characteristics of the micromechanical structures driven by electrostatic force," Journal of Fudan University (Natural Science), Vol. 38, No. 3, pp. 282, 1999.

[7] H. Rong, Q. A. Huang, M. Nie, *et al.* "An analytical model for pull-in voltage of doubly-clamped multi-layer beams," Chinese Journal of Semiconductors, Vol. 24, No. 11, pp. 1185, 2003.

[8] D. H. Sun, Y. Q. Huang, W. Zheng, *et al.* "On the modeling methodology of MEMS system-level simulation," Journal of Xiamen University (Natural Science), Vol. 40, No. 2, pp. 297, 2001.

[9] M. Nie, Q. A. Huang, J. H. Wang, *et al.* "Analysis of deflection and pull-in voltage of a multi-layer cantilever under an electrostatic load," Chinese Journal of mechanical engineering, Vol. 40, No. 8, pp. 72, 2004.

[10] Y. Y. Wang, T. I. Kamins, B. Y. Zhao, *et al.* Polysilicon Film and its Application in IC (2nd Version), Science Press, Beijing, 2001.

[11] C. S. Wang, W. B. Zhang, J. Fang, *et al.* "Research on coupled electro-mechanical analysis and application for typical components in MEMS," Journal of Mechanical Strength, Vol. 23, No. 4, pp. 503, 2001.

[12] P. M. Osterberg, S. D. Senturia, "M-test: A test chip for MEMS material property measurement using electrostatically actuated test structures," Journal of Microelectromechanical Systems, Vol. 6, No. 2, pp. 107, 1997.

Engineering, **2010, 2, 60-64**
doi:10.4236/eng.2010.21008 Published Online January 2010 (http://www.SciRP.org/journal/eng/).

Research on Early Fault Self-Recovery Monitoring of Aero-Engine Rotor System

Zhongsheng WANG, Shiwe MA
School of Aeronautics Northwestern Polytechnical University Xi'an, China
Email: *sazs_wang@nwpu.edu.cn*
Received August 19, 2009; *revised September* 14, 2009; *accepted September* 20, 2009

Abstract

In order to increase robustness of the AERS (Aero-engine Rotor System) and to solve the problem of lacking fault samples in fault diagnosis and the difficulty in identifying early weak fault, we proposed a new method that it not only can identify the early fault of AERS but also it can do self-recovery monitoring of fault. Our method is based on the analysis of the early fault features on AERS, and it combined the SVM (Support Vector Machine) with the stochastic resonance theory and the wavelet packet decomposition and fault self-recovery. First, we zoom the early fault feature signals by using the stochastic resonance theory. Second, we extract the feature vectors of early fault using the multi-resolution analysis of the wavelet packet. Third, we input the feature vectors to a fault classifier, which can be used to identify the early fault of AERS and carry out self-recovery monitoring of fault. In this paper, features of early fault on AERS, the zoom of early fault characteristics, the extraction method of early fault characteristics, the construction of multi-fault classifier and way of fault self-recovery monitoring are studied. Results show that our method can effectively identify the early fault of AERS, especially for identifying of fault with small samples, and it can carry on self-recovery monitoring of fault.

Keywords: AERS, Early Fault, Support Vector Machine, Classification Identification of Fault, Self-Recovery Monitoring of Fault

1. Introduction

With the development of the modern aviation industry, the safety of aircraft and reliability are more and more attracted. The engine is heart of aircraft and AERS is central part of engine. If AERS has fault in flight, the aircraft would be severely threatened in safety. Because the structure of AERS is complex, load is bigger and the change of operation conditions is frequency, it will be the more difficult in the fault identification. Especially in the early failures occurred, if we can catch timely the fault information, and can effectively identify it and carry on self-recovery monitoring of fault, it will have important significance to eradicate or eliminate the potential fault caused of an accident. Fault identification on AERS is widely studied and many results are obtained [1−3]. At present now, effective identification of early fault has more difficulty. In particular, obtaining of the early fault information is more difficulty, and effectiveness of fault identification is not satisfactory. Because operation of AERS is in strong noise density, the early fault information is very weak and signals are easily flooded in noises, satisfactory results are very difficult by general methods. Our method can not only fast identify the early fault of the AERS but also carry on self-recovery monitory of fault.

*Project (50675178, 60472116) Supported by National Natural Science Foundation of China

2. Features of Early Fault on AERS

In the flight, the wear, deformation, corrosion and fracture of structure components and effects of work stress, external environment and human factors will lead to faults of AERS. The early fault of AERS often shows a form in micro-cracks, micro-creeping, micro-corrosion and micro-wear. These faults, with the exception of little sudden faults, the majority have a development process from nothing to fault, minor to general, evolution to fast. In this process, the structure of system, properties and internal energy will change. We can monitor the early faults through catching these fault information in time,

and it can be identified and be self-recovery. These early faults of AERS have the following features:

1) Fault signal is very weak. When fault is at early stage or just sprouted, the changes of the fault signal is very weak in amplitude, phase and time-frequency characteristics, and the little fault characterization is often difficult to detect.

2) Fault signal will be drowned by noise signal. During the flight, the noise signal is usually among the fault signal. When the fault signal is very weak and the noise signal is very strong, the early fault signal will be drowned by noise signal. In order to detect the early fault, we must reduce noise or extracted early fault information from the noise signal.

3) Fault signal is often a transient one. The damage structure components of aircraft is generated by impact loading, and this fault is manifested by transient signal, such as expansion of early cracks, is a process from gradual to mutation.

4) Fault occurred in the area of stress concentration. When area of stress concentration is acted by strong regional load, the fault of structural part is easy occurrence in the creeping. For the ferromagnetic metal part, the magnetic memory method can be used for the early fault detection and localization [4].

5) Fault has volatile. That means sometimes we can not find the fault trace in fault condition. But with the change of action or time, the system can recover automatically. Volatile fault can be shown by eventuality fault, transient fault and interruption fault.

3. Extraction of Early Fault Characreristic

Because the vibration of engine is larger in operation, fault characteristics of the AERS will be submerged in the strong background noise in the early. In order to extract characteristics of early fault from being submerged signals by noise, we use the stochastic resonance theory to zoom characteristic signals of early fault [5].

In the failure, the energy of AERS will change in all frequency bands and different faults have different effects to the signal energy in each frequency band. So, we can use wavelet packet to decompose the output signal in stochastic resonance and select signal energy in the characteristics frequency band as a feature vector.

If the data length of original signal $x(t)$ is N, the data length of discrete signals $x^{k,m}(i)$ is reduced to $2^{-k}N$ by decomposition of wavelet packet and its energy can be expressed as:

$$E_n\left(x^{k,m}\right)=\frac{1}{2^{-k}N-1}\sum_{i=1}^{2^{-k}N}\left(x^{k,m}(i)\right)^2 \tag{1}$$

where, N = the length of original data, k = layer number by wavelet packet decomposition, m = the serial number of decomposition frequency band location = 0, 1, 2, ... 2^k-1.

The energy within the each frequency band can be calculated by Equation (1) and the feature vector can be constructed by the energy.

4. Classification Identification of Fault

The SVM can do the classification very well in the few numbers of fault samples and it can solve the identification problems on nonlinear and high-dimensional pattern.

The fault identification of the AERS belongs to problem of multi-classification and it needs to construct multi-fault classifier. In this paper, we adopt improved classification algorithm of the "one-to-many".

To the classification of K type, way of "one-to-many" is need to construct K two classifiers. In this way constructing every two classifiers, all n training samples of K type should be operated. In the testing and classification, the scale of classifiers is larger and speed is slower.

An improved the "one-to-many" classification algorithm is to construct two classifiers of K, and the training sample of m type in the K classifiers is $y_i^m=1$ and other types marking is $y_i^m=-1$. Established output function in M classifiers can be expressed as:

$$f^m(x)=\operatorname{sgn}\left\{\sum_{SVM}\alpha_i^m y_i^m \mathrm{K}(x_i\cdot x)+b^m\right\} \tag{2}$$

The algorithm overcomes defects that "one-on-one" approach needs to establish many number classifiers and it can control the training samples number in traditional "one-to-many". Its operate speed is fast and effect of fault classification is well.

5. Self-Recovery Monitoring of Fault

Fault monitor process of AERS is shown in Figure 1.

Fault self-recovery database consist of many intelligent models. It can carry on self-recovery monitoring based on different fault sources and fault characteristics.

Self-recovery monitoring of fault can be realized by smart structure [6], affixing magnetic field [7], regenerating materials [8], etc.

When AERS has fault, the system can not work normally. Self-recovery model based on the fault compensation can restore original function of system by means of the fault self-recovery compensator. If system equations under normal condition are

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{3}$$

$$y(t)=Cx(t) \tag{4}$$

Compensator equations are

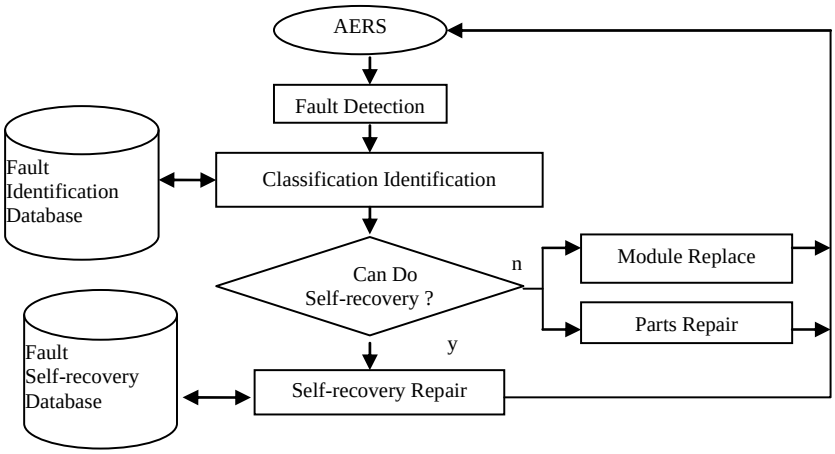


Figure 1. Fault self-recovery monitor process of AERS.

$$\dot{z}(t) = Dz(t) + Ey(t) \tag{5}$$

$$u(t) = Fz(t) + Hy(t) \tag{6}$$

The system loop equations depicted in (3) and (4) are

$$\dot{x}(t) = (A + BHC)x(t) + BFz(t) \tag{7}$$

$$\dot{z}(t) = ECx(t) + Dz(t) \tag{8}$$

When fault of component or sub-system occurred during the flight, the loop feedback equations are

$$\dot{x}(t) = (A_f + B_f \overline{H} C_f)x(t) + B_f \overline{F} z(t) \tag{9}$$

$$\dot{z}(t) = \overline{E} C_f x(t) + \overline{D} z(t) \tag{10}$$

To make the performance of fault system as close as possible to the performance of the original system, we can design the appropriate self-recovery compensator $\{\overline{D}, \overline{E}, \overline{F}, \overline{H}\}$ to achieve the fault self-recovery compensation. During flight, when the aircraft cockpit, wings and other important components have severe vibration or chatter, distributed piezoelectric driver compensator can weaken or offset the impact of vibration by vibration control and active vibration absorber.

Self-recovery monitoring based on intelligent structure is shown in Figure 2.

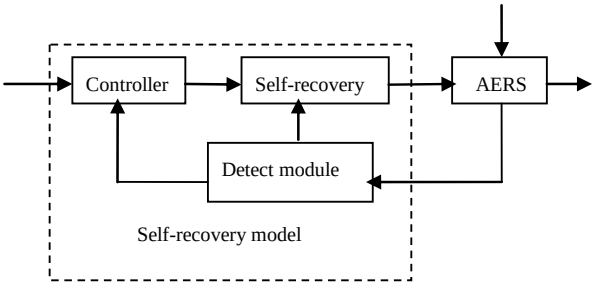


Figure 2. Self-recovery monitoring based on smart structure.

When the sensor module detects the fault information, the signals will transmit to self-recovery module and act to the controller. Fault of AERS will be restored by smart structure. It is based on different fault sources and fault feature, and the fault self-recovery tactics is adopted.

6. Experimental Results and Its Analysis

In order to verify the usefulness of the method, we choose the four conditions of the AERS. They are normal condition, early rotor misalignment, early rotor unbalance and early rotor crack. These signals are preprocessed and the fault features are extracted. Fault is identified by the multi-fault classifiers and fault is monitored by self-recovery module.

In the experiment, according to the character of AERS, we collected 10 group data by the acceleration sensor. They respectively correspond to the over four conditions in the 1800 rpm. The sampling frequency is 256 Hz and the rotation frequency is 30 Hz. The characteristic frequency of fault and its concomitant frequency are shown in Table 1.

Figure 3 shows the output waveform of the stochastic resonance system and its spectrum. The frequency component in 30 Hz is obvious in Figure 2. Because the fundamental frequency is 30 Hz, the rotor misalignment fault is often accompanied by 1xω (30 Hz), 2 xω (60 Hz) and 3xω (90 Hz).

Table 1. 2xω (60Hz) crossed concomitant frequency and unbalance.

Type	Misalignment	Unbalance	Crack
Characteristic frequency	1xω (30Hz), 2 xω (60Hz)	1xω (30Hz)	2xω (60Hz)
Concomitant frequency	3 xω (90Hz)		4xω (120Hz)

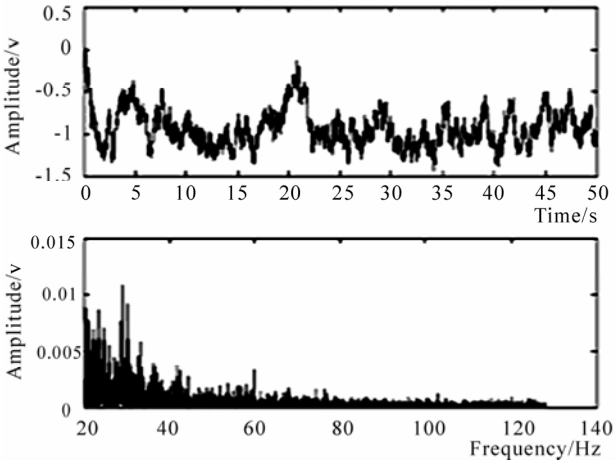


Figure 3. The output waveform of stochastic resonance system and its spectrum.

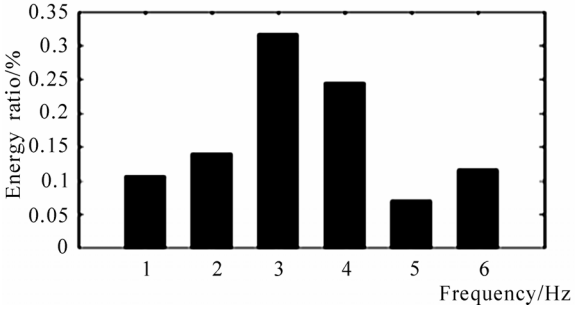
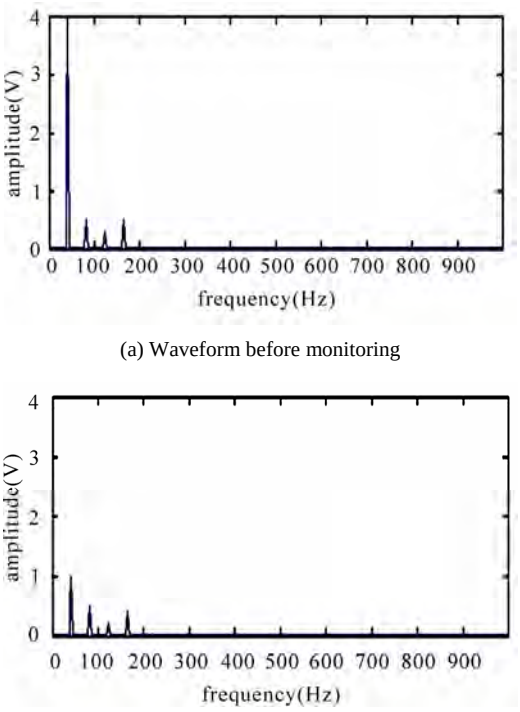


Figure 4. Energy distribution of rotor misalignment.

Table 2. Test results of stochastic resonance system.

Fault type	Misjudgment samples	Diagnosis samples	Diagnosis rate
Misalignment	1	40	97.5 %
Unbalance	0	40	100 %
Crack	2	40	95 %

The energy feature extraction is done by the wavelet packet analysis. It is based on the signal waveform of stochastic resonance system. Seven-layer wavelet packet is resolved in the db3 wavelet and we obtained 64 frequency bands. In order to reduce the amount of computation, we divide the frequency bands as 6 segments: 0 ~ 0.4X, 0.4 ~ 0.8X, 0.8 ~ 1.2 X, 1.8 ~ 2.2X, 2.8 ~ 3.2X and greater than 3.6X. Thus 64 frequency bands will be composed of the 6 groups: 1 ~ 4 bands, 5 ~ 8 bands, 9 ~ 11 bands, 12 ~ 15 bands, 16 ~ 30 bands and 55 ~ 64 bands. Energy value of each group is added together and they are processed on the normalization. The energy distribution of rotor misalignment fault will be acquired. It is shown in Figure 4.



(a) Waveform before monitoring

(b) Waveform after monitoring

Figure 5. Frequency domain waveform of rotor misalignment.

Repeated the above process, the energy distribution of each condition can be obtained and it is taken as training samples of SVM.

According to the training samples obtained from four conditions, we take 40 groups energy distributions from each condition as training samples and input to the improved fault classifier of "one-to-many". We choose Gaussian RBF kernel function as a classification function and make the parameters $\sigma = 0.01$, punishment factor C = 100. The classification results are shown in Table 2.

We can see that classification results by the stochastic resonance system are significantly high than classification result by direct wavelet packet feature extraction. The classification time of the each testing samples is smaller (in 0.05s, 1.8 GHZ computers). Its accuracy is higher and speed is quick in early fault identification.

In order to monitor misalignment, we adopt a principle of the electromagnetic effect [9]. The number of misalignment is detected by four acceleration sensors, and four electromagnetic sets are controlled by the output signal of four sensor. When misalignment occurred, the misalignment force F by rotor produced can be adjusted by alignment force F' of electromagnet produced. The alignment force F' is equal to the misalignment force F in number and they are contrary in direction. Figure 5 shows the result of self-recovery monitoring on rotor misalignment.

We can find out that vibration of rotor is obviously reduced in Figure 5(a) and Figure 5(b). Because of the misalignment fault of rotor is counteracted by the electromagnet force, the rotor misalignment is inhibited and the normal operation condition is restored well.

7. Conclusions

1) Our method can effectively extract the early fault feature of the AERS by combination the stochastic resonance with the wavelet packet resolving. Energy eigenvectors of constructed by this methods can accurately reflect the condition changes of AERS.

2) Multi-fault classifier based on the SVM has characteristics that its algorithm is simpler, the classification effect is well and identification efficiency is higher. It particularly suits to the classification identification of small sample and self-recovery monitoring of early fault on AERS.

3) The fault diagnosis is aimed at finding failure in time and to ensure safe operation of plant. The method of fault self-recovery monitoring provides an effective way.

8. References

[1] D. Simon, "A comparison of filtering approaches for aircraft engine health estimation," Aerospace Science and Technology, Vol. 12, No. 4, pp. 276–284, 2008.

[2] S. Borguet and O. Léonard, "Coupling principal component analysis and Kalman filtering algorithms for on-line aircraft engine diagnostics," Control Engineering Practice, Vol. 17, No. 4, pp. 494–502, 2009.

[3] T. Ramesh Babu and A. S. Sekhar, "Detection of two cracks in a rotor-bearing system using amplitude deviation curve," Journal of Sound and Vibration, Vol. 314, No. 3–5, pp. 457–464, 2008.

[4] A. Dubov and S. Kolokoinikov, "Review of welding problems and allied processes and their solution using the metal magnetic memory effect," Welding in the World, Vol. 49, No .9, pp. 306–313, 2005.

[5] K. Tanaka and M. Kawakatsu, "Stochastic resonance in auditory steady-state responses in a magnetoencephalogram," Clinical Neurophysiology, Vol. 119, No. 9, pp. 2104–2110, 2008.

[6] S. Hurlebaus and L. Gaul, "Smart structure dynamics," Mechanical Systems and Signal Processing, Vol. 20, No. 2, pp. 255–281, 2006.

[7] A. Ignatios and B. Alexey, "Anomaly induced effects in a magnetic field," Nuclear Physics B, Vol. 793, No. 1–2, pp. 246–259, 2008.

[8] Mueller and S. N. Sokolova, "Characteristics of lightweight aggregate from primary and recycled raw materials," Construction and Building Materials, Vol. 22, No. 4, pp. 703–712, 2008.

[9] Y. Asher, P. Yosef, and L. Yuri, "Spectral and variational principles of electromagnetic field excitation in wave guides," Physics Letters A, Vol. 344, No. 1, pp. 18–28, 2005.

Engineering, **2010, 2, 65-77**
doi:10.4236/eng.2010.21009 Published Online January 2010 (http://www.SciRP.org/journal/eng/).

An Adaptive Differential Evolution Algorithm to Solve Constrained Optimization Problems in Engineering Design

Youyun AO[1], Hongqin CHI[2]
[1]*School of Computer and Information, Anqing Teachers College, Anqing, China*
[2]*Department of Computer, Shanghai Normal University, Shanghai, China*
Email: *youyun.ao@gmail.com, chihq@shnu.edu.cn*
Received July 28, 2009; *revised August* 23, 2009; *accepted August* 28, 2009

Abstract

Differential evolution (DE) algorithm has been shown to be a simple and efficient evolutionary algorithm for global optimization over continuous spaces, and has been widely used in both benchmark test functions and real-world applications. This paper introduces a novel mutation operator, without using the scaling factor *F*, a conventional control parameter, and this mutation can generate multiple trial vectors by incorporating different weighted values at each generation, which can make the best of the selected multiple parents to improve the probability of generating a better offspring. In addition, in order to enhance the capacity of adaptation, a new and adaptive control parameter, i.e. the crossover rate *CR*, is presented and when one variable is beyond its boundary, a repair rule is also applied in this paper. The proposed algorithm ADE is validated on several constrained engineering design optimization problems reported in the specialized literature. Compared with respect to algorithms representative of the state-of-the-art in the area, the experimental results show that ADE can obtain good solutions on a test set of constrained optimization problems in engineering design.

Keywords: Differential Evolution, Constrained Optimization, Engineering Design, Evolutionary Algorithm, Constraint Handling

1. Introduction

Many real-world optimization problems involve multiple constraints which the optimal solution must satisfy. Usually, these problems are also called constrained optimization problems or nonlinear programming problems. Engineering design optimization problems are constrained optimization problems in engineering design. Like a constrained optimization problem, an engineering design optimization problem can be generally defined as follows [1–4]:

Minimize $f(\vec{x})$, $\vec{x} = [x_1, x_2, ..., x_n] \in \Re^n$

$$\text{Subject to } \begin{aligned} & g_j(\vec{x}) \le 0, \ j = 1,2,...,q \\ & h_j(\vec{x}) = 0, \ j = q+1, q+2, ..., m \end{aligned} \quad (1)$$

where $L_i \le x_i \le U_i, i = 1,2,...,D$

Here, n is the number of the decision or parameter variables (that is, $\vec{x}$ is a vector of size D), the ith variable x_i varies in the range $[L_i, U_i]$. The function $f(\vec{x})$ is the objective function, $g_j(\vec{x})$ is the jth inequality constraint and $h_j(\vec{x})$ is the jth equality constraint. The decision or search space S is written as $S = \prod_{i=1}^{D} [L_i, U_i]$, the feasible space expressed as $F = \{\vec{x} \in S \mid g_j(\vec{x}) \le 0, j = 1,2,...,q; h_j(\vec{x}) = 0, j = q+1, q+2,...,m\}$ is one subset of the decision space S (obviously, $F \subseteq S$) which satisfies the equality and inequality constraints.

Population-based evolutionary algorithm, mainly due to its ease to implement and use, and its less susceptibleness to the characteristics of the function to be optimized, has been very popular and successfully applied to constrained optimization problems [5]. And many successful applications of evolutionary algorithms to solve engineering design optimization problems in the specialized literature have been reported. Ray and Liew [6] used a swarm-like based approach to solve engineering optimization problems. He *et al.* [7] proposed an improved particle swarm optimization to solve mechanical design

optimization problems. Zhang *et al.* [8] proposed a differential evolution with dynamic stochastic selection to constrained optimization problems and constrained engineering design optimization problems. Akhtar *et al.* [9] proposed a socio-behavioural simulation model for engineering design optimization. He and Wang [10] proposed an effective co-evolutionary particle swarm optimization for constrained engineering design problems. Wang and Yin [11] proposed a ranking selection-based particle swarm optimizer for engineering design optimization problems. Differential evolution (DE) [12,13], a relatively new evolutionary technique, has been demonstrated to be simple and powerful and has been widely applied to both benchmark test functions and real-world applications [14]. This paper introduces an adaptive differential evolution (ADE) algorithm to solve engineering design optimization problems efficiently.

The remainder of this paper is organized as follows. Section 2 briefly introduces the basic idea of DE. Section 3 describes in detail the proposed algorithm ADE. Section 4 presents the experimental setup adopted and provides an analysis of the results obtained from our empirical study. Finally, our conclusions and some possible paths for future research are provided in Section 5.

2. The Basic DE Algorithm

Let's suppose that $\vec{x}_i^t = [x_{i,1}^t, x_{i,2}^t, ..., x_{i,D}^t]$ are solutions at generation t, $P^t = \{\vec{x}_1^t, \vec{x}_2^t, ..., \vec{x}_N^t\}$ is the population, where D denotes the dimension of solution space, N is the population size. In DE, the child population P^{t+1} is generated through the following operators [12,15]:

1) Mutation Operator: For each $\vec{x}_i^t$ in parent population, the mutant vector $\vec{v}_i^{t+1}$ is generated according to the following equation:

$$\vec{v}_i^{t+1} = \vec{x}_{r_1}^t + F \times (\vec{x}_{r_2}^t - \vec{x}_{r_3}^t) \tag{2}$$

where $r_1, r_2, r_3 \in \{1,2,...,N\} \setminus i$ are randomly chosen and mutually different, the scaling factor F controls amplification of the differential variation $(\vec{x}_{r_2}^t - \vec{x}_{r_3}^t)$.

2) Crossover Operator: For each individual $\vec{x}_i^t$, a trial vector $\vec{u}_i^{t+1}$ is generated by the following equation:

$$u_{i,j}^{t+1} = \begin{cases} v_{i,j}^{t+1}, & \text{if } (rand \le CR \,\|\, j = rand[1,D]) \\ x_{i,j}^t, & \text{otherwise} \end{cases} \tag{3}$$

where $rand$ is a uniform random number distributed between 0 and 1, $rand[1,D]$ is a randomly selected index from the set $\{1,2,...,D\}$, the crossover rate $CR \in [0,1]$ controls the diversity of the population.

3) Selection Operator: The child individual $\vec{x}_i^{t+1}$ is selected from each pair of $\vec{x}_i^t$ and $\vec{u}_i^{t+1}$ by using greedy selection criterion:

$$\vec{x}_i^{t+1} = \begin{cases} \vec{u}_i^{t+1}, & \text{if } (f(\vec{u}_i^{t+1}) < f(\vec{x}_i^t)) \\ \vec{x}_i^t, & \text{otherwise} \end{cases} \tag{4}$$

where the function f is the objective function and the condition $f(\vec{u}_i^{t+1}) < f(\vec{x}_i^t)$ means the individual $\vec{u}_i^{t+1}$ is better than $\vec{x}_i^t$.

Therefore, the conventional DE algorithm based on scheme DE/rand/1/bin is described in Figure 1 [15].

3. The Proposed Algorithm ADE

3.1. Generating Initial Population Using Orthogonal Design Method

Usually, the initial population $P^0 = \{\vec{x}_1^0, \vec{x}_2^0, ..., \vec{x}_N^0\}$ of evolutionary algorithms is randomly generated as follows:

$$\forall i \le N, \forall j \le D : x_{i,j}^0 = L_j + \mathrm{r}_j \times (U_j - L_j) \tag{5}$$

where N is the population size, D is the number of variables, r_j is a random number between 0 and 1, the jth variable of $\vec{x}_i^0$ is written as $x_{i,j}^0$, which is initialized in the range $[L_j, U_j]$. In order to improve the search efficiency, this paper employs orthogonal design method to generate the initial population, which can make some points closer to the global optimal point and improve the diversity of solutions. The orthogonal design method is described as follows [16]:

For any given individual $\vec{x} = [x_1, x_2, ..., x_D]$, the ith

1: Generate initial population $P^0 = \{\vec{x}_1^0, \vec{x}_2^0, ..., \vec{x}_N^0\}$
2: Let $t = 0$
3: **repeat**
4: **for** each individual $\vec{x}_i^t$ in the population P^t **do**
5: Generate three random integers r_1, r_2 and
6: $r_3 \in \{1,2,...,N\} \setminus i$, with $r_1 \ne r_2 \ne r_3$
7: Generate a random integer $j_{rand} \in \{1,2,...,D\}$
8: **for** each parameter j **do**
9: $u_{i,j}^{t+1} = \begin{cases} x_{r_3,j}^t + F \times (x_{r_1,j}^t - x_{r_2,j}^t), \\ \quad \text{if } (rand \le CR \,\|\, j = rand[1,D]) \\ x_{i,j}^t, \quad \text{otherwise} \end{cases}$
10: **end for**
11: Replace $\vec{x}_i^t$ with the child $\vec{u}_i^{t+1}$ in the population P^{t+1},
12: if $\vec{u}_i^{t+1}$ is better, otherwise $\vec{x}_i^t$ is retained
13: **end for**
14: $t = t + 1$
15: **until** the termination condition is achieved

Figure 1. Pseudocode of differential evolution based on scheme DE/rand/1/bin.

decision variable x_i varies in the range $[L_i, U_i]$. Here, each x_i is regarded as one factor of orthogonal design. Suppose that each factor holds Q levels, namely, quantize the domain $[L_i, U_i]$ into Q levels $\alpha_1, \alpha_2, ..., \alpha_Q$. The jth level of the ith factor is written as $\alpha_{i,j}$, which is defined as follows:

$$a_{i,j} = \begin{cases} L_i & , j = 1 \\ L_i + (j-1)(\frac{U_i - L_i}{Q-1}) & , 2 \le j \le Q-1 \\ U_i & , j = Q \end{cases} \quad (6)$$

And then, we create the orthogonal array $M = (b_{i,j})_{N \times D}$ with D factors and Q levels, where N is the number of level combinations. The procedure of constructing one orthogonal array $M = (b_{i,j})_{N \times D}$ is described in Figure 2.

Therefore, the initial population $P^0 = (x_{i,j}^0)_{N \times D}$ is generated by using the orthogonal array $M = (b_{i,j})_{N \times D}$, where the jth variable of individual $\vec{x}_i^0$ is $x_{i,j}^0 = a_{j,b_{i,j}}$.

3.2. Multi-Parent Mutation Scheme

According to the different variants of mutation, there are several different DE schemes often used, which are formulated as follows [12]:

"DE/rand/1/bin": $\vec{v}_i^{t+1} = \vec{x}_{r_1}^t + F \times (\vec{x}_{r_2}^t - \vec{x}_{r_3}^t)$ (7)

"DE/best/1/bin": $\vec{v}_i^{t+1} = \vec{x}_{best} + F \times (\vec{x}_{r_1}^t - \vec{x}_{r_2}^t)$ (8)

"DE/current to best/2/bin":

$$\vec{v}_i^{t+1} = \vec{x}_i^t + F \times (\vec{x}_{best} - \vec{x}_i^t) + F \times (\vec{x}_{r_1}^t - \vec{x}_{r_2}^t) \quad (9)$$

"DE/best/2/bin":

$$\vec{v}_i^{t+1} = \vec{x}_{best} + F \times (\vec{x}_{r_1}^t - \vec{x}_{r_2}^t) + F \times (\vec{x}_{r_3}^t - \vec{x}_{r_4}^t) \quad (10)$$

"DE/rand/2/bin":

$$\vec{v}_i^{t+1} = \vec{x}_{r_1}^t + F \times (\vec{x}_{r_2}^t - \vec{x}_{r_3}^t) + F \times (\vec{x}_{r_4}^t - \vec{x}_{r_5}^t) \quad (11)$$

1: **for** $(i = 1; i \le N; i++)$
2: $\{ b_{i,1} =$ **int**$((i-1)/Q)$ **mod** Q ; $b_{i,2} = (i-1)$ **mod** Q $\}$
3: **for** $(j = 3; j \le D; j++)$
4: **for** $(i = 1; i \le N; i++)$
5: $\{ b_{i,j} = (b_{i,1} \times (j-2) + b_{i,2})$ **mod** Q $\}$
6: Increment $b_{i,j}$ by one for $1 \le i \le N, 1 \le j \le D$

Figure 2. Procedure of constructing one orthogonal array $M = (b_j^i)_{N \times D}$.

where $\vec{x}_{best}$ is the best individual of the current population. Usually, based on both the control parameter *F* and the selected multiple parents, using these DE schemes can only generate a vector after a single mutation. Tsutsui *et al.* [17] proposed a multi-parent recombination with simplex crossover in real coded genetic algorithms to utilize the selected multiple parents and improve the diversity of offspring. Inspired by multi-parent recombination with simplex crossover, this paper proposes a novel multi-parent mutation in differential evolution. The multi-parent mutation is described in the following.

For each individual $\vec{x}_i^t$ from the population P^t with population size N, $i = 1, 2, ..., N$. A perturbed vector $\vec{v}_i^{t+1}$ is generated according to the following formula:

$$\vec{v}_i^{t+1} = \vec{x}_i^t + \sum_{k=1}^{K} w_k \times (\vec{x}_{r_k}^t - \vec{x}_{r_{k+1}}^t) \quad (12)$$

where $r_1, r_2, ..., r_K \in \{1, 2, ..., N\} \setminus i$, K randomly chosen integers are mutually different, and $\vec{x}_{r_{K+1}}^t = \vec{x}_{r_1}^t$. The weighted value w_k is defined as follows:

$$\vec{\xi} = randn(1, K), \vec{w} = \vec{\xi} / sum(\vec{\xi}) \quad (13)$$

where $randn(1, K)$ is a 1-by-K matrix with normally distributed random numbers, $sum(\vec{\xi})$ is used for calculating the sum of all components of the vector $\vec{\xi}$, and $\vec{w} = [w_1, w_2, ..., w_K]$.

According to the varying $\vec{w}$, repeat Formulas (13) and (12) for K times, K new vectors $\vec{v}_i^{t+1}\{1\}$, $\vec{v}_i^{t+1}\{2\}$, ..., $\vec{v}_i^{t+1}\{K\}$ are generated from these K selected parents. And then K vectors $\vec{x}_i^{t+1}\{1\}$, $\vec{x}_i^{t+1}\{2\}$, $\cdots$, $\vec{x}_i^{t+1}\{K\}$ are created by crossover, repair and constraint handling described in Subsections 3.3-3.5 respectively. Finally, an offspring individual $\vec{x}_i^{t+1}$ of the $(t+1)$th generation population P^{t+1} is obtained by selecting the best individual from these K offspring and their common parent $\vec{x}_i^t$.

3.3. Adaptive Crossover Rate *CR*

In conventional DE, the crossover rate CR is a constant value between 0 and 1. This paper proposes an adaptive crossover rate CR, which is defined as follows:

$$CR = CR_0 \times \exp(-a(\frac{t}{T})\wedge b) \quad (14)$$

where the initial crossover rate CR_0 is a constant value and usually is set to 0.8 or 0.85, t is the current genera-

tion number and T is the maximal generation number, b is a shape parameter determining the degree of dependency on the generation number, a and b are positive constants, usually a is set to 2, b is set to 2 or 3. At the early stage, DE uses a bigger crossover rate CR to preserve the diversity of solutions and prevent premature; at the later stage, DE employs a smaller crossover rate CR to enhance the local search and prevent the better solutions found from being destroyed.

3.4. Repair Method

After crossover, if one or more of the variables in the new vector $\vec{u}_i^{t+1}$ are beyond their boundaries, the violated variable value $\vec{u}_{i,j}^{t+1}$ is either reflected back from the violated boundary or set to the corresponding boundary value using the repair rule as follows [18,19]:

$$u_{i,j}^{t+1}=\begin{cases}\dfrac{L_j+u_{i,j}^{t+1}}{2}, & \text{if } (p\le 1/3)\wedge(u_{i,j}^{t+1}<L_j)\\ L_j, & \text{if } (1/3<p\le 2/3)\wedge(u_{i,j}^{t+1}<L_j)\\ 2L_j-u_{i,j}^{t+1}, & \text{if } (p>2/3)\wedge(u_{i,j}^{t+1}<L_j)\\ \dfrac{U_j+u_{i,j}^{t+1}}{2}, & \text{if } (p\le 1/3)\wedge(u_{i,j}^{t+1}>U_j)\\ U_j, & \text{if } (1/3<p\le 2/3)\wedge(u_{i,j}^{t+1}>U_j)\\ 2U_j-u_{i,j}^{t+1}, & \text{if } (p>2/3)\wedge(u_{i,j}^{t+1}>U_j)\end{cases} \quad (15)$$

where p is a probability and uniformly distributed random number in the range $[0,1]$.

3.5. Constraint Handling Technique of Feasibility-Based Rule

In evolutionary algorithms for solving constrained optimization problems, the most common method to handle constraints is to use penalty functions. In general, the constraint violation function of one individual $\vec{x}$ is transformed by m equality and inequality constraints as follows [4]:

$$G(\vec{x})=\sum_{j=1}^{q} w_j \max(0, g_j(\vec{x}))^{\beta}+\sum_{j=q+1}^{m} w_j \max(0,|h_j(\vec{x})|-\varepsilon)^{\beta} \quad (16)$$

where the exponent β is usually set to 1 or 2, ε is a tolerance allowed (a very small value) for the equality constraints and the coefficient w_j is greater than zero. If $\vec{x}$ is a feasible solution, $G(\vec{x})=0$, otherwise $G(\vec{x})>0$. The function value $G(\vec{x})$ shows that the degree of constraints violation of individual $\vec{x}$. β is set to 2 and w_j is set to 1 in this study.

In this study, a simple and efficient constraint handling technique of feasibility-based rule is introduced, which is also a constraint handling technique without using parameters. When two solutions are compared at a time, the following criteria are always applied [1]:

1) If one solution is feasible, and the other is infeasible, the feasible solution is preferred;

2) If both solutions are feasible, the one with the better objective function value is preferred;

3) If both solutions are infeasible, the one with smaller constraint violation function value is preferred.

3.6. Algorithm Framework

The general framework of the proposed algorithm ADE is described in Figure 3.

4. Experimental Study

4.1. Constrained Optimization Problems in Engineering Design

In order to validate the proposed algorithm ADE, we use six benchmark test problems, which are commonly used

1: Generate initial population $P^0=\{\vec{x}_1^0,\vec{x}_2^0,...,\vec{x}_N^0\}$ using
2: orthogonal design method, set CR_0 and let $t=0$
3: **repeat**
4: **for** each individual $\vec{x}_i^t$ in the population P^t **do**
5: Generate K random integers r_1, r_2, …, r_K
6: $\in\{1,2,...,N\}\setminus i$, they are also mutually different
7: **for** each $k\in\{1,2,...K\}$ **do**
8: Apply multi-parent mutation to generate new
9: vector $\vec{v}_i^{t+1}\{k\}$
10: **for** each parameter j **do**
11: $u_{i,j}^{t+1}\{k\}=\begin{cases}\vec{v}_{i,j}^{t+1}\{k\}, & \text{if } (rand\le CR \,\|\, j=rand[1,D])\\ x_{i,j}^t, & \text{otherwise}\end{cases}$
12: If $u_{i,j}^{t+1}$ is beyond its lower or upper boundaries,
13: repair rule is enforced
14: **end for**
15: **end for**
16: Find out the best one $\vec{u}_i^{t+1}$ of the children
17: $\{\vec{u}_i^{t+1}\{1\},\vec{u}_i^{t+1}\{2\},...,\vec{u}_i^{t+1}\{K\}\}$ /*apply the
18 feasibility–based rule */
19: Replace $\vec{x}_i^t$ with $\vec{u}_i^{t+1}$ in the population P^{t+1},
20: if $\vec{u}_i^{t+1}$ is better, otherwise $\vec{x}_i^t$ is retained
21: **end for**
22: $t=t+1$
23: **until** the termination condition is achieved

Figure 3. The general framework of the ADE algorithm.

in the specialized literature, and which are described in the following.

1) Three-bar truss design [8]:

Minimize $f(\vec{x}) = (2\sqrt{2}x_1 + x_2) \times l$

Subject to $g_1(\vec{x}) = \dfrac{\sqrt{2}x_1 + x_2}{\sqrt{2}x_1^2 + 2x_1x_2} P - \sigma \le 0$,

$$g_2(\vec{x}) = \frac{x_2}{\sqrt{2}x_1^2 + 2x_1x_2} P - \sigma \le 0,$$

$$g_3(\vec{x}) = \frac{1}{x_1 + \sqrt{2}x_2} P - \sigma \le 0$$

where $0 \le x_1 \le 1$ and $0 \le x_2 \le 1$; $l = 100\text{cm}$, $P = 2\text{KN/cm}^2$, and $\sigma = 2\text{KN/cm}^2$.

2) Spring design [8]:

Minimize $f(\vec{x}) = (x_3 + 2)x_1x_2^2$

Subject to $g_1(\vec{x}) = 1 - \dfrac{x_1^3 x_3}{71785x_2^4} \le 0$,

$$g_2(\vec{x}) = \frac{4x_1^2 - x_1x_2}{12566(x_1x_2^3 - x_2^4)} + \frac{1}{5108x_2^2} - 1 \le 0,$$

$$g_3(\vec{x}) = 1 - \frac{140.45x_2}{x_1^2 x_3} \le 0,$$

$$g_4(\vec{x}) = \frac{x_1 + x_2}{1.5} - 1 \le 0$$

where $0.25 \le x_1 \le 1.3$, $0.05 \le x_2 \le 2.0$, and $2 \le x_3 \le 15$.

3) Pressure vessel design [9,20]:

Minimize $f(\vec{x}) = 0.6224x_1x_3x_4 + 1.7781x_2x_3^2 + 3.1661x_1^2x_4 + 19.84x_1^2x_3$

Subject to $g_1(\vec{x}) = -x_1 + 0.0193x_3 \le 0$,

$$g_2(\vec{x}) = -x_2 + 0.00954x_3 \le 0,$$

$$g_3(\vec{x}) = -\pi x_3^2 x_4 - \frac{4}{3}\pi x_3^3 + 1{,}296{,}000 \le 0,$$

$$g_4(\vec{x}) = x_4 - 240 \le 0$$

where $x_1 = 0.0625n_1$, $x_2 = 0.0625n_2$, $1 \le n_1 \le 99$, $1 \le n_2 \le 99$, $10 \le x_3 \le 200$, $10 \le x_4 \le 200$.

4) Welded beam design [9]:

Minimize

$$f(\vec{x}) = 1.10471x_1^2x_2 + 0.04811x_3x_4(14.0 + x_2)$$

Subject to $g_1(\vec{x}) = \tau(x) - \tau_{\max} \le 0$,

$$g_2(\vec{x}) = \sigma(x) - \sigma_{\max} \le 0,$$

$$g_3(\vec{x}) = x_1 - x_4 \le 0$$

$$g_4(\vec{x}) = 0.10471x_1^2 + 0.04811x_3x_4(14.0 + x_2) - 5.0 \le 0,$$

$$g_5(\vec{x}) = 0.125 - x_1 \le 0$$

$$g_6(\vec{x}) = \delta(x) - \delta_{\max} \le 0,$$

$$g_7(\vec{x}) = P - P_c(x) \le 0$$

The other parameters are defined as follows:

$$\tau(\vec{x}) = \sqrt{(\tau')^2 + \frac{2\tau'\tau''x_2}{2R} + (\tau'')^2}, \quad \tau' = \frac{P}{\sqrt{2}x_1x_2}, \quad \tau'' = \frac{MR}{J},$$

$$M = P(L + \frac{x_2}{2}), \quad R = \sqrt{\frac{x_2^2}{4} + (\frac{x_1 + x_3}{2})^2},$$

$$J = 2\left\{\frac{x_1x_2}{\sqrt{2}}\left[\frac{x_2^2}{12} + \left(\frac{x_1 + x_3}{2}\right)^2\right]\right\}, \qquad \sigma(\vec{x}) = \frac{6PL}{x_4x_3^2},$$

$$\delta(\vec{x}) = \frac{4PL^3}{Ex_4x_3^3},$$

$$P_c(\vec{x}) = \frac{4.013\sqrt{EGx_3^2x_4^6/36}}{L^2}\left(1 - \frac{x_3}{2L}\sqrt{\frac{E}{4G}}\right),$$

where $P = 6000$ lb., $L = 14$ in, $\delta_{\max} = 0.25$ in., $E = 30 \times 10^6$ psi, $G = 12 \times 10^6$ psi, $\tau_{\max} = 13{,}600$ psi, $\sigma_{\max} = 30{,}000$ psi, $0.1 \le x_1 \le 2.0$, $0.1 \le x_2 \le 10.0$, $0.1 \le x_3 \le 10.0$, and $0.1 \le x_4 \le 2.0$.

5) Speed reducer design [8]:

Minimize

$$f(\vec{x}) = 0.7854x_1x_2^2(3.3333x_3^2 + 14.9334x_3 - 43.0934) - 1.508x_1(x_6^2 + x_7^2) + 7.4777(x_6^3 + x_7^2) + 0.7854(x_4x_6^2 + x_5x_7^2)$$

Subject to $g_1(\vec{x}) = \dfrac{27}{x_1x_2^2x_3} - 1 \le 0$,

$$g_2(\vec{x}) = \frac{397.5}{x_1x_2^2x_3^2} - 1 \le 0,$$

$$g_3(\vec{x}) = \frac{1.93x_4^3}{x_2x_3x_6^4} - 1 \le 0,$$

$$g_4(\vec{x}) = \frac{1.93x_5^3}{x_2x_3x_7^4} - 1 \le 0,$$

$$g_5(\vec{x}) = \frac{[(745x_4/(x_2x_3))^2 + 16.9 \times 10^6]^{1/2}}{110.0x_6^3} - 1 \le 0,$$

$$g_6(\vec{x}) = \frac{[(745x_5/(x_2x_3))^2 + 157.5 \times 10^6]^{1/2}}{85.0x_7^3} - 1 \le 0,$$

$$g_7(\vec{x}) = \frac{x_2x_3}{40} - 1 \le 0, \; g_8(\vec{x}) = \frac{5x_2}{x_1} - 1 \le 0,$$

$$g_9(\vec{x}) = \frac{x_1}{12x_2} - 1 \le 0,$$

$$g_{10}(\vec{x}) = \frac{1.5x_6 + 1.9}{x_4} - 1 \le 0,$$

$$g_{11}(\vec{x}) = \frac{1.1x_7 + 1.9}{x_5} - 1 \le 0$$

where $2.6 \le x_1 \le 3.6$, $0.7 \le x_2 \le 0.8$, $17 \le x_3 \le 28$, $7.3 \le x_4 \le 8.3$, $7.3 \le x_5 \le 8.3$, $2.9 \le x_6 \le 3.9$, $5.0 \le x_7 \le 5.5$.

6) Himmelblau's Nonlinear Optimization Problem [21]:

This problem was proposed by Himmelblau and similar to problem $g04$ [22] of the benchmark except for the second coefficient of the first constraint. There are five design variables. The problem can be stated as follows:

$$\text{Minimize } f(\vec{x}) = 5.3578547x_3^2 + 0.8356891x_1x_5 + 37.293239x_1 - 40792.141$$

$$\text{Subject to } g_1(\vec{x}) = 85.334407 + 0.0056858x_2x_5 + 0.00026x_1x_4 - 0.0022053x_3x_5, \quad -92 \le 0$$

$$g_2(\vec{x}) = -85.334407 - 0.0056858x_2x_5 - 0.00026x_1x_4 + 0.0022053x_3x_5 \le 0,$$

$$g_3(\vec{x}) = 80.51249 + 0.0071317x_2x_5 + 0.0029955x_1x_2 + 0.0021813x_3^2, \quad -110 \le 0$$

$$g_4(\vec{x}) = -80.51249 - 0.0071317x_2x_5 - 0.0029955x_1x_2 - 0.0021813x_3^2, \quad +90 \le 0$$

$$g_5(\vec{x}) = 9.300961 + 0.0047026x_3x_5 + 0.0012547x_1x_3 + 0.0019085x_3x_4, \quad -25 \le 0$$

$$g_6(\vec{x}) = -9.300961 - 0.0047026x_3x_5 - 0.0012547x_1x_3 - 0.0019085x_3x_4, \quad +20 \le 0$$

where $78 \le x_1 \le 102$, $33 \le x_2 \le 45$, and $27 \le x_i \le 45$ $(i = 3,4,5)$.

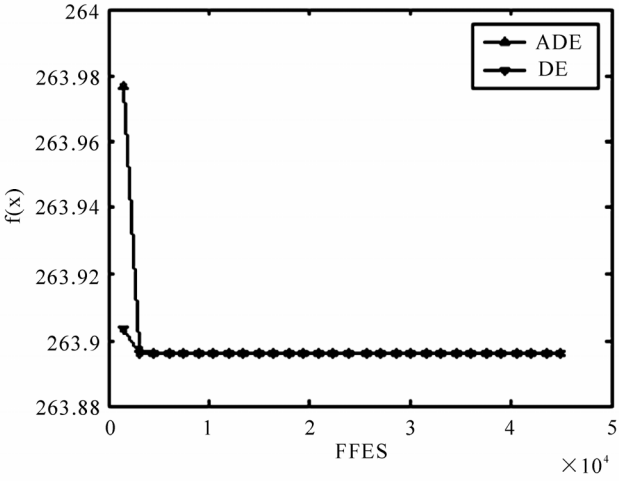


Figure 4. Convergence graph for three-bar truss design.

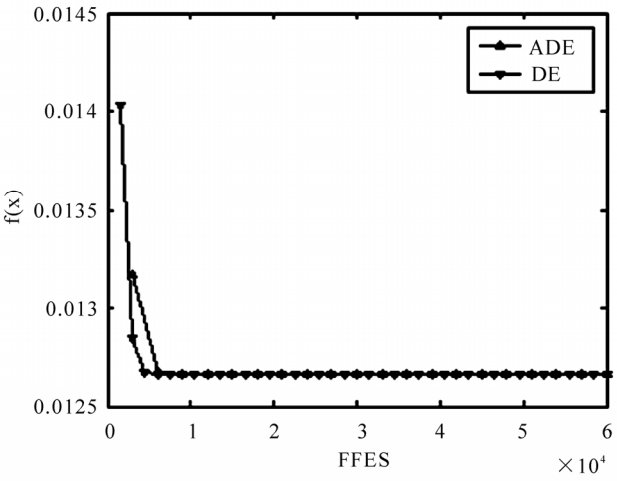


Figure 5. Convergence graph for spring design.

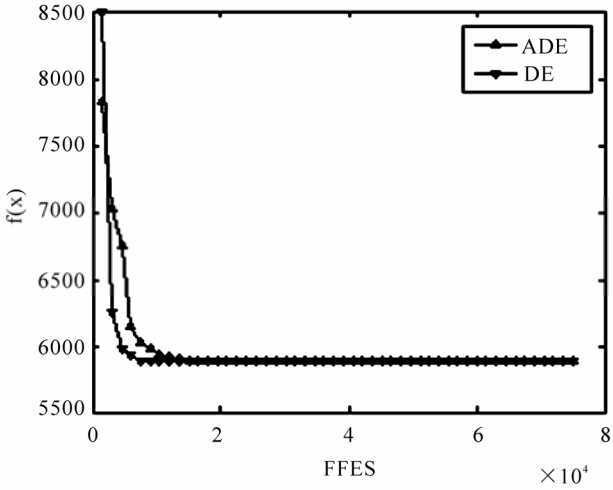


Figure 6. Convergence graph for pressure vessel design.

4.2. Convergence of ADE

In this section, Figures 4-9 depict the convergence graphs for 6 engineering optimization problems described above respectively. From Figures 4-6, we know that ADE and DE all can be quickly convergent. In the figures, FFES is the number of fitness function evaluations.

4.3. Comparing ADE with Respect to Some State-of-the-Art Algorithms

In this experimental study, the parameter values used in ADE are set as follows: the population size $N = 50$, the maximal generation number $T = 300$, the level number $Q = \lfloor \sqrt{N} \rfloor$, the mutation parent number $K = D + 1$, the

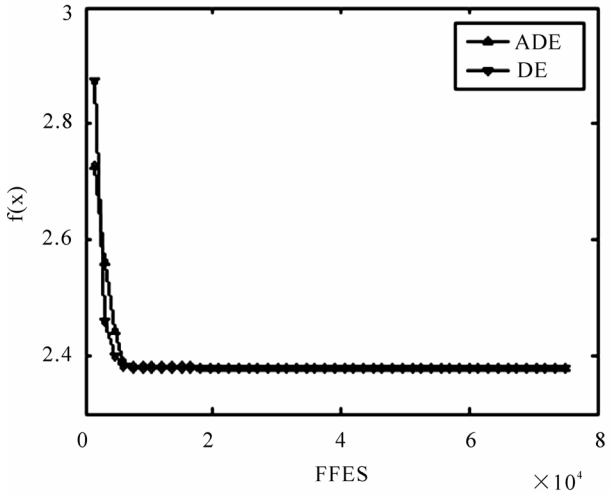


Figure 7. Convergence graph for welded beam design.

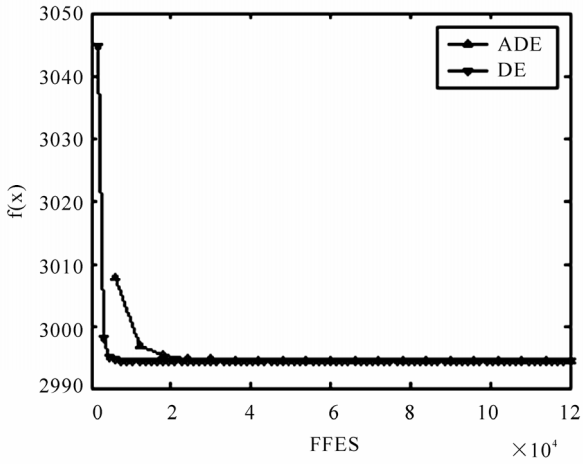


Figure 8. Convergence graph for speed reducer design.

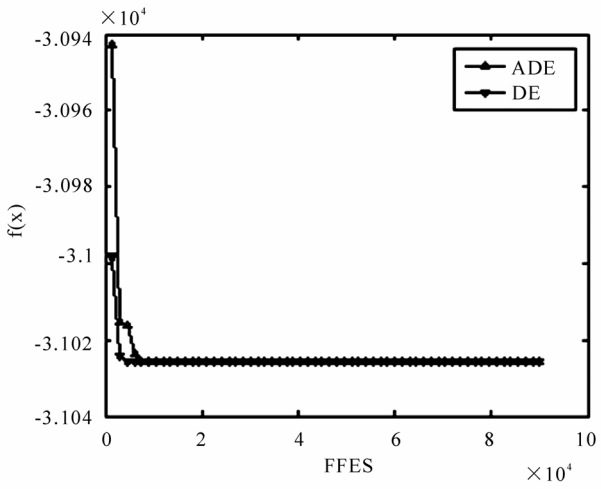


Figure 9. Convergence graph for Himmelblau's nonlinear optimization problem.

initial crossover rate $CR_0 = 0.8$, the coefficient $a = 2$, the shape parameter $b = 3$, the exponent $\beta = 2$. The number of fitness function evaluations (FFES) is equal to $N \times T \times K$. The achieved solution at the end of $N \times T \times K$ FFES is used to measure the performance of ADE. ADE is independently run 30 times on each test problem above. The optimized objective function values (of 30 runs) arranged in ascending order and the 15th value in the list is called the median optimized function value. Experimental results are presented in Tables 1-12. And NA is the abbreviation for "Not Available".

For three-bar truss design problem, the experimental results are given in Tables 1-2. According to Table 1, ADE and DSS-MDE [8] can obtain the approximate best and median values, which are slightly better than those obtained by Ray

Table 1. Comparison of statistical results for three-bar truss design over 30 runs.

Algorithms	Best	Median	Mean	Worst	Std	FFES
ADE	263.89584338	263.89584338	263.89584338	263.89584338	4.72e-014	45,000
DSS-MDE [8]	263.8958434	263.8958434	263.8958436	263.8958498	9.72e-07	15,000
Ray and Liew [6]	263.8958466	263.8989	263.9033	263.96975	1.26e-02	17.610

Table 2. Comparison of best solutions found for three-bar truss design.

Function	ADE	DSS-MDE [8]	Ray and Liew [6]	ECT [23]	Ray and Saini [24]
x_1	0.7886751376014	0.7886751359	0.7886210370	0.78976441	0.795
x_2	0.4082482819599	0.4082482868	0.4084013340	0.40517605	0.395
$f(x)$	263.895843376	263.8958434	263.8958466	263.896710000	264.300
FFES	45,000	15,000	17,610	55,000	2712

Table 3. Comparison of statistical results for spring design over 30 runs.

Algorithms	Best	Median	Mean	Worst	Std	FFES
ADE	0.0126652328	0.0126652458	0.0129336018	0.02064372078	1.46e-03	60,000
SiC-PSO [20]	0.012665	NA	0.0131	NA	4.1e-04	24,000
FSA [25]	0.012665285	NA	0.012665299	0.012665338	2.2e-08	49,531
DSS-MDE [8]	0.012665233	0.012665304	0.012669366	0.012738262	1.25e-05	24,000
Ray and Liew [6]	0.01266924934	0.012922669	0.012922669	0.016717272	5.92e-04	25,167
Coello [26]	0.01270478	0.01275576	0.01276920	0.01282208	NA	900,000

Table 4. Comparison of best solutions found for spring design.

Function	ADE	SiC-PSO [20]	DSS-MDE [8]	FSA [25]	He *et al*. [7]
x_1	0.35674653865	0.354190	0.3567177469	0.35800478345599	0.356750
x_2	0.05169025814	0.051583	0.0516890614	0.05174250340926	0.051690
x_3	11.28727756428	11.438675	11.2889653382	11.21390736278739	11.287126
$f(x)$	0.0126652328	0.012665	0.01265233	0.012665285	0.012665
FFES	60,000	24,000	24,00	49,531	15,000

Table 5. Comparison of statistical results for pressure vessel design over 30 runs.

Algorithms	Best	Median	Mean	Worst	Std	FFES
ADE	5885.3327736	5885.3327785	5885.3349564	5885.3769425	8.66e-03	75,000
SiC-PSO [20]	6059.714335	NA	6092.0498	NA	12.1725	24,000
Ray and Liew [6]	6171.00	NA	6335.05	NA	NA	20,000
He *et al.* [7]	6059.714	NA	6289.929	NA	3.1e+2	30,000
Montes *et al.* [3]	6059.702	6059.702	6059.702	6059.702	1.0e-12	24,000

and Liew [6] respectively. The mean and worst values obtained by ADE are the best among three algorithms, while the FFES (45,000) of ADE is also the highest. And we also find that these algorithms can find the near-optimal solutions. From Table 2, we can see that ADE can find the best value when compared with respect to DSS-MDE [8], Ray and Liew [6], ECT [22] and Ray and Saini [23]. The best result obtained by ADE is

$$f(\vec{x}) = 263.8958433764684,$$

corresponding to

$$\vec{x} = [x_1, x_2] = [0.78867513760142, 0.40824828195990]$$

and constraints

$$[g_1(\vec{x}),\ g_2(\vec{x}),\ g_3(\vec{x})\ ]$$

$$= [0, -1.46410162480516, -0.53589837519484].$$

For spring design problem, the experimental results are given in Tables 3-4. According to Table 3, ADE, Sic-PSO [20], FSA [24], DSS-MDE [8] can find out the best value when compared with respect to Ray and Liew [6] and Coello [25]. The median value obtained by ADE is better than obtained by other methods, but the mean and worst values are worse, this is because that ADE can only find 29 near-optimal solutions in 30 runs and the other is an exception solution (i.e., the worst value is 0.02064372078). Table 4 presents the detail of each best value obtained by ADE, SiC-PSO [20], DSS-MDE [8], FSA [24] or He *et al.* [7] respectively. The best result obtained by ADE is

$$f(\vec{x}) = 0.01266523278832,$$

corresponding to

$$\vec{x} = [x_1, x_2, x_3]$$

$$= [0.35671785021031, 0.05168906567225, 11.28895927857073]$$

and constraints

$[g_1(\vec{x}), g_2(\vec{x}), g_3(\vec{x}), g_4(\vec{x})]$
=[-2.220446049250313e-016, -4.4408 92098500626e-016, 4.05378584839796, -0.72772872274496].

For pressure vessel design problem, the experimental results are given in Tables 5-6. According to Table 5, the best, median, mean, worst and standard deviation of values obtained by ADE are the best when compared with respect to Sic-PSO [20], Ray and Liew [6], He *et al.* [7], and Montes *et al.* [3], while the FFES (75,000) of ADE is also the highest. Table 6 presents the detail of each best value obtained by ADE, SiC-PSO [20], Ray and Liew [6], He *et al.* [7] or Montes *et al.* [3] respectively. The best result obtained by ADE is

$$f(\vec{x}) = 5885.332773616458,$$

corresponding to

$$\vec{x} = [x_1, x_2, x_3, x_4]$$
$$= [0.778168641375, 0.384649162628, 40.319618724099, 200]$$

and constraints

$[g_1(\vec{x}), g_2(\vec{x}), g_3(\vec{x}), g_4(\vec{x})]$
=[-1.110223024625157e-016,0,0,-40].

For welded beam design problem, the experimental results are provided with Tables 7-8. According to Table 7, the best, median, mean, worst and standard derivation of values obtained by ADE are slightly worse than those obtained by DSS-MDE [8] and are better than those obtained by Ray and Liew [6], FSA [25] and Deb [1]. However, the FFES (75,000) of ADE is the highest. Table 8 presents the detail of each best value obtained by DSS-MDE [8], He *et al.* [7], FSA [25], Ray and Liew [6], and Akhtar *et al.* [9] respectively. The best result obtained by ADE is

$$f(\vec{x}) = 2.3809565\ 8032252,$$

corresponding to

$$\vec{x} = [x_1, x_2, x_3, x_4]$$
$$= [0.24436897580173, 6.21751971517460, 8.29147139048\ 684, 0.24436897580173]$$

and constraints

$[g_1(\vec{x}), g_2(\vec{x}), g_3(\vec{x}), g_4(\vec{x}), g_5(\vec{x}), g_6(\vec{x}), g_7(\vec{x})]$
=[-1.091393642127514e-011, -3.310560714453459e-010, -1.387778780781446e-016, -3.02295458760400, -0.11936897580173, -0.23424083488769, -1.273292582482100e-011].

Table 6. Comparison of best solutions found for pressure vessel design.

Function	ADE	Sic-PSO [20]	Ray and Liew [6]	He *et al.* [7]	Montes *et al.* [3]
x_1	0.7781686414	0.812500	0.8125	0.8125	0.8125
x_2	0.3846491626	0.437500	0.4375	0.4375	0.4375
x_3	40.319618724	42.098445	41.9768	42.098446	42.098446
x_4	200	176.636595	182.9768	176.636052	176.636047
$f(x)$	5885.3327736	6059.714335	6171.0	6059.7143	6059.701660
FFES	75,000	24,000	20,000	30,000	24,000

Table 7. Comparison of statistical results for welded beam design over 30 runs.

Algorithms	Best	Median	Mean	Worst	Std	FFES
ADE	2.380956580	2.380956580	2.380956585	2.380956708	2.35e-08	75,000
DSS-MDE [8]	2.38095658	2.38095658	2.38095658	2.38095658	3.19e-10	24,000
Ray and Liew [6]	2.3854347	3.2551371	3.0025883	6.3996785	0.959078	33,095
FSA [25]	2.381065	NA	2.404166	2.488967	NA	56.243
Deb [1]	2.38119	2.39289	NA	2.64583	NA	40,080

Table 8. Comparison of best solutions found for welded beam design.

Function	ADE	DSS-MDE [8]	He *et al.* [7]	FSA [25]	Ray and Liew [6]	Akhtar *et al.* [9]
x_1	0.24436897580	0.2443689758	0.244369	0.24435257	0.244438276	0.2407
x_2	6.21751971517	6.2175197152	6.217520	6.2157922	6.237967234	6.4851
x_3	8.29147139049	8.2914713905	8.291471	8.2939046	8.288576143	8.2399
x_4	0.24436897580	0.2443689758	0.244369	0.24435258	0.244566182	0.2497
$f(x)$	2.38095658032	2.38095658	2.380957	2.381065	2.3854347	2.4426
FFES	75,000	24,000	30,000	56,243	33,095	19,259

Table 9. Comparison of statistical results for speed reducer design over 30 runs.

Algorithms	Best	Median	Mean	Worst	Std	FFES
ADE	2994.4710662	2994.4710662	2994.4710662	2994.4710662	1.85e-012	120,000
DSS-MDE [8]	2994.471066	2994.471066	2994.471066	2994.471066	3.58e-012	30,000
Ray and Liew [6]	2994.744241	3001.758264	3001.7582264	3009.964736	4.0091423	54,456
Montes *et al.* [27]	2996.356689	NA	2996.367220	NA	8.2e-03	24,000
Akhtar *et al.* [9]	3008.08	NA	3012.12	3028	NA	19,154

Table 10. Comparison of best solutions found for speed reducer design.

Function	ADE	DSS-MDE [8]	Ray and Liew [6]	Montes *et al.* [27]	Akhtar *et al.*[9]
x_1	3.5	3.5	3.50000681	3.500010	3.506122
x_2	0.7	0.7	0.70000001	0.7	0.700006
x_3	17	17	17	17	17
x_4	7.3	7.3	7.32760205	7.300156	7.549126
x_5	7.715319911478	7.7153199115	7.71532175	7.800027	7.859330
x_6	3.350214666096	3.3502146661	3.35026702	3.350221	3.365576
x_7	5.286654464980	5.2866544650	5.28665450	5.286685	5.289773
$f(x)$	2994.4710662	2994.471066	2994.744241	2996.356689	3008.08
FFES	120,000	30,000	54,456	24,000	18,154

Table 11. Comparison of statistical results for himmelblau's nonlinear optimization problem.

Algorithms	Best	Median	Mean	Worst	Std	FFES
ADE	-31025.56024	-31025.56024	-31025.56024	-31025.56024	5.91e-010	90,000
COPSO [28]	-31025.56024	NA	-31025.56024	NA	0	200,000
HU-PSO [29]	-31025.56142	NA	-31025.56142	NA	0	200,000

Table 12. Comparison of best solutions found for himmelblau's nonlinear optimization problem.

Function	ADE	COPSO [28]	HU-PSO [29]	Colleo [21]	Homaifar *et al.* [30]
x_1	78.00000000000000	78	78.0	78.0495	78.0000
x_2	33.00000000000000	33	33.0	33.0070	33.0000
x_3	27.07099710517604	27.070997	27.070997	27.0810	29.9950
x_4	45.00000000000000	45	45.0	45.0000	45.0000
x_5	44.96924255010549	44.969242	44.96924255	44.9400	36.7760
$f(x)$	-31025.56024249794	-31025.56024	-31025.56142	-31020.859	-30665.609
FFES	90,000	200,000	200,000	NA	NA

For speed reducer design problem, the experimental results are given in Tables 9-10. According to Table 9, the best, median, mean, worst and standard derivation of values obtained by ADE and DSS-MDE [8] are superior to those obtained by Ray and Liew [6], Montes *et al.* [27] and Akhtar *et al.* [9] respectively, while the FFES (120,000) of ADE is the highest. Table 10 shows the detail of each best value obtained by ADE, DSS-MDE [8], Ray and Liew [6], Montes *et al.* [27] and Akhtar *et al.* [9] respectively. The best result obtained by ADE is

$$f(\vec{x}) = 2994.47106614682020,$$

corresponding to

$$\vec{x} = [x_1, x_2, x_3, x_4, x_5, x_6, x_7]$$
$$= [3.5, 0.7, 17, 7.3, 7.71531991147825, 3.35021466609645, 5.28665446498022]$$

and constraints

$$[g_1(\vec{x}), g_2(\vec{x}), g_3(\vec{x}), g_4(\vec{x}), g_5(\vec{x}), g_6(\vec{x}), g_7(\vec{x}), g_8(\vec{x}), g_9(\vec{x}), g_{10}(\vec{x}), g_{11}(\vec{x})]$$
$$=[-0.07391528039787, -0.19799852714195, -0.49917224810242, -0.90464390455607, -6.661338147750939e-016, 0, -0.70250000000000, -2.220446049250313e-016, -0.58333333333333, -0.05132575354183, -8.881784197001252e-016].$$

For Himmelblau's nonlinear optimization problem, the best, median, mean, worst and standard derivation of values is shown in Tables 11-12, it is clearly seen that ADE, COPSO [28], and HU-PSO [29] all can find one near-optimal solution after a single run. Additionally, ADE only requires 90,000 FFES, which is superior to other several algorithms, such as COPSO [28] 200,000 FFES and HU-PSO [29] 200,000 FFES. The best result obtained by ADE is

$$f(\vec{x}) = -3.1025.56024249794,$$

corresponding to

$$\vec{x} = [x_1, x_2, x_3, x_4, x_5]$$
$$= [78, 33, 27.07099710517604, 45, 44.96924255010549]$$

and constraints

$$[g_1(\vec{x}), g_2(\vec{x}), g_3(\vec{x}), g_4(\vec{x}), g_5(\vec{x}), g_6(\vec{x})]$$
$$=[0, -92, -9.59476568762383, -10.40523431237617, -5, 0].$$

In sum, compared with respect to several state-of-the-art algorithms, ADE can perform better on six benchmark test problems. It is clearly shown that ADE is feasible and effective to solve constrained optimization problems in engineering design. The reason is that ADE uses multi-parent mutation to generate a better offspring, and applies self-adaptive control parameter and effective repair rule etc.

5. Conclusions and Future Work

This paper proposes an adaptive differential evolution (ADE) algorithm for constrained optimization in Engineering Design. Firstly, ADE employs the orthogonal design method to generate the initial population to improve the diversity of solutions. Secondly, a multi-parent mutation scheme is developed to improve the capacity of exploration and the convergence speed of ADE. Thirdly, in order to improve the adaptive capacity of crossover operator, a new approach to adjusting the crossover rate is presented. In addition, ADE introduces a new repair rule and a constraint handling technique of the feasible-based rule is also applied when comparing two solutions at a time. Finally, ADE is tested on six constrained engineering design optimization problems taken from the specialized literature. Compared with respect to several state-of-the-art algorithms, the experimental results show that ADE is highly competitive and can obtain good results in terms of a test set of constrained optimization

problems in engineering design. However, there are still some things to do in the future. Firstly, we will further validate ADE in the case of higher dimensions. Secondly, we also will take some measures to improve the convergence speed during the evolutionary process. Additionally, testing some initial parameters of ADE is another future work.

6. References

[1] K. Deb, "An efficient constraint handling method for genetic algorithms," Computer Methods in Applied Mechanics and Engineering, Vol. 186, No. 2, pp. 311–338, 2000.

[2] E. Mezura-Montes and A. G. Palomeque-Qrtiz, "Parameter control in differential evolution for constrained optimization," 2009 IEEE Congress on Evolutionary Computation (CEC'2009), pp. 1375–1382, 2009.

[3] E. Mezura-Montes, C. A. Coello Coello, J. Velázquez-Reyes, and L. Muñoz-Dávila, "Multiple trial vectors in differential evolution for engineering design," Engineering Optimization, Vol. 39, No. 5, pp. 567–589, 2007.

[4] C. A. Coello Coello, "Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: A survey of the state of the art," Computer Methods in Applied Mechanics and Engineering, Vol. 191, No. 11-12, pp. 1245–1287, 2002.

[5] R. Landa-Becerra and C. A. Coello Coello, "Cultured differential evolution for constrained optimization," Computer Methods in Applied Mechanics and Engineering, Vol. 195, No. 33–36, pp. 4303–4322, 2006.

[6] T. Ray and K. M. Liew, "Society and civilization: An optimization algorithm based on the simulation of social behavior," IEEE Transactions on Evolutionary Computation, Vol. 7, No. 4, pp. 386–396, 2003.

[7] S. He, E. Prempain, and Q. H. Wu, "An improved particle swarm optimizer for mechanical design optimization problems," Engineering Optimization, Vol. 36, No. 5, pp. 585–605, 2004.

[8] M. Zhang, W. J. Luo, and X. F. Wang, "Differential evolution with dynamic stochastic selection for constrained optimization," Information Sciences, Vol. 178, pp. 3043–3074, 2008.

[9] S. Akhtar, K. Tai, and T. Ray, "A socio-behavioural simulation model for engineering design optimization," Engineering Optimization, Vol. 34, No. 4, pp. 341–354, 2002.

[10] Q. He and L. Wang, "An effective co-evolutionary particle swarm optimization for constrained engineering design problems," Engineering Applications of Artificial Intelligence, Vol. 20, No. 1, pp. 89–99, 2007.

[11] J. H. Wang and Z. Y. Yin, "A ranking selection-based particle swarm optimizer for engineering design optimization problems," Structural and Multidisciplinary Optimization, Vol. 37, No. 2, pp. 131–147, 2008.

[12] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," Journal of Global Optimization, Vol. 11, pp. 341–359, 1997.

[13] K. Price, R. Storn, and J. Lampinen, "Differential evolution: A practical approach to global optimization," Berlin: Springer-Verlag, 2005.

[14] Z. Y. Yang, K. Tang, and X. Yao, "Self-adaptive differential evolution with neighborhood search," 2008 Congress on Evolutionary Computation (CEC'2008), pp. 1110–1116, 2008.

[15] H. A. Abbass, R. Sarker, and C. Newton, "PDE: A Pareto-frontier differential evolution approach for multiobjective optimization problems," Proceedings of IEEE Congress on Evolutionary Computation, Vol. 2, pp. 971–978, 2001.

[16] Y. W. Leung and Y. P. Wang, "An orthogonal genetic algorithm with quantization for global numerical optimization," IEEE Transactions on Evolutionary Computation, Vol. 5, No. 1, pp. 40–53, 2001.

[17] S. Tsutsui, M. Yamamure, and T. Higuchi, "Multi-parent recombination with simplex crossover in real coded genetic algorithms," Proceedings of the Genetic and Evolutionary Computation Conference, pp. 657–664, 1999.

[18] J. Brest, V. Zumer, and M. S. Maucec, "Self-adaptive differential evolution algorithm in constrained real-parameter optimization," 2006 IEEE Congress on Evolutionary Computation (CEC'2006), pp. 919–926, 2006.

[19] Y. Wang and Z. X. Cai, "A hybrid multi-swarm particle swarm optimization to solve constrained optimization problems," Frontiers of Computer Science in China, Vol. 3, No. 1, pp. 38–52, 2009.

[20] L. C. Cagnina, S. C. Esquivel, and C. A. Coello Coello, "Solving engineering optimization problems with the simple constrained particle swarm optimizer," Informatica, Vol. 32, pp. 319–326, 2008.

[21] C. A. Coello Coello, "Use of a self-adaptive penalty approach for engineering optimization problems," Computers in Industry, Vol. 41, No. 2, pp. 113–127, 2000.

[22] T. P. Runarsson and X. Yao, "Stochastic ranking for constrained evolutionary optimization," IEEE Transactions on Evolutionary Computation, Vol. 4, No. 3, pp. 284–294, 2000.

[23] K. Hans Raj, R. S. Sharma, G. S. Mishra, A. Dua, and C. Patvardhan, "An evolutionary computational technique for constrained optimisation in engineering design," Journal of the Institution of Engineers India Part Me Mechanical Engineering Division, Vol. 86, pp. 121–128, 2005.

[24] T. Ray and P. Saini, "Engineering design optimization using a swarm with intelligent information sharing among individuals," Engineering Optimization, Vol. 33, No. 33, pp. 735–748, 2001.

[25] A. R. Hedar and M. Fukushima, "Derivative-free filter simulated annealing method for constrained continuous global optimization," Journal of Global Optimization, Vol. 35, No. 4, pp. 521–549, 2006.

[26] C. A. Coello, "Self-adaptive penalties for GA- based optimization," Proceedings of the Congress on Evolutionary Computation 1999 (CEC'99), Vol. 1, pp. 573–580, 1999.

[27] E. Mezura-Montes, C. A. Coello, and J. V. Reyes, "Increasing successful offspring and diversity in differential evolution for engineering design," Proceedings of the Seventh International Conference on Adaptive Computing in Design and Manufacture (ACDM 2006), pp. 131–139, 2006.

[28] A. E. Muñoz Zavala, A. Hernández Aguirre, E. R. Villa Diharce, and S. Botello Rionda, "Constrained optimization with an improved particle swarm optimization algorithm," International Journal of Intelligent Computing and Cybernetics, Vol. 1, No. 3, pp. 425–453, 2008.

[29] X. H. Hu, R. C. Eberhart, and Y. H. Shi, "Engineering optimization with particle swarm," Proceedings of the 2003 IEEE Swarm Intelligence Symposium, pp. 53–57, 2003.

[30] A. Homaifar, S. H. Y. Lai, and X. Qi, "Constrained optimization via genetic algorithms," Simulation, Vol. 62, No. 4, pp. 242–254, 1994.

Engineering, **2010, 2, 78-84**
doi:10.4236/eng.2010.21010 Published Online January 2010 (http://www.SciRP.org/journal/eng/).

`

Water Film in Saturated Sand

Xiaobing LU[1], X. H. ZHANG[1], Cui PENG[2]
[1]*Institute of Mechanics, Chinese Academy of Sciences, Beijing, China*
[2]*Institute of Mountain Hazard and Environment, Chinese Academy of Sciences, Chengdu, China*
Email: *xblu@imech.ac.cn*
Received June 16, 2009; *revised July* 22, 2009; *accepted August* 4, 2009

Abstract

In order to investigate the initiation mechanism of landslide and debris flow occurring on a gentle slope. The formation and evolution of water film (or crack) in saturated sand is analyzed by numerical and theoretical simulations under given conditions. First a psudo-three-phase model is presented considering the movement of skeleton and water and the erosion. Secondly, difference method is used to analyze the formation conditions and the evolution of the velocity of water and pore pressure and porosity. Thirdly, a simplified theoretical method is presented based on the consolidation theory to analyze the initiation, expansion and close. It is shown that there are stable water films when some point is blocked and the state keeps unchangeable or there exists a thin layer with very low permeability. Once the blocked point is open, the water film will disappear gradually. The evolution of water film may be calculated by a simplified method. The analytical results are agreement well with that of Kokusho.

Keywords: Water Film, Saturated Sand, Liquefaction, Percolation, Erosion

1. Introduction

It is often observed that sand deposit on slope spreads laterally or even turns into landslide or debris flow, especially after earthquakes. When the deposit is multilayered, a water film will form once it is liquefied [1], which may serve as a sliding surface for the postliquefaction failure. As a result, landslide or debris flow may happen on very gentle slope. Seed [2] was the first to suggest that the existence of "water film" in sand bed is the reason of slope failures in earthquakes. Later, some researchers [3–5] performed some experiments to investigate the formation of "water film" in layered sand or in sand containing a seam of non-plastic silt. However, the mechanism of cracks or "water film" in sand with porosity distributed continuously is not very clear.

A theoretical and numerical analysis is presented in this paper. Firstly, we present a pseudo-three-phase model describing the moving of liquefied sand and give some theoretical analysis. Then we give the numerical simulations under four initial and boundary conditions. Secondly, we present a simplified method to analyze the evolution of the water film.

2. Formulation of the Problem

Consider a saturated horizontal sand stratum, with the porosity changing only vertically. The fine grains may be eroded from the skeleton and the eroding relation is assumed as follows [6–7]

$$\frac{1}{\rho_s}\left(\frac{\partial Q}{\partial t}+u_s\frac{\partial Q}{\partial x}\right)=\frac{\lambda}{T}\left(\frac{u-u_s}{u^*}-q\right)$$

$$\text{if}\quad -\varepsilon(x,0)\le\frac{Q}{\rho_s}\le\frac{Q_c(x)}{\rho_s} \tag{1}$$

$$\frac{1}{\rho_s}\left(\frac{\partial Q}{\partial t}+u_s\frac{\partial Q}{\partial x}\right)\le 0\quad\text{otherwise, including}\quad u=0 \tag{2}$$

where the x axis is upward, Q is the mass of sand eroded per unit volume of the sand/water mixture, ρ_s is the density of the grains, u and u_s are the velocities of the percolating fluid containing fine sand particles and the sand grains, q is the volume fraction of sand carried in the percolating fluid, T and u^* are physical parameters, λ is a small dimensionless parameter, $\varepsilon(x,t)$ is the porosity, $Q_c(x)$ is the maximum Q that can be eroded at x.

Considering the eroding of the fine grains, a psudo-three-phase model is presented here. The mass conservation equations are:

$$\frac{\partial(\varepsilon-q)\rho}{\partial t}+\frac{\partial(\varepsilon-q)\rho u}{\partial x}=0 \tag{3}$$

$$\frac{\partial q\rho_s}{\partial t}+\frac{\partial q\rho_s u}{\partial x}=\frac{\partial Q}{\partial t}+u_s\frac{\partial Q}{\partial x} \tag{4}$$

$$\frac{\partial(1-\varepsilon)\rho_s}{\partial t}+\frac{\partial(1-\varepsilon)\rho_s u_s}{\partial x}=-\frac{\partial Q}{\partial t}-u_s\frac{\partial Q}{\partial x} \tag{5}$$

in which ρ is the density of water. From these follows a general equation:

$$\varepsilon u+(1-\varepsilon)u_s=U(t) \tag{6}$$

in which $U(t)$ is the total mass of fluid and grains at a transect. The momentum equations may be written as

$$[(\varepsilon-q)\rho+q\rho_s]\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}\right)=-\varepsilon\frac{\partial p}{\partial x}-\frac{\varepsilon^2(u-u_s)}{k(\varepsilon,q)}-[(\varepsilon-q)\rho+q\rho_s]g \tag{7}$$

$$\begin{aligned}&[(\varepsilon-q)\rho+q\rho_s]\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}\right)+(1-\varepsilon)\rho_s\left(\frac{\partial u_s}{\partial t}+u_s\frac{\partial u_s}{\partial x}\right)\\&=-\frac{\partial p}{\partial x}-\frac{\partial\sigma_e}{\partial x}-[(\varepsilon-q)\rho+q\rho_s]g-(1-\varepsilon)\rho_s g\\&\quad-\left(\frac{\partial Q}{\partial t}+u_s\frac{\partial Q}{\partial x}\right)(u-u_s)\end{aligned} \tag{8}$$

in which Equation (7) denotes the momentum conservation of grains, Equation (8) denotes the total momentum conservation, the last term on the right hand side of Equation (8) denotes the momentum caused by the eroded fine grains, p is the pore pressure, k is the permeability, σ_e is the effective stress, θ is the slope. Here k is assumed to be a function of ε and q in the following form

$$k(\varepsilon,q)=k_0 f(q,\varepsilon)=k_0(-\alpha q+\beta\varepsilon) \tag{9}$$

in which α,β are parameters and $1<\beta<<\alpha$, we choose to let α much greater than β, so that changes in q overweighs that of ε.

3. Numerical Simulations

Based on the model presented above, we will analyze the occurrence of the crack in saturated sand, This case is about the cracks in a liquefied sand ($\sigma_e=0,\tau\approx 0$) where the grains sink while the water is pressed to move upward just like the consolidation. Here, the sand column is assumed to be long enough to neglect the boundary effects. At the same time, we neglect some factors which may be important in other cases. We will simulate this problem by different methods.

Being an appropriate constant, the mass conservation Equation (6) yield

$$\varepsilon u+(1-\varepsilon)u_s=U(t)=0 \tag{10}$$

assuming both u and u_s to be zero at x=0.

Let T be the appropriate characteristic time in Equation (10), u_t the characteristic velocity and L the characteristic length, and let

$$\bar{u}=\frac{u}{u_t},\ \tau=\frac{t}{T},\ \xi=\frac{x}{Tu_t} \tag{11}$$

Then Equations (3) and (4) can be rewritten in the non- dimensional form:

$$\begin{aligned}\frac{\partial\varepsilon}{\partial\tau}+\frac{\partial\varepsilon\bar{u}}{\partial\xi}&=\bar{u}\frac{u_t}{u^*(1-\varepsilon)}-q\\\frac{\partial q}{\partial\tau}+\frac{\partial q\bar{u}}{\partial\xi}&=\bar{u}\frac{u_t}{u^*(1-\varepsilon)}-q\end{aligned} \tag{12}$$

For $Tg/u_t>>1$, the inertia terms are negligible and the last equation of Equation (8) becomes

$$\begin{aligned}\bar{u}&=\left(\frac{1-\varepsilon}{\varepsilon}\right)^2(\varepsilon-q)f(q,\varepsilon)\frac{k_0\rho_s g(1-\rho/\rho_s)}{u_t}\\&=\left(\frac{1-\varepsilon}{\varepsilon}\right)^2(\varepsilon-q)f(q,\varepsilon)\end{aligned} \tag{13}$$

where u_t is

$$u_t=k_0\rho_s g(1-\rho/\rho_s)\ . \tag{14}$$

Now the problem reduces to finding $\varepsilon(\xi,\tau)$ and $q(\xi,\tau)$ as solutions to Equation (12). The initial conditions are clearly

$$\varepsilon(\xi,0)=\varepsilon_0(\xi),\ q(\xi,0)=0 \tag{15}$$

In order to de-couple the problem from the complication arising from the effect of the consolidation wave due to the bottom of the sand column, we assume that the sand column is very tall so that cracks may develop before the consolidation wave arrives.

4. Numerical Results and Analysis

In this section, we will solve Equation (17) by using of the finite differential method under two types of initial conditions to study the evolution of water film.

Parameters adopted in simulation are as follows(Lu *et al.*, 2007): β=47~56, $\rho_s=2400kg/m^3$, $\rho_w=1000kg/m^3$, $u^*=0.04$, $k_0=4\times10^{-6}m/s$, $\alpha=1$, κ=50.0, $a=0.08$, time step $\Delta\tau=9\times10^{-4}$, step length $\Delta x=0.01$, critical pore pressure $P_{cr}=0.25MPa$, L is the length of the sand column.

The results are given as follows:

1) Condition 1: The initial porosity changes continuously, and assuming that once a place is choked, the choked state keeps unchangeable.

It is shown from Figure 1 that when some position is choked, the porosity just below this position increases gradually till it equals 1.0. We think there is a water film full of water occurring here. Once it occurs, the water film expands gradually.

Figure 2 shows the velocity development of pore water. It is shown that the velocity of pore water first increases because of the high hydraulic gradient. Nevertheless, with the decrease of porosity and the permeability, the velocity decrease to near zero. Outside the water film, the velocity of pore water changes little. We may think there is discontinuity exists between inside and outside the water film.

Figure 3 shows the development of the fine grains eroded from the skeleton. It is shown that at the choked position, the sum of fine grains decreases fast because of the small velocity difference between the pore water and grains, while above the choked position, the sum of fine grains increases fast. The eroded fine grains are related with the velocity difference between pore water and

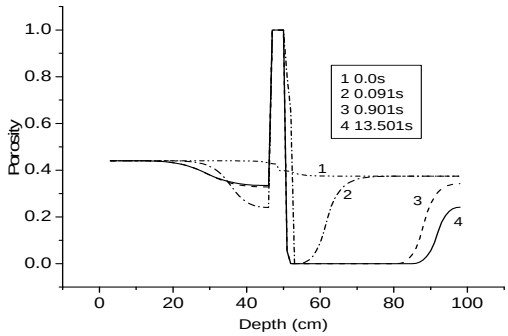


Figure 1. Development of porosity under assumption 1.

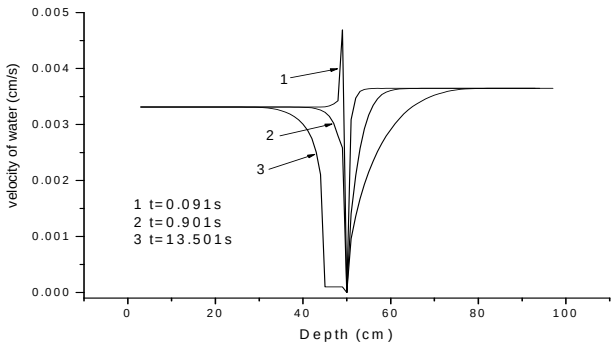


Figure 2. Development of velocity of pore water under assumption 1.

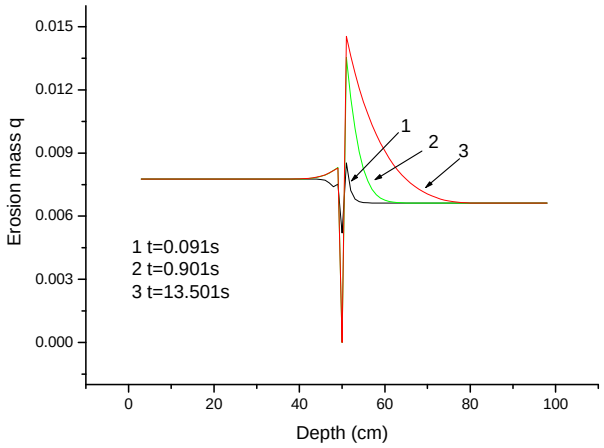


Figure 3. Development of q under assumption 1.

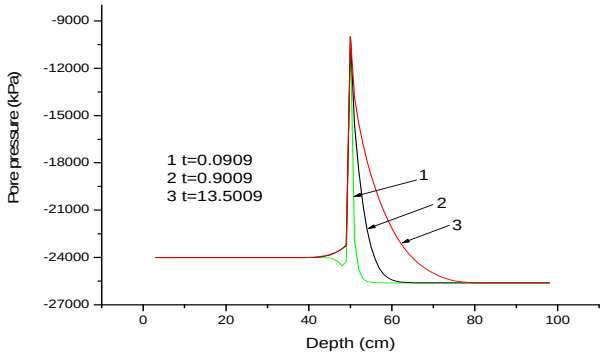


Figure 4. Development of pore pressure under assumption 1 (minus denotes compressive pore pressure).

grains according to Equations (1) and (2). Therefore, in the water film, the eroded fine gains q is very small because the velocity difference is near zero, while it is big above the water film because the velocity difference is big. Some distance away from the water film, the eroded fine grains is very small and does not change with position.

Figure 4 shows the distribution of pore pressure. It can be seen that the pore pressure is discontinuity. The peak of the pore pressure is at the choked point. It causes the pore water percolate upward and the porosity above the choked point increase. The pore pressure below the choked position becomes from the initial continuous to increase suddenly. The reason is that the chocked position leads the pore water flow difficult and thus causes the pore pressure increase.

The reason may be explained as follows: when the sand column is choked, the velocity of pore water decreases to near zero and a high hydraulic gradient forms at this position because the porosity is very small. Thus the discontinuity of the pore water velocity, grain velocity and the amount of fine grains eroded from the skele-

ton occurs. The pore water beneath the choked position flows upwards but passes through the choked position very difficult, which leads a net inflow of pore water just beneath the choked position. Therefore, water film forms gradually at this position. With the pore water beneath the choked position flows upwards while the grains sink gradually, the water film becomes wider and wider. The expanding velocity of water film is equal to that of the discontinuity

$$D = \left[\left(1-\varepsilon^{+}\right)u_s^{+} - \left(1-\varepsilon^{-}\right)u_s^{-}\right] / \left[\left(1-\varepsilon^{+}\right) - \left(1-\varepsilon^{-}\right)\right] = u_s^{-},$$

where u_s^{+}, ε^{+} denotes the velocity of grains and porosity in the water film and thus $u_s^{+} = 0, \varepsilon^{+} = 1$, u_s^{-}, ε^{-} denotes the velocity of grains and porosity beneath the water film. It can be seen that the expanding velocity of water film is the same as the settlement velocity of grains. This velocity will change with the other parameters of the sand column.

2) Condition 2: The porosity distribution is the same as that in Condition 1, but there is no choking-state-keeping assumption.

It is shown from Figure 5 that when some position is choked, the porosity at this position increases gradually till water film occurs. If the choked position is dredged again when the pore pressure is over the critical pore pressure, the pore pressure at the chocked position decreases gradually, and the pore water and grain velocities becomes smooth from discontinuity states. At last, water film disappears.

Figure 6 gives the development of velocity of pore water under Condition 2. It is shown that the velocity of pore water develops to be discontinuity at the first stage after water film occurring. When the choked position is dredged, the velocity of pore water below the chocked position increases while the velocity in water film is still zero. The increase may be that the pore water flows through the once chocked position and the pore pressure decreases gradually. At last, the velocity of pore water becomes smoothly.

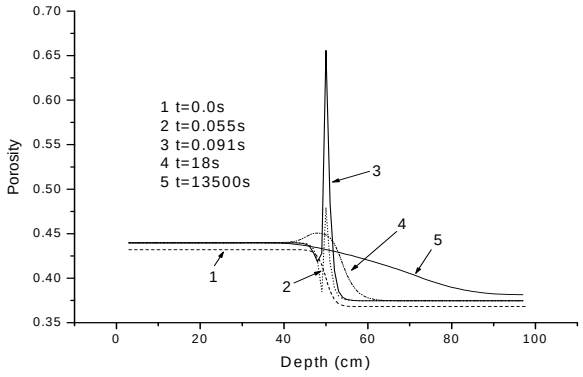


Figure 5. Development of porosity under assumption 2.

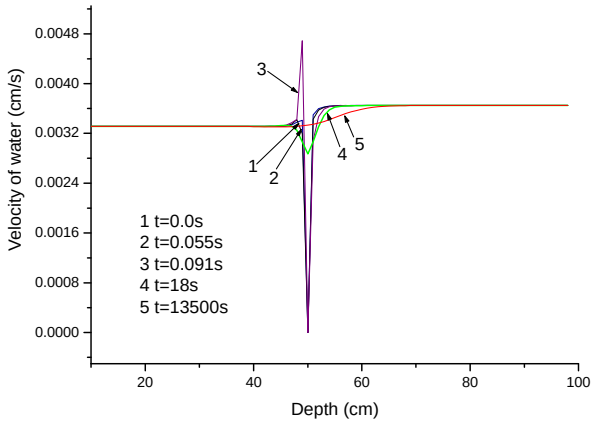


Figure 6. Development of velocity of pore water under assumption 2.

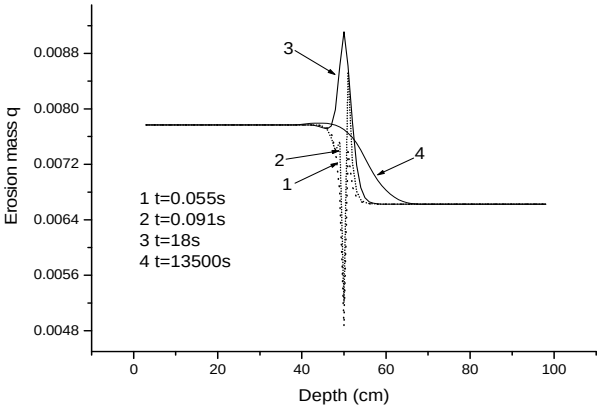


Figure 7. Development of q under assumption 2.

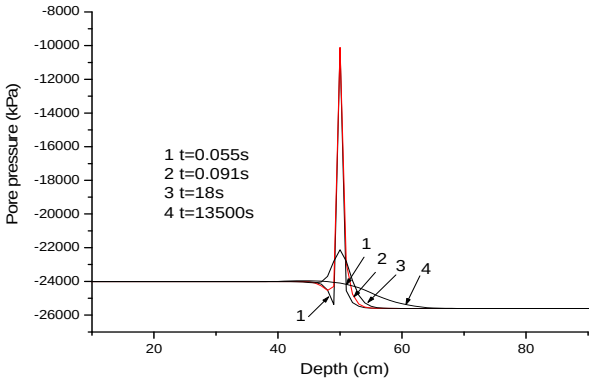


Figure 8. Development of pore pressure under assumption 2.

Figure 7 gives the distribution and development of the eroded fine grains under Condition 2. It is shown that the distribution of eroded fine grains is sharply decreased in the water film, but with the disappearing of water film, the velocity of pore water increases first and then de-

creases, so the eroded fine grains increase first at the position once being water film, and then the distribution of eroded fine grains becomes smooth.

Figure 8 gives the development of pore pressure under Condition 2. It is shown that the development of pore pressure increases fast once the chocked position forms. Nevertheless, with the dredging, the water film disappears gradually and the pore pressure decreases gradually.

At the first stage, the water film forms similar to that under Condition 1. Nevertheless, once the pore pressure is over the critical value, water can flow through the choked position again. The grains above the choked position catch up the grains beneath the choked position. The pore pressure becomes smoothly gradually and the water film disappears.

5. A Simplified Evaluation Method

Although the numerical simulation may give more information about the evolution of the water film, a simplified method is needed. Florin *et al*. [8] pointed out that when the settling particles reach solid material, which is usually the non-liquefied underlying soil, or the container base in an experiment, they accumulate to form a solidified zone which increases in thickness with time. A solidification front therefore moves upward until it reaches the surface or the overlying non-liquefied material. Zhang *et al*. [5] and Scott *et al.* [9] had analyzed the development of the solidification. Here we present a simplified analytical method of the water film in saturated sand with initial non-uniform grade series.

Assuming that the whole mass reaches its terminal velocity, *k*, which is the permeability, instantaneously at the end of liquefaction, Florin gave an expression for the constant velocity, $\dot{z}$, of the solidification front:

$$\dot{z} = \frac{\rho}{\rho_w}\frac{1-n_1}{n_2-n_1}k \tag{16}$$

in which $\rho' = \rho_s - \rho_w$ is the buoyant unit weight of the liquefied soil, n_1 is the porosity of the liquefied soil, n_2 is the porosity of the solidified soil.

From Equation (16), we can obtain the duration of liquefaction and subsequent excess pore pressure decline for any point in the soil column.

$$t = \frac{\rho_w}{\rho}\frac{n_2-n_1}{1-n_1}\frac{h}{k} \tag{17}$$

h is the height of any point in the soil column.

The final settlement of the top surface of the sand layer is

$$\Delta L = \frac{n_0-n_1}{1+n_0}H \tag{18}$$

This occurs at time given by Equation (17), so the rate of settlement is

$$\dot{s}_a = \frac{\gamma' k}{\gamma_w} \tag{19}$$

The settlement at any time is

$$s_a = \frac{\gamma' k}{\gamma_w}t \tag{20}$$

The settlement velocity v_e of the elements above the water film is determined by the combined permeability k_{es} of the middle layer and the upper layer [10]:

$$k_{es} = \sum_1^m L_i \Big/ \sum_1^m L_i/k_i \tag{21}$$

The upward seepage flow has the same velocity:

$$v = k_{es} i_e \tag{22}$$

i_e is the average hydraulic gradient.

In fact, the particles and solidified soil are compressible, the skeleton may consolidate by the geostatic stress after the solidification. If the permeability is small, the excess pore pressure does not disperse immediately with the increase of solidification zone. The deform of the skeleton of the sand by the geostatic stress in the solidification zone may be expressed as [5]

$$s_2 = \frac{1}{2}\frac{\rho' g}{m_s}x^2(t) \tag{23}$$

in which the compressible modulus m_s is assumed as a constant. The percolation is assumed to obey the Darcy lay. The total deformation is:

$$\Delta s = \Delta s_1 + \Delta s_2 \tag{24}$$

Instituting Equations (16) and (23) into (24), it yields

$$\Delta s = \frac{n_{20}-n_{21}}{1-n_{20}}(\Delta z + \Delta s) + \frac{\rho' g}{m_s} z\Delta z \tag{25}$$

The settling velocity of the surface is

$$\frac{\Delta s}{\Delta t} = k\frac{\rho'}{\rho_w} \tag{26}$$

The increase velocity of the thickness of solidification is

$$\frac{\Delta z}{\Delta t} = \frac{k\rho^{'}}{\rho_w} \Bigg/ \left(\frac{n_0 - n_1}{1 - n_1} + \frac{1 - n_1}{1 - n_0} \frac{\rho^{'} g}{m_s} z \right) \tag{27}$$

The time for the solidification front to arrive at any location is

$$t = \frac{\rho_w}{k\rho^{'}} \left(\frac{n_0 - n_1}{1 - n_1} z + \frac{1}{2} \frac{1 - n_1}{1 - n_0} \frac{\rho^{'} g}{m_s} z^2 \right) \tag{28}$$

The side friction, if influential in the situation, may be expressed as

$$\sigma_s = \mu K_0 \sigma_z \tag{29}$$

This effect should be considered in the pore-pressure-gradient.

The effect of the changes of porosity on the permeability is considered here using a linear relation:

$$k = k_0 \left[1 - \alpha (n_0 - n) \right] \tag{30}$$

in which k_0 is the initial porosity, α is a parameter, n_0 is the initial porosity, k is the permeability when the porosity is n.

Thus, by considering Equations (29) and (30) in the pore pressure gradient and the consolidation of the solidification zone, we can compute the course that the water film increases and closes.

6. Comparison with the Experimental Results

The results are compared with the experimental data of Kokusho [10] (Figure 5). In Kokusho's experiment, a saturated loose sand layer of 200cm depth sandwiches a seam of nonplastic silt in the middle (96 cm from the bottom). The initial void ratios of the upper, the middle and the lower sand layers are 0.924, 1.5 and 0.831, respectively; and the initial permeability coefficients are 0.04 cm/s, 0.00018 cm/s and 0.04 cm/s, respectively. The silt seam is about 4mm thick. The saturated sand is in a tube with an inner diameter of 13 cm and a height of 211.5 cm. The one-dimensional sand layer is instantaneously liquefied under a loading caused by a steel hammer. In the computation for comparison, we adopted the data in Kokusho [10] (Table 1). In Figure 9, we can see that the two results are agreement with each other in some extent.

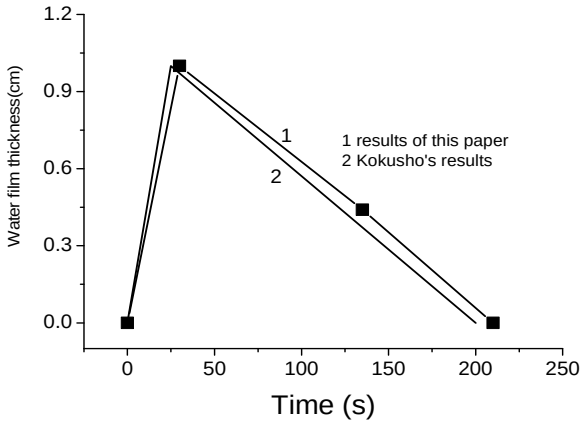


Figure 9. The comparison of our results with the experimental of Kokusho [10].

7. Conclusions

Numerical simulations under four conditions have shown that the stable water films occur only in the conditions that: (1) the porosity of the upper part of the sand column is smaller than that of the lower, and (2) the keeping of the jamming state or the effective stress to prevent the free dropping of the grain or the skin friction in Kokusho's experiments is needed. A simplified method for evaluating the thickness of water film is presented and the computed results are close to the experimental results.

Although some main characteristics of the formation of water film are obtained based on the presented model, some improvements such as the erosion relation Equation (1) and the permeability function Equation (9) are needed in order to make the model more practical. The analyze on the evolution of the sand column in two and three dimensions will be processed in the future.

8. Acknowledgements

This paper is supported by National Basic Research Program of China "Activity characteristics and formation rules of secondary mountain hazard of earthquake" (No. 2008CB425802) and Key Program of Chinese Academy of Sciences (No. KZCX2-YW-302-02).

Table 1. The data of the parameters.

sand	Thickness (cm)	Relative density (%)	Permeability coefficient (cm/s)	Initial pore ratio	The max. strain (%)
The upper layer of sand	103.6	14	0.04	0.924(0.48)	2.4
seam	0.4		1.8E-4	1.5(0.6)	2.4
The lower layer of sand	96	39	0.04	0.831(0.454)	0.95

9. References

[1] T. Kokusho, K. Watanabe, and T. Sawano, "Effect of water film on lateral flow failure of liquefied sand," Proceedings in 11th European Conference on Earthquake Engineering., Paris, CD publication, ECEE/T2/kokeow.pdf, 1998.

[2] H. B. Seed, "Design problems in sand liquefaction," Journal of Geotechnical Engineering, ASCE, Vol. 113, No. 8, pp. 827–845, August 1987.

[3] G. L. Fiegel and B. L. Kutter, "Liquefaction mechanism for layered sands," Journal of Geotechnical Engineering., ASCE, Vol. 120, No. 4, pp. 737–755, Spring 1994.

[4] T. Kokusho, "Water film in liquefield sand and its effect on lateral spread," Journal of Geotechnical and Geoenvironmental Engineering, Vol. 125, No. 10, pp. 817–826, October 1999.

[5] J. F. Zhang, "Experimental study on the strengthening of percolation and the damage of structure under impact loading," Dissertation for Ph. D, Institute of Mechanics, Chinese Academy of Sciences, 1998.

[6] C. M. Cheng, Q. M. Tan, and F. J. Peng, "On the mechanism of the formation of horizontal cracks in a vertical column of saturated sand," ACTA Mechanica Sinica (English Serials), Vol. 17, No. 1, pp. 1–9, March 2000.

[7] X. B. Lu and Z. M. Zheng, "Formation of water film in saturated sand," ACTA Mechania Sinica, Vol. 22, No.3, pp. 377–383, June 2006.

[8] V. A. Florin and P. L. Ivanov, "Liquefaction of saturated sandy soils," Proceedings in 5th International Conference on Soil Mechanic Foundation Engineering, Paris, 1, pp. 107–111, 1961.

[9] R. F. Scott, "Solidification and consolidation of a liquefied sand column," Soils and Foundations, Vol. 26, No. 4, pp. 23–31, 1986.

[10] T. Kokusho and T. Kojima, "Mechanism for postliquefaction water film generation in layered sand," Journal of Geotechnical and Geoenvironmental Engineering, ASCE, Vol. 128, No. 2, pp. 129–137, February 2002.

Call for Papers

ENGINEERING

A Journal Published by Scientific Research Publishing, USA

www.scirp.org/journal/eng

ENGINEERING is an international journal dedicated to the latest advancement of engineering. The goal of this journal is to provide a platform for engineers and academicians all over the world to promote, share, and discuss various new issues and developments in different areas of engineering. All manuscripts must be prepared in English, and are subject to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by printed hard copy. The journal publishes original papers including but not limited to the following fields:

- Aerospace Engineering
- Agricultural Engineering
- Chemical Engineering
- Civil Engineering
- Electrical Engineering
- Environmental Engineering
- Industrial Engineering
- Materials Engineering
- Mechanical Engineering
- Mining Engineering
- Nanotechnology
- Nuclear Engineering
- Power Engineering
- Test Engineering
- Transportation Engineering

We are also interested in: 1) Short Reports—2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data; 2) Book Reviews—Comments and critiques.

☆ Notes for Intending Authors

Submitted papers should not be previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. For more details, please access the website.

☆ Website and E-Mail

http://www.scirp.org/journal/eng eng@scirp.org

Engineering, 2010, 2, 1-84
Published Online January 2010 in SciRes (www.scirp.org/journal/eng/)

TABLE OF CONTENTS

Volume 2 Number 1 **January 2010**

9771947393001 04