

Estimation of the Stapes-Bone Thickness in Stapedotomy Surgical Procedure Using a Machine-Learning Technique *

Vassilis G. Kaburlasos⁽¹⁾, Vassilios Petridis⁽¹⁾, Peter Brett⁽²⁾, Dave Baker⁽²⁾

⁽¹⁾Department of Electrical and Computer Engineering
Aristotle University of Thessaloniki
GR-54006 Thessaloniki, GREECE
e-mail: vgekabs@egnatia.ee.auth.gr, petridis@vergina.eng.auth.gr

⁽²⁾Department of Mechanical Engineering
University of Bristol, Queen's Building, University Walk
Bristol, BS8 1TR, U.K.

Abstract

Stapedotomy is a surgical procedure aiming at the treatment of hearing impairment, where the latter is due to otosclerosis, by drilling a hole through the stapes bone in the inner ear in order to insert a prosthesis. Safety requires knowledge of the non-measurable stapes thickness. The technical goal herein has been the design of high level controls for an intelligent mechatronics drilling tool in order to enable the estimation of stapes thickness from measurable drilling data. The goal has been met by learning a map between drilling features, hence no model of the physical system has been necessary. Learning has been achieved as explained in this paper by a scheme, namely *d-S Fuzzy Lattice Neurocomputing (dS-FLN)* scheme for classification, within the framework of *Fuzzy Lattices*. The successful application of the *dS-FLN* scheme is demonstrated in estimating the thickness of a stapes bone “on-line” using drilling data obtained experimentally in the laboratory.

Index Terms - Intelligent mechatronics, stapedotomy surgery, neural networks, pattern classification, fuzzy lattice framework, skillful tasks.

1 INTRODUCTION

As surgical procedures are increasing in complexity, the precision to which non-invasive measurements can be obtained has increased in many cases beyond the capacity for accurate deployment of surgical tools by manual methods. There is a need to be able to work with comparable accuracy in the uncertain surgical environment. Dexterous and intelligent electro-mechanical, that is *mechatronical*, surgical tools have been envisaged for assisting in the surgical theater.

In the above context a concrete surgical procedure is dealt with herein, that is *stapedotomy*. Stapedotomy is a corrective surgical procedure whose aim is to overcome *otosclerosis*, where *otosclerosis* is the name for excessive bone growth within the mid-ear cavity which typically leads to hearing impairment. A stapedotomy may be performed by drilling a hole through the base, namely *footplate*, of the stapes bone. Finally hearing is restored by inserting a prosthesis through the footplate of the stapes and connecting the prosthesis appropriately. Note that the term *stapedotomy* may also be used for the actual hole drilled through the stapes.

* This work was supported in part by Brite Euram Project BE7470.

Drilling through the stapes could be dangerous because of the potentially excessive tool protrusion, at the end of drilling, due to deflection of the flexible stapes bone under tool action. Excessive tool protrusion could hurt the inner ear irreparably. Such a danger can be eliminated should the thickness of stapes is known. Nevertheless the thickness of the stapes bone cannot be determined before drilling. Herein the problem of determining “on-line” the stapes thickness has been formulated as a classification problem. In particular a machine-learning scheme has been employed, namely *d-S Fuzzy Lattice Neurocomputing* scheme or *dS-FLN* scheme for short. To put this work in perspective, selected approaches to learning and classification from the literature are reviewed in the sequel.

An older approach to classification in an unpredictable environment is considered in [21] by proposing theoretically a hierarchical scheme of increasing intelligence and decreasing precision. In [17] a learning application is shown which involves the use of a robot in brain surgery; in particular Kalman filtering is employed for estimating with precision a robot’s geometric parameters at the time of surgery in order to effect submillimeter accuracy in positioning the robot’s end effector.

The performance of various learning and classification techniques is shown comparatively in [13]. A similar comparative study of different machine learning, neural, and statistical classification algorithms has been undertaken in the context of the European research project Statlog [4]. A distributed model for learning, recognition, and prediction is provided in [7] together with a summary of various adaptive resonance theory (ART) models. In addition, there already exist web cites summarizing comparatively various classification algorithms [10].

Decision-making has been treated in various contexts. Lately, in the context of fuzzy set theory, the work in [22] proposes a set-theoretic similarity measure between fuzzy sets. In the context of the Dempster-Shafer theory of evidence, the work in [18] following a review, treats information which is uncertain, imprecise, and occasionally inaccurate, by the fusion of sensors at the symbol level. Herein, decision making has been dealt with in a novel framework as it is explained towards the end of this section.

The selection of the features to be fed to a pattern classifier may be a more important problem than the classifier architecture itself; for instance the work in [12] employs the Bienenstock, Cooper, and Munro (BCM) unsupervised network for feature extraction and it demonstrates an improved classification of underwater acoustic signals using a conventional back propagation classifier.

The notion *lattice* has been employed in the literature as well for learning and decision-making. We remark that the term *lattice* is not always used in the literature with respect to its strict mathematical definition. For instance in [16] the word *lattice* is employed to denote a regular arrangement of points in R^n . In what follows by *lattice* we mean a *mathematical lattice* [9]. Lattices may arise naturally when dealing with knowledge in computer science; for instance in [2] a lattice of rules is employed for the analysis of techniques focused on detecting termination of rule execution in a knowledge base. Lattices have also arisen in the study of hybrid systems [6], these are systems in which the state set possesses both continuous and discrete components. The mathematically sound *framework of fuzzy lattices*, or *FL-framework*, for learning and decision making has been introduced in [15], [19]. A novelty of the *FL-framework* is that *FL-framework* promotes the processing of lattice data elements. An advantage of the *FL-framework* is its inherent capacity to cope with disparate types of data including numbers, symbols, fuzzy sets, etc.; such a capacity is expected by the authors to be important in skillful tasks including surgical procedures. In this work we have employed the *FL-framework* for learning and decision making.

The layout of this article is as follows. Section 2 reviews the state-of-the-art technology in stapledotomy. Section 3 details a feature extraction technique, it outlines aspects of the underlying theory, it reviews the *dS-FLN* scheme for learning, and it exposes the equations needed for the application of the latter scheme. Section 4 details both data acquisition and the experiments carried out, and finally it shows the corresponding results. Section 5 concludes this article with a discussion of the merits of the technique proposed.

2 STATE-OF-THE-ART IN STAPEDOTOMY

The transmission of sound from the outside world to the sense organs in the inner ear is mediated via the three ossicular bones of the middle ear, namely Malleus, Incus, and Stapes, shown in Fig.1(a). *Otosclerosis* is the excess bone growth within the mid-ear cavity especially on the stapes footplate (Fig.1(a)) and can lead to hearing impairment by partial or complete fixation of elements in the ossicular chain. The *stapedotomy* surgical procedure aims at overcoming otosclerosis by opening a hole through the stapes footplate in order to insert a prosthesis and connecting the prosthesis either to the remaining ossicular bones or directly to the tympanic membrane as shown in Fig.1(b). Note that, safety requires a non-excessive protrusion of the drilling tool past the far medial surface of the flexible stapes bone because an excessive protrusion might risk further hearing impairment to the patient.

Due to its size, the flexible stapes bone is delicate and can be broken easily. Note that typically the dimensions of the stapes are approximately 2.5 by 1.5 mm² with a thickness of about 0.3 mm for a healthy bone, and with thickness in the range 0.2-2.5 mm for a diseased bone. Currently there exist three main tool types to create a stapedotomy.

1. Metal pick. Usually triangular in cross section; it is rotated back and forth to grind away the underlying bone. The hole created is often of poor quality, and there is also a risk of mechanical trauma to the inner ear.
2. Electric drill. It produces a good quality opening, and it is run at a subacoustic speed to prevent acoustic trauma to the inner ear. Nevertheless care must be taken to avoid drilling past the medial surface of the stapes footplate and the possible rupture of the underlying fluid sac.
3. Lasers. Visible light [8] and long wavelength (*LW*) lasers have been employed. Both types of lasers are short of the capacity to control the depth of penetration of the beam. Additional drawbacks include the durability of the stapedotomy as well as the risk of both thermal and acoustic traumata to the inner ear.

Our research team has opted for an electric drill to perform a stapedotomy because of the high quality hole it produces for inserting and sustaining the prosthesis. The main disadvantage of the drilling procedure is the potentially excessive and therefore dangerous overshoot of the drill bit at the stapes breakthrough. Such an overshoot can be avoided if the thickness (*t*) of the stapes footplate is known. Nevertheless the thickness *t* cannot be determined before the drilling operation.

In the past, a mathematical model of the following form has been employed with a mechatronic surgical drill to estimate the thickness of a stapes.

$$\frac{dF}{dx} = f_1(F, \mathbf{g}, P, K, \mathbf{q})$$

$$T = f_2(F, \mathbf{b}, P, \mathbf{q})$$

where functions f_1 and f_2 are algebraic expressions of their arguments, F and T stand respectively for *force* and *torque*, γ is the *cutting coefficient*, β is the *loss coefficient*, P is a vector of the drilling parameters including the *angular velocity*, the *feed rate*, and the *drill bit radius*, K is the bone *stiffness parameter*, and $\theta=(\theta_1, \theta_2)$ is the vector of the *near-* and *far-* cutting surfaces *angle*. Estimation of the thickness (*t*) is achieved by calculating iteratively improved estimates of the thickness *t*. For additional details the reader may refer to [1].

The previous, model-based, method is successful in estimating the stapes thickness suffices certain simplifying assumptions hold regarding the materials involved in the drilling. Nevertheless difficulties have been experienced in practice because the aforementioned simplifying assumptions are not always valid. Moreover, the above mathematical model cannot be used in the face of unexpected events like, for instance, the existence of a tiny blood vessel within the stapes bone.

In this work a machine-learning technique has been employed in order to estimate a stapes thickness during the stapedotomy surgical procedure by classification as explained in section 3.3. Preliminary results have been presented in [14]. In the following chapter specific techniques are shown for learning the stapes thickness in the context of the *framework of fuzzy lattices*, or *FL-framework*; it is also explained how the estimation of stapes thickness can be cast as a classification problem.

3 A MACHINE-LEARNING APPROACH FOR ESTIMATING A STAPES THICKNESS

3.1 A Feature Extraction Technique

Specific drilling features have been extracted, as explain below, and then an estimation of a stapes thickness has been effected by classification as it is outlined next. The drilling data from a stapes of unknown thickness are assigned to a previously learned class of drilling data all of whom correspond to the same stapes thickness. That is, the thickness label of the learned class in question is attached to the “unknown thickness” stapes. The remaining of this section details the specific feature extraction technique employed.

Experimentation has shown that for various combinations of force and torque features it has not been possible to locate clusters of drilling data of the same thickness. Moreover, experimentation has shown that such a clustering was feasible by a linear map of force features to torque features. In particular, force features and torque features have been arranged in two different 2×2 matrices F and T respectively. The datum considered for “learning” has been a 2×2 matrix M so as $MF=T$; that is the linear operators (matrices) M have been the data employed for learning. The entries of matrices F and T have been selected as it is explained below. A decision to employ square matrices was taken because a useful norm-induced metric can be defined in square matrices [5]. Herein the *spectral norm* $\| \cdot \|_s$ of square matrices has been considered which is defined as $\|A\|_s = \max\{\sqrt{\lambda}, \lambda \text{ is an eigenvalue of } A'A\}$, where A' is the transpose of a matrix A with real elements. Apparently the space M of 2×2 matrices M constitutes a linear metric space, with the spectral metric $d_s(M_1, M_2) = \|M_1 - M_2\|_s$.

Entries have been selected for matrices $F = \begin{bmatrix} f_a & f_b \\ f_c & f_d \end{bmatrix}$ and $T = \begin{bmatrix} t_a & t_b \\ t_c & t_d \end{bmatrix}$ so as to maximize an index of performance in the data. Specifically, among various “candidate” entries for matrices F and T we have selected the ones which maximize the distance between matrices M corresponding to different stapes thickness. To illustrate the procedure of selecting suitable entries for matrices F and T consider the spectral distance $d_s(M_1, M_2) = \|M_1 - M_2\|_s$ between two matrices M_1 and M_2 . Assuming that matrix $\begin{bmatrix} d_1 & d_2 \\ d_3 & d_4 \end{bmatrix}$ denotes the difference $M_1 - M_2$ it can be shown by straight-forward algebraic manipulations that $\lambda_{\max} = \frac{(d_1^2 + d_2^2 + d_3^2 + d_4^2) + \sqrt{(d_1^2 + d_2^2 + d_3^2 + d_4^2)^2 - (d_1 d_4 - d_2 d_3)^2}}{2}$. Hence from a set of tentative entries for matrices F and T , entries have been selected by trial-and-error so as the quantity $(\delta_1^2 + \delta_2^2 + \delta_3^2 + \delta_4^2)^2$ is maximized and $\delta_1 \delta_4$ becomes approximately equal to $\delta_2 \delta_3$. In conclusion the following matrices have been employed.

$$F = \begin{bmatrix} \frac{\text{slope}(y_f)}{\sqrt{\sum_{i=1}^l (y_f - f)^2}} & \text{std}(f) \\ d(l) & \end{bmatrix}, \text{ and } T = \begin{bmatrix} \frac{\text{slope}(y_q)}{\sqrt{\sum_{i=1}^l (y_q - q)^2}} & \text{std}(q) \\ K^2 & \end{bmatrix}$$

where l is the number of *force* (f) and *torque* (q) data, y_f and y_q are the least square linear regressors of f and q respectively, $slope(\cdot)$ is the slope of its linear argument, $std(\cdot)$ is the standard deviation of its argument, the square root term indicates the overall error of approximating the data by a straight line, $d(l)$ is the current displacement of the drilling tool, and K is the stiffness parameter of the stapes. All the aforementioned quantities are in standard System International (SI) units but the displacement $d(l)$ is in [mm] and the stiffness K is in [N/mm].

In conclusion, it has been matrices M which have been employed for learning and classification. On the one hand during learning, clusters of known matrices M corresponding to the same stapes thickness have been located. On the other hand during classification, a matrix $M(\mathbf{t})$ corresponding to a stapes of unknown thickness being drilled kept being assigned to a cluster, as it is explained in sections 3.3 and 4.2, and consequently the thickness label of the cluster in question was assigned to $M(\mathbf{t})$. Therefore the measurable input data to the $d\mathbf{s}$ -FLN have been : (1) the drilling force versus displacement, (2) the drilling torque versus displacement, and (3) the stapes stiffness. The output from the $d\mathbf{s}$ -FLN scheme has been the stapes thickness.

3.2 Theoretical Perspectives

This section constitutes a “detour” from our focal point of stapedotomy surgery in order to delineate the mathematical tools used herein for data processing.

3.2.1 Hidden-lattices

The *framework of fuzzy lattices*, or *FL-framework* for short, has been introduced in [15], [19] for dealing with lattice elements including conventional numbers, fuzzy sets, symbols, etc. The novel notion *hidden-lattice* is introduced herein as an enhancement of the notion *lattice*.

Let, P be a *partly ordered set*, L be a *lattice*, and $\mathbf{y}: P \rightarrow L$ be a *monotone* map [9]. The implied partial ordering relation in P and L will be denoted respectively by \leq_P and \leq_L . Recall also that a map $\mathbf{y}: P \rightarrow L$ is called *monotone* if $x \leq_P y$ implies $\mathbf{y}(x) \leq_L \mathbf{y}(y)$. For $x, y \in P$ consider an upper bound, say z , of x and y ; if there exists a unique upper bound, namely z_0 , of x and y in P such that $\mathbf{y}(z_0) \leq_L \mathbf{y}(z)$ for all upper bounds $z \in P$ then z_0 is called the *join* of x and y in P . Note that *join* z_0 does not have to be included in every other upper bound of x and y in P . The notion *meet* of $x, y \in P$ is defined dually; that is the *meet* is the unique lower bound, say w_0 , of x and y in P such that $\mathbf{y}(w) \leq_L \mathbf{y}(w_0)$ for all lower bounds w of x and y in P . Likewise, note that *meet* w_0 does not have to include every other lower bound of x and y in P . The definition of the notion *hidden-lattice* follows.

Definition 1 A partly ordered set P is called *hidden-lattice* if there exists a lattice L , namely *the underlying lattice*, and a monotone map $\mathbf{y}: P \rightarrow L$, namely *the corresponding map*, such that any two elements $x, y \in P$ have a *join* denoted by $x \vee_P y$, and a *meet* denoted by $x \overline{\wedge}_P y$. A hidden-lattice is *complete* when each of its subsets has a meet and a join in P .

By considering the identity map on a lattice L it is implied that lattice L is a hidden-lattice; nevertheless the converse is not always true.

The utility of a *metric* and an *inclusion measure* in learning- and decision-making- involving lattice elements has been shown in [15], [19]. Likewise it could be useful to have both a metric and an inclusion measure in a hidden-lattice. A *positive valuation* function has been employed in [15] for defining both a *metric* (or, *distance function*) and an *inclusion measure* in a lattice. Recall that a *valuation* on a lattice L is a real function $v: L \rightarrow \mathbf{R}$ which satisfies $v(x) + v(y) = v(x \vee_L y) + v(x \wedge_L y)$, $x, y \in L$; a valuation is called *monotone* if

and only if $x \leq_L y$ implies $v(x) \leq v(y)$, and *positive* if and only if $x <_L y$ implies $v(x) < v(y)$. On the one hand, in a lattice L with a positive valuation v , a *metric* is defined by $d(x,y) = v(x \vee_L y) - v(x \wedge_L y)$. On the other hand, in [15] an *inclusion measure* σ is defined with regards to a complete lattice. Underneath the definition of an inclusion measure is extended to any lattice, be it complete or not.

Definition 2 Let L be a lattice. An *inclusion measure* in L is a map $\sigma: L \times L \rightarrow [0,1]$ such that $\sigma((x,u)) \equiv \sigma(x \leq_L u)$ satisfies the following two conditions:

$$(E1) \sigma(u \leq_L u) = 1, \forall u \in L.$$

$$(E2) u \leq_L w \Rightarrow \sigma(x \leq_L u) \leq \sigma(x \leq_L w), x, u, w \in L - \text{Consistency Property.}$$

Consider a hidden-lattice P with underlying lattice L and corresponding map \mathbf{y} . Suppose that there exists a positive valuation function v in L . Then a metric and an inclusion measure can be defined in hidden-lattice P as follows : (1) a *metric* is defined by $d(x,y) \doteq d(\mathbf{y}(x \vee_P y), \mathbf{y}(x \wedge_P y))$, and (2) an *inclusion measure* is defined by $\sigma(x \leq_P y) \doteq \frac{v(\mathbf{y}(y))}{v(\mathbf{y}(x \vee_P y))}$. It can be shown that $\mathbf{y}(x \wedge_P y) \leq_L \mathbf{y}(x \vee_P y)$ in L , therefore the metric in P equals $d(x,y) = v(\mathbf{y}(x \vee_P y)) - v(\mathbf{y}(x \wedge_P y))$.

3.2.2 The hidden-lattice of generalized spheres in a linear metric space

Let (L,d) denote a linear metric space, where L is a linear space and $d(.,.)$ is a metric in L . In this section it is shown how learning can be effected in (L,d) by employing a concrete hidden-lattice defined in (L,d) . Specifically it is shown that experience can be collected and a meaningful generalization can be attained by defining *balls* in L . Note that a *ball* centered on a vector $c \in L$ with a (non-negative) radius $r \geq 0$ is a set of the form $\{x \in L: d(x,c) \leq r\}$, and it will be denoted by “ball(c,r)”, where $c \in L$ and $r \geq 0$. Apparently “ball($c,0$)” is a *trivial ball* which contains only a single element $c \in L$; a trivial ball will also be called *point* or *atom*. Let B denote the collection of balls in (L,d) ; we write “ball(k,a) \leq_B ball(c,r)” if and only if all the points of ball(k,a) are in ball(c,r).

The utility of balls for learning in (L,d) lies on the fact that crisp sets in L can be approximated by collections of balls. Hence learning crisp- and mutually-disjoint sets in a linear metric space (L,d) can be pursued by a procedure of defining and enhancing collections of balls. We remark that balls have been considered either explicitly or implicitly by various learning schemes including the adaptive resonance theory [25], learning vector quantization (LVQ) [11], radial basis functions (RBF) [3], etc. Moreover the Vapnik-Chervonenkis dimension of the set of balls has been calculated [24].

The set B of balls in a linear metric space is partly ordered. Nevertheless the set B is not a lattice for a pair (ball(c,r), ball(k,a)) there is no upper bound ball contained in every other upper bound ball, that is there is no *least upper bound*. Likewise for a pair (ball(c,r), ball(k,a)) there is no lower bound ball which contains every other lower bound ball, that is there is no *greatest lower bound*. However, it is shown in this section that it is possible to introduce both a metric and an inclusion measure in the partly ordered set B by mapping B to a hidden-lattice Φ whose underlying lattice L is equipped with a positive valuation function.

More specifically, consider the *set of spheres* S in a linear metric space (L,d) , where a *sphere* centered on a vector $c \in L$ with a (non-negative) radius $r \geq 0$, is a set of the form $\{x \in L: d(x,c) = r\}$, and it will be denoted by “sphere(c,r)”. Apparently there exists a *bijection* (that is an one-to-one map and onto) $b_B: B \leftrightarrow S$ between the set B of balls and the set S of spheres. Hence, a partial ordering relation is implied in S , symbolically sphere(k,a) \leq_S sphere(c,r), if and only if ball(k,a) \leq_B ball(c,r). The *set Φ of generalized spheres* is introduced as follows.

Definition 3 The set Φ of generalized spheres in a linear metric space (L,d) is defined by $\Phi=\{(c,r): c\in L, r\in\mathbf{R}\}$.

Note that the members of Φ are only ordered pairs (c,r) , where $c\in L$ and $r\in(-\infty,+\infty)$; in other words a member of Φ is not a set of members of L . The two elements c and r of a generalized sphere (c,r) are called respectively *center* and *radius* of the generalized sphere.

Let Φ_0^+ denote a subset of Φ such that $\Phi_0^+=\{(c,r)\in\Phi: r\geq 0\}$. Apparently there exists a *bijection* (that is an one-to-one map and onto) $b_0^+: \Phi_0^+\leftrightarrow S$ between the sets Φ_0^+ and S and consequently there exists a bijection between the sets Φ_0^+ and B . Note that to a generalized sphere (c,r) there corresponds one ball, that is $\text{ball}(c,|r|)$ namely *the corresponding ball of the generalized sphere* (c,r) . A partial ordering relation $(c_1,r_1)\leq_{\Phi}(c_2,r_2)$ is defined in the set Φ of generalized spheres as follows.

Case-1 : $r_1, r_2 \geq 0$

Then $(c_1,r_1)\leq_{\Phi}(c_2,r_2) \Leftrightarrow \text{ball}(c_1,r_1)\leq_B\text{ball}(c_2,r_2)$.

Case-2 : $r_1, r_2 < 0$

Then $(c_1,r_1)\leq_{\Phi}(c_2,r_2) \Leftrightarrow \text{ball}(c_2,r_2)\leq_B\text{ball}(c_1,r_1)$.

Case-3 : $r_1 < 0, r_2 \geq 0$

Then $(c_1,r_1)\leq_{\Phi}(c_2,r_2) \Leftrightarrow \text{ball}(c_1,|r_1|)\cap\text{ball}(c_2,r_2)\neq\emptyset$. The latter relation signifies that a generalized sphere (c_1,r_1) with a negative radius is smaller than a generalized sphere (c_2,r_2) with a positive radius if and only if the corresponding balls of the two generalized spheres intersect each other.

It can be shown that the partly ordered set Φ of generalized spheres is a hidden-lattice with *underlying lattice* the set \mathbf{R} of real numbers and *corresponding map* r given by the radius of a generalized sphere. Moreover the function $v(x)=x$ constitutes a positive valuation in \mathbf{R} , hence both a *metric* and an *inclusion measure* are implied in the hidden-lattice Φ .

In summary, the previous considerations imply an injective monotone map $f: B\leftrightarrow\Phi$, from the partly ordered set B of balls to the hidden-lattice Φ ; the aforementioned injective monotone map f is given by $f(\text{ball}(c,r))=(c,r)$. More specifically the injective monotone map f implies “ $\text{ball}(k,a)\leq_B\text{ball}(c,r) \Rightarrow f(\text{ball}(k,a))\leq_{\Phi}f(\text{ball}(c,r)) \Rightarrow (k,a)\leq_{\Phi}(c,r)$ ”. Therefore both the *metric* and the *inclusion measure* of Φ are meaningful in B . Due to the practical orientation of this work detailed theoretical illustrations will be given elsewhere.

3.3 The *ds-FLN* Scheme

A learning scheme has been employed, namely *d-s Fuzzy Lattice Neurocomputing* scheme, or *ds-FLN* scheme for short, for the on-line estimation of stapes thickness. Note that the term “*Neurocomputing*” implies a neural implementation of a learning scheme which could also be implemented otherwise, e.g. algorithmically in software.

We remark that the *ds-FLN* scheme has been introduced under the name “two-level-fuzzy-lattice (*2L-FL*) scheme” in [20], where it was applied in another surgical procedure, that is the epidural puncture, for the recognition of soft tissues. Nevertheless we have decided to switch to the name “*ds-FLN*” herein so as to adhere to the terminology established subsequently within the *FL*-framework. In particular, the letters “*d*” and “ σ ” show that the training phase and the testing phase are carried out by employing a distance function (*d*) and an inclusion measure (σ), respectively, the initials “*FL*” denote a scheme applicable within the *FL*-framework, and the initial “*N*” signifies a neural implementation. We summarize qualitatively in the sequel the *ds-FLN* scheme for learning and classification in a hidden-lattice Φ whose underlying lattice is equipped with a positive valuation function. A quantitative description of the *ds-FLN* scheme will be shown using specific equations in the second part of this section.

Training-Phase of the $d\mathbf{s}$ -FLN scheme

Step-S1. A “lower level” supervised clustering is carried out in the training-set of hidden-lattice Φ elements by assuming that two training data are in the same cluster suffices their distance be less than a distance threshold T_d . Note that the initial value of T_d is defined by the user. A *contradiction* occurs when two data from different classes are put in the same cluster; then a decrease of T_d is triggered, followed by a new repetition of step-S1. Step-S1 keeps repeating until a contradiction-free cycle occurs which results in *homogeneous clusters*, these are clusters of data from a single class.

Step-S2. The “lower level” supervised clustering concludes by the *join* operation (\cup_{Φ}) of all hidden-lattice Φ data in the same cluster. Hence, one hidden-lattice Φ element is defined for each lower level cluster.

Step-S3. An “upper level” supervised learning is conducted by attaching the label of a concrete stapes thickness to each lower level cluster. The labels are attached by an external teacher. Note that more than one clusters may be assigned the same thickness label.

The Training-Phase of the $d\mathbf{s}$ -FLN scheme described above bases its decision making on a metric d . In the sequel, to test $d\mathbf{s}$ -FLN scheme’s capacity for generalization there follows the Testing-Phase which bases its decision making on an inclusion measure σ .

Testing-Phase of the $d\mathbf{s}$ -FLN scheme

All the data are fed one-by-one to the $d\mathbf{s}$ -FLN scheme and their degree of inclusion is calculated in all classes. The class which provides with the largest degree of lattice inclusion σ_{\max} is selected as the “winner class”, provided that σ_{\max} is larger than a user-defined threshold T_{σ} ; otherwise, if $\sigma_{\max} < T_{\sigma}$, no class is assigned to the current input data and the latter data are declared to be of an “undefined” class.

In conclusion, “learning” by the $d\mathbf{s}$ -FLN scheme is effected by *supervised clustering*. It has been found experimentally that “learning” requires only a few passes through the training data, and furthermore, “learning” claims certain neurocomputing qualities like parallel processing and generalization. The “worse case training scenario” would be to employ training data of different classes very near to each other. Then step-S1 of $d\mathbf{s}$ -FLN scheme’s training-phase will trigger repetitive decreases of T_d until all homogeneous learned clusters correspond to a single datum. The aforementioned “worse case training scenario” might arise as a result of “poor” data preprocessing and never occurred in the experiments described in the next section.

In the sequel are cited the equations involved in both training and testing by the $d\mathbf{s}$ -FLN scheme. The data to be processed are in the linear metric space M of 2×2 matrices presented in section 3.1.

Equations used in the Training-Phase of the $d\mathbf{s}$ -FLN scheme

• For each force-torque pair of drilling data profiles used for training matrices $F_j = \begin{bmatrix} f_{j,1} & f_{j,2} \\ f_{j,3} & f_{j,4} \end{bmatrix}$ and

$T_j = \begin{bmatrix} t_{j,1} & t_{j,2} \\ t_{j,3} & t_{j,4} \end{bmatrix}$ are calculated, where $j=1, \dots, n$, and n is the number of training data. Note that the entries of matrices F_j and T_j have been detailed in section 3.1 and they will not be repeated again.

• For each pair of matrices (F_j, T_j) , $j=1, \dots, n$ calculate matrix $M_j = T_j F_j^{-1}$, $j=1, \dots, n$. Matrices M_j are the data to be clustered.

• A “lower level” supervised clustering is carried out in the set of matrices M_j , $j=1, \dots, n$ by employing the nearest neighbor clustering algorithm illustrated in step-S1 of the Training-Phase of the $d\mathbf{s}$ -FLN scheme.

According to the details in section 3.1 note that the distance used between two matrices $M_i = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ and

$M_j = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ has been the *spectral distance* which is calculated by applying the formula $d(M_i, M_j) =$

$$d\left(\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}\right) = \sqrt{\frac{\sum_{k=1}^4 (a_k - b_k)^2 + \sqrt{[\sum_{k=1}^4 (a_k - b_k)^2]^2 - 4[(a_1 - b_1)(a_4 - b_4) - (a_2 - b_2)(a_3 - b_3)]^2}}{2}}$$

• The aforementioned formulas are employed repetitively as illustrated in step-S1 of the Training-Phase of the *ds-FLN* scheme until all clusters resulting in are homogeneous.

• Let the j^{th} homogeneous cluster include the matrices $M_{ji}, i=1, \dots, n_j$ where n_j is the number of matrices in the j^{th} (homogeneous) cluster. Consider the trivial balls, $\text{ball}(M_{j1}, 0), \dots, \text{ball}(M_{jn_j}, 0)$. Replace the aforementioned balls by their circumscribed $\text{ball}(c_j, r_j)$, where $c_j = \frac{1}{n_j} \sum_{i=1}^{n_j} M_{ji}$ and $r_j = \max_{i=1}^{n_j} \{d(c_j, M_{ji})\}$.

By the conclusion of the Training-Phase, N homogeneous balls $(c_j, r_j), j=1, \dots, N$ have been calculated, where c_j is a 2×2 matrix of real numbers and r_j is a non-negative real number. Recall that each $\text{ball}(c_j, r_j), j=1, \dots, N$ has been assigned a stapes thickness label in step-S3.

Equations used in the Testing-Phase of the *ds-FLN* scheme

• For a stapes being drilled, whose thickness is to be estimated by classification, two matrices $F_k = \begin{bmatrix} f_{k,1} & f_{k,2} \\ f_{k,3} & f_{k,4} \end{bmatrix}$ and $T_k = \begin{bmatrix} t_{k,1} & t_{k,2} \\ t_{k,3} & t_{k,4} \end{bmatrix}$ are calculated are every 20 samples. Note that the entries of matrices F_k and T_k have been detailed in section 3.1 and they will not be repeated again.

• Matrix $M_k = T_k F_k^{-1}$ is computed. In the sequel the degree of inclusion of $\text{ball}(M_k, 0)$ to $\text{ball}(c_j, r_j), j=1, \dots, N$ is calculated as follows.

• For each pair $(\text{ball}(M_k, 0), \text{ball}(c_j, r_j)), j=1, \dots, N$ calculate the circumscribed $\text{ball}(c'_j, r'_j)$, where $c'_j = \frac{1}{2}(M_k + c_j)$, and $r'_j = \frac{1}{2}(d(M_k, c_j) + r_j), j=1, \dots, N$.

• The degree of inclusion of $\text{ball}(M_k, 0)$ in $\text{ball}(c_j, r_j)$ is calculated by $\sigma(\text{ball}(M_k, 0) \leq \text{ball}(c_j, r_j)) = \frac{r_j}{r'_j}, j=1, \dots, N$.

In conclusion matrix M_k , and consequently the stapes it has stemmed from, is classified to one ball among $\text{ball}(c_j, r_j), j=1, \dots, N$, that is in particular the ball which corresponds to the largest degree of inclusion

$\frac{r_j}{r'_j}, j=1, \dots, N$, and finally the label (thickness) of the winner ball is assigned to the stapes bone being drilled.

As it will explained in the next section a winner's label might change as the drilling progresses. The decisive classification is made at the moment of breakthrough. Note that a stapes breakthrough can be easily detected by the rapid changes in the values of force and torque.

4 EXPERIMENTS AND RESULTS

4.1 Data Acquisition

The data have been obtained at the University of Bristol. Different sample materials have been considered such as metals, ceramics, various composites, plastics, and animal bones. Animal bone has been used as the most appropriate material due to its inherent similarity to human bone. The thickness (t) of a test piece had been in the range $0.2-2\text{ mm}$, whereas the dimensions of a test piece other than thickness are unimportant.

To acquire the data, a drill with force, torque, and position sensors was fed at a constant feed rate through a test piece representing the stapes footplate. The drill had a spherical burr drill bit of 0.6 mm diameter, and a linear feed actuator repeatable to within $\pm 0.1\text{ mm}$. The force was measured to an accuracy of $\pm 0.05\text{ N}$, and the torque to an accuracy of $\pm 0.01\text{ Nmm}$. Note that units of System International (SI) have been used, these are in particular *Newton (N)* for the force, and *Newton-millimeter (Nmm)* for the torque. Both a bone sample to be drilled and the drill unit were mounted on a test rig simulating the compliance of an in-situ stapes. For details regarding the design and operation of the mechatronic drilling tool the reader may refer to [1], [23].

A series of experiments was carried out for an array of parameters as explained in the following section. The drilling data have been obtained as follows. The drill was advanced until it was in contact with and perpendicular to a test piece. Then the drill was set in motion under computer control with a rotational subacoustic speed of 2 Hz . The feed rate was kept constant in the range $0.1-10\text{ mm/min}$. Drilling was continuous and stopped when the drill bit was clear of the far surface. The displacement, in the range $0-6\text{ mm}$, the force, in the range $0-10\text{ N}$, and the torque, in the range $0-5\text{ Nmm}$, were logged throughout the drilling. Note that the measured displacement has been in a larger range ($0-6\text{ mm}$) than the range of the stapes thickness ($0.2-2\text{ mm}$) because the flexible stapes has yielded under drilling tool action.

4.2 Experiments

Twelve groups of drilling data have been obtained for all combinations of 3 values of stapes thickness (t) and 4 values of stapes stiffness (K). Regarding thickness, in particular, the typical values 0.5 mm , 1 mm , and 1.5 mm have been dealt with, whereas the corresponding stiffness values had been 1030 N/m , 2440 N/m , 4900 N/m , and 11900 N/m . Each one of the 12 groups of data included 5 force/torque pairs of drilling profiles, resulting in a total of 60 pairs of data profiles. Three outliers, that is 5 % of the data, were removed in all subsequent experiments. Note that a group of drilling data was characterized by two labels, these labels are (1) the corresponding stapes thickness, and (2) the corresponding stapes stiffness.

One group of drilling measurements has been treated as one class. That is 12 classes had to be learned by the *dS-FLN* scheme, where a class has been characterized by two labels : (1) a stapes thickness, and (2) a stapes stiffness. In the sequel during testing, and since the stapes stiffness (K) is measurable before drilling, the drilling data of a stapes had to be classified to one of the three classes which correspond to stapes thickness 0.5 mm , 1.0 mm , and 1.5 mm . Consequently the problem of estimating the thickness of a stapes has been treated as a pattern classification problem.

Both in the training- and in the testing- phase of the *dS-FLN* scheme the measurable input data have been : (1) drilling force versus displacement, (2) drilling torque versus displacement, (3) the stapes stiffness, and (4) the displacement at a stapes breakthrough. Regarding the value of “displacement at a stapes breakthrough” we remark that in the training-phase the displacement at breakthrough was given, whereas in the testing-phase the breakthrough could be detected automatically by the rapid changes in the values of force and torque.

Learning has been effected during the training-phase in the linear metric space M of 2×2 matrices which map the entries of a matrix F to the entries of a matrix T as it has been explained in section 3.1. Recall

that in this work the stiffness K of a stapes had to be measured before the drilling operation in order the stiffness value to be employed as one of the entries of matrix T (for details see in section 3.1). Moreover, we remark that an estimate of the stapes stiffness is not considered “dependable” for determining a stapes thickness, because “... a (stapes) compliance value determined from a stiffness test is not always representative of that for the drilling process as a whole ... [1]”.

For each one of the 12 classes, a “training part” of data in a class has been specified comprising at least 1 pair of force/torque drilling profiles. On the one hand, the overall training-set was the collection of the “training parts” from all 12 classes. On the other hand, the overall testing-set comprised all the pairs of force/torque drilling profiles. In this way, during the testing-phase it has been possible to grade both $d\mathbf{s}$ - FLN 's ability to learn and its capacity to infer sensible conclusions with regards to new and hitherto unknown inputs.

Six learning experiments have been carried out using the $d\mathbf{s}$ - FLN scheme on the 57 pairs of drilling force/torque profiles by employing one, two, and all the force/torque profiles from each of the 12 classes of drilling profiles in two, three, and one experiments, respectively (where applicable, the training data have been selected randomly). After the training-phase, in each learning experiment, 12 homogeneous clusters had been resulted in that is in particular one homogeneous cluster per group of drilling profiles. Therefore each cluster resulting from “training” corresponded to a specific combination of stapes stiffness and stapes thickness. During the subsequent testing-phase, a matrix $M(t)$ was being defined on-line every 20 samples from whatever drilling data were available up to that moment and then the degree of inclusion of $M(t)$ in the learned classes was being calculated. Since the stiffness of the “testing stapes” was known, the set of learned clusters to be attended has been restricted to the clusters labeled by the same stiffness as the stiffness of the “testing stapes”.

4.3 Results

Fig.2 shows typical testing examples for the on-line classification of stapes thickness using the $d\mathbf{s}$ - FLN scheme. The data in Fig.2 correspond to stapes thickness 0.5 mm (column a), 1 mm (column b), and 1.5 mm (column c). None of the examples in Fig.2 had been in the training set. Figures 2 (a1), (b1), (c1) and Figures 2 (a2), (b2), (c2) show, respectively, force and torque drilling profiles. The three curves in each one of the Figures 2 (a3), (b3), (c3) correspond to three thickness classes, in particular 0.5 mm (dotted curve), 1.0 mm (dashed curve), and 1.5 mm (solid curve). Each curve indicates the degree of inclusion of the “currently” registered force/torque profile in the three known stapes thicknesses versus displacement.

The behavior of the classification decisions during “drilling” changes as the drilling progresses (Fig.2 (a3), (b3), (c3)). In the latter figures the dotted line (thickness 0.5 mm) leads the way, followed by the dashed line (thickness 1.0 mm), which is followed by the solid line (thickness 1.5 mm). Nevertheless the “winner class” at the time of stapes breakthrough, denoted in Fig.2 by a vertical dash-dotted line, assigns its label (thickness) to the thickness of the stapes being currently drilled.

In all the experiments only one misclassification has been recorded. It occurred when the stiffness parameter $K=1030 N/m$ was dealt with, and one training datum per class was employed. Then one of the 57 testing pairs of profiles has been misclassified in class “0.5 mm” instead of being classified to the correct class “1.0 mm”. The aforementioned misclassification has been attributed by the authors to the small number of data used for training. By increasing the number of training data the performance improved to an 100 % correct classification. Note that in all the experiments and during testing, the training data have been classified correct in their corresponding classes.

It should be stressed that the work herein is only a feasibility study and further trials are needed for assessing the system's performance under more realistic conditions. In any case, the presented $d\mathbf{s}$ - FLN scheme is able to produce an estimate of the stapes thickness and furthermore the scheme in question provides a degree of certainty to the estimate it has produced by issuing a number between zero and one, that is the inclusion measure σ . Nevertheless the final decision should be taken by the operating surgeon.

5 DISCUSSION AND CONCLUSION

High level controls have been designed and tested for an intelligent mechatronics drill to be utilized in stapedotomy surgery. The specific technical goal has been the on-line estimation of a stapes bone thickness from drilling force- and torque- data.

The aforementioned technical goal has been met without employing a model of the physical system but, instead, a machine-learning scheme for classification has been employed, namely *d-s Fuzzy Lattice Neurocomputing (ds-FLN)* scheme. The on-line estimation of a stapes thickness has been demonstrated on force- and torque- drilling profiles obtained experimentally.

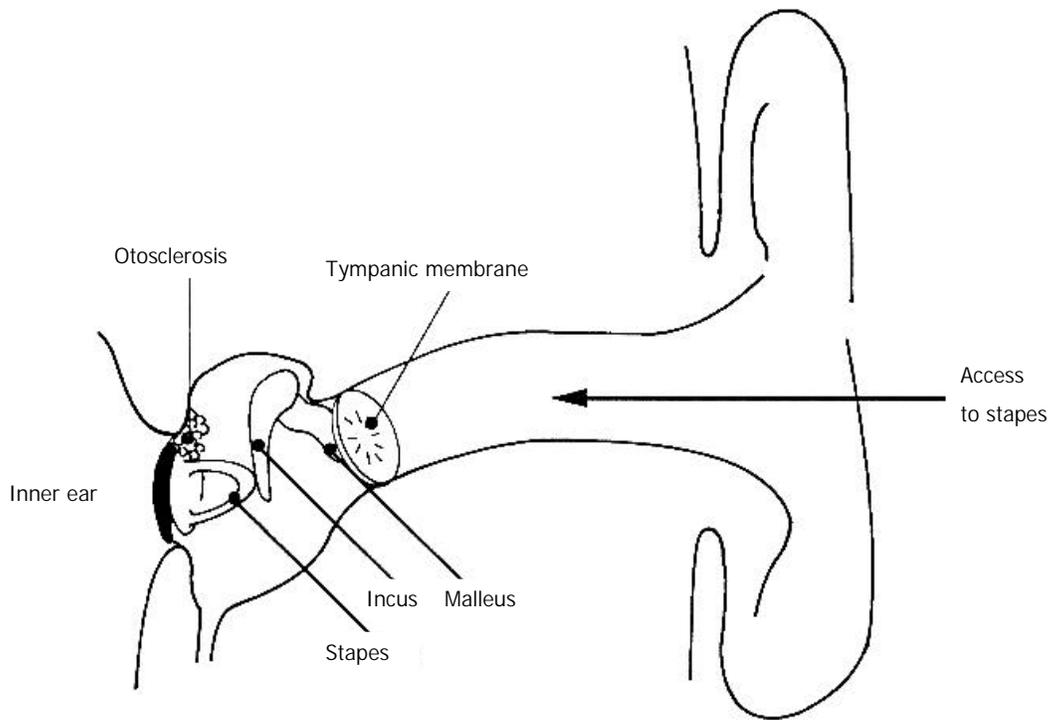
A theoretical contribution of the work herein has been the introduction of the notions *generalized sphere* and *hidden-lattice*. From a practical point of view, note that even though the inputs to the *ds-FLN* in this work have been solely “trivial balls”, the latter are individual point measurements, the *ds-FLN* can also accommodate non-trivial balls. Hence, “uncertainty” in the measurements can be accommodated by feeding the *ds-FLN* scheme a “ball”, that is a neighborhood of data, instead of feeding it an individual point measurement.

A significant advantage of the *ds-FLN* scheme is that disparate data can be dealt with and jointly. For instance the data in question might include waveforms, images, etc. stemming from measurement equipment as well as fuzzy set-, linguistic- data and/or logic statements issued by a human, thus implying a quite sophisticated capacity for learning and decision-making by a machine. Applications other than surgical ones could also be considered.

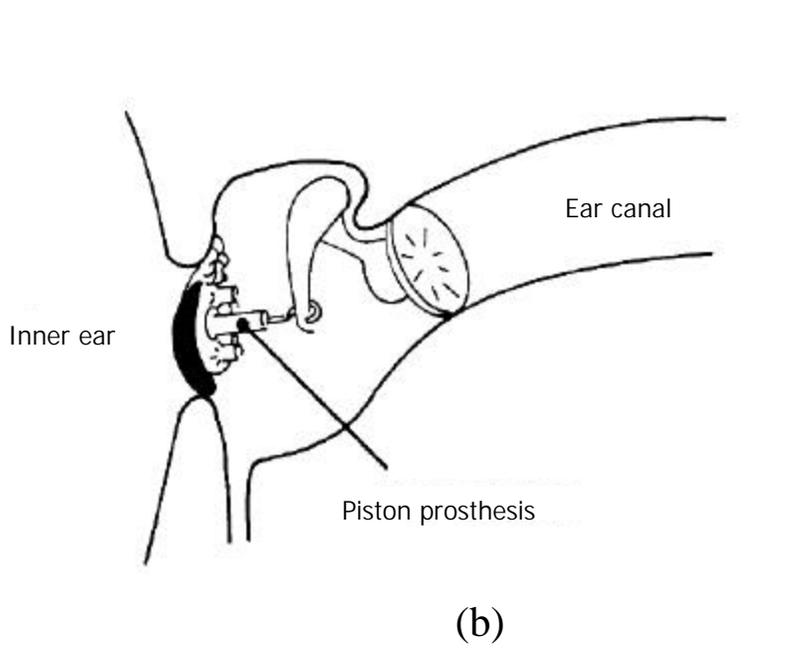
REFERENCES

- [1] D. Baker, P.N. Brett, M.V. Griffiths, and L. Reyes, “A Mechatronic Drilling Tool for Ear Surgery : A Case Study of Some Design Characteristics”, *Mechatronics*, vol. 6 (4), pp. 461-477, 1996.
- [2] E. Baralis, S. Ceri, and S. Paraboschi, “Compile-Time and Runtime Analysis of Active Behaviors”, *IEEE Trans. Knowledge Data Engineering*, vol. 10 (3), pp. 353-370, 1998.
- [3] M. Bianchini, P. Frasconi, and M. Gori, “Learning Without Local Minima in Radial Basis Function Networks”, *IEEE Trans. Neural Networks*, vol. 6 (3), pp. 749-756, 1995.
- [4] P.B. Brazdil, and J. Gama, “Continued Comparative Studies of Machine Learning, Neural, and Statistical Classification Algorithms”. Available at <ftp://ftp.ncc.up.pt/pub/statlog/>.
- [5] R. Bronson, *Schaum's Outline Series - Matrix Operations*, New York, N.Y.: McGraw-Hill, 1989.
- [6] P.E. Caines, and Y-J Wei, “Hierarchical Hybrid Control Systems: A Lattice Theoretic Formulation”, *IEEE Trans. Automatic Control*, vol. 43 (4), pp. 501-508, 1998.
- [7] G. Carpenter, “Distributed Learning, Recognition, and Prediction by ART and ARTMAP Neural Networks”, *Neural Networks*, vol. 10 (8), pp. 1473-1494, 1997.
- [8] J.B. Causse, S. Gherini, K.L. Horn, “Surgical Treatment of Stapes Fixation by Fiberoptic Argon Laser Stapedotomy with Reconstruction of the Annular Ligament”, *Otolaryngologic Clinics of North America*, vol. 26, no. 3, pp. 395-416, 1993.
- [9] B.A. Davey, and H.A. Priestley, *Introduction to Lattices and Order*, Cambridge University Press, 1990.
- [10] DELVE (Data for Evaluating Learning in Valid Experiments), Copyright (c) 1995-1996 by The University of Toronto, Toronto, Ontario, Canada. Available at <http://www.cs.utoronto.ca/~delve>.
- [11] R.M. Gray, “Vector Quantization”, *IEEE ASSP Magazine*, vol. 1, pp. 4-29, 1984.
- [12] Q.Q. Huynh, L.N. Cooper, N. Intrator, and H.Shouval, “Classification of Underwater Mammals Using Feature Extraction Based on Time-Frequency Analysis and BCM Theory”, *IEEE Trans. Signal Processing*, vol. 46 (5), pp. 1202-1207, 1998.

- [13] A. Joshi, N. Ramakrishnan, E.N. Houstis, and J.R. Rice, "On Neurobiological, Neuro-Fuzzy, Machine Learning, and Statistical Pattern Recognition Techniques", *IEEE Trans. Neural Networks*, vol. 8, no. 1, pp. 18-31, 1997.
- [14] V.G. Kaburlasos, V. Petridis, P. Brett, and D. Baker, "Learning a Linear Association of Drilling-Profiles in Stapedotomy Surgery", *Proc. of the IEEE 1998 International Conference on Robotics & Automation (ICRA'98)*, Leuven, Belgium, 16-20 May 1998, pp. 705-710.
- [15] V.G. Kaburlasos, and V. Petridis, "Fuzzy Lattice Neurocomputing (FLN) : A Novel Connectionist Scheme for Versatile Learning and Decision Making by Clustering", *International Journal of Computers and Their Applications*, vol. 4, no. 2, pp. 31-43, 1997.
- [16] W-H Kim, T. Nguyen, and Y-H Hu, "Adaptive Wavelet Packet Based Image Coding with Optimal Entropy-Constrained Lattice Vector Quantizer (ECLVQ)", *IEEE Trans. Signal Processing*, vol. 46 (7), pp. 2019-2026, 1998.
- [17] Y.S.Kwoh, J. Hou, E.A. Jonckheere, and S. Hayti, "A Robot with Improved Absolute Positioning Accuracy for CT Guided Stereotactic Brain Surgery", *IEEE Trans. Biomedical Engineering*, vol 35 (2), pp. 153-160, 1988.
- [18] R.R. Murphy, "Dempster-Shafer Theory for Sensor Fusion in Autonomous Mobile Robots", *IEEE Trans. Robotics and Automation*, vol. 14 (2), pp. 197-206, 1998.
- [19] V. Petridis, and V.G. Kaburlasos, "Fuzzy Lattice Neural Network (FLNN): A Hybrid Model for Learning", *IEEE Trans. Neural Networks*, vol. 9 (5), pp. 877-890, 1998.
- [20] V. Petridis, V.G. Kaburlasos, P. Brett, T. Parker, and J.C.C. Day, "Two Level Fuzzy Lattice (2L-FL) Supervised Clustering : A New Method for Soft Tissue Identification in Surgery", *Proc. CESA'96 IMACS Multiconference*, Lille, France, 9-12 July 1996, pp. 232-237.
- [21] G.N. Saridis, "Intelligent Robotic Control", *IEEE Trans. Automatic Control*, vol. 28 (5), pp. 547-557, 1983.
- [22] M. Setnes, R. Babuška, U. Kaymak, and H.R. van Nauta Lemke, "Similarity Measures in Fuzzy Rule Base Simplification", *IEEE Trans. Systems, Man, and Cybernetics B*, vol. 28 (3), pp. 376-386, 1998.
- [23] R.S. Stone, P.N. Brett, B.S. Evans, "An Automated Handling-System for Soft Compact Shaped Non-rigid Products", *Mechatronics*, vol. 8 (2), pp. 85-102, 1998.
- [24] R.S. Wenocur, and R.M. Dudley, "Some Special Vapnik-Chervonenkis Classes", *Discrete Mathematics*, vol. 33, pp. 313-318, 1981.
- [25] J.R. Williamson, "Gaussian ARTMAP: A Neural Network for Fast Incremental Learning of Noisy Multidimensional Maps", *Neural Networks*, vol. 9 (5), pp. 881-897, 1996.



(a)



(b)

Figure 1 (a) Human ear-gross anatomy.
(b) Ossicular chain.

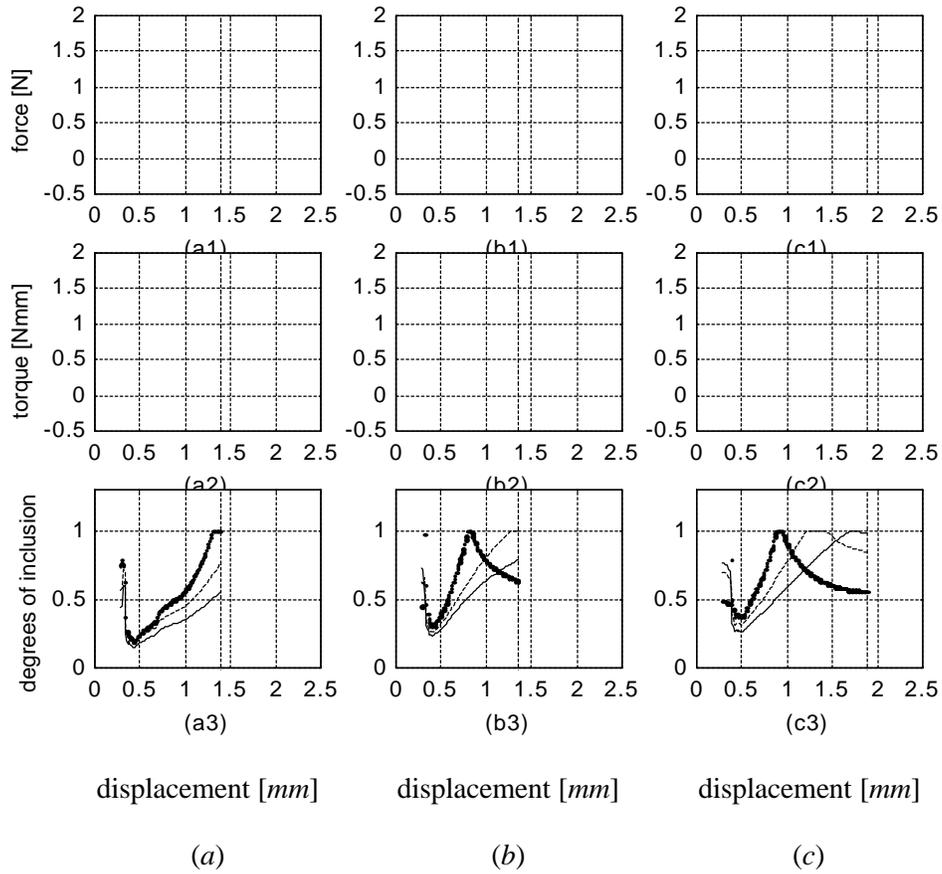


Figure 2 Force/torque drilling profiles and inclusion measures for thickness (a) 0.5 mm , (b) 1.0 mm , and (c) 1.5 mm . The first and second figures in a column show respectively a force and a torque drilling profile. The third figure in a column shows the degrees of inclusion to the three classes: 0.5 mm (dotted line), 1.0 mm (dashed line), and 1.5 mm (solid line). The maximum inclusion measure at the time of stapes-breakthrough, the latter time is denoted in every figure by a vertical dash-dotted line, determines the “winner class” and hence it determines the thickness value of the stapes being drilled.