Image reconstruction algorithms for electrical capacitance tomography

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2003 Meas. Sci. Technol. 14 R1
(http://iopscience.iop.org/0957-0233/14/1/201)
View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 134.151.40.2
This content was downloaded on 31/01/2014 at 10:39

Please note that terms and conditions apply.
REVIEW ARTICLE

Image reconstruction algorithms for electrical capacitance tomography

W Q Yang\(^1\),\(^3\) and Lihui Peng\(^2\)

\(^1\) Department of Electrical Engineering and Electronics, UMIST, PO Box 88, Manchester M60 1QD, UK
\(^2\) Department of Automation, Tsinghua University, Beijing 100084, China

E-mail: w.yang@umist.ac.uk

Received 24 July 2002, in final form 9 October 2002, accepted for publication 4 November 2002
Published 11 December 2002
Online at stacks.iop.org/MST/14/R1

Abstract

Electrical capacitance tomography (ECT) is used to image cross-sections of industrial processes containing dielectric material. This technique has been under development for more than a decade. The task of image reconstruction for ECT is to determine the permittivity distribution and hence material distribution over the cross-section from capacitance measurements. There are three principal difficulties with image reconstruction for ECT: (1) the relationship between the permittivity distribution and capacitance is non-linear and the electric field is distorted by the material present, the so-called ‘soft-field’ effect; (2) the number of independent measurements is limited, leading to an under-determined problem and (3) the inverse problem is ill posed and ill conditioned, making the solution sensitive to measurement errors and noise. Regularization methods are needed to treat this ill-posedness. This paper reviews existing image reconstruction algorithms for ECT, including linear back-projection, singular value decomposition, Tikhonov regularization, Newton–Raphson, iterative Tikhonov, the steepest descent method, Landweber iteration, the conjugate gradient method, algebraic reconstruction techniques, simultaneous iterative reconstruction techniques and model-based reconstruction. Some of these algorithms are examined by simulation and experiment for typical permittivity distributions. Future developments in image reconstruction for ECT are discussed.

Keywords: electrical capacitance tomography, image reconstruction, iterative algorithm, inverse problem

List of symbols

\(f\) function
\(F\) transform function from permittivity distribution to capacitance
\(g\) normalized permittivity vector
\(\hat{g}\) reconstructed permittivity vector
\(\bar{g}, \bar{\hat{g}}\) mean values of \(g\) and \(\hat{g}\) respectively
\(\bar{g}\) estimated solution
\(\tilde{g}\) solution of \(g\) in null space

\(^3\) Author to whom any correspondence should be addressed.
$$I$$ identity matrix  
$$J$$ Jacobian matrix  
$$k$$ iteration step number  
$$M$$ number of measurements  
$$n$$ number of electrodes  
$$N$$ number of pixels  
$$p$$ rank of $$S$$  
$$P$$ projection operator  
$$Q$$ electric charge  
$$s$$ sensitivity of capacitance transducer  
$$s_k$$ $$k$$th row of $$S$$  
$$S$$ normalized sensitivity matrix  
$$u_k$$ identity normalized capacitance vector  
$$u_m$$ $$m$$th column of $$U$$  
$$U, \Sigma, V$$ orthogonal and diagonal matrices for singular value decomposition  
$$v_n$$ $$n$$th column of $$V$$  
$$V$$ potential difference between two electrodes  
$$w_i$$ weighting factor  
$$x$$ variable vector  
$$\alpha, \beta, \gamma$$ relaxation factors  
$$\delta_i$$ $$i$$th singular value of $$S$$  
$$\epsilon(x, y)$$ permittivity distribution  
$$\epsilon_0$$ permittivity of vacuum space  
$$\epsilon_r$$ relative permittivity of material  
$$O((\Delta \epsilon)^2)$$ second order term in a differential equation  
$$O((\Delta \epsilon)^3)$$ and higher order terms  
$$\phi(x, y)$$ potential distribution  
$$\Gamma$$ electrode surface  
$$\lambda$$ normalized capacitance vector  
$$\lambda_k$$ $$k$$th component of $$\lambda$$  
$$\lambda_o$$ noise-contaminated $$\lambda$$  
$$\lambda^*$$ calculated vector of normalized capacitance  
$$\mu$$ Tikhonov regularization parameter  
$$\theta, L, r$$ model parameters  
$$\xi$$ function

1. Introduction

Electrical capacitance tomography (ECT) is based on measuring capacitances of a multi-electrode sensor surrounding an industrial vessel or pipe containing two materials of different permittivities. These measurements are used to reconstruct the permittivity distribution and hence the material distribution over the cross section, using a suitable algorithm. Investigations into ECT date back to the 1970s. The first reported ECT system was developed by the US Department of Energy in Morgantown, to image the gas/solids distribution in fluidized beds (Fasching and Smith 1991, Halow et al 1993, Fasching et al 1994). The first real-time ECT system was developed at UMIST in the late 1980s to image gas/oil pipelines (Huang et al 1988, 1992, Xie et al 1992). Subsequently, the UMIST ECT system has been used successfully in a number of research investigations for imaging other two-phase processes, such as

- combustion flames in engine cylinders (He et al 1994, Waterfall et al 1996);
- a water hammer (Yang et al 1996).

In the past decade, research into ECT has been carried out in several countries (Beck and Williams 1996, Beck et al 1997, Hammer and Johansen 1997, Beck et al 1998, Reinecke and Mewes 1996, Reinecke et al 1998, York 2001). The institutions involved in development of hardware and/or software of ECT systems are given in Table 1.

A typical ECT system, as shown in Figure 1, comprises three main units:

1. a multi-electrode sensor;
2. the sensing electronics and image reconstruction;
3. a computer for hardware control and data processing, including image reconstruction.

A set of electrodes, usually eight or 12, is mounted around a pipe or vessel and the capacitance values between all single-electrode combinations are measured. The sensing electronics provides excitation signals and converts the capacitances into voltage signals, which are conditioned and then digitized for data acquisition. The computer controls the system hardware and implements image reconstruction to show the permittivity distribution.

There are two major computational problems in ECT: the forward problem and the inverse problem. The forward problem is to determine inter-electrode capacitances from the permittivity distribution, e.g. by solving the partial differential equations, governing the sensing domain. The inverse problem is to determine the permittivity distribution from capacitance measurements. The result is usually presented as a visual image, and hence this process is called image reconstruction. Current image reconstruction methods, particularly iterative methods, also involve solving the forward problem.

The relationship between capacitance and permittivity distribution is governed by the following equation:

$$C = \frac{Q}{V} = \frac{1}{V} \int_{\Gamma} \epsilon(x, y) \nabla \phi(x, y) \, d\Gamma$$

(1)

where $$\epsilon(x, y)$$ is the permittivity distribution in the sensing field, $$V$$ is the potential difference between two electrodes forming the capacitance, $$\phi(x, y)$$ is the potential distribution and $$\Gamma$$ is the electrode surface.

In some cases, equation (1) can be simplified. For example, for an ideal parallel-plate sensor with a homogeneous permittivity distribution, it becomes

$$C = \epsilon_0 \epsilon_r \frac{A}{d}$$

(2)

where $$\epsilon_0$$ is the permittivity of vacuum, $$\epsilon_r$$ is the relative permittivity of the material inside the sensor, $$A$$ is the area of the plates and $$d$$ is the distance between the two plates.

Equation (2) implies such a simple relationship that capacitance is proportional to permittivity. However, the geometry of an ECT sensor is much more complicated than that of a parallel-plate sensor and the permittivity distribution is generally not uniform. In this case, equation (1) cannot be simplified. In equation (1), the potential distribution $$\phi(x, y)$$ also depends on the permittivity distribution $$\epsilon(x, y)$$. 

$$\epsilon_r$$ relative permittivity of material
Figure 1. Typical ECT system with an eight-electrode sensor, with sensing field boundaries shown.

Table 1. Institutions involved in development of ECT systems.

<table>
<thead>
<tr>
<th>Country</th>
<th>Institution</th>
<th>Hardware</th>
<th>Algorithma</th>
<th>Application</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Northeast University</td>
<td>AC-based circuit</td>
<td>LBP, linear regression and regularization</td>
<td></td>
<td>Yan et al (2001)</td>
</tr>
<tr>
<td></td>
<td>Zhejiang University</td>
<td>Charge transfer</td>
<td>LBP, ART, SIRT, electrical field lines</td>
<td>Oil pipeline</td>
<td>Tikhonov and Arsenin (1977)</td>
</tr>
<tr>
<td>Germany</td>
<td>University of Hannover</td>
<td>Capacitance meter/impedance analyser</td>
<td>LBP, ART, SIRT, electrical field lines</td>
<td>Trickle bed</td>
<td>Loser et al (2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holland</td>
<td>Delft University of Technology</td>
<td>Active differentiator</td>
<td>LBP</td>
<td>Fluidized bed</td>
<td>Kuhn and van Halderen (1997)</td>
</tr>
<tr>
<td>Mexico</td>
<td>Institute Mexico Petroleum</td>
<td>AC-based circuit</td>
<td>LBP</td>
<td>Oil pipeline</td>
<td>Gamio (2002)</td>
</tr>
<tr>
<td>Norway</td>
<td>University of Bergen</td>
<td>Self-balancing AC-based circuit</td>
<td>Model-based iteration</td>
<td>Oil pipeline, oil separator</td>
<td>Hammer and Johansen (1997)</td>
</tr>
<tr>
<td>UK</td>
<td>Manchester Metropolitan University</td>
<td>AC-based circuit</td>
<td>LBP</td>
<td></td>
<td>Deloughtry et al (2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>US Department of Energy, Morgantown</td>
<td>High voltage AC bridge</td>
<td></td>
<td>Fluidized bed</td>
<td>Fasching and Smith (1991)</td>
</tr>
</tbody>
</table>

a The abbreviations are in the main text.

Therefore, the capacitance between electrode combinations can be considered as a functional of permittivity distribution.

\[ C = \xi(\varepsilon). \]  

(3)

The change in capacitance in response to a perturbation of the permittivity distribution is given by

\[ \Delta C = \frac{d\xi}{d\varepsilon} (\Delta \varepsilon) + O((\Delta \varepsilon)^2). \]  

(4)
where $\frac{dC}{d\varepsilon}$ is the sensitivity of the capacitance versus permittivity distribution and $O((\Delta \varepsilon)^2)$ represents $(\Delta \varepsilon)^2$ and higher order terms.

In ECT applications, $\Delta \varepsilon$ is usually small. By neglecting $O((\Delta \varepsilon)^2)$, equation (4) can be simplified to the linear form.

$$\Delta C = s \Delta \varepsilon$$ (5)

where $s = \frac{dC}{d\varepsilon}$ is the sensitivity of the capacitance transducer to changes in permittivity.

Equation (5) has to be discretized for implementation. In order to visualize the permittivity distribution, the sensing area is divided into $N$ elements or pixels, typically of the order of 1000. For example, a $32 \times 32$ grid generates 1024 pixels in the image area of a square sensor and about 800 pixels for a circular sensor.

Using an eight-electrode system as an example, capacitance measurements are obtained in the following steps. A voltage signal is first applied to electrode 1, and then the electric charges on electrodes 2–8 respectively are measured, to obtain the capacitances between electrode 1 and the other seven electrodes. Next electrodes 2–7 are energized in sequence, so that capacitances of all single-electrode pairs are measured, giving a total of 28 independent electrode combinations. There are $M = n(n - 1)/2$ independent capacitance measurements for an $n$-electrode sensor, giving $M$ equations in the form of equation (5). The linearized and discrete form of the forward problem can now be expressed as

$$\Delta C = J \Delta \varepsilon$$ (6)

where $J$ is a Jacobian matrix, i.e. the sensitivity distribution matrix, giving a sensitivity map for each electrode pair.

By this treatment, the non-linear forward problem has been simplified to a linear approximation. This satisfies most applications with a small permittivity contrast or perturbation. Commonly, equation (6) is written in a normalized form (Xie et al 1992).

$$\lambda = S g$$ (7)

where $\lambda$ is the normalized capacitance vector, $S$ is the Jacobian matrix of normalized capacitance with respect to normalized permittivity, i.e. the transducer sensitivity, and $g$ is the normalized permittivity vector, i.e. the grey level of pixels for visualization.

The matrix $S$ contains $M$ sensitivity distributions, but with a symmetrical electrode system there will be only $n/2$ unique distributions and other distributions can be obtained by rotating and mirroring. Figure 2 shows two typical sensitivity distributions for an adjacent electrode pair and an opposing electrode pair respectively, which have been obtained using a finite element software package—PC-Opera from Vector Fields Ltd (1992), based on a sensitivity theorem (Xie et al 1992).

The task of image reconstruction for ECT is to determine the permittivity distribution $\varepsilon(x, y)$ from the measured capacitance vector $C$. In the discrete form, it is to find the unknown $g$ from the known $\lambda$ using equation (7), while $S$ is treated as a constant matrix for simplicity. Note that $S$ will change with permittivity distribution.

There are two major difficulties associated with equation (7). Firstly, it is under-determined because the number of unknown variables $N$ (i.e. the number of pixels) is usually much larger than the number of equations $M$ (i.e. the number of capacitance measurements). Therefore, the solution is not unique. Secondly, equation (1) is an integral equation, which is ill posed, and its corresponding discrete form, equation (7), is ill conditioned. This means that the solution of equation (7) is sensitive to small perturbations of $\lambda$ (Tikhonov and Arsenin 1977, Engl et al 1996). In addition, the matrix $S$ is not truly constant, but varies with the actual permittivity distribution.

Image reconstruction algorithms for ECT were reviewed some years ago (Isaksen 1996). Since then a number of new or different algorithms have been developed to address the ill posed and ill conditioned problems. In general, they can be categorized into two groups, non-iterative (or single step) algorithms and iterative algorithms.

2. Non-iterative algorithms

2.1. Linear back-projection (LBP)

If the inverse of $S$ existed, equation (7) would be solved directly by

$$g = S^{-1} \lambda.$$ (8)

However, the inverse of $S$ does not exist and so other methods for solution must be used. The first attempt to reconstruct images for ECT was made using linear back-projection (Xie et al 1992). If $S$ is considered to be a linear mapping from the permittivity vector space to the capacitance vector space, $S^T$ can be considered as a related mapping from the capacitance vector space to the permittivity vector space, giving an approximated solution.

$$\hat{g} = S^T \lambda.$$ (9)

A normalized or weighted form of equation (9) can be expressed as
where \( u_n \) is an identity vector. The division of the two vectors \( S^T \lambda \) and \( S^T u_n \) is defined as one numerator component being divided by the corresponding denominator component. Note that mathematically \( \hat{g} \) is only an approximate solution of equation (7).

If the sensor is filled first with a lower permittivity material and then with a higher permittivity material, the normalized change in capacitance must be \( u_n \), and the corresponding change in permittivity is \( S^T u_n \). If the normalized measured capacitance is \( \lambda \), the corresponding change in permittivity is given by equation (10). Although mathematically not accurate, the LBP algorithm is still widely used for on-line image reconstruction because of its simplicity. However, it produces poor-quality images and can only provide qualitative image reconstruction because of its simplicity. However, it is not unique, because the inverse problem is ill posed and under-determined. Suppose that \( \hat{g} \) is in the null space of \( S \), then

\[
S\hat{g} = 0
\]

(18)

where \( \hat{g} \) can be expressed as a linear combination of column vector \( v_{p+1}, \ldots, v_N \), i.e., the column vector of orthogonal matrix \( V \).

If \( \hat{g} \) is the solution of equation (7) or (11), \( \hat{g} + \hat{g} \) must also be a solution. The solution according to equations (14)–(16) is the minimum norm solution. However, it is only acceptable mathematically.

Truncated singular value decomposition (TSVD) may improve this situation, especially when an obvious gap exists between the singular values. Equation (17) may be modified as

\[
\Sigma^{-1} = \text{diag} \left[ \frac{1}{\delta_1^2}, \frac{1}{\delta_2^2}, \ldots, \frac{1}{\delta_{p-1}^2}, \frac{1}{\delta_p^2} \right]
\]

(19)

\[
w_i = \frac{\delta_i^2}{\delta_i^2 + \mu} \quad i = 1, \ldots, p
\]

(20)

where \( w_i \) can be considered as a filter and \( \mu \) is a regularization parameter, which must be positive.

The solution according to equations (16), (19) and (20) is actually the same as Tikhonov regularization (to be discussed in section 2.3). A simple way of choosing \( w_i \) is to let \( w_i = 1 \), when \( \delta_i^2 \geq \mu \), and \( w_i = 0 \) when \( \delta_i^2 < \mu \). This is the so-called TSVD. Compared to the conventional SVD, TSVD may be considered as having a filter, and hence it is less sensitive to high frequency noise in the measurements.

2.2. Direct method based on singular value decomposition

When measurement errors are considered, equation (7) becomes

\[
Sg = \lambda + e
\]

(11)

where \( e \) is the capacitance measurement error vector.

A common method to obtain the solution of equation (11) is to minimize \( \frac{1}{2}||Sg - \lambda||^2 \), giving

\[
S^T Sg = S^T \lambda.
\]

(12)

If the inverse of \( S^T S \) existed, the solution of equation (7) in the least squares sense would be

\[
\hat{g} = (S^T S)^{-1} S^T \lambda.
\]

(13)

However, the inverse of \( S^T S \) does not exist in most cases, because the number of pixels is usually much larger than the number of measurements. Singular value decomposition (SVD) provides a means to obtain the solution of equation (7) directly. If its rank is \( p \), \( S \) can be decomposed as

\[
S = U \Sigma V^T
\]

(14)

where

\[
U = [u_1, u_2, \ldots, u_M]
\]

\[
V = [v_1, v_2, \ldots, v_N]
\]

(15)

\[
\Sigma = \text{diag}[\delta_1, \delta_2, \ldots, \delta_{p-1}, \delta_p]
\]

and \( U \) is an \( M \times M \) orthogonal matrix, \( V \) is an \( N \times N \) orthogonal matrix (note that any two columns of \( U \) or \( V \) are orthogonal), \( \Sigma \) is an \( M \times N \) matrix with all components zero except for the diagonal components and \( \delta_1, \delta_2, \ldots, \delta_{p-1}, \delta_p \) are \( p \) singular values of \( S \) (note that \( \delta_1 \geq \delta_2 \geq \cdots \geq \delta_{p-1} \geq \delta_p > 0 \)).

The solution of equation (7) can now be written as

\[
\hat{g} = V \Sigma^{-1} U^T \lambda
\]

(16)

where \( V \Sigma^{-1} U^T \) is the pseudo-inverse of \( S \), and \( \Sigma^{-1} \) is given by

\[
\Sigma^{-1} = \text{diag} \left[ \frac{1}{\delta_1^2}, \frac{1}{\delta_2^2}, \ldots, \frac{1}{\delta_{p-1}^2}, \frac{1}{\delta_p^2} \right]
\]

(17)

The solution of equation (7) or (11) is an \( N \times M \) diagonal matrix. However, it is not unique, because the inverse problem is ill posed and under-determined. Suppose that \( \hat{g} \) is in the null space of \( S \), then

\[
S\hat{g} = 0
\]

(18)

2.3. Tikhonov regularization

Regularization tools have been developed for many years to solve ill posed inverse problems (Tikhonov and Arsenin 1977, Engl et al. 1996, Bertero and Boccacci 1998, Hansen 1998). They are used to determine a set of solutions using prior constraint information and then one solution from this set is chosen. Tikhonov method is one of the most commonly used universal regularization tools for solving ill posed inverse problems and has been applied to ECT for image reconstruction (Peng et al. 2000, Lionheart 2001). Based on the standard Tikhonov regularization procedure, the solution of equation (7) or (11) is expressed as

\[
\hat{g} = (S^T S + \mu I)^{-1} S^T \lambda
\]

(21)

where \( I \) is an identity matrix, and \( \mu \) is the regularization parameter, which must be positive.

Equation (21) may be considered as a modified version of equation (13). To make its inverse exist, matrix \( S^T S \) must be modified into \( S^T S + \mu L \).

A more general form of Tikhonov regularization is to find \( g \), which minimizes the following function:

\[
\frac{1}{2}(||Sg - \lambda||^2 + \mu ||L(g - \hat{g})||^2)
\]

(22)

R5
where \( \hat{g} \) is the estimated solution according to some prior information and \( \mu \| L(g - \hat{g}) \|^2 \) is used as a constraint.

Minimizing this function means that the discrepancy between the measured capacitance and the estimated capacitance is minimized while the estimated solution is kept reasonably close to the true solution. Actually, this principle underlies most regularization methods. In practice, it is difficult to obtain \( \hat{g} \) when prior information is not available. In general, choosing \( \hat{g} \) as zero and \( L \) as an identity matrix gives the standard Tikhonov regularization, which is equivalent to solving the following equation:

\[
(S^T S + \mu I)g = S^T \lambda. \tag{23}
\]

It has been proved that the inverse of \( S^T S + \mu I \) always exists, so equation (21) can be obtained from equation (23).

The quality of Tikhonov regularization strongly depends on the regularization parameter \( \mu \). It is crucial to choose an optimal regularization parameter \( \mu \), so that a solution as close to the true solution as possible can be obtained. Some mathematical methods, e.g., the discrepancy principle method (Engl and Pfluger 1988), need prior information on the noise in the measurements. Other methods, e.g., the generalized cross-validation method (Golub et al 1979) and the L-curve method (Hansen 1992, Hansen and O’Leary 1993), need less prior information but the related calculations are laborious. In general, a small value of \( \mu \) gives a good approximation to the original problem but the influence of errors may make the solution physically unacceptable. Conversely, a large value of \( \mu \) suppresses the data errors but increases the approximation error. At present, in almost all cases, \( \mu \) is chosen empirically.

2.4. Multiple linear regression and regularization

A multiple linear regression and regularization algorithm for ECT is reported by Yan et al (2001). In principle, it is similar to Tikhonov regularization. The main difference is that the sensitivity matrix is replaced by a system matrix obtained through multi-linear regression.

3. Iterative algorithms

Because of the non-linear relationship between the permittivity distribution and the capacitance, it is almost impossible to find an accurate solution by any single-step (i.e. non-iterative) algorithm with a simplified linear model. To improve the image quality, the inverse problem has to be solved iteratively. Iterative algorithms developed and applied to ECT are based on calculating capacitance values from the permittivity distribution of the current image, and then producing a new image using the discrepancy between the measured capacitance and the calculated capacitance. This process is repeated until a satisfactorily low discrepancy is achieved.

Calculating capacitance values (i.e. solving the forward problem) is carried out using finite element techniques or by superimposing sensitivity maps according to the estimated image. This is a time-consuming process and consequently iterative algorithms are slow compared with the non-iterative algorithms. Therefore, iterative algorithms can only be used for off-line data processing at present. Different iterative algorithms have been developed, according to the methods used to modify the image.

3.1. Newton–Raphson and iterative Tikhonov methods

Newton–Raphson method is an iterative approach, which was initially devised to find the root of a non-linear function. For image reconstruction with ECT, suppose \( F \) is the transform function from permittivity distribution \( g \) to capacitance \( \lambda \); the relationship between them without measurement error can be described as

\[
\lambda = F(g). \tag{24}
\]

The solution of equation (24) can be deduced by minimizing the square error between the measured capacitance and the capacitance estimated from the image, i.e. \( F(g) \):

\[
e = \frac{1}{2} \| F(g) - \lambda \|_2^2. \tag{25}\]

Differentiating the above error with respect to permittivity gives

\[
e' = [F'(g)]^T [F(g) - \lambda] \tag{26}
\]

where \([F'(g)]_j = \frac{\partial F}{\partial g_j}\) is the so-called Jacobian matrix.

To find the solution of equation (24), let equation (26) be zero. Expanding equation (26) in Taylor series and neglecting the non-linear or higher order term gives a solution of equation (24) as

\[
\hat{g}_{k+1} = \hat{g}_k - [(F'(\hat{g}_k))^T (F'(\hat{g}_k))]^{-1} [(F'(\hat{g}_k))^T (F(\hat{g}_k) - \lambda)] \tag{27}
\]

where \( F'(\hat{g}_k) \) and \( F(\hat{g}_k) \) are the Jacobian and the capacitance which are associated with the current permittivity distribution \( \hat{g}_k \).

Note that in principle \( F'(\hat{g}_k) \) and \( F(\hat{g}_k) \) should be updated at each iteration step. This is the so-called Newton–Raphson method. A simplified form of equation (27) is

\[
\hat{g}_{k+1} = \hat{g}_k - (S_k^T S_k)^{-1} S_k^T (\lambda_k - \lambda) \tag{28}
\]

where \( S_k \) is the sensitivity matrix (i.e. Jacobian) and \( \lambda_k \) is the capacitance.

However, a problem with equation (28) is that the inverse of \( S_k^T S_k \) usually does not exist because \( S_k^T S_k \) is not a full-rank matrix. A modified Newton–Raphson iterative algorithm can be implemented by adding a term \( \gamma I \) to \( S_k^T S_k \), making its inverse exist.

\[
\hat{g}_{k+1} = \hat{g}_k - (S_k^T S_k + \gamma I)^{-1} S_k^T (\lambda_k - \lambda) \tag{29}
\]

where \( I \) is an identity matrix, and \( \gamma \) is a positive scalar.

In optimization theory, this method is also called the Levenburg–Marquardt method when \( \gamma \) changes with each iteration.

Updating the sensitivity matrix \( S_k \) and capacitance \( \lambda_k \) at each iteration is very time consuming. Therefore, a fixed sensitivity matrix is usually used from the beginning of the iteration process, and capacitance \( \lambda_k \) is simply calculated by multiplying the sensitivity matrix with the current permittivity distribution. Now, equation (29) becomes

\[
\hat{g}_{k+1} = \hat{g}_k - (S^T S + \gamma I)^{-1} S^T (S \hat{g}_k - \lambda). \tag{30}
\]
In literature about ill posed inverse problems, this method is called iterative Tikhonov regularization (Hansen 1998). Although the Newton–Raphson method is introduced as a non-linear method in this section, it should be pointed out that iterative Tikhonov regularization has actually been used, which simplifies image reconstruction as a linear problem.

3.2. Landweber iteration and steepest descent method

Landweber iteration is a variation of the steepest gradient descent method, which is widely used in optimization theory. Suppose that our goal is to minimize\( f(g) = \frac{1}{2} \| S \cdot g - \lambda \|^2 \), i.e. to find \( g \) which makes the following function minimal.

\[
 f(g) = \frac{1}{2}(Sg - \lambda)^T(Sg - \lambda) = \frac{1}{2}(g^T S^T S g - 2g^T S^T \lambda + \lambda^T \lambda).
\]

(31)

The gradient of \( f(g) \) can be simply calculated as

\[
 \nabla f(g) = S^T S g - S^T \lambda = S^T (Sg - \lambda).
\]

(32)

The steepest gradient descent method chooses the direction in which \( f(g) \) decreases most quickly as the new search direction for the next iteration. This direction is opposite to the gradient of \( f(g) \) at the current point. The iteration procedure is therefore

\[
 \hat{g}_{k+1} = \hat{g}_k - \alpha_k \nabla f(\hat{g}_k) = \hat{g}_k - \alpha_k S^T (S\hat{g}_k - \lambda)
\]

(33)

where \( \alpha_k \) is a positive scalar, which decides the \( \alpha_k \)th step size. \( \alpha_k \) can be chosen so that \( \nabla f(\hat{g}_{k+1}) \) and \( \nabla f(\hat{g}_k) \) are orthogonal. However, choosing an optimal \( \alpha_k \) needs more calculations. A simple way is to choose \( \alpha_k \) as a fixed number in the beginning of the iteration process. If necessary, its value may be changed during iteration (Peng et al 2000, Liu et al 1999).

With a fixed relaxation factor \( \alpha \), the Landweber iterative algorithm can be expressed as

\[
 \hat{g}_k + 1 = \hat{g}_k - \alpha S^T (S\hat{g}_k - \lambda).
\]

(34)

The initial guess \( \hat{g}_0 \) can be calculated using a simple algorithm, e.g. LBP. A problem with Landweber iteration is that its convergence property is poor. A simple way to improve its convergence is to modify it to give the so-called projected Landweber iteration.

\[
 \hat{g}_k + 1 = P[\hat{g}_k - \alpha S^T (S\hat{g}_k - \lambda)]
\]

(35)

where \( P \) is a projection operator. The solution is projected to a convex set after each iteration.

\[
 P[f(x)] = \begin{cases} 
 0 & \text{if } f(x) < 0 \\
 f(x) & \text{if } 0 \leq f(x) \leq 1 \\
 1 & \text{if } f(x) > 1.
\end{cases}
\]

(36)

This means that the constraint is imposed so that the reconstructed image is non-negative and has an upper boundary of the material being imaged. Landweber iteration is faster, because it only uses the first order derivative. However, this does not necessarily mean that the Landweber method converges quickly. Usually, it needs many iteration steps before reaching a minimum point. Another drawback of the Landweber method rests with its semi-convergence characteristics. The image error decreases very rapidly in the beginning, but after reaching a minimal point, the image error increases as the iteration process continues. The optimal number of iterations can be determined if there is some prior information about the capacitance measurements, but usually a fixed number of iterations is chosen empirically, up to a few hundred. Landweber iteration is the most widely used iterative method for ECT (Yang et al 1999, Liu et al 1999), and also in most cases produces the best images as demonstrated in section 4.

The conjugate gradient method, which makes use of conjugate direction, is similar to the steepest descent method. It is widely used for solving equations and for optimization. Although the conjugate method converges more quickly than the steepest descent method, it does not give better results because it is less regularizing than the steepest descent method.

3.3. Algebraic reconstruction technique and simultaneous iterative reconstruction technique

The algebraic reconstruction technique (ART) and simultaneous iterative reconstruction technique (SIRT) are commonly used for image reconstruction in x-ray computerized tomography (Kak and Slaney 1988). They are simple and effective, especially when the system matrix is very large. ART and SIRT have been applied to ECT (Reinecke and Mewes 1996, Sa et al 2000).

ART uses only one set of projection data in each iteration step. The ART iterative formula for ECT image reconstruction based on equation (7) is

\[
 \hat{g}_k = \hat{g}_{k-1} - \frac{(\lambda_k \hat{g}_{k-1} - \lambda_k)}{s_k \hat{g}_{k-1}^T} \cdot s_k^T
\]

(37)

where \( \lambda_k \) is the \( k \)th row vector of the sensitivity matrix.

The \( k \)th iteration of the ART algorithm is based on the \( k \)th row in equation (7). In the \( k \)th iteration, only the \( k \)th normalized capacitance is used to update the normalized permittivity distribution vector. If the last row of equation (7) is used in the \( k \)th iteration, the \((k+1)\)th iteration will be based on the first row of equation (7).

The problem with ART is that it suffers from noise in the measurement data. Because only one set of measurement data is used in each iteration step, it may not converge if the measurement contains a significant error. SIRT overcomes this disadvantage of ART. Assuming that equation (7) is composed of \( M \) rows, \( M \) vectors can be obtained according to equation (37) using ART. \( \hat{g}_k \) is updated using the average of those \( M \) vectors. The iterative formula of the SIRT algorithm is

\[
 \hat{g}_{k+1} = \hat{g}_k - \beta S^T \cdot \frac{S\hat{g}_k - \lambda}{\text{diag}(S^T S)}
\]

(38)

where \( \beta \) is a relaxation factor, \( \text{diag}(S^T S) \) is a vector composed of diagonal components of \( S^T S \) and the division means that each numerator is divided by the corresponding denominator.

If necessary, \( \beta \) can be replaced by a vector. This means that the \( M \) vectors based on ART are not treated equally when trying to obtain the average vector. For ECT as a non-linear system, this is important for improvement of image quality.

Actually, equation (38) can be deduced from equation (33) by replacing \( \alpha_k \) with a weighting vector. Basically, SIRT is also
A descent gradient method. Its principle is the same as that of Landweber iteration.

A common issue with the above iterative algorithms is that they can only guarantee converging to a local minimum. For example, the steepest descent method will give a least squares solution, which is unlikely to present a good image, if the iteration process is not stopped after certain number of iterations. In principle, a suitable solution can be found during the iterative process. The problem is when and how to decide the iterative process should be stopped. This is the key issue for most of the iterative methods for solving ill posed problems.

3.4. Model-based iteration

A model-based iterative method was proposed by Isaksen and Nordtvedt (1993). A few typical models are categorized and described by a few parameters. The iteration process is carried out by searching for the optimal parameters according to the error between the measured capacitance vector and the capacitance vector calculated from the current permittivity distribution. If some prior information on permittivity distribution is known, explicit models can be used. As an example, figure 3 shows annular and stratified distributions in a pipeline. The model for the annular distribution can be described by a single parameter \( r \), and that for the stratified distribution by two parameters \( \theta \) and \( d \).

When prior information on distribution is not available, a threshold is instead searched for to minimize the discrepancy between the measured capacitance and the calculated capacitance. This is called implicit model-based iterative image reconstruction (Isaksen and Nordtvedt 1993).

4. Evaluation of algorithms by simulation and experiment

Five image reconstruction algorithms, namely LBP, SVD, Tikhonov, iterative Tikhonov and projected Landweber iteration, were evaluated by both simulation and experiment. Three evaluation criteria were used:

1. relative image error,
2. relative capacitance residual and
3. correlation coefficient between the test object and the reconstruction (Xie et al 1994).

Because image and capacitance are treated as vectors, their norms were used to calculate the error and the residual respectively.

\[
\text{Image error} = \frac{\| \hat{g} - g \|}{\| g \|} \quad (39)
\]

\[
\text{Capacitance residual} = \frac{\| \lambda - S \cdot \hat{g} \|}{\| \lambda \|} \quad (40)
\]

\[
\text{Correlation coefficient} = \frac{\sum_{i=1}^{N} (\hat{g}_i - \bar{\hat{g}})(g_i - \bar{g})}{\sqrt{\sum_{i=1}^{N} (\hat{g}_i - \bar{\hat{g}})^2 \sum_{i=1}^{N} (g_i - \bar{g})^2}} \quad (41)
\]

where \( g \) is the true permittivity distribution of the test object, \( \hat{g} \) is the reconstructed permittivity distribution and \( \bar{g} \) and \( \bar{\hat{g}} \) are the mean values of \( g \) and \( \hat{g} \) respectively.

4.1. Evaluation by simulation

To compare the algorithms, four typical permittivity distributions were chosen for simulation with an eight-electrode square ECT sensor. Sets of capacitance values for all single-electrode combinations were calculated with each of these permittivity distributions, using a finite element method (FEM). The lower and higher permittivity materials chosen were air and Perspex with permittivity values of 1.0 and 2.6 respectively. The algorithms were implemented using Matlab\textsuperscript{TM} on a PC with a Pentium\textsuperscript{TM} III 750 MHz CPU and 128 Mbytes memory.

Figure 4 shows the simulation results. In order to compare the speed of the algorithms, the elapsed time was also recorded. The image errors, capacitance residuals, correlation coefficients and elapsed time of the five algorithms are listed in tables 2–5. For the iterative methods, the number of iterations and the relaxation coefficients are also given.

Figure 4 shows that, as expected, the iterative algorithms, especially the projected Landweber method, produce superior
images. This is confirmed by table 2, where in all cases the projected Landweber iteration gives the smallest image error.

Table 3 shows that SVD produces capacitance residuals very close to zero, although the corresponding image errors as listed in table 2 are very significant. This indicates that a small capacitance error does not necessarily mean a good image. The least squares solution, although acceptable mathematically for under-determined and ill-posed problems, can be physically meaningless. Except for SVD, the projected Landweber algorithm gives the smallest capacitance errors in most cases.

From table 4, it can be seen that in all cases the projected Landweber iteration gives the greatest correlation coefficient, again indicating the best images.

Obviously, the elapsed time for iterative algorithms is much greater than that for non-iterative algorithms. It is interesting to see from table 5 that, although the number of iterations for the projected Landweber algorithm is much larger than that for the iterative Tikhonov method, it takes a shorter time than the iterative Tikhonov algorithm to converge.

4.2. Experimental evaluation

Static experiments were carried out using an ECT system with an eight-electrode square sensor of dimensions 68 cm × 68 cm. The signal-to-noise ratio (SNR) of the ECT system is about 35 dB. Four test distributions were created using Perspex rods and Perspex beads:

1. a Perspex rod of 25 cm in diameter, positioned in the centre of the sensor,
(2) a Perspex rod of 25 cm in diameter, positioned near the sensor wall,
(3) two Perspex rods of 25 and 15 cm in diameter respectively, positioned near the sensor wall, and
(4) stratified Perspex beads.

Figure 5 shows the true distributions and the images reconstructed from the experimental data. The capacitance errors were evaluated and are given in table 6.

Both the simulation and experimental results show that SVD improves the image resolution in the central area, but introduces false images elsewhere. Image reconstruction based on Tikhonov regularization strongly depends on the regularization parameter. With an appropriate regularization parameter, it gives better results than LBP. In general, the projected Landweber iteration gives slightly better reconstruction than the iterative Tikhonov method. Although it requires more iteration steps, it takes less computing time.

As discussed in section 4.1, a small capacitance residual does not necessarily imply good image quality. One of the reasons is that the capacitance corresponding to the true image was simply calculated using a sensitivity matrix multiplied by the reconstructed permittivity distribution. This is a typical characteristic of ill posed inverse problems.

4.3. Effect of errors in capacitance measurements

The influence of errors in capacitance measurements on image reconstruction has been investigated, by adding a random error into the capacitance vector. A normalized capacitance vector $\tilde{\lambda}$ was calculated using FEM, using the test permittivity distribution shown in figure 4(a). A capacitance vector contaminated by noise $\lambda_e$ was generated by adding a random noise to $\tilde{\lambda}$. The relative norm error between $\lambda_e$ and $\tilde{\lambda}$, which is defined as $\|\lambda_e - \tilde{\lambda}\|/\|\tilde{\lambda}\|$, was 8.9%. Figure 6 shows the capacitance and noise vectors. Figure 7 shows the images reconstructed using different methods. Table 7 gives the evaluation criteria for these images.

Figure 7 and table 7 indicate that image reconstruction for ECT is not very sensitive to errors in capacitance measurements. The main reason is that the condition number of the sensitivity matrix derived from the $32 \times 32$ grid mesh for the eight-electrode ECT sensor used in the simulations is only about 50. In this case, the non-linearity dominates image reconstruction. The error in image reconstruction is caused mainly by the linear approximation of sensitivity matrix.

5. Discussion and future development

Besides the algorithms categorized above, some other techniques have been reported, e.g. threshold filtering by Xie et al (1992) and maximum entropy filtering by Mwambela et al (1997). The threshold method was successfully used to improve the results based on LBP. In the future, filtering methods including photo-processing methods may be combined with the current algorithms to improve the image quality (Ge and Yang 2001). Actually, most regularization methods may be considered as spectral filters. For example, SVD and Tikhonov regularization according to equation (23) may be expressed as

$$\hat{g} = \sum_{i=1}^{p} \frac{\delta_i^2}{\delta_i^2 + \mu} \cdot u_i^T \cdot \lambda \cdot v_i. \quad (42)$$
The equivalent formula in SVD of the Landweber iteration given in equation (36) is

\[ \hat{g}_k = \sum_{i=1}^{p} \left[ 1 - (1 - \alpha \delta_i^2)^k \right] \cdot \frac{u_i^T \cdot \lambda}{\delta_i} \cdot v_i \]  

(43)

where \( k \) is the iteration step number. A generalized filtering solution can be expressed as

\[ \hat{g} = \sum_{i=1}^{p} w_i \cdot \frac{u_i^T \cdot \lambda}{\delta_i} \cdot v_i \]  

(44)

where \( w_i \) forms a spectral filter.

In Tikhonov regularization, \( w_i = \frac{\delta_i}{\delta_i + \alpha} \), and in Landweber iteration, \( w_i = 1 - (1 - \alpha \delta_i^2)^k \). For future development in image reconstruction for ECT, other spectral filters may be considered. The iterative Tikhonov method is almost the same as Landweber iteration when an appropriate parameter \( \gamma \) is chosen.

As discussed in section 4.2, the condition number of the sensitivity matrix for an eight-electrode ECT sensor with \( 32 \times 32 \) mesh is only about 50. For a 12-electrode ECT system and the same meshing strategy, the condition number will be about 2000. This means that a 12-electrode ECT system is inferior to an eight-electrode system in terms of ill-conditioning. In the future, further modelling work should be carried out so that an optimal sensor structure can be designed, including the number and length of electrodes, the meshing strategy and the measurement protocol.

Selecting the number of electrodes requires a trade-off between the number of independent measurements and the measurement uncertainty. Fewer electrodes give fewer independent measurements, and make the solution more under-determined. Too many electrodes would result in too small sensing area, i.e. too small capacitance to be measured accurately. This will increase the uncertainty of capacitance measurements.

Selecting the length of electrodes also requires a trade-off between signal bandwidth and measurement uncertainty. Longer electrodes produce an average signal over a greater axial length. Shorter electrodes may result in capacitance too small to be measured accurately. Typically, the electrodes are 10 cm in length.

For industrial applications, problems with scaling of electrode size must be considered. In principle, the inter-electrode capacitance of an ECT sensor would be independent on the sensor diameter if the fringe effect could be ignored. This may be the case if the sensor diameter is small, say less than 5 cm. However, if the sensor diameter is larger than the usual electrode length (10 cm), the fringe effect will be significant. In this case, either long electrodes should be used or driven guards should be employed. At present, most ECT investigations involve sensors of less than 4\" (i.e. 10 cm) in diameter although a huge sensor of 1 m in diameter was used to measure interface levels in an oil separation tank (Isaksen et al 1994).

For research investigations, comparatively slow but accurate image reconstruction techniques implemented off line should be adequate. However, for monitoring and control purposes, fast on-line techniques are desirable. The main aim of future work on algorithms will be to provide both more accurate and faster image reconstruction algorithms. Parallel data processor arrays, e.g. with DSPs, would enable real-time iterative image reconstruction.

Currently, most algorithms are based on a simplified linear mathematical model. Because ECT systems are essentially non-linear, future work should include investigation into non-linear techniques for both forward problem modelling and image reconstruction. Some interesting work has already started, such as image reconstruction based on electrical field lines (Loser et al 2001), and algorithms based on artificial neural networks (York and Duggan 1997, Nooralahiyan and Hoyle 1997, Tapson 1999, Warsito and Fan 2001).

To enable ECT technology to be used in a real industry environment, both hardware and software of ECT systems have to be improved. On the hardware side, improved SNR and increased data acquisition rates are requested. On the software
side, improved image reconstruction algorithms, e.g. real-time iterative, feature extraction from tomographic images and user-friendly Windows interface are essential.

Acknowledgments

The Royal Society of UK and the Chinese Natural Science Foundation are acknowledged for supporting Dr Peng’s visit to UMIST under the Co-operation Study scheme (ART/CN/XSI/14026). The authors would like to thank Geng Lu and Professor Bao Fen Zhang at Tsinghua University (China), Dr Bill Lionheart, Dr Roger Waterfall and Nick Polydorides at UMIST (UK), Dr Cheng Gang Xie at Schlumberger Cambridge Research Ltd (UK), Professor Kathleen Hennessy at ISOA Inc. (USA) and Dr Les Stott (UK) for their help in preparing the manuscript.

References

Bertero M and Boccacci P 1998 Introduction to Inverse Problems in Imaging (Bristol: Institute of Physics Publishing)
Hammer E A and Johansen G A 1997 Process tomography in the oil industry—state of the art and future possibilities Meas. Control 30 212–16
Reinecke N and Mewes D 1996 Recent developments and industrial research applications of capacitance tomography Meas. Sci. Technol. 7 325–37
Tapson J 1999 Neural networks and stochastic search methods applied to industrial capacitive tomography Control Eng. Pract. 7 117–21
Tikhonov A N and Arsenin V Y 1977 Solutions of Ill-Posed Problems (Washington, DC: Winston)
White R B and Zakhar A 1999 Internal structures in fluid beds of different scales: an application of electrical capacitance tomography Proc. 1st World Congr. on Industrial Process Tomography (Buxton, 1999) pp 39–46
Yang W Q 1996 Hardware design of electrical capacitance tomography systems Meas. Sci. Technol. 7 233–46