SUMMARY The Karhunen-Loeve technique of random process representation was investigated as a method of quantitatively characterizing body surface potential maps. One hundred ninety-two lead body surface potential maps from 124 normal subjects and 97 patients with independently documented heart disease were used in the study. Each map frame in QRS and ST-T of 34 maps in a test set was represented as a linear sum of orthonormal distributions derived from the covariance matrix estimated from all QRS frames in the 221 training maps. A 16:1 reduction in spatial data of the test set was achieved with rms errors of 45 and 21 pV in QRS and ST-T, respectively. Results suggest that 12 independent waveforms, derived from the 192 measured ECGs, may be used in place of those 192 ECGs. In addition to providing a convenient and familiar method of display for map data, the technique puts the data in an appropriate form for quantitative statistical analysis.

THE USE of body surface potential mapping in electrocardiography has shown promise in improving the accuracy and resolution with which cardiac states may be classified. Specifically, improved recognition of myocardial infarction, estimation of size and severity of ischemic lesions, and the detection and sizing of ventricular hypertrophy have been reported (Holt et al., 1969; Vincent et al., 1977; Holt et al., 1978; Muller et al., 1978). Recent studies have demonstrated the utility of surface potential mapping in estimating the location of anomalous preexcitation pathways in Wolff-Parkinson-White syndrome (DeAmbroggi et al., 1976; Spach et al., 1978). Evidence has been reported that supports the use of maps in assessing the vulnerability of the heart to ventricular arrhythmias (Urie et al., 1978). Finally, application of mapping to stress electrocardiography has shown promise of improving the sensitivity and specificity of the test for detecting and localizing coronary artery disease (Fox et al., 1979). In general, assessment of cardiac state on a regional basis is suggested through the use of mapping.

More widespread use of body surface potential mapping has not been realized, in large part because of the specialized hardware and computer support necessary for acquiring, processing, and displaying the large amount of data required. Additionally, interpretation of maps has not been a simple task because of the extensive amount of redundancy in the data which may mask subtle features characteristic of specific disease states. Typical mapping systems utilize up to 250 electrocardiographic leads covering the torso. Electrocardiograms recorded from each site are sampled at rates up to 2000 Hz, thus requiring extremely high data acquisition rates and storage capacities.

Considerable redundancy exists in map data, however, and several independent studies have verified that arrays of between 20 and 32 appropriately placed body surface electrodes acquire the data necessary to estimate accurately total body surface potentials (Barr et al., 1971, Lux et al., 1979). In spite of this reduction in the number of leads required for data acquisition, a considerable amount of data remains for analysis. Traditional methods of ECG analysis, such as amplitude and interval measurement and waveform morphology, could be applied but would be impractical and cumbersome to apply to large numbers of leads. Moreover, such an approach would not directly utilize the information contained in the interrelationships between sites; i.e., the scalar waveforms are samples of a spatial distribution. The work reported in this paper describes the application of a technique for eliminating the spatial redundancy in recorded surface potential distributions whether measured from many leads or estimated from limited lead arrays. This technique yields an alternative display of surface potential data. The new display is appealing...
from the standpoints that the resulting representation accurately characterizes the original map data and is in a familiar format similar to conventional electrocardiographic displays. Moreover, pathologically characteristic patterns obvious in serial isopotential maps can be easily inferred from the new display. Since the reduced data accurately characterize the measured data, they form a basis for applying statistical decision-making techniques to improving classification of heart disease.

**Methods**

One hundred ninety-two lead body surface potential maps from single cardiac complexes were obtained from 221 adults, including 124 subjects with no known heart disease and 97 patients with documented heart disease. Diseases represented in the sample included 76 patients with old myocardial infarction, 16 with intraventricular conduction defects, and 5 with ventricular hypertrophy. All 192 ECG leads for each map were recorded simultaneously at a 1000-Hz rate, multiplexed, and recorded on instrumentation tape as described in an earlier paper (Wyatt and Lux, 1974). ECGs were digitized, gain and baseline adjusted, and displayed as serial isopotential maps on both 16-mm film and hardcopy paper. Selected QRS potential distributions of a normal subject are shown in Figure 1. A complete “map” of a single cardiac complex would consist of approximately 600 such frames, one for each millisecond of time throughout cardiac activation and recovery.

Applying the mathematical theory of signal representation, it is possible to characterize the spatial, potential distribution of any single map frame by a weighted sum of orthonormal distributions. Specifically, if the 192 potentials in a single frame are formed into a 192 dimensional vector, \( P \),

\[
P = (p_1, p_2, \ldots, p_{192})'
\]

where \( p_i \) is the voltage at site 1, \( p_2 \) is the voltage at site 2, etc., then \( P \) may be exactly represented by a linear sum of basis vectors

\[
P = \sum_{i=1}^{192} a_i \phi_i
\]

where \( \{\phi\} \) is a set of orthonormal basis vectors which span the 192 dimensional vector space. The set of coefficients \( \{a_i\} \) is unique for each \( P \) and \( \{\phi\} \), and is defined by

\[
a_i = P \cdot \phi_i
\]

In practice, any orthonormal vector set could be used, e.g., Fourier sequences. Moreover, not all 192 basis vectors might be required for an acceptably accurate representation of \( P \), particularly when de-
pendency exists between dimensions. Thus,

\[ P_N = \sum_{i=1}^{N} \alpha_i \phi_i \quad N < 192 \]  

is an approximate representation of \( P \) where the error in using less than the total number of basis vectors is

\[ E_N = |P_N - P| \]  

Since \( P \) is a random vector, we can apply the theory developed by Karhunen and Loeve (K-L) for the statistical expansion of a random process along a set of basis functions which are optimal for that process in the sense of minimizing the representation error (Chen, 1973). It may be shown that for a given \( N \) in Equation 4, the mean squared error \( E_N^2 \) may be minimized by constraining the expansion coefficients to be uncorrelated, i.e.,

\[ E[a_{i,j} = \delta_{i,j} \quad i, j = 1, 2, \ldots, 192 \]  

where \( E \) is the statistical expectation operator, \( \delta \) is the Kronecker delta function and \( \lambda \) is a constant. Substituting Equation 2 into Equation 5 yields

\[ E(P \cdot \phi_i)(P \cdot \phi_j) = \delta_{i,j} \lambda_j \]  

or

\[ \phi_i \cdot K \phi_j = \lambda_i \delta_{i,j} \]  

where \( K \) is the covariance matrix of the random vector \( P \). Since \( \{\phi_i\} \) is an orthonormal set, Equation 6 may be satisfied if

\[ K \phi_i = \lambda_i \phi_i \]  

Equation 7 is a statement of the classical eigenvalue problem which has as solutions, the set \( \{\lambda_i, \phi_i\} \), \( i = 1, N \) where the \( \lambda_i \)'s are eigenvalues and the \( \phi_i \)'s are the corresponding basis vectors or eigenvectors of the matrix \( K \). The resulting set of basis vectors, when taken in decreasing order of the magnitudes of their associated eigenvalues, forms an optimal basis in the sense that, for a given number of basis vectors, \( N \), on the average they better represent the vector \( P \) than any other basis set of \( N \) vectors.

This representation, which also is referred to as principal components analysis, was applied by Barr et al. (1971) and later by Warren (1977) for purposes of estimating complete maps from a limited number of sites. We have included details of the technique, since we wish to emphasize the independence from sites. We have included details of the technique, of estimating complete maps from a limited number of sites. We have included details of the technique.

For purposes of illustrating the spatial redundancy in body surface potential maps, correlation coefficients between all sites were determined from the covariance matrix estimated using 12 eigenvectors. For each map, peak errors, rms errors, and correlations for QRS and ST-T were calculated and statistics accumulated. Represented and measured map frames were retained for purposes of documenting the worst case errors of the technique. Additionally, represented and measured map leads similar to \( V_3 \), \( V_4 \), and \( V_5 \) were compared for purposes of showing typical waveform representation.

**Results**

For purposes of illustrating the spatial redundancy in body surface potential maps, correlation coefficients between all sites were determined from the covariance matrix estimated using all QRS frames from all subjects. Specifically, for the \( i \)-th site, the \( i \)-th row of the covariance matrix is

\[ [\rho_{11}, \sigma_1^2, \rho_{12}, \sigma_2^2, \ldots, \rho_{192}, \sigma_{92}^2] \]

where \( \rho_{ij} \) is the correlation coefficient of potentials between sites \( i \) and \( j \), and \( \sigma_i \) is the standard deviation of potentials at site \( i \). Figure 2 shows the distribution of correlation coefficients for a site (circled in the figure) plotted as isocorrelation lines. Clearly, sites close to this site are highly correlated to it (\( \rho > 0.9 \)) and additionally there are many sites which are highly negatively correlated (\( \rho < -0.8 \)). The impli-
ISOCORRELATION MAP - SITE 90

**FIGURE 2** Isocorrelation contour map of site 90. The display shows the distribution of correlation coefficients with respect to the circled site and was calculated from all QRS frames in the training set. Isocorrelation lines are separated by 0.1 unit. Correspondence of electrode sites and anatomic regions is the same as that described in Figure 1.

cation of this distribution, the features of which are typical of those observed at other sites, is that spatial redundancy exists in 192 lead maps.

For purposes of spatially representing the QRS and T portions of maps, several possible approaches could be followed. First, all QRS and T map frames could be lumped together and an estimate of the spatial basis, \( \{ \phi \}_{QT} \), calculated from the data. Using this approach, all map frames during both QRS and T would be represented using the same basis. Second, separate bases for QRS and T could be determined, \( \{ \phi \}_{QRS} \) and \( \{ \phi \}_{T} \), by using separate QRS and T training data, respectively. With this approach, QRS frames would be represented using \( \{ \phi \}_{QRS} \) and T frames would be represented using \( \{ \phi \}_{T} \). Finally, if it could be demonstrated that \( \{ \phi \}_{QRS} \) spans \( \{ \phi \}_{T} \), it would be efficient to represent QRS and T frames using \( \{ \phi \}_{QRS} \).

The first and third approaches have the advantage that only one basis set would be required for representation. The second approach is appealing from the standpoint that QRS and T portions of the map would be represented using bases derived from data during physiologically similar events and, hence, efficient representation would be achieved. These considerations were investigated by calculating covariance matrices for QRS, T and QRST, \( K_{QRS} \), \( K_{T} \), and \( K_{QRST} \), respectively. In each case, all map frames in the training set were used. From each covariance, the first 15 eigenvectors and their corresponding eigenvalues were estimated. The curves in Figure 3 show the estimated rms voltage error, \( e_{N} \), and the explained variance expressed as a percent of the trace (\( K \)), as a function of the number of basis vectors for QRS-, ST-T-, and QRST-derived bases. These curves are helpful in determining the number of basis functions required to represent the data from which they were derived. For example, the ST-T curve flattens out near the fifth eigenvector. The QRS and QRST curves flatten beyond the ninth or tenth eigenvector.

From theoretic considerations evident in the K-L theory, minimum error of representation using a fixed number of basis functions would be achieved by representing QRS data with QRS-derived eigenvectors and ST-T data with ST-T-derived eigenvectors. Use of QRST-derived basis vectors to represent QRS and ST-T data would require a greater number of vectors than in individual cases. Since the QRS portion of surface potential maps shows the greatest diversity of signal amplitude and pattern complexity, we considered it better to sacrifice a small amount of T representation accuracy for improved QRS representation accuracy. Add to this the convenience of using one basis set for all representation and one should choose the QRS- or the QRST-derived basis set.

Table 1 shows the cumulative sum of eigenvalues, and corresponding percent of the trace of the covariance matrix estimated from QRS data alone and then QRS plus ST-T data. The percent of trace

**FIGURE 3** The percent of the trace of the covariance, % TRACE (\( K \)), in the left panel and the RMS ERROR in the right panel associated with a given number of eigenvectors are shown for QRS, ST-T-, and QRST-derived bases.
column yields a figure of merit which relates the sum squared signal accounted for by the first \( n \) eigenvalues to the total squared signal. The table shows that the QRS-derived basis is slightly better than the QRST-derived basis in the sense that the first \( n \) QRS-derived eigenvectors explain a greater percentage of the total variation than the first \( n \) QRST-derived eigenvectors (\( n < 15 \)).

If it can be shown that the QRS basis spans the required \( T \) basis, then one would be justified in using the QRS basis for all data representation. The rms error of representing \( T \) data using six \( T \)-derived eigenvectors was calculated from the covariance matrix estimated from \( T \) data to be 21.4 mV. The first six \( T \) eigenvectors were expanded (represented) using the first 12 QRS-derived eigenvectors with an error of 21.8 mV. The additional rms error incurred was only 0.4 mV; hence we chose to utilize the QRS-derived eigenvectors to represent the information in the recorded map of isopotential maps. The next step in evaluating the relationships between isopotential maps and the coefficient waveforms, one normal and two abnormal maps were reduced along the 12 basis vectors shown in Figure 4. Each patient map was a sequence of 192 dimensional vectors, each of which was represented as a linear sum of the basis vectors as shown in Equation 3. Using 12 basis vectors, each frame of 192 potentials was reduced to 12 coefficients which specified how much of each characteristic pattern existed in the original measured frame. Mathematically, let \( a_n \) be the coefficient of the \( k \)-th spatial eigenvector for the \( k \)-th map frame, \( P_k \)

\[
a_n = P_k \cdot \phi_n. \quad (10)
\]

Each coefficient associated with a particular basis vector was then plotted as a function of time, \( k \), to yield the time-varying amount of the characteristic pattern existing in the measured map. Figure 5 shows the 12 coefficient "waveforms" obtained from a map recorded from a normal subject. The waveforms provide the weighting of each basis vector and could be used along with the basis vectors to resynthesize the maps to a high degree of accuracy. These waveforms provide a 16:1 reduction in the map data. Note that the amplitude of the waveforms decreases for the higher order basis vectors, as predicted from the curves of Figure 3. Apart from the quantitative accuracy of the representation, the waveforms give qualitative information regarding the dominant distribution. For example, the large amplitude of waveform 1 corresponds to a strong pattern 1 which is a characteristic isopotential pattern in the early normal QRS map.

The waveforms in Figure 6 were derived by reducing a map from a patient diagnosed as having had an old anterior wall infarction. The large, negative deflection in waveform 2 reflects an intense pattern 2 with opposite polarity. This distribution is characteristic of those observed in patients with anterior wall infarction.

Finally, Figure 7 shows waveforms derived from a patient with an old inferior wall myocardial infarction. The negative deflection in waveform 3 reflects the dominance of pattern 3 with opposite polarity and is similar to patterns observed early in QRS in patients with inferior wall infarction.

These examples were shown merely to indicate that the coefficient waveforms exhibit features which reflect the dominant characteristics of the isopotential maps. The next step in evaluating the representation technique was to apply the representation scheme on test maps and accumulate error statistics to show that these waveforms accurately represent the information in the recorded map of 192 leads. A set of thirty-four 192 lead maps not used in the training set was reduced along the 12 QRS basis vectors. For each map frame, peak error,
rms error, and correlation of measured to represented frame were determined. Peak error was the largest error at any of the 192 sites for that frame. The rms error for each frame was calculated by the equation

\[ e = \sqrt{\frac{\sum (P - P_N)^2}{192}} \]  

(11)

where \( P \) is the measured frame and \( P_N \) is the expanded frame. Correlation coefficients between measured and represented frames were calculated by the equation

\[ \rho = \frac{P \cdot P_N}{|P||P_N|}. \]  

(12)

The average and peak rms error voltages and correlation coefficients were determined for QRS and ST-T portions of the cardiac cycle for each of the 34 maps in the test set. In addition, averages of these statistics over all maps were determined and the measured and represented frames at the times of peak error during QRS and ST-T printed as examples of worst case representation. Also, body surface leads approximating leads \( V_2, V_3, \) and \( V_5 \) for both measured and expanded maps were plotted to verify retention of waveform integrity by the representation process. For purposes of illustration, performance of the representation is documented for a normal subject (Figs. 8–10) and an abnormal patient (Figs. 11–13). Included in the figures are peak error voltage, rms representation error, and correlation coefficient vs. time throughout the Q-T interval, comparison of measured and represented map frames at times of worst case peak error and worst case rms error in both QRS and ST-T, and comparison of measured and represented leads \( V_2, V_3, \) and \( V_5 \). These figures illustrate the typical errors incurred by spatial representation of maps using the K-L technique.

Since the K-L technique minimizes the mean
The data in panel B show a measured frame and its representation obtained from a patient’s map at a time mid-way in the ST segment. Note that the scale is 25 μV/contour. In this example, the signal...
is extremely small (peak of 100 μV) but with a simple distribution. The rms error of 8 μV and correlation of 0.94 suggest a reasonably good representation. In spite of the low level nature of this frame, the locations and magnitudes of the extrema, as well as the shape of the contours, are accurately characterized by the representation.

Table 2 summarizes the errors for all QRS and ST-T maps from all 34 patients in the test group. The small errors suggest the technique is a practical and accurate approach for representing body surface potential maps, reducing map storage requirements, providing features for map classification, and displaying maps in a convenient format for interpretation.

**Discussion**

The application and utility of body surface potential mapping to improved electrocardiographic diagnosis of heart disease has been hampered by...
the requirements of extensive hardware for data collection and the lack of a clearly defined scheme for utilizing the vast amount of data accumulated from each patient. The complexities of surface potential distributions and the often subtle features characteristic of underlying cardiac states have made the quantitative interpretation of maps a difficult task. In light of the difficulty of utilizing the extensive amount of map data and their inherent redundancy, the work reported in this paper was aimed at reducing the spatial redundancy in maps, while maintaining, on the average, a specified accuracy of representation. Specifically, the Karhunen-Loeve approach of random process representation along a statistically derived orthonormal basis was applied to maps with a resulting rms reconstruction error of 45 μV rms when 12 basis vectors are used. The representation permitted a 16:1 reduction in the amount of data required to characterize accurately body surface potential map data. The end result of the technique was to reduce 192 body surface electrocardiograms to 12 waveforms, each associated with a fixed body surface potential distribution which was derived from statistics of 192 lead maps on 221 patients. It should be stressed that increasing the number of eigenvectors used in the representation would improve its accuracy. However, in light of the fact that no standards exist for lower limits of potentials for map interpretation, we selected 12 to illustrate the technique.

The representation errors obtained in this study compare favorably with those reported previously, although exact comparisons are difficult to make...
because of differences in lead sets, criteria for inclusion of maps in both training and testing sets, and patient populations. Barr et al. (1971) reported 1 mV^2 sum squared error (equivalent to 82 µV rms error or 1% mean square error) using 30 spatial patterns. Warren (1977) showed an error of 0.67 mV^2 sum squared error (67 µV rms or 2.5%) for 12 patterns in the presence of 26.6 mV^2 sum squared signal (421 µV rms). In comparison, we observed a 0.36 mV^2 sum squared error (44 µV rms or 1.8%) in the presence of 19.8 mV^2 sum squared signal (321 µV rms) for 12 patterns. These are average values but reflect consistency with the findings of others.

The relationship of this work to work by Gesselowitz (1960) who used the multipole expansion as a representation of the cardiac generator should be pointed out. The multipole expansion utilizes measured body surface potentials, torso geometry, and physical properties to arrive at a solution of the multipole cardiac generators. These generators (dipole, quadrupole, etc.) which are temporal waveforms of pole strengths are then the representation of the cardiac activity in the sense that they may be used to reconstruct the body surface potentials. In contrast, our approach utilizes the statistical relationships of body surface potentials to arrive at a set of basis distributions for characterizing the cardiac activity. No assumptions need be made with

---

**Figure 13** Same format as Figure 10 for patient described in Figure 11. Panels A and B are during the T wave in this example.

**Figure 14** Panel A: Measured and represented frames from a patient with an unusual late QRS pattern. Panel B: Measured and represented map frames from the ST segment of a patient with an old myocardial infarction. Note that contours are at 25-µV increments.
regard to geometry or physical properties of the torso. Our approach provides a natural basis founded on the measured surface potential activity, whereas the multipole approach utilized lead sets and source orientations selected from physical considerations. By the nature of the Karhunen-Loève approach, it requires fewer basis vectors for a fixed level of representation accuracy than any other representation scheme, including the multipole expansion which is an orthogonal expansion.

The significance of this work lies in the reduction in the amount of data required to characterize accurately body surface potential maps. This should simplify the difficulty of interpreting maps, although the utility of the method has not been evaluated. Another benefit of the technique is that the representation format is similar to conventional electrocardiographic displays. Namely, the waveforms showing the amount of each basis distribution in a map are similar to standard electrocardiographic waveforms. Finally, the representation provides a basis for still further data reduction by removing temporal redundancy. The ultimate utility of the technique will not be realized before extensive, prospective studies have been carried out on adequate numbers of patients with a much greater range of diseases represented. In addition, the diagnostic capability of the representation parameters will have to be carefully evaluated.

References


Holt JH, Barnard ACI, Lynn MS (1969) A study of the human heart as a multiple dipole electrical source. II. Diagnosis and quantification of left ventricular hypertrophy. Circulation 40: 697-710


Redundancy reduction for improved display and analysis of body surface potential maps. I. Spatial compression.
R L Lux, A K Evans, M J Burgess, R F Wyatt and J A Abildskov

doi: 10.1161/01.RES.49.1.186

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/49/1/186

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation Research can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation Research is online at:
http://circres.ahajournals.org/subscriptions/