Early detection of Myocardial Infraction by Blind Source Separation of ECG Signals

By

Reem Ashkar                      Kamil Khoury

Supervised by

Eldad Klaiman

May 2004
Thank you…

We would like to thank:

The lab staff:
   Johanan Erez.
   Ina Krinski
   Ahron Yacoby

Dr. Robber Drago who helped us much.

Eldad Klaiman our supervisor.
1. Objective ............................................................................. 3

2. Heart and ECG
   2.1 What is ECG? ................................................................ 4
   2.2 The electricity of the heart ........................................... 5
   2.3 What is a heart attack? ................................................. 7
   2.4 The effect of a heart attack on ECG ......................... 9

3. Protocols
   3.1 PCA ........................................................................ 11
   3.2 ICA ......................................................................... 14

4. Work approach:
   4.1 Blocks diagram ......................................................... 18
   4.2 The work stages ...................................................... 19
      4.2.1 Building up database ......................................... 19
      4.2.2 Choosing one phase ......................................... 19
      4.2.3 Performing PCA ............................................... 21
      4.2.4 Performing ICA ................................................ 22
      4.2.5 Comparing the signals .................................... 23

5. Results
   5.1 QRS signal ............................................................. 24
   5.2 P and T signals ....................................................... 25
   5.3 The signals number ................................................ 28

6. Future work ............................................................... 30

7. Appendix
   7.1 ConverFile.m .......................................................... 31
   7.2 Comparison GUI ..................................................... 33
1. Objective:

The objective of this project is to try and manipulate the ECG signals that we have in order to achieve an easy way for identifying heart disease. A prominent and fatal heart disease is called Myocardial Infraction, more commonly known as “Heart Attack”. Modern medicine can identify the heart attacks in the majority of cases, most of them only after they have occurred. Our project is a step in the goal for identifying heart attacks in an automated way, and the result may provide a tool for identifying heart attacks before they happen. This project will try to give an answer for: "Is there a way to identify heart attacks by looking at the separated sources of the ECG signal?"
2. Heart and ECG

2.1 What is ECG?

ECG (or EKG) stands for electrocardiogram; which is a measurement of the electrical signals that control the rhythm of the heartbeat. The heart is a muscular organ that beats in rhythm to pump the blood through the body. The signals that make the heart's muscle fibers contract come from the senatorial node, which is the natural pacemaker of the heart.

In an ECG procedure, the electrical impulses made while the heart is beating are recorded and usually printed out on a piece of paper. This is known as an electrocardiogram, and records any problems with the heart's rhythm, and the conduction of the electrical signal through the heart, which may be affected by underlying heart disease.

How does ECG help doctors?

* An ECG can be used to assess if the patient has had a heart attack or evidence of a previous heart attack.
* An ECG can be used to monitor the effect of medicines used for coronary artery disease

These are the main issues that ECG helps doctors with heart attacks.

How is an ECG performed?

The ECG test is made usually when the patient is resting, but when there is concern that the patient's symptoms may be caused by coronary artery disease the test is made while the patient is on an exercise bike or treadmill. Some times the ECG is performed on the patient in order to monitor his condition after a surgery.

Up to 12 electrodes will be attached to selected locations of the skin on the arms, legs and chest (as shown down). Every electrode is called lead. After the test the electrodes is removed and the ECG is printed on paper that the doctor will preview (usually).
There is a modern ECG machines that records the ECG to a file on a computer, the number of these machines in Israel is limited.

2.2 The electricity of the heart:

The contraction of any muscle is associated with electrical changes called depolarization, and these changes can be detected by electrodes attached to the surface of the body. Since all muscular contraction will be detected, the electrical changes associated with contraction of the heart muscle will only be clear if the patient is fully relaxed and no skeletal muscles are contracting.

The heart has four chambers (as shown in the figure), but from the electrical point of view is can be thought of as having only two, because the two atria (or atrium) contract together and then the two ventricles contract together.
The wiring diagram of the heart:

The electrical discharge for each cardiac cycle normally starts in a special area of the right atrium called the 'sinoatrial (SA) node'. Depolarization then spreads through the atrial muscle fibres. There is a delay while the depolarization spreads through other special area in the atrium, the 'atrioventricular node' (AV or 'the node'). Thereafter, the electrical discharge travels very rapidly, down specialized conduction tissue: first a single pathway, the 'bundle of His', which then divides in the septum between ventricles into right and left bundle branches. The left bundle brunch itself divides into two. Within the mass of ventricular muscle, conduction spreads somewhat more slowly, through specialized tissue called 'Purkinje fibers'.

The shape of the ECG:

The muscle mass of the atria is small compared with that of the ventricles, and the electrical change accompanying the contraction of the atria is therefore small. Contraction of the atria is associated with the ECG wave called 'P'. The ventricular mass is large, and so there is a large
deflection of the ECG when the ventricles are depolarized. This is called the 'QRS' complex. The 'T' wave of the ECG is associated with the return of the ventricular mass to its resting electrical state ('repolarization').

2.3 What is a heart attack?

When the heart muscle is damaged, because blood was blocked from reaching it, a heart attack occurs. This is also called a myocardial infarction (MI) because the middle layer of the heart muscle (myocardium) is damaged or is non-functional (infarct).

"Myocardial Infarction" (abbreviated as "MI") means there is death of some of the muscle cells of the heart as a result of a lack of supply of oxygen and other nutrients. This lack of supply is caused by closure of the artery ("coronary artery") that supplies that particular part of the heart muscle with blood. This occurs 98% of the time from the process of arteriosclerosis ("hardening of the arteries") in coronary vessels.

Heart attacks, as already noted, are the result of a clot in an artery to the heart. However, while they seem to occur suddenly and often without warning the process underlying the event has been going on for many years.

After a small heart attack the patient might still be at a high risk. The risk is "Sub-Endocardial MI" ("non-Q MI"); a small MI, as judged from the amount of enzyme released, may occur from one of two situations:
* A large artery closed long enough to cause some damage, but later re-opened enough to prevent all of the muscle in its territory from dying, or
* A small artery closed completely, with essentially all of the muscle in its territory dying.

The amount of enzyme released in each case will be similar. However, in the first case, a significant blockage may still be present, and more damage is still possible if it closes again -- this is the high risk situation that is being referred to. In the second case, all the damage has already been done, and the chances of a bad outcome are less.

There are ways to treat heart attacks. One of them is by giving a medicine to the patient; this is used usually when the heart attack is small. Other treatment is 'Angioplasty and Stenting'. Coronary angioplasty is a nonsurgical treatment for blocked or narrowed passages in one or more of the coronary arteries. This procedure allows a normal supply of blood to flow through the heart muscle.

Angioplasty can be done in one of several ways, depending on the type of plaque (fatty deposits) blocking the artery and the size and shape of the artery. It may be completed by balloon, stent, laser rotational or directional atherectomy.

* Balloon angioplasty. A catheter (thin tube) with a balloon on the end is threaded through the groin up to the narrowed artery. The balloon is then inflated at high pressure one or more times to widen the artery and allow blood to flow more easily though it.

* Laser angioplasty. A laser-equipped catheter is inserted into the narrow artery. The laser vaporizes the plaque, and the artery is once again opened to normal width.

* Stent implantation. A thin metal mesh-like structure is attached to a balloon catheter and moved into the blocked heart artery. The balloon is inflated, allowing the stent to be expanded and permanently implanted in the artery. The expanded stent acts as a scaffold in the artery to keep it open. Sometimes arteries again become narrow or blocked after stent implantation due to overgrowth of the smooth muscle cells. When that happens, radiation is delivered to the area to shrink the overgrowth. This procedure is called intracoronary brachytherapy.
* Directional coronary atherectomy. A small sharp blade inside a catheter is placed against the plaque. The interventional cardiologist then cuts and removes part of the plaque from the wall of the artery.

* Rotational atherectomy. A burr coated with diamond dust is advanced through the plaque at 150,000 rotations per second. The plaque is pulverized into small particles that are washed into the bloodstream.

The result of the treatment can be checked by monitoring the ECG of the patient.

Figures:

2.4 The effect of a heart attack on ECG:
As were mentioned before ECG is the record of the electrical activity in the heart muscles, and so it's very obvious that heart attacks would have an effect on the ECG.

Heart attack changes the ECG signals that are measured, the effect is immediate and it lasts for few weeks – at least, an illustration of those changes can be seen in the next figure:

* **Hyperacute phase** begins immediately after a heart attack
* **Fully evolved phase** starts a few hours to days after a heart attack
* **Resolution phase** appears a few weeks after a heart attack
* **Stabilized chronic phase** is the last phase and typically has permanent pathological changes compared to a normal ECG tracing

As we see the effect of a heart attack would last for long time, and we would use this propriety to examine our method for identifying heart attacks in ECG. More details will be given further.
3. Protocols

3.1 PCA - Principal Components Analysis:

PCA is a technique for identifying patterns in data, and expressing the data such a way as to highlight their similarities and differences. Since patterns in data can be hard to find in data of high dimension, where the luxury of graphical representation is not available, PCA is a powerful tool for analyzing data.

The other main advantage of PCA is that once you have found these patterns in the data, and you compress the data, ie. by reducing the number of dimensions, without much loss of information. This technique used in image compression.

PCA method:

Step 1 – Set our data:

For PCA work properly, you have to subtract the mean from each of the data dimensions, which means that you give PCA data with mean zero. For example we'll use two dimension data – it's easier to show plots for it.

Example:

![Original Data and after setting mean zero](image)

Step 2 – Calculate the covariance matrix:
After the calculation we have:

\[
\text{cov} = \begin{pmatrix}
0.616555556 & 0.615444444 \\
0.615444444 & 0.716555556
\end{pmatrix}
\]

The covariance gives us a hint on the relation between our dimensions, here expect that both the \(x\) and \(y\) variable increase together.

**Step 3 – Calculate the eigenvectors and eigenvalues of the covariance matrix:**

Since the covariance matrix is square, we can calculate the eigenvectors and eigenvalues for this matrix. These are rather important, as they tell us useful information about our data:

\[
eigenvalues = \begin{pmatrix}
0.04908339 \\
1.28402771
\end{pmatrix}
\]

\[
eigenvectors = \begin{pmatrix}
-0.735178656 & -0.677873399 \\
0.677873399 & -0.735178656
\end{pmatrix}
\]

We see how the data has a strong pattern. And as expected from the covariance matrix, the two variables do indeed increase together.

By this process of taking the eigenvectors of the covariance matrix, we have been able to extract lines that characterise the data. The rest of the
steps involve transforming the data so that it is expressed in terms of these lines.

**Step 4 – Choosing components and forming a feature vector:**

Here is where the notion of data compression and reduced dimensionality comes into it. If you look at the eigenvectors and eigenvalues from the previous step, you will notice that the eigenvalues are quite different values. In fact, it turns out that the eigenvector with the highest eigenvalue is the principle component of the data set. In our example, the eigenvector with the largest eigenvalue was the one that pointed down the middle of the data. It is the most significant relationship between the data dimensions.

In general, once eigenvectors are found from the covariance matrix, the next step is to order them by eigenvalue, highest to lowest. This gives you the components in order of significance. Now, if you like, you can decide to ignore the components of lesser significance. You do lose some information, but if the eigenvalues are small, you don't lose much.

We will build a FeatureVector from the eigenvectors that we want to keep, and then we calculate the new data:

\[
FeatureVector = (eig_1, eig_2, eig_3, ... eig_n)
\]

\[
FinalData = RowFeatureVector \times RowDataAdjust
\]

And in our example:

![Final Result of PCA, no eigenvectors have been removed.](image-url)
3.2 ICA – Independent Components Analysis:

**Motivation – cocktail-party problem:**

Imagine that you are in a party with other people in a room, and every one is speaking simultaneously. You have number of microphones (as the number of the guests), which you hold in different locations. The signals that you record are a linear equation of the original signals. The question is, can we get the original signals from the recorded ones?

Actually if you knew the linear equation parameters you could solve the linear equation by classical methods. The point is, however, that if you don't know the parameters, the problem is considerably more difficult.

**What is Independent Component Analysis?**

Independent Component Analysis (ICA) is a statistical technique for decomposing a complex dataset into independent sub-parts. Here, we demonstrate ICA for solving the Blind Source Separation (BSS) problem.

We are given two linear mixtures of two source signals which we know to be independent of each other, i.e. observing the value of one signal does not give any information about the value of the other. The BSS problem is then to determine the source signals given only the mixtures.

Putting this into mathematical notation, we model the problem by

\[ \mathbf{x} = \mathbf{A}\mathbf{s} \]

Where \( \mathbf{s} \) is a two-dimensional random vector containing the independent source signals, \( \mathbf{A} \) is the two-by-two mixing matrix, and \( \mathbf{x} \) contains the observed (mixed) signals.

This first plot (below) shows the signal mixtures on the left and the corresponding joint density plot on the right. That is, at a given time instant, the value of the top signal is the first component of \( \mathbf{x} \), and the value of the bottom signal is the corresponding second component. The plot on the right is then simply constructed by plotting each such point \( \mathbf{x} \). The marginal densities are also shown at the edge of the plot.
A first step in many ICA algorithms is to whiten (sphere) the data. This means that we remove any correlations in the data, i.e. the signals are forced to be uncorrelated. Again putting the words in mathematical terms, we seek a linear transformation $V$ such that when $y = Vx$ we now have $E\{yy'\} = I$. This is easily accomplished by setting $V = C^{-1/2}$, where $C = E\{xx'\}$ is the correlation matrix of the data, since then we have $E\{yy'\} = E\{Vxx'V'\} = C^{-1/2}CC^{-1/2} = I$.

The figure below shows the signals $y$ and the joint density $p(y)$ after such an operation.

After sphering, the separated signals can be found by an orthogonal transformation of the whitened signals $y$ (this is simply a rotation of the joint density). The appropriate rotation is sought by maximizing the non-normality of the marginal densities (shown on the edges of the density plot). This is because of the fact that a linear mixture of independent random variables is necessarily more Gaussian than the original variables. (This is the same phenomenon as is stated by the central limit
theorem.) This implies that in ICA we must restrict ourselves to at most one Gaussian source signal.

There are many algorithms for performing ICA, but the most efficient to date is the FastICA (Fixed-point) algorithm. The plot below shows the result after one step of the FastICA algorithm.

Separated signals after 1 step of FastICA

The rotation continues...

Separated signals after 2 steps of FastICA

...and continues...

Separated signals after 3 steps of FastICA
Until it starts to converge

**Separated signals after 4 steps of FastICA**

Convergence! The source signals (components of \(s\)) in this example were a sinusoid and impulsive noise, as can be seen in the left part of the plot below. The right plot shows the joint density which can be seen to be the product of the marginal densities,

**Separated signals after 5 steps of FastICA**

\[ p(s) = p(s_1)p(s_2) \]

i.e. \( p(s) = p(s_1)p(s_2) \). This is of course the definition of independence.
4. Work approach:

4.1 Blocks diagram:

- Sick
- ECG
- Choose one phase
- PCA
- Remove 4 signals
- ICA
- Comparison

- Healthy
- ECG
- Choose one phase
- PCA
- Remove 4 signals
- ICA

Results:
4.2 The work stages:

4.2.1 Building up database:

We have gathered ECG signals from healthy subjects and sick subjects after treatment. We need two kinds of ECG data because we would like to try and identify the heart attacks using the differences between the healthy ECG and the sick ones.

We built our database from Dr. Robert Drago’s data (cardiac ICU – Rambam Hospital). We received three ECG signals:

#1: A healthy person ECG; length 3.5 minutes. "the normal".

#2: A sick person ECG; length 3.5 hours. This ECG is for person after he had angioplasty operation. "hanipas".

#3: Another sick person ECG; length 3.5 hours. This ECG is for person after he had angioplasty operation. "breitman".

The ECG files that we have must be converted to Matlab format in order we can work with them in Matlab environment. More details on the conversion can be found at the appendix under ConvertFile.

4.2.2 Choosing one phase:

Now we have the ECG data set as a matrix, where every line is one signal of the ECG (one of the twelve signals we get from the ECG machine), therefore we chose to cut the line in order to get one cycle.

This operation has been made manually, We take 2 seconds data length (its 1000 samples) and find the peaks positions:
We find the middle positions between every two peaks, and we cut the signals there, and we get our one cycle:

Example for normal ECG:

In the next figure we see one sick ECG; we can notice immediately the ripples in the signals (“noise”). This noise appears in all sick ECG and not in this one only.
4.2.3 Performing PCA:

The ECG machine is connected to the patient by 10 electrodes which are mixed to give 12 signals. Two from these electrodes are used as reference, and this leaves us with 8 signals which are linearly nondependent.

The patient can’t remain still during the treatment so the ECG also measures muscles electrical activity and this can be seen as a noise that we have in the ECG.

Because of all the reasons mentioned before we would use PCA to determine which 4 signals we would remove from the ECG data before passing it to the ICA.

After many trials we found out that we get the best results when we remove the first two signals and the last two signals after performing PCA. We can consider the first two signals (with the highest energy) to be the muscle noise, especially that the electrodes is close to the muscles than to the heart, and the last two signals (with the smallest energy) to be the electrical networks noise.
Normal ECG example: (ECG $\Rightarrow$ PCA)

Sick ECG example: (ECG $\Rightarrow$ PCA)
4.2.4 Performing ICA:

We used ICA with Relative-Newton method, provided by Michael Zibulevsky from the Department of Electrical Engineering at the Technion.

Now we perform ICA on the 8 signals that we have got from the PCA and we get as a result new 8 signals.

Normal ECG example: (ECG $\Rightarrow$ PCA $\Rightarrow$ ICA)

Sick ECG example: (ECG $\Rightarrow$ PCA $\Rightarrow$ ICA)
4.2.5 Comparing the signals:

After performing ICA we get new 8 signals. This is the last stage we compare between the signals that we got in order to find differences between the normal ECG and the sick ECG. We would try to find a way to show that the sick ECG gets better with time.

The results and the conclusions are in the next chapter.
5. Results:

We got three ways to identify the sick ECG from the normal ECG. And they are:

5.1 QRS signal:

We have a signal that appears in both ECG; normal and sick ECG. This signal can be thought on as representing the RQS signal in the ECG. The most obvious thing we see immediately is that the signal itself is very "clean", it does not have ripples.

The QRS signal for the normal ECG:

![Normal QRS](image1)

We found out that the similarity between the sick ECG and the normal ECG in the QRS signal gets better with the time, in other words the similarity of the last sample of the sick ECG is better than the similarity of the first sample of the sick ECG. The results for hanipas and breitman ECG:

![Comparison between Normal and Hanipas](image2)

Zoom
According to the graphs that we saw before, we can use the QRS signal to monitor a patient condition after heart attack. This can help building a program that monitors the patient ECG, and it alerts when there is any progress in the ECG or if the patient situation is getting worse.

From the original sick ECG (a figure is found before), we can't see the QRS signal clearly, cause the patient is not laying without any movement, and also he might be feeling pains that make his muscles produce electrical effort.

We should notice that we got here a signal that similar to the QRS signal that ECG contains but we don’t see them clearly in the original ECG, and so we can consider the process that we did as a process for cleaning the original ECG data from noise.

5.2 P and T signals:

We have another signal that appears in both ECG; normal and sick ECG. This signal can be considered as the signal that represents the P and T signals in the ECG. This signal is similar in some way to the P and T signals that supposed to be in the original ECG, but we don’t see them clearly in the original ECG. The signal contains two main parts:
We got two main peaks in the signal:

Part 1: The first peak can be thought of as the P signal of the ECG. It's almost in the middle of the cycle and also it's the smallest between the two peaks we have. We should notice that the position of the P signal in the original ECG isn’t at the middle but at the first part of the cycle, therefore we should not take it as the P signal.

Part 2: The second peak can be thought of as the T signal of the ECG. It’s in the last part of the cycle and this is what we have in the ECG, therefore we can consider it as the T signal.

The difference between this signal and the QRS signal is that we don’t see improvement in the sick ECG in three hours as we saw in the QRS signal, therefore we can conclude that this signal will take more time to return to it’s “normal” shape. This signal can be used to detect smaller heart attacks. Its can be used to monitor a patient situation for longer time. This signal might be used to predict heart attacks.

In the next figure we see the results of the two patients ECG. We can see that the main deference between the sick ECG and the normal ECG is in the P peak, while the T peak is almost the same.
Comparison between Normal and Hanipas

Comparison between Normal and Breitman
5.3 The signal number:

The last comparison result is the number of signals in the sick result that appears to be similar to the normal result signals.

In the beginning we have 3 common signals:
And in the end we have 4 common signals:

The important point in this result that, we can’t find such result in the original ECG, cause in the original ECG we have 12 signals and that’s it. But here we could detect sick ECG using the number of signals that is similar to the normal ECG.
6. Future work:

Our project is only one step in the path for predicting heart attacks. As have been seen from the results before we can identify heart attacks using blind source separation method, and in one of the results we got a way for identifying small heart attacks. But much work still lies in front of us to reach the ultimate goal – predicting heart attacks.

- A bigger and longer research has to be made in order to validate the results that we found. This require a bigger database than what we have and can be integrated in hospitals in addition to the traditional methods, and with time it can be proved as a valid method or not, and also it can be improved this way.

- Another research can be made to try and test the method on another kind of heart diseases, which might have predicate symptoms which are difficult to be identified in the traditional ECG.

- Building an automated heart attack identifier, which would identify heart attack (or other diseases) according to the parameters of the signal we got.
7. Appendix:

7.1 ConvertFile.m:

In order to work with the ECG data files, they must be converted to matlab matrices. This operation have been made by a function we build called “ConvertFile”. The function prototype:

ConvertFile(Filename, StartLimit, EndLimit)

Filename – The name of the file we want to convert.
StartLimit – The time we want to start reading from, in seconds.
EndLimit – The end time we want to stop reading at, in seconds.

The function will return the data as a matrix of 12 lines. If the EndLimit is nonpositive the function reads to the end of the file.

The data in the file is arranged as follows:

<table>
<thead>
<tr>
<th>Byte Number</th>
<th>Lead Number</th>
<th>Sample Number</th>
<th>Second Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 + (n-1)*24</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 + (n-1)*24</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 + (n-1)*24</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 + (n-1)*24</td>
<td>aVr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 + (n-1)*24</td>
<td>aVl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 + (n-1)*24</td>
<td>aVf</td>
<td>n</td>
<td>n/500</td>
</tr>
<tr>
<td>13 + (n-1)*24</td>
<td>V1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 + (n-1)*24</td>
<td>V2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 + (n-1)*24</td>
<td>V3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 + (n-1)*24</td>
<td>V4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 + (n-1)*24</td>
<td>V5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 + (n-1)*24</td>
<td>V6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We have 500 samples per a second, and every sample is represented by 2 bytes, which are a signed integer.
7.2 Comparison GUI:

In order to compare the last results we got, we built up a GUI which can helps us in this task:

1: This is the two signals that we want to compare, the blue graph is controlled by 2 and 4, and the red graph is controlled by 3 and 5.

2: This box contains the data in the workspace of the matlab and it’s updated at the start up and when we reset the data by “Reset Data”. We can choose the matrix we want to work with. It controls the blue graph.

3: This box is the same as 2 and it controls the red graph.

4: This controls the line number we want to plot from the matrix, it’s limited to be between zero and the number of the lines in the matrix. It can be updated manually or using the + and -. This area controls the blue graph.

5: This area is the same as 4 and it controls the red graph.

6: This area controls the red graph; it moves it up, down, right left. These controls are useful especially because the choosing of the cycles was manually.

7: This area controls the x-axes for the red graph. And also it can multiply the y-axes of the red graph by -1 (Invert Graph).
8: This area is for resetting the data. “Reset View” resets all the parameters we set using 6, 7 and 9. “Reset Data” rereads the data from the workspace and reset all the parameters we have.

9: This area is for displaying part of the ripples we have in the signals, it uses the matlab function “smooth” which makes average of the data. We can set the number of samples it uses to make the average.

10: This area displays all the parameters we set.