A Maximum Likelihood Framework for Iterative Eigendecomposition

A. Robles-Kelly and E. R. Hancock
The University of York
Heslington, York YO105DD, UK
{arobkell,erh}@cs.york.ac.uk

Abstract

This paper presents an iterative maximum likelihood framework for perceptual grouping. We pose the problem of perceptual grouping as one of pairwise relational clustering. The method is quite generic and can be applied to a number of problems including region segmentation and line-linking. The task is to assign image tokens to clusters in which there is strong relational affinity between token pairs. The parameters of our model are the cluster memberships and the link weights between pairs of tokens. Commencing from a simple probability distribution for these parameters, we show how they may be estimated using an EM-like algorithm. The cluster memberships are estimated using an eigendecomposition method. Once the cluster memberships are to hand, then the updated link-weights are the expected values of their pairwise products. The new method is demonstrated on region segmentation and line-segment grouping problems where it is shown to outperform a non-iterative eigenclustering method.

1. Introduction

Recently, there has been considerable interest in the use of matrix factorisation methods for perceptual grouping. These methods can be viewed as drawing their inspiration from spectral graph theory [2]. The basic idea is to commence from an initial characterisation of the perceptual affinity of different image tokens in terms of a matrix of link-weights. Once this matrix is to hand then its eigenvalues and eigenvectors are located. The eigenmodes represent pairwise relational clusters which can be used to group the raw perceptual entities together. There are several examples of this approach described in the literature. At the level of image segmentation, several authors have used algorithms based on the eigenmodes of an affinity matrix to iteratively segment image data. One of the best known is the normalised cut method of Shi and Malik [12]. Recently, Weiss [13] has shown how this, and other closely related methods, can be improved using a normalised affinity matrix. At higher level, both Sarkar and Boyer [10] and Perona and Freeman [9] have developed matrix factorisation methods for line-segment grouping. These non-iterative methods both use the eigenstructure of a perceptual affinity matrix to find disjoint subgraphs that represent the main arrangements of segmental entities.

Although elegant by virtue of their use of matrix factorisation to solve the underlying optimization problem, one of the criticisms which can be leveled at these methods is that their foundations are not statistical in nature. The aim in this paper is to overcome this shortcoming by developing a maximum likelihood framework for perceptual grouping. We pose the problem as one of pairwise clustering which is parameterised using two sets of indicator variables. The first of these are cluster membership variables which indicate to which perceptual cluster a segmental entity belongs. The second set of variables are link weights which convey the strength of the perceptual relations between pairs of nodes in the same cluster. We use these parameters to develop a mixture model which represents the pairwise clustering of the perceptual entities. We iteratively maximise the likelihood of the configuration of pairwise clusters using an EM-like algorithm. By casting the log-likelihood function into a matrix setting, we are able to estimate the cluster-memberships using matrix factorisation. Once these memberships are to hand, then the link-weights may be estimated.

The resulting iterative process may be regarded as the high-level analogue of a number of low-level iterative processes for perceptual grouping. Here several authors have explored the use of iterative relaxation style operators for edgel grouping. This approach was pioneered by Shashua and Ullman [11] and later refined by Guy and Medioni [3] among others. Parent and Zucker have shown how co-circularity can be used to gauge the compatibility of neighbouring edges [7]. Our method differs from these methods by virtue of the fact that it uses a statistical framework rather than one dictated by considerations from neurobiology.

*Supported by CONACYT, under grant No. 146475/151752.
The outline of this paper is as follows. Section 2 reviews previous work on how matrix factorisation may be applied to the link-weight matrix to perform perceptual grouping. In Section 3 we develop our maximum likelihood framework and show how the parameters of the model, namely the cluster membership probabilities and the pairwise link-weights can be estimated using an iterative EM-like algorithm. In Section 4 we describe how the method can be applied to line-grouping. Here we present a simple model which can be used to initialise the link-weights. The Section 3 also provides a sensitivity study on synthetic data and furnishes some examples on real world images. In Section 5 we provide a preliminary study of the application of the method to region segmentation. Finally, Section 6 concludes the paper by summarising our contributions and offering directions for future research.

2. Grouping by Matrix Factorisation

We pose the problem of perceptual grouping as that of finding the pairwise clusters which exist within a set of image tokens. These objects may be pixels, or segmental entities such as corners, lines, curves or regions. However, in this paper we focus on the two problems of region segmentation and grouping line-segments. The process of pairwise clustering is somewhat different to the more familiar one that represent the objects and the set of weighted edges partition the node-set into disconnected subgraphs. If \( \mathcal{S}_i \) represents one of these subsets and \( \mathcal{S}_i \neq \mathcal{S}_j \) if \( i, j \) are not part of the same cluster. We denote the set of nodes assigned to the cluster with modal index \( \omega \) as \( V_\omega = \{ i \} \). The elements of this matrix convey the following meaning in the hard limit:

\[
A_{ij} = \begin{cases} 
1 & \text{if there exists a partition } V_\omega \text{ such that } i \in V_\omega \text{ and } j \in V_\omega, \\
0 & \text{otherwise}
\end{cases}
\]  

(1)

In this paper we are interested in how matrix factorisation methods can be used to locate the set of edges which partition the nodes. One way of viewing this is as the search for the permutation matrix which re-orders the elements of \( A \) into non-overlapping blocks. However, when the elements of the matrix \( A \) are not binary in nature, then this is not a straightforward task. However, Sarkar and Boyer [10] have shown how the positive eigenvectors of the matrix of link-weights can be used to assign nodes to perceptual clusters. Using the Rayleigh-Ritz theorem, they observe that the scalar quantity \( \chi^T A \chi \) where \( \chi \) is the weighted adjacency matrix, is maximised when \( \chi \) is the leading eigenvector of \( A \). Moreover, each of the subdominant eigenvectors corresponds to a disjoint perceptual cluster. We confine our attention to the same-sign positive eigenvectors (i.e. those whose corresponding eigenvalues are real and positive, and whose components are either all positive or are all negative in sign). If a component of a positive eigenvector is non-zero, then the corresponding node belongs to the perceptual cluster associated with the associated eigenvalue of the weighted adjacency matrix. The eigenvalues \( \lambda_1, \lambda_2, ..., \lambda_k \) are the solutions of the equation \( |A - \lambda I| = 0 \) where \( I \) is the \( N \times N \) identity matrix. The corresponding eigenvectors \( \mathbf{x}_\lambda, \mathbf{x}_{\lambda_2}, ..., \mathbf{x}_{\lambda_k} \) are found by solving the equation \( A \mathbf{x}_\lambda = \lambda \mathbf{x}_\lambda \). Let the set of positive same-sign eigenvectors be represented by \( \Omega = \{ \omega | \lambda_\omega > 0 \land \{ \mathbf{x}_\omega(i) > 0 \forall i \vee \mathbf{x}_\omega(i) < 0 \forall i \}\} \). Since the positive eigenvectors are orthogonal, this means that there is only one value of \( \omega \) for which \( \mathbf{x}_\omega(i) \neq 0 \). In other words, each node \( i \) is associated with a unique cluster. We denote the set of nodes assigned to the cluster with modal index \( \omega \) as \( V_\omega = \{ i \} \).

3. Maximum Likelihood Framework

In this paper, we are interested in exploiting the factorisation property of Sarkar and Boyer [10] to develop a maximum likelihood method for updating the link-weight matrix \( A \) with the aim of developing a more robust perceptual grouping method. We commence by factorising the likelihood of the observed arrangement of objects over the set of modal clusters of the link-weight matrix. Since the set of modal clusters are disjoint we can write,

\[
P(A) = \prod_{\omega \in \Omega} P(\Phi_\omega)
\]  

(2)

where \( P(\Phi_\omega) \) is the probability distribution for the set of link-weights belonging to the modal-cluster indexed \( \omega \). To
model the component probability distributions, we introduce a cluster membership indicator which models the degree of affinity of the object indexed $i$ to the cluster with modal index $\omega$. This is done using the magnitudes of the modal co-efficients and we set

$$s_{i\omega} = \frac{|x'_i(i)|}{\sum_{i \in V_\omega}|x'_i(i)|}$$

(3)

Using these variables, we develop a model of probability distribution for the link-weights associated with the individual clusters. We commence by assuming that there are putative edges between each pair of nodes $(i, j)$ belonging to the cluster. The set of putative edges is $\Phi_\omega = V_\omega \times V_\omega - \{\{i, i\} | i \in V\}$. We further assume that the link-weights belonging to each cluster are independent of one another and write

$$P(\Phi_\omega) = \prod_{(i,j) \in \Phi_\omega} P(A_{i,j})$$

(4)

To model the probability distribution for the individual link-weights, we adopt the Bernoulli distribution and write

$$p(A_{i,j}) = A_{i,j}^{s_{i\omega} s_{j\omega}} (1 - A_{i,j})^{1 - s_{i\omega} s_{j\omega}}$$

(5)

This distribution takes on its largest values when either the link weight $A_{i,j}$ is unity and $s_{i\omega} = s_{j\omega} = 1$, or if the link weight $A_{i,j} = 0$ and $s_{i\omega} = s_{j\omega} = 0$.

With these ingredients the log-likelihood function for the observed pattern of link weights is

$$\mathcal{L} = \sum_{\omega \in \Omega} \sum_{(i,j) \in \Phi_\omega} \left\{ s_{i\omega} s_{j\omega} \ln A_{i,j} + (1 - s_{i\omega} s_{j\omega}) \ln (1 - A_{i,j}) \right\}$$

(6)

After some algebra to collect terms, the log-likelihood function simplifies to

$$\mathcal{L} = \sum_{\omega \in \Omega} \sum_{(i,j) \in \Phi_\omega} \left\{ s_{i\omega} s_{j\omega} \ln \frac{A_{i,j}}{1 - A_{i,j}} + \ln (1 - A_{i,j}) \right\}$$

(7)

Posed in this way the structure of the log-likelihood function is reminiscent of that underpinning the expectation-maximisation algorithm. The modes of the link-weight matrix play the role of mixing components. The product of cluster-membership variables $s_{i\omega} s_{j\omega}$ plays the role of an a posteriori measurement probability. Secondly, the link-weights are the parameters which must be estimated. However, there are important differences. The most important of these is that the modal clusters are disjoint. As a result there is no mixing between them.

Based on this observation, we will exploit an EM-like process to update the link-weights and the cluster-membership variables. In the “M” step we will locate maximum likelihood link-weights. In the “E” step we will use the revised link-weight matrix to update the modal clusters. To this end we index the link weights and cluster memberships with iteration number and aim to optimise the quantity

$$Q(A^{(n+1)} | A^{(n)}) = \sum_{\omega \in \Omega} \sum_{(i,j) \in \Phi_\omega} \left\{ s_{i\omega}^{(n)} s_{j\omega}^{(n)} \ln \frac{A_{i,j}^{(n+1)}}{1 - A_{i,j}^{(n+1)}} + \ln (1 - A_{i,j}^{(n+1)}) \right\}$$

(8)

The revised link weight parameters are indexed at iteration $n+1$ while the cluster-memberships are indexed at iteration $n$.

### 3.1. Expectation

To update the cluster-membership variables we have used a gradient-based method. We have computed the derivatives of the expected log-likelihood function with respect to the cluster-membership variable

$$\frac{\partial Q(A^{(n+1)} | A^{(n)})}{\partial s_{i\omega}^{(n)}} = \sum_{j \in V_\omega} s_{j\omega}^{(n)} \ln \frac{A_{i,j}^{(n+1)}}{1 - A_{i,j}^{(n+1)}}$$

(9)

Since the associated saddle-point equations are not tractable in closed form, we use the soft-assign ansatz of Bridle [1] to update the cluster membership assignment variables. This involves exponentiating the partial derivatives of the expected log-likelihood function in the following manner

$$s_{i\omega}^{(n+1)} = \frac{\exp \left[ \frac{\partial Q(A^{(n+1)} | A^{(n)})}{\partial s_{i\omega}^{(n)}} \right]}{\sum_{i \in V_\omega} \exp \left[ \frac{\partial Q(A^{(n+1)} | A^{(n)})}{\partial s_{i\omega}^{(n)}} \right]}$$

(10)

As a result the update equation for the cluster membership indicator variables is

$$s_{i\omega}^{(n+1)} = \frac{\exp \left[ \sum_{j \in V_\omega} s_{j\omega}^{(n)} \ln \frac{A_{i,j}^{(n+1)}}{1 - A_{i,j}^{(n+1)}} \right]}{\sum_{i \in V_\omega} \exp \left[ \sum_{j \in V_\omega} s_{j\omega}^{(n)} \ln \frac{A_{i,j}^{(n+1)}}{1 - A_{i,j}^{(n+1)}} \right]}$$

(11)

### 3.2. Modal Sharpening

Although the update equation given in the previous section is effective for updating the cluster membership variables, we have found that significant improvements can be achieved by refining the modal structure of the clusters prior to updating the link matrix. To proceed, we introduce some...
matrix notation. We commence by representing the cluster-
memberships of the cluster indexed \( \omega \) using the column vec-
tor \( z_{\omega}^T = (s_{1\omega}, \ldots, s_{n\omega})^T \). We also define a \( [V] \times [V] \)
weight-matrix \( W^\omega \), whose elements are

\[
W^\omega_{i,j} = \zeta_{i,j,\omega} \ln \frac{A_{ij}^{(n)}}{1 - A_{ij}^{(n)}}
\]

where

\[
\zeta_{i,j,\omega} = \begin{cases} 1 & \text{if } s_{i\omega}^{(n)} \neq 0 \text{ and } s_{j\omega}^{(n)} \neq 0 \\ 0 & \text{otherwise} \end{cases}
\]

With this notation, the algorithm focuses on the quantity

\[
\tilde{Q}(A^{(n+1)} | A^{(n)}) = \sum_{\omega \in \Omega} S^T \Sigma W^\omega S_{\omega}
\]

In this way the log-likelihood function is decomposed into contributions from the distinct modal clusters. Moreover, each cluster weight matrix \( W^\omega \) is disjoint. For each such matrix, we will perform a further eigendecomposition to identify the foreground and background modal structure. Recall that Sarkar and Boyer [10] have shown that the scalar quantity \( \bar{z}^T A \bar{z} \), where \( A \) is the weighted adjacency matrix and \( \bar{z} \) is a vector of cluster-membership variables, is maximised when \( \bar{z} \) is the leading eigenvector of \( A \). Unfortunately, we can not exploit this property directly. The reasons for this are twofold. First, the utility measure underpinning our maximum likelihood algorithm is a sum of terms of the form \( S^T W^\omega S_{\omega} \). Second, the elements of \( W^\omega \) may be negative (since it is computed by taking logarithms) and hence its eigenvalues will not be real. We overcome the first of these problems by applying the Rayleigh-Ritz theorem to each weight matrix \( W^\omega \) in turn. Each such matrix represents a distinct cluster and its leading eigenvector represents the individual cluster-membership affinities of the nodes. To overcome the second problem we make use of the fact that the directions of the eigenvectors of the matrices \( A \) and \( \ln A \) are identical. We therefore commence from the matrix \( W^\omega \) whose elements are

\[
W'_{ij} = \frac{A_{ij}^{(n+1)}}{1 - A_{ij}^{(n+1)}}
\]

We use the components of the leading eigenvector \( z_{\omega} \) of \( W^\omega \) to perform “modal sharpening” on the cluster memberships. They are re-assigned according to the following formula

\[
s_{i\omega}^{(n+1)} = \frac{|z_{\omega}(i)|}{\sum_{i=1}^N |z_{\omega}(i)|}
\]

3.3. Maximisation

Once the revised cluster membership variables are to hand then we can apply the maximisation step of the algo-

rithm to update the link-weight matrix. The updated link-
weights are found by computing the derivatives of the expected log-likelihood function

\[
\frac{\partial Q(A^{(n+1)} | A^{(n)})}{\partial A_{ij}^{(n+1)}} = \sum_{\omega \in \Omega} \zeta_{i,j,\omega} \left\{ \frac{s_{i\omega}^{(n)} s_{j\omega}^{(n)}}{A_{ij}^{(n+1)}(1 - A_{ij}^{(n+1)})} - \frac{1}{1 - A_{ij}^{(n+1)}} \right\}
\]

and solving the saddle-point equations

\[
\frac{\partial Q(A^{(n+1)} | A^{(n)})}{\partial A_{ij}^{(n+1)}} = 0
\]

As a result the updated link-weights are given by

\[
A_{ij}^{(n+1)} = \sum_{\omega \in \Omega} s_{i\omega}^{(n)} s_{j\omega}^{(n)}
\]

In other words, the link-weight for the pair of nodes \((i, j)\) is simply the average of the product of individual node cluster memberships over the different perceptual clusters. Since each node is associated with a unique cluster, this means that the updated affinity matrix is composed of non-overlapping blocks. Moreover, the link-weights are guaranteed to be in the interval \([0, 1]\).

3.4. Algorithm description

To summarise, the iterative steps of the algorithm are as follows:

- (1) Compute the eigenvectors of the current link-weight matrix \( A^{(n)} \). Each same-sign eigenvector whose eigenvalue is positive represents a disjoint pairwise cluster. The number of such eigenvectors determines the number of clusters for the current iteration. This number may vary from iteration to iteration.
- (2) Compute the updated cluster-membership variables using the E-step. At this stage modal sharpening may be performed to improve the cluster-structure if desired. This sharpening process may be iterated to refine the current set of clusters.
- (3) Update the link-weights using the M-step to compute the updated link weight matrix \( A^{(n+1)} \).
- Goto step (1).

4. Line Grouping

In this section we provide the first example application of our new clustering method. This involves the grouping or linking of line-segments.
4.1. Initial Line-Grouping Field

We are interested in locating groups of line-segments that exhibit strong geometric affinity to one-another. In this section we provide details of a probabilistic linking field that can be used to gauge geometric affinity. This problem has attracted considerable interest in the literature. For instance, Heitiger and von der Heydt [4] have shown how to model the line extension field using directional filters whose shapes are motivated by studies of the visual field of monkeys. Parent and Zucker [7] use edges co-circularity compatibility. Williams and his co-workers [15][14] have taken a different approach using the stochastic completion field. Here the completion field of curvilinear features is computed using Monte Carlo simulation of particle trajectories between the end-points of contours.

Here we follow the former approach and to provide an initial characterisation of the matrix of link-weights using a grouping field. To be more formal suppose we have a set of line-segments \( \mathcal{L} = \{ \Lambda_i; i = 1, \ldots, n \} \). Consider two lines \( \Lambda_i \) and \( \Lambda_j \) drawn from this set. Their respective lengths are \( l_i \) and \( l_j \). Our model of the linking process commences by constructing the line \( \Gamma_{i,j} \) which connects the closest pair of endpoints for the two lines. The geometry of this connecting line is represented using the polar angle \( \theta_{ij} \) of the line \( \Gamma_{i,j} \) with respect to the base-line \( \Lambda_i \) and its length \( \rho_{ij} \). We measure the overall scale of the arrangement of lines using the length of the shorter line \( \rho_{i,j} = \min[l_i, l_j] \).

The relative length of the gap between the two line-segments is represented in a scale-invariant manner using the dimensionless quantity \( \xi_{i,j} = \frac{\rho_{i,j}}{\rho_{i,j}} \).

Following Heitiger and Von der Heydt [4] we model the linking process using an elongated polar grouping field. To establish the degree of geometric affinity between the lines we interpolate the end-points of the two lines using the polar lemniscate \( \xi_{i,j} = k \cos^2 \theta_{ij} \).

The value of the constant \( k \) is used to measure the degree of affinity between the two lines. For each linking line, we compute the value of the constant \( k \) which allows the polar locus to pass through the pair of endpoints. The value of this constant is

\[
k = \frac{\rho_{i,j}}{\rho_{i,j} \cos \theta_{ij}}
\]

It is important to note that the polar angle is defined over the interval \( \theta_{ij} \in (-\pi/2, \pi/2) \) and is rotation invariant.

We use the parameter \( k \) to model the linking probability for the pair of line-segments. When the lemniscate envelope is large, i.e. \( k \) is large, then the grouping probability is small. On the other hand, when the envelope is compact, then the grouping probability is large. To model this behaviour, we assign the linking probability using the exponential distribution

\[
A_{ij}^{(0)} = \exp[-\lambda k]
\]

where \( \lambda \) is a constant whose best value has been found empirically to be unity. As a result, the linking probability is large when either the relative separation of the endpoints is small i.e. \( \rho_{i,j} \ll \rho_{i,j} \) or the polar angle is close to zero or \( \pi \), i.e. the two lines are colinear or parallel. The linking probability is small when either the relative separation of the endpoints is large i.e. \( \rho_{i,j} \gg \rho_{i,j} \) or the polar angle is close to \( \pi/2 \), i.e. the two lines are perpendicular.

4.2. Experiments

In this Section we provide some experiments to illustrate the utility of our new perceptual grouping method when applied to line-linking. There are two aspects to this study. We commence by providing some examples for synthetic images. Here we investigate the sensitivity of the method to clutter and compare it with an eigen-decomposition method. The second aspect of our study focuses on real world images with known ground-truth.
4.2.1 Synthetic Images

The first sequence of synthetic images is shown in Figure 1. Here the foreground structure is an approximately circular arrangement of line-segments. In the first column of Figure 1 we show the arrangement of lines with increasing amounts of added clutter. In the subsequent columns we show the results of grouping for each of the images of the left-hand column in turn. In each row the second image is the pattern of foreground line segments extracted by applying the non-iterative eigendecomposition method described in [6] to the grouping field already detailed in Section 4.1. This method can be viewed as a variant of the Sarkar and Boyer [10] eigenclustering method, which uses entropic thresholding of the affinity matrix and uses positivity constraints on the eigenvectors to locate perceptual clusters. The third image in each row shows the result obtained with the EM-like algorithm when modal sharpening is used. Here the line segments are coded according to the value of the cluster membership weights \(w_{ij}\), where \(w_i\) is the foreground cluster label (i.e. the one associated with the largest eigenvalue). In each case, the foreground cluster located by the EM algorithm contains less noise contamination than the result delivered by eigendecomposition. Moreover, none of the line segments leaks into the background. The final column in each row shows the result obtained with the EM-like algorithm when the cluster memberships are updated using soft-assign. Although the results are better than those obtained using non-iterative eigendecomposition, they are poorer than those obtained with modal sharpening.

We have repeated the experiments described above for a sequence of synthetic images in which the density of distractors increases. For each image in turn we have computed the number of distractors merged with the foreground pattern and the number of foreground line-segments which leak into the background. Figures 5 a and b respectively show the fraction of nodes merged with the foreground and the number of nodes which leak into the background as a function of the number of distractors. The three curves shown in each plot are for the non-iterative eigendecomposition method and for the EM-like algorithm when both soft-assign and modal sharpening are used to update the cluster membership weights. In both cases, the shoulder of the response curve for the two variants of the EM-like algorithm occurs at a significantly higher error rate than that for the non-iterative eigen-decomposition method. Of the two alternative methods for updating the cluster membership weights in the EM algorithm, modal sharpening works best.

Finally, we present results on a real-world image in 3. The edges shown in Figure 3b have been extracted from the raw image using the Canny edge-detector. Straight-line segments have been extracted using the method of Yin [8]. The resulting clusters obtained with the EM method can be seen in Figure 3c.

5. Grey Scale Image Segmentation

The second application of our new pairwise clustering method involves segmenting grey-scale images into regions. To compute the initial affinity matrix, we use the difference in grey-scale values at different pixel sites. Suppose that \(g_i\) is the grey-scale value at the pixel indexed \(i\) and \(g_j\) is the grey-scale value at the pixel indexed \(j\). The corresponding entry in the affinity matrix is

\[
A_{ij}^{(0)} = \exp[-k_j (g_i - g_j)^2]
\] (22)

If we are segmenting an \(R \times C\) image of \(R\) rows

![Figure 3. Real world images: a) raw image, b) results of Canny edge detection and c) the result of applying the Eigendecomposition algorithm.](image-url)
and C columns, then the affinity matrix is of dimensions $RC \times RC$. This initial characterisation of the affinity matrix is similar to that used by Shi and Malik [12] in their normalised-cut method of image segmentation.

By applying the clustering method to this initial affinity matrix we iteratively segment the image into regions. Each eigenmode corresponds to a distinct region.

![Image](image.png)

**Figure 4.** Results obtained on synthetic images using the Eigendecomposition method and the Softassign.

![Image](image.png)

**Figure 5.** Number of segmentation errors versus noise standard deviation for non-iterative eigendecomposition (top curve), soft-assign EM (middle curve) and modal sharpening EM (lower curve).

5.1. Experiments

We have conducted experiments on both synthetic and real images. We commence considering the effect of added random noise. In Figure 4 a), b), c), and d) we show a sequence of images in which we have added Gaussian noise of zero mean and known standard deviation to the grey-scale values in an image containing three rectangular regions. In the sequence, each triple of images represents the original noise corrupted image (on the top), the segmentation result obtained with the EM algorithm when cluster-memberships are updated using modal sharpening (middle) and the segmentation obtained using the EM algorithm with cluster membership update using soft-assign (bottom). For the different images, standard deviation of the added Gaussian noise is 35%, 50%, 65% and 95% of the grey-scale difference between the regions. The final segmentations are obtained with an average of 2.3 iterations per cluster. The final segmentations contain 3, 3, 4 and 4 clusters respectively. The method begins to fail once the noise exceeds 60%. It is also worth noting that the region boundaries and corners are well reconstructed.

Figure 5 offers a more quantitative evaluation of the segmentation capabilities of the method. Here we compute the fraction of mislabelled pixels in the segmented images as a function of the standard deviation of the added Gaussian noise. The plot shows three performance curves obtained with a) the non-iterative eigendecomposition algorithm, b) the EM-like method with soft-assign and c) the EM-like algorithm with modal sharpening. The non-iterative method fails abruptly at low noise-levels. The two variants of the

![Image](image.png)

**Figure 6.** Segmentations of grey-scale images.
EM-like algorithm perform much better, with the modal sharpening method offering a useful margin of advantage over the soft-assign method.

To conclude this section, in Figure 6 we provide some example segmentations on real-world images. In the top row we show the original image, the middle row is the segmentation obtained with EM and modal sharpening, while the bottom row shows the segmentation obtained with EM and soft-assign. On the whole the results are quite promising. The segmentations capture the main region structure of the images. Moreover, they are not unduly disturbed by brightness variations or texture. The modal sharpening method gives the cleanest segmentations. It should be stressed that these results are presented to illustrate the scope offered by our new clustering algorithm and not to make any claims concerning its utility as a tool for image segmentation. To do so would require comparison and sensitivity analysis well beyond the scope of this paper.

One of the interesting properties of our method is that the number of modes or clusters changes with each iteration of the algorithm. This is because we perform a new modal analysis each time the link-weight matrix $A^{(m)}$ is updated. For the segmentation results shown in fig 6, we have investigated how the number of modal clusters varies with iteration number. In Figure 7 we show the number of active clusters as a function of iteration number for each of the real-world images. In each case the number of clusters increase with iteration number. In the best case the number of clusters stabilizes after 2 iterations, and in the worst case after 6 iterations.

6. Conclusions

In this paper, we have presented a new perceptual clustering algorithm which uses an EM-like algorithm to estimate link-weights and cluster membership probabilities. The method is based on an iterative modal decomposition of the link-weight matrix. The modal cluster membership probabilities are modeled using a Bernoulli distribution for the link-weights. We apply the method to the problems of region segmentation and of line-segment grouping. In the case of line-segment grouping, the method appears robust to severe levels of background clutter. Although more preliminary, the results obtained for region segmentation are promising and underline the flexibility of the new method.

References