Minimum distance constrained nonnegative matrix factorization for the endmember extraction of hyperspectral images

Yue YU*, Shan GUO, Weidong SUN
Remote Sensing Laboratory
Institute of Image and Graphics
Department of Electronic Engineering
Tsinghua University, Beijing, 100084

ABSTRACT
Endmember extraction and spectral unmixing is a very challenging task in multispectral/hyperspectral image processing due to the incompleteness of information. In this paper, a new method for endmember extraction and spectral unmixing of hyperspectral images is proposed, which is called as minimum distance constrained nonnegative matrix factorization (MDC-NMF). After being compared with a newly developed method named MVC-NMF, MDC-NMF not only has been demonstrated more reasonable in theory but also shows promising results in the experiments.

Keywords: Endmember extraction, hyperspectral image, nonnegative matrix factorization, distance constrained.

1. INTRODUCTION
Endmember extraction and spectral unmixing, which has received a considerable interest in remote sensing community during recent years, is a very challenging task in multispectral/hyperspectral image processing.

It is always assumed that the measured spectrum of a single pixel is a mixture of several constituent spectra known as endmembers, weighted by their corresponding fractions known as abundances. This assumption is true because the spatial resolution of most nowadays spectral imaging sensors is large enough compared with the size of the objects being observed and the observing area is usually a mixing object. Thus, the objective of spectral unmixing is to recover the endmembers as well as their abundances, given such observed mixed pixels only. Numerous algorithms have already been proposed to solve this problem, either from the concepts of convex geometry or from the perspectives of blind source separation. However, due to the incompleteness of information, endmember extraction and spectral unmixing still remains an open problem.

Nonnegative Matrix Factorization (NMF) is a newly developed approach for image analysis and computer vision, which is originally designed to extract a set of nonnegative basis vectors from natural images. As a matter of fact, such nonnegative basis vectors play a similar role as endmembers in linear mixing model (LMM). Later this connection was noticed by remote sensing community and the conception of NMF was applied to spectral unmixing.

However, it is worth pointing out that the solution of the NMF problem is not unique. In order to make this problem well-posed, some extra constraint other than nonnegativity is required. Most recently, some researchers claim to use volume as an additional constraint. However, as it will be made clear in section 2, incorporating a volume constraint into the formulation of NMF is not the best choice. In this paper, a new constrained NMF algorithm for spectral unmixing of hyperspectral images is proposed, which is called as minimum distance constrained nonnegative matrix factorization (MDC-NMF).

The rest of this paper is organized as follows. In section 2, linear mixing model (LMM) and nonnegative matrix factorization (NMF) will be briefly introduced first. Then the details of MDC-NMF will be discussed. In section 3, the performance of MDC-NMF will be evaluated using a set of experiments. Conclusion and future work will be included in section 4.

* yuyue05@mails.tsinghua.edu.cn
2. METHODOLOGY

2.1 Linear mixing model

Linear mixing model (LMM) is widely used in endmember extraction and spectral unmixing for its simplicity, tractability and effectiveness. An LMM is valid when the reflecting surface can be portrayed as a checkerboard mixture, and any given package of incident radiation only interacts with one component\(^1\). Mathematically, LMM is expressed as:

$$\mathbf{x} = \sum_{i=1}^{M} a_i \mathbf{s}_i + \mathbf{w} = \mathbf{S} \mathbf{a} + \mathbf{w}$$  \hspace{1cm} (1)

In the above equation, \( \mathbf{x} \in \mathbb{R}^L \) denotes the measured spectrum of a single pixel with \( L \) spectral bands, and the subscript \( _i \) means that every component of \( \mathbf{x} \) should be nonnegative. \( \mathbf{S} \in \mathbb{R}^{N \times M} \) denotes the endmember matrix whose columns (i.e. \( \mathbf{s}_i \in \mathbb{R}^L, i = 1, \ldots, M \) ) are the corresponding endmembers’ spectra, and \( M \) is the number of endmembers. \( \mathbf{a} \in \mathbb{R}^M \) denotes the abundance vector and in order to make this model physically plausible, the components of \( \mathbf{a} \) have to satisfy two constraints, namely the abundance nonnegative constraint (i.e. \( a_j \geq 0, j = 1, \ldots, M \) ) and the abundance sum-to-one constraint (i.e. \( \sum_{j=1}^{M} a_j = 1 \)). The possible error of this model and noises are denoted by \( \mathbf{w} \in \mathbb{R}^L \).

When \( N \) pixels are considered together, block notation can be utilized which results in the following model:

$$\mathbf{X} = \mathbf{S} \mathbf{A} + \mathbf{W}$$ \hspace{1cm} (2)

with \( \mathbf{X} = [\mathbf{x}(1), \ldots, \mathbf{x}(N)] \), \( \mathbf{A} = [\mathbf{a}(1), \ldots, \mathbf{a}(N)] \) and \( \mathbf{W} = [\mathbf{w}(1), \ldots, \mathbf{w}(N)] \).

2.2 Nonnegative matrix factorization

As indicated by equation (1), it is obvious that, if all the endmembers are extracted correctly, then all the pixels in the hyperspectral image should be represented by these extracted endmembers as perfectly as possible (provided that the assumption of linear mixing is valid), to put it another way, which means that the error term \( \mathbf{w} \) in the linear mixing model should be as small as possible under some measure. Thus, this is the first condition which must be satisfied.

If Euclidean distance is adopted as this measure (for its simplicity and some good mathematical properties that we need), the condition described above can be directly formulated as an optimization problem given by

$$\text{minimize} \quad f(S, A) = \frac{1}{2} \left\| \mathbf{X} - \mathbf{S} \mathbf{A} \right\|_F^2$$ \hspace{1cm} (3)

subject to \( \mathbf{S} \geq 0 , \quad \mathbf{A} \geq 0 , \quad \mathbf{1}_M \mathbf{A} = \mathbf{1}_N \mathbf{r} \)

where the symbol \( \geq \) denotes componentwise inequality, the operator \( \| \|_F \) represents the Frobenius norm and \( \mathbf{1}_M (\mathbf{1}_N \mathbf{r}) \) is an \( M \) (\( N \))-dimensional column vector of all 1s.

If we omit the equality constraint \( \mathbf{1}_M \mathbf{A} = \mathbf{1}_N \mathbf{r} \) for a while, then it is readily seen that, the optimization problem formulated in equation (3) is exactly the same as the nonnegative matrix factorization (NMF) problem proposed by Lee and Seung\(^8\). Thus, in the rest of this paper, the optimization problem formulated in equation (3) (with equality constraint) will be referred to as modified NMF, or still NMF for short, if not indicated explicitly.

2.3 Minimum distance constraint

However it should be pointed out that, only trying to minimize the representation error in linear mixing model (see equation (3)) is not enough, which is only necessary but not sufficient, for the success of endmember extraction and spectral unmixing. This is because the solution of the NMF problem is not unique which can be easily seen in the following manner, either algebraically or geometrically.
For an algebraic interpretation, let \( \hat{\mathbf{S}} \) and \( \hat{\mathbf{A}} \) be a solution of the NMF problem, let \( \mathbf{D} \in \mathbb{R}^{M \times M} \) be an invertible matrix, then it is obviously true that \( \hat{\mathbf{S}} \mathbf{D} \) and \( \mathbf{D}^{-1} \hat{\mathbf{A}} \) will still be a solution of this NMF problem, as long as \( \hat{\mathbf{S}} \mathbf{D} \succeq \mathbf{0} \), \( \mathbf{D}^{-1} \hat{\mathbf{A}} \succeq \mathbf{0} \) and \( \mathbf{1}_M^T \mathbf{D}^{-1} \hat{\mathbf{A}} = \mathbf{1}_M^T \), which can be achieved by carefully choosing some proper invertible matrix \( \mathbf{D} \).

For an alternative geometric interpretation, first assume that, for the purpose of clarity, the number of endmembers is known, say \( n \), then it is apparent that any simplex with \( n \) vertices which can enclose the data cloud completely in a high dimensional space (where the dimension must be greater than or equal to \( n-1 \)) will be a solution of the NMF problem geometrically, and such simplex will be referred to as a qualified simplex hereafter in this paper. Thus, theoretically speaking, there are infinitely many qualified simplexes which are also solutions of the NMF problem. A two-dimensional case is illustrated in Fig. 1.

![Fig. 1. Geometric interpretation of the nonuniqueness of the solution. The number of spectral bands is two, the number of endmembers is three, which are represented by the circles \( \mathbf{E}_1 - \mathbf{E}_3 \) and the filled circles can be considered as mixed pixels. Every triangle which can enclose all the filled circles (e.g. the two larger triangles represented by dashed lines and dot-dashed lines respectively) is a qualified simplex with its three vertices being the corresponding extracted endmembers.](image)

In summary, in order to make the NMF problem well-posed (to make the solution of which unique), some extra constraint other than the abundance nonnegative constraint and the abundance sum-to-one constraint is required. In fact, there is a widely accepted belief in remote sensing community that, the volume of the simplex determined by the true endmembers should be the minimum among that of all the simplexes enclosing the data cloud. Thus, it is implied that the volume of the simplex determined by the extracted endmembers could be used as an additional constraint, which is exactly what has been used by Miao and Qi. However, incorporating a volume constraint into the formulation of NMF is not the best choice, which will be made clear shortly after. Actually, to find a simplex with minimum volume is in essence to find a simplex which is as compact as possible and this compactness can be measured by other means, which are more suitable for the formulation of NMF.

Here an alternative new approach to measure this compactness is proposed. Technically, the distance is measured and summed up from every endmember to their centroid, which is defined as Endmember Distance (ED) and is closely related to the volume of the simplex determined by these endmembers. Mathematically, ED can be expressed as:

\[
ED(\mathbf{S}) = \sum_{i=1}^{M} \left\| \tilde{s}_i - \bar{s} \right\|^2_2, \quad \bar{s} = \frac{1}{M} \sum_{i=1}^{M} \tilde{s}_i
\]  

or equivalently in matrix form:

\[
ED(\mathbf{S}) = \text{trace}(\mathbf{P}^T \mathbf{S}^T \mathbf{SP}) = \text{trace}(\mathbf{SPS}^T), \quad \mathbf{P} = \mathbf{I} - \frac{1}{M} \mathbf{1}_M \mathbf{1}_M^T
\]
where the operator $\text{trace}(\bullet)$ represents the trace of a matrix, $\mathbf{I}$ denotes an $M \times M$ identity matrix, $\mathbf{1}_M$ is an $M$-dimensional column vector of all 1s by convention and $\mathbf{P}$ denotes a projection matrix which make the columns of $\mathbf{S}$ zero-mean.

Now, combining the goal of minimizing the representation error in equation (3) with the ED constraint (or penalty term, regularization term) in equation (5), we arrive at the following multicriterion optimization problem, namely minimum distance constrained nonnegative matrix factorization (MDC-NMF):

$$
\begin{align*}
\text{minimize} \quad & f(\mathbf{S}, \mathbf{A}) = \frac{1}{2} \| \mathbf{X} - \mathbf{S} \mathbf{A} \|_F^2 + \lambda \times \text{trace}(\mathbf{S} \mathbf{P}^T) \\
\text{subject to} \quad & \mathbf{S} \geq 0, \quad \mathbf{A} \geq 0, \quad \mathbf{1}_M^T \mathbf{A} = 1_n^T 
\end{align*}
$$

with $\lambda \in \mathbb{R}_+$ a predefined constant known as regularization parameter used to control the tradeoff between the representation error and the ED constraint.

### 2.4 Theoretical comparison

A point should be made clear. MDC-NMF is closely related to a newly developed method for endmember extraction and spectral unmixing, named minimum volume constrained nonnegative matrix factorization (MVC-NMF). One of the main differences between MVC-NMF and MDC-NMF as mentioned in section 2.3, is that they claim to use simplex volume as a constraint, while we propose to use ED instead.

Compared with simplex volume, ED has at least the following advantages:

First, using ED has lower computational load, that is, only matrix multiplication is needed. While using simplex volume needs to calculate both matrix determinant and matrix inversion (matrix inversion is required in the optimization algorithm of MVC-NMF).

Secondly, ED is a convex (to be exact, quadratic) function of $\mathbf{S}$ whereas simplex volume is not. Thus, although neither of MVC-NMF and MDC-NMF is convex in both variables together (i.e. $\mathbf{S}$ and $\mathbf{A}$) and local minimum exists, it is still noticed that if one of the variables (say $\mathbf{S}$) is fixed, then it is convex for MDC-NMF in the other variable (say $\mathbf{A}$). However, this is only partly true for MVC-NMF, because when $\mathbf{A}$ is fixed, MVC-NMF is not convex in $\mathbf{S}$ due to the existence of the simplex volume term. To sum up, the above argument implies that when an alternating minimization procedure is adopted for such NMF problems (MVC-NMF or MDC-NMF), that is, we alternatively update one variable, holding the other one fixed. Then at each iteration step, the subproblem of MDC-NMF has only global minimum and can be solved by highly efficient and well developed convex optimization algorithms (e.g. interior-point method$^{10}$) , while the subproblem of MVC-NMF has local minimum and is more dependent on gradient-based methods which can suffer from slow convergence.

Thirdly, compared with MDC-NMF, MVC-NMF has some problem of numerical instability, because the calculation of the simplex volume is carried out using the determinant of the augmented matrix of endmembers as follows:

$$
\text{Volume}(\mathbf{S}) = \frac{1}{(M-1)!} |\text{det}(\mathbf{Z})|, \quad \mathbf{Z} = \begin{bmatrix} 1 & \ldots & 1 \\ s_1 & \ldots & s_M \end{bmatrix}
$$

or equivalently

$$
\text{Volume}(\mathbf{S}) = \frac{1}{(M-1)!} |\text{det}([s_2-s_1, \ldots, s_M-s_1])|
$$

Thus, to make the simplex volume in (8) small is in essence equivalent to make the $M-1$ column vectors $s_2-s_1, \ldots, s_M-s_1$ linear dependent, which means that MVC-NMF always tends to force the $M-1$ translated endmembers (for which the new origin is $s_1$, accordingly) to reside in a lower-dimensional subspace which may not yield satisfying results. Furthermore, MVC-NMF needs to calculate the inverse of the augmented matrix $\mathbf{Z}$ at each iteration step, however, as the simplex volume becomes small, $\mathbf{Z}$ will become near singular, which causes numerical instability.
Lastly, as mentioned above, simplex volume is calculated through equation (7) or (8), however in most cases, the number of endmembers is much smaller than the number of spectral bands, which means that, the augmented matrix $Z$ is non-square most of the time and the determinant is not defined on non-square matrices. Under the circumstances, an approximation is usually made by performing some dimension reduction technique first, typically Principal Component Analysis (PCA)\(^{11}\) or Minimum Noise Fraction transform (MNF)\(^{12}\). Since MNF is just a special form of PCA\(^ {13}\), we will take PCA for example.

When a certain amount of noise exists in the hyperspectral image, or when the number of the pixels of some endmember is much smaller than that of others, then there is a large probability that such endmember may not reside in the subspace spanned by the first few principal components, which means that, after PCA, the minimum volume constraint almost does not take effect in the direction determined by such endmember. The basic idea is illustrated in Fig. 2, from which it can be seen that $\text{volume} [\hat{E}, \vec{E}_{\text{else}}] = \text{volume} (\vec{E}_{\text{else}})$, where $\hat{E}$ represents the projection of $E$ on the subspace $S$, $E$ is a statistically insignificant endmember and $\vec{E}_{\text{else}}$ represents the projection of the other endmembers on the same subspace. On the contrary, MDC-NMF does not encounter such problem.

2.5 Optimization algorithm

It is necessary to reiterate that, the optimization problem formulated in equation (6) is not convex in both variables together and local minimum exists. Thus, it is critical to design an appropriate optimization algorithm so that some satisfying results of endmember extraction and spectral unmixing can be obtained.

Here we first and basically resort to an alternating minimization procedure which is actually the block coordinate descent method in bound-constrained optimization\(^ {14}\). This method decomposes the original optimization problem into two subproblems as follows:

$$S^{k+1} = \arg \min_{S} f(S, A^{k})$$
$$A^{k+1} = \arg \min_{A} f(S^{k}, A)$$

(9)

Fortunately, as mentioned in section 2.4, each subproblem in equation (9) is convex and can be accurately solved at each iteration step by some well developed algorithm (e.g., interior-point method).

However, it is found out later that, there is no need to accurately solve each subproblem at every iteration step, more surprisingly, doing so can even lead to a slower convergence (here convergence means finding the right endmembers) and poorer results.
Thus, after a slight modification of equation (9), we finally arrive at the following optimization algorithm:

\[
\begin{align*}
S^{k+1} &= \arg \min_{S \geq 0} f(S, A^k) \\
A^{k+1} &= ((S^{k+1})^T S^{k+1})^{-1}(S^{k+1})^T X \\
\tilde{A}^{k+1} &= \max(0, A^{k+1}) \\
A_{lm}^{k+1} &= \frac{A_{lm}^{k+1}}{\sum_{i=1}^{M} A_{im}^{k+1}}, \quad 1 \leq l \leq M, \quad 1 \leq m \leq N
\end{align*}
\]

(10)

That is, for \( S \), the updating rule remains unchanged as in equation (9), while for \( A \), an unconstrained least-squares solution is first calculated, then the negative components of \( A \) are simply forced to zeros and at last each column of \( A \) is rescaled to meet the abundance sum-to-one constraint.

It should be noted that, although a rigid mathematical proof for the convergence of the optimization algorithm in equation (10) is not available now, it turns out that this algorithm converges well at least in all our experiments and the updating rule for \( A \) in equation (10) is the key step to assure promising results of endmember extraction.

Nevertheless, a heuristic interpretation of this algorithm can be given as follows: If the simplex determined by the extracted endmembers (i.e. \( S \)) can enclose the data cloud completely, then, as can be easily shown, \( A \) will remain unchanged before and after the last two steps in equation (10), thus in the next iteration step, \( S \) only needs to be further adjusted to make the simplex more compact; However, if this is not true, that is, the simplex determined by the extracted endmembers can not enclose the data cloud completely, then \( A \) must have changed after the last two steps in equation (10) and this change of \( A \) will cause an increase in the representation error which will be decreased by the updating rule for \( S \) in the next iteration step, and doing so will make the extracted endmembers updated and approximate the real ones step by step.

3. EXPERIMENTAL RESULTS

3.1 Experiment 1

In this experiment, MDC-NMF has been tested on two simple datasets to demonstrate its feasibility first. For the purpose of illustrating, the number of spectral bands is fixed set to 2. As can be verified from Fig. 3, MDC-NMF can indeed find the vertices of the simplexes (i.e. a line segment or a triangle), despite the presence of some small amounts of error, which is introduced as a tradeoff between the minimum representation error and the ED constraint as mentioned in section 2.3.

![Fig. 3. Illustration of the feasibility of MDC-NMF using two simple datasets with (a) the number of endmembers set to 2 and (b) the number of endmembers set to 3, respectively. Here, green 'point' represents the mixed pixels, red 'asterisk' represents the initializations of endmembers, red 'circle' represents the extracted endmembers at each iteration step follow the arrow directions, and blue 'square' represents the true endmembers. The number of iterations is set to 200 and \( \lambda = 1 \).](image-url)
3.2 Experiment 2

In this experiment, the performances of MVC-NMF and MDC-NMF have been compared with each other, using another simple dataset which is specially designed for this purpose and the result is shown in Fig. 4. For the purpose of illustrating, the number of spectral bands is still fixed to 2 and the number of endmembers is fixed to 3.

![Fig. 4](image1.png)

Fig. 4. Comparison of (a) the performance of MVC-NMF and (b) the performance of MDC-NMF. The meanings of the markers in this figure are the same as those in Fig. 3 and the configuration of the parameters in this experiment is the same as that in experiment 1.

As can be seen from Fig. 4, though maybe not obviously enough, MVC-NMF has always a tendency to force the translated endmembers (i.e. $s_1 - s_2, s_1 - s_3$) to reside in a lower-dimensional subspace, as argued in section 2.4. In this experiment, such assertion means that MVC-NMF always tends to make the extracted endmembers collinear, the validity of which can be made more apparent if $\lambda$ is set relatively large, say 20 and the corresponding result is shown in Fig. 5.

![Fig. 5](image2.png)

Fig. 5. Comparison of (a) the performance of MVC-NMF and (b) the performance of MDC-NMF on the same dataset as in Fig. 4. Again, the meanings of the markers in this figure are the same as those in Fig. 3 and the configuration of the parameters is the same as that in experiment 1 except that $\lambda$ is set to 20 at this time.

From Fig. 5, it can be seen that, under such extreme configuration (i.e. $\lambda = 20$ ), MVC-NMF has totally failed to estimate the position of the vertices of this simplex (triangle), in fact, all the endmembers extracted by MVC-NMF have nearly fallen on the same black dashed line, while the results given by MDC-NMF can still be accepted to some extent, even though the extracted endmembers by MDC-NMF have already fallen in the data cloud.

3.3 Experiment 3

In this experiment, a synthetic hyperspectral image with the size of $64 \times 64$ is used, which is generated by randomly mixing four endmember spectra under the abundance sum-to-one constraint and the abundance nonnegative constraint. These four endmember spectra are selected from the spectral libraries of the ENVI software, and namely, they are buddingtonite, kaolinite, nontronite and sphene.
For the purpose of comparison, one measure adopted to evaluate the similarity between the true endmember spectrum \( s \) and its estimate, the extracted endmember spectrum \( \hat{s} \), is the well-known spectral angle mapper (SAM), defined in equation (11). Obviously, the smaller SAM is, the more similar the two spectra will be.

\[
SAM = \cos^{-1}\left(\frac{s^T \hat{s}}{\|s\| \|\hat{s}\|}\right)
\]  

(11)

In order to eliminate the influence of the mean value of each spectrum, the spectral correlation coefficient is used as another measure defined in equation (12), with the mean values of \( s \) and \( \hat{s} \) denoted by \( \bar{s} \) and \( \bar{\hat{s}} \), respectively and \( \mathbf{1}_L \) is an \( L \)-dimensional column vector of all 1's by convention. However, for CORR, the larger the value is, the more similar the two spectra will be.

\[
CORR = \left(\bar{s} - \bar{\hat{s}}\right)^T \left(\hat{s} - \bar{\hat{s}}\right) / \left\|\bar{s} - \bar{\hat{s}}\right\| \left\|\hat{s} - \bar{\hat{s}}\right\|
\]  

(12)

The result of this experiment shows that, MDC-NMF performs much faster than MVC-NMF does (see Table 1), and also gets better results in the sense that, compared with MVC-NMF, the spectra of the endmembers extracted by MDC-NMF have higher correlation with those of the true endmembers, when the similarity is measured through both equation (11) and (12) (see Table 2 and Fig. 6).

Table 1. Comparison of the execution time (in seconds) of MDC-NMF and MVC-NMF. The result is obtained on a Pentium 1500 MHz Windows-based laptop and the number of iterations is set to 400.

<table>
<thead>
<tr>
<th></th>
<th>MDC-NMF</th>
<th>MVC-NMF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Execution time (sec)</td>
<td>319.69</td>
<td>1084</td>
</tr>
</tbody>
</table>

Table 2. Comparison of the similarity measured by both SAM and CORR between the spectra of the true endmembers and those of the endmembers extracted by MVC-NMF and MDC-NMF, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Sphene</th>
<th>Nontronite</th>
<th>Kaolinite</th>
<th>Buddingtonite</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>MVC-NMF</td>
<td>0.0253</td>
<td>0.0210</td>
<td>0.0179</td>
<td>0.0537</td>
</tr>
<tr>
<td></td>
<td>MDC-NMF</td>
<td>0.0358</td>
<td>0.0341</td>
<td>0.0089</td>
<td>0.0063</td>
</tr>
<tr>
<td>CORR</td>
<td>MVC-NMF</td>
<td>0.9976</td>
<td>0.9988</td>
<td>0.9913</td>
<td>0.9711</td>
</tr>
<tr>
<td></td>
<td>MDC-NMF</td>
<td>0.9990</td>
<td>0.9995</td>
<td>0.9977</td>
<td>0.9998</td>
</tr>
</tbody>
</table>

Fig. 6. Comparison of (a) the endmembers’ spectra extracted by MVC-NMF and (b) the endmembers’ spectra extracted by MDC-NMF. Red dashed lines represent the extracted spectra of the endmembers, while blue solid lines represent the spectra of the true endmembers which are most correlated to the spectra extracted.
From Table 2 and Fig. 6, when the similarity is measured by SAM, it can be seen that, for the first two endmembers (i.e. Sphene and Nontronite), the results obtained by MVC-NMF and MDC-NMF are almost the same, or to be honest, the result obtained by MVC-NMF is even a little better. But, for the last two endmembers (i.e. Kaolinite and Buddingtonite), the result obtained by MDC-NMF is much better than that obtained by MVC-NMF. On the average, MDC-NMF also outperforms MVC-NMF, as can be seen from the last column of Table 2. However when the similarity is measured by CORR, it can be seen that MDC-NMF totally outperforms MVC-NMF for all the four endmembers.

4. CONCLUSION AND FUTURE WORK

Nonnegative matrix factorization (NMF) is a powerful approach for basis extraction. However due to the lack of uniqueness of the solution, it can not be directly used in endmember extraction and spectral unmixing. Some other constraint other than nonnegativity needs to be added. In this paper a new constraint NMF algorithm for endmember extraction and spectral unmixing has been proposed, which is called as minimum distance constrained nonnegative matrix factorization (MDC-NMF). After a comparison with a newly developed method named MVC-NMF, MDC-NMF not only has been demonstrated more reasonable in theory, but also shows promising results in the experiments. However, as a new method, there are still some important issues that are in need of further research. First, MDC-NMF needs to be tested on some real hyperspectral images to fully evaluate its performance. Second, some rigid mathematical reasoning for the convergence of the optimization algorithm proposed in section 2.5 needs to be investigated.

5. ACKNOWLEDGEMENTS

This work is supported by the Cultivation Found of the Key Scientific and Technical Innovation Project, Ministry of Education of China (No. 706004), and also supported by National Natural Science Foundation of China (No. 60472029). The authors would also like to thank Doc. Miao for providing the source code of MVC-NMF.

REFERENCES