IMAGE INPAINTING VIA WEIGHTED SPARSE NON-NEGATIVE MATRIX FACTORIZATION[*]

Yu-Xiong Wang, Yu-Jin Zhang

Department of Electronic Engineering, Tsinghua University
Tsinghua National Laboratory for Information Science and Technology, Tsinghua University
Beijing 100084, China
albertwyx@gmail.com, *zhang-yj@tsinghua.edu.cn*

ABSTRACT

This paper proposes a novel patch propagation inpainting algorithm based on Weighted Sparse Non-negative Matrix Factorization (WSNMF). Unlike existing methods, we cast the inpainting task as a sequential low-rank matrix recovery and completion problem, where the incomplete data matrix consists of the image patch to be inpainted and several similar intact candidate patches under the assumption that they can be described using a low-dimensional linear model. Besides, the non-negativity and sparsity constraints are enforced for the additive sparse linear combination. The WSNMF, based on the Expectation-Maximization (EM) procedure, is then introduced to predict missing values. Experimental results show that this approach exploits the available information from the source region more adequately and thus has capabilities to recover both structure and composite textures more effectively as well as preventing unwanted artifacts compared to current exemplar-based techniques.

Index Terms— Image inpainting, Non-negative Matrix Factorization (NMF), matrix completion, weighted low-rank approximation

1. INTRODUCTION

How can we make computer modify an image in a visually undetectable way analogous to sophisticated artists? This is quite an amazing topic in both art and image processing areas, with applications from the restoration of damaged paintings and photographs to the removal of selected objects. Image inpainting exactly aims to reconstitute the missing region (called the target region) using information from the remaining image areas (called the source region).

The challenge of image inpainting roots deeply in the nature of real-world scene photographs, which often consists of 1-D or 0-D linear structures, such as edges and corners, and 2-D pure or composite textures [3, 4]. The boundaries (or called the fill front) between the target and source regions are a complex product of the mutual influences of these factors. In this sense, propagating the spatially interacted multiple textures while preserving the structures becomes the core concern.

The existing image inpainting techniques can be divided into two categories: diffusion-based approach [1, 2] and exemplar-based approach [3, 4, 11, 13, 14]. They differ from each other on the focused image level, and thus have disparate performance. The former tackles the filling-in problem by diffusing the image from the known surrounding regions into the missing region at the pixel level by using the variational principles and partial differential equations (PDE). Hence, it is superior for structure propagation or relatively smaller missing region, yet poor in handling textured or large region due to the introduction of smoothing effect. The latter propagates the image information at the patch level based on the texture synthesis technique. By incorporating the patch priorities to determine the filling order, the exemplar-based method can deal with structure propagation as well as texture propagation, and hence outperforms the diffusion-based one with respect to large missing region.

One of the core stages in exemplar-based inpainting algorithms is how to synthesize the needed texture by exploiting the known candidate patches from the source region. Criminisi et al. selected the single best match patch in taking the risk of putting all eggs in one basket [3, 4]. Wong and Orchard proposed a nonlocal-means approach to infer the target patch by weighting a set of similar candidate patches [13]. This indeed reduces the greediness; nevertheless it also introduces unwished blurring effect. Shen et al. borrowed the signal sparse representation theory to fit the missing patch by sparse combination of a redundant dictionary constructed by source patches [11]. Xu and Sun furthered the sparse representation model by introducing more regulation terms [14]. In both cases, the sharp recovery results are achieved together with less greedy procedure. However, in our experiments, it is found that the most similar candidate patch always takes a dominant role with coefficient up to 0.8 to 0.9 far outweighing other patches, even if there exist some patches of nearly equal similarity. Meanwhile, the candidate patches with smaller similarity have little effect in the combination and the helpful information from them is lost, though the dictionary itself is over-complete. For that matter, this approach is still greedy, leading to unwanted object in the recovered region. The foregoing strategies are alike in that they handle the fitting problem in the original image domain. To utilize the available information more adequately, we change to the transformed domain, and treat the inpainting problem under the framework of sequential low-rank matrix recovery and completion. To be specific, we assume the image patch to be inpainted and the top several similar candidate patches as random samples from the same source to construct an

[*] This work was partially supported by National Nature Science Foundation (NNSF: 60872084).

978-1-4577-1302-6/11/$26.00 ©2011 IEEE

incomplete data matrix. This is thus a matrix completion problem, which has been well studied and can be solved by weighted low-rank approximation approaches [12]. Here the modified Weighted Sparse Non-negative Matrix Factorization (WSNMF) algorithm is applied to accord with the characteristics of inpainting problem.

The remainder of this paper is organized as follows. In Section 2 the weighted low-rank approximation and non-negative matrix factorization theory are reviewed briefly. The proposed inpainting algorithm is elaborated in Section 3. Section 4 shows the experimental results. Discussions and conclusions are drawn in Section 5.

2. WEIGHTED NMF

Low-rank approximation (LRA), which tries to find a parsimonious representation, is a fundamental tool in multivariate data analysis. The formulation of LRA can be regarded as decomposing the original data matrix into two or three low-rank factor matrices. By imposing the non-negativity constraint, Lee and Seung initiated a new LRA paradigm called Non-negative Matrix Factorization (NMF) [7]. Due to the purely additive combination, NMF obtains the parts-based representation and thus enhances the interpretability of the issue.

In the case of incomplete data matrix with some entries missing or unobserved, matrix completion is imperative to predict missing elements while obtaining the low-rank representation [12]. By introducing a weight matrix with binary weights 1 or 0 to differentiate between the observed and unobserved values, the matrix completion problem can be solved through Weighted Non-negative Matrix Factorization (WNMF) [9]. Such approach has been applied in collaborative filtering successfully [6, 15].

Generally speaking, given an M-D random vector $\boldsymbol{x}$ with non-negative elements, whose N observations are denoted as $\boldsymbol{x}_{j,\, j=1,2,\ldots,N}$, let data matrix be $\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N]\in\mathbb{R}_{\geq 0}^{M\times N}$, NMF seeks non-negative basis matrix $\boldsymbol{U}\in\mathbb{R}_{\geq 0}^{M\times L}$ and coefficient matrix $\boldsymbol{V}\in\mathbb{R}_{\geq 0}^{L\times N}$, such that $\boldsymbol{X}\approx\boldsymbol{UV}$. Using Frobenius norm as the measurement, it minimizes the following objective function

$$F_{NMF}(\boldsymbol{X},\boldsymbol{UV})=\frac{1}{2}\left\|\boldsymbol{X}-\boldsymbol{UV}\right\|_F^2=\frac{1}{2}\sum_{ij}(\boldsymbol{X}_{ij}-[\boldsymbol{UV}]_{ij})^2 \tag{1}$$

And WNMF seeks to minimize the following objective function

$$J_{WNMF}(\boldsymbol{X},\boldsymbol{UV})=\frac{1}{2}\sum_{ij}\boldsymbol{W}_{ij}(\boldsymbol{X}_{ij}-[\boldsymbol{UV}]_{ij})^2 \tag{2}$$

where $\boldsymbol{W}_{ij}$ are non-negative weights.

WNMF can be solved by introducing the weight matrix and modifying the standard NMF iterative update rules [6, 9]. An alternative is to employ the EM algorithm where missing entries are replaced by the corresponding values in the current model estimation at the E-step, and the unweighted NMF is applied on the filled-in matrix at the M-step [6, 15].

3. INPAINTING ALGORITHM

We now elaborate the inpainting algorithm via WSNMF. Given an input image I, the user selects a target region Ω to be removed and filled. The remaining or manually specified areas can be defined as the source region Φ. The boundary of the target region is indicated by $\delta\Omega$. The M-D vector $\boldsymbol{\Psi}_p$, denoting the $k\times k$ image patch centered at pixel p, is the basic processing unit in the exemplar-based approach. The whole inpainting procedure consists of the propagation of patches inward sequentially from the continuously updated boundary according to predefined filling order.

3.1. Filling order

A crucial technique in exemplar-based inpainting algorithm is how to determine the filling order so as to balance the recovery of both texture and structure. Here we follow the method proposed by Criminisi et al. [3, 4], which encourages the filling-in of patches on the high-confidence structure. At each step, the patch priority $P(p)$ for every pixel p on the boundary $\delta\Omega$ is computed, and then the patch $\boldsymbol{\Psi}_{pm}$ with the highest priority is selected as the target patch in the current iteration. The details of the patch priority computation can be found in [3, 4].

3.2. Construction of data matrix and weight matrix

Once the target patch $\boldsymbol{\Psi}_{pm}$ has been found, it can be filled by using the available information from the source region as much as possible. Similar to [11, 13, 14], we search N-1 patches denoted as $\boldsymbol{\Psi}_{qj,\, j=2,\ldots,N}$ in the source region, which are most similar to $\boldsymbol{\Psi}_{pm}$. Formally

$$\boldsymbol{\Psi}_{qj}=\arg\min_{\Psi_q\in\Phi\setminus\Psi_{qk,k=2,\ldots,j-1}} d(\boldsymbol{\Psi}_{pm},\boldsymbol{\Psi}_q) \tag{3}$$

where the distance $d(\boldsymbol{\Psi}_a,\boldsymbol{\Psi}_b)$ between two patches is still measured by the sum of squared differences (SSD) defined in the already filled parts of both patches. Taking account of the computational consumption, the original source region in the whole image can be shrunk into sub-source region defined in a window of certain size centered at the target pixel.

So the data matrix is constructed as

$$\boldsymbol{X}=[\boldsymbol{\Psi}_{pm},\boldsymbol{\Psi}_{q2},\ldots,\boldsymbol{\Psi}_{qj},\ldots\boldsymbol{\Psi}_{qN}]=[\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_N]\in\mathbb{R}_{\geq 0}^{M\times N} \tag{4}$$

whose column vectors are assumed as N observations of the same random vector. It should be pointed out that the components of $\boldsymbol{\Psi}_{pm}$ corresponding to the unknown pixels located in the target region are simply replaced with 0. Since the patch $\boldsymbol{\Psi}_{pm}$ is an incomplete signal with some elements lost while $\boldsymbol{\Psi}_{qj,\, j=2,\ldots,N}$ are all intact, the data matrix $\boldsymbol{X}$ needs to be completed.

Before utilize the following NMF approach for matrix completion, we are supposed to define the corresponding weight matrix $\boldsymbol{W}=[\boldsymbol{W}_1,\boldsymbol{W}_2,\ldots,\boldsymbol{W}_N]\in\mathbb{R}_{\geq 0}^{M\times N}$ first.

Different weights assigning strategies have been adopted for the incomplete and complete signals in $\boldsymbol{X}$, respectively. For $\boldsymbol{W}_1$, the binary weights are given by

$$\boldsymbol{W}_{i1}=\begin{cases}1 & \text{if } \boldsymbol{X}_{i1} \text{ is in the source region}\\ 0 & \text{if } \boldsymbol{X}_{i1} \text{ is in the target region}\end{cases} \tag{5}$$

Since the similarity between the target patch and the candidate patch decreases from $\boldsymbol{\Psi}_{q2}$ to $\boldsymbol{\Psi}_{qN}$, this implies the decay in the confidence of these candidate patches. So proper weights are needed in response to the relative consequences of $\boldsymbol{X}_{j,\, j=2,..,N}$. This is equivalent to choose a decreasing function of $d(\boldsymbol{\Psi}_a,\boldsymbol{\Psi}_b)$, such as $W(\boldsymbol{\Psi}_j)=\exp(-d(\boldsymbol{\Psi}_p,\boldsymbol{\Psi}_j)/h)$ or $W(\boldsymbol{\Psi}_j)=c/d(\boldsymbol{\Psi}_p,\boldsymbol{\Psi}_j)$. Here we select the latter, and let

$$\boldsymbol{W}_{ij}=\frac{\min(d(\boldsymbol{\Psi}_{pm},\boldsymbol{\Psi}_{qj}))}{d(\boldsymbol{\Psi}_{pm},\boldsymbol{\Psi}_{qj})}=\frac{d(\boldsymbol{\Psi}_{pm},\boldsymbol{\Psi}_{q2})}{d(\boldsymbol{\Psi}_{pm},\boldsymbol{\Psi}_{qj})},\ \text{for } i=1,\ldots,M,\ j=2,\ldots,N \tag{6}$$

where the components in the same patch have the equal weights, and the coefficient in the numerator is introduced to scale the weights between 0 and 1, which will be clarified in Section 3.3.

3.3. EM procedure based WSNMF

Now the original inpainting task has been converted into a matrix completion problem, which can be solved by WNMF as discussed in Section 2. Here additional sparseness constraint similar to [5] is imposed on the coefficient matrix $\boldsymbol{V}$ to enforce the sharp inpainting result, referred to as Weighted Sparse Non-negative Matrix Factorization (WSNMF), which is different from the existing WNMF model. The objective function to be minimized becomes

$$J_{WSNMF}(\boldsymbol{X},\boldsymbol{UV}) = \frac{1}{2}\sum_{ij}\boldsymbol{W}_{ij}(\boldsymbol{X}_{ij} - [\boldsymbol{UV}]_{ij})^2 + \lambda\sum_{ij}\boldsymbol{V}_{ij} \tag{7}$$

Our aim is to predict the missing values through low-rank matrix factorization. Recall Section 2, here we still adopt the idea of EM based WLRA like [6, 12, 15], which views the WLRA problem as a maximum-likelihood problem with missing values; nevertheless the specific optimization approach will be changed to meet the new objective function. In brief, a filled-in matrix $\boldsymbol{Y}$ is computed from the current model estimation at the Expectation step, and unweighted Sparse NMF (SNMF) is utilized on $\boldsymbol{Y}$ to re-estimate the decomposition model at the Maximization step. The details are as follows.

E-step

The update rule is similar to [12, 15]. Formally

$$\boldsymbol{Y} \leftarrow \boldsymbol{W} \otimes \boldsymbol{X} + (\boldsymbol{I}_{M\times N} - \boldsymbol{W}) \otimes (\boldsymbol{UV}) \tag{8}$$

where $\boldsymbol{I}_{M\times N} \in \mathbb{R}^{M\times N}$ is the matrix with all entries equaling to 1, and $\otimes$ is Hadamard multiplication. The weight matrix needs normalization such that all elements are in interval [0, 1]. In addition, the estimation of missing pixel values in $\boldsymbol{Y}$ at the initial iteration can simply be the means of the counterparts of the candidate patches.

M-step

At this stage, a standard SNMF for matrix $\boldsymbol{Y}$ is required, and several effective SNMF algorithms have already been developed [5, 8]. Here we apply the SENSC algorithm proposed by Li and Zhang [8], considering that it obtains relatively low reconstruction error while preserving certain sparsity. The effectiveness of SENSC has been demonstrated in image clustering [10]. For details of this algorithm, please refer to [8]. To speed up, the simple trick of partial M-step which avoids determining optimal solutions at earlier iterations can also be adopted [6].

As the EM iterations proceed, the current target patch $\boldsymbol{\Psi}_{pm}$ is recovered according to the learnt low-rank matrix. Then the boundary of the target region is updated, and the previous steps repeats until all pixels have been filled.

3.4. Overall algorithm

As discussed above, a pseudo-code description of the overall algorithm is given in Table 1.

4. EXPERIMENTAL RESULTS

To evaluate the proposed algorithm, experiments on a variety of natural images, especially the ones containing large and complex holes with different neighborhood topologies, have been carried out. We also make comparisons with previous single exemplar-based (SE) [3, 4] and sparse representation (SR) based inpainting algorithms [11]. In the following experiments, the patch size is selected as 9×9 , and the patch number N is set to 25.

Table 1. Proposed inpainting algorithm.

```
Input: Image $I$ with target region $\Omega$, source region $\Phi$.
Repeat until all the pixels in $\Omega$ are filled.
1.  Compute patch priority for every pixel $p$ on the
    boundary $\delta\Omega$, and select the patch with the highest
    priority as the target.
2.  Find $N$-1candidate patches most similar to the target
    patch, and construct the incomplete data matrix and
    corresponding weight matrix according to (3) ~ (6).
3.  Recover the incomplete data matrix using WSNMF
    proposed in Section 3.3.
4.  Copy the pixel values in the target region of selected
    patch from the recovered data matrix.
5.  Update the target region $\Omega$ and the boundary $\delta\Omega$.
Output: The inpainted image.
```

Some typical results of different scenarios by our algorithm and other two algorithms for comparisons are shown in Fig.1 to Fig.3. In each figure, the first row gives the original image and the image to be inpainted with selected target region marked in green. The second row gives the inpainting results of these three algorithms. From left to right, are SE, SR, and our proposed algorithms, respectively. In Fig.1 and Fig.2, the additional third row shows the zoom-in details for fine comparisons.

Fig.1 manifests that the proposed approach is capable of inferring both structure and texture of large missing region. The horizontal structures are recovered successfully, and both the textures of the sky and the grass are padded appropriately by our algorithm, whereas there are some structural artifacts produced by the other two algorithms. Notice that the horizontal railings in Fig.1 (c) and (d) are dislocated, where the parts with light color are intersected by the parts with deep color improperly. Conversely, the color transition obtained by our approach is more naturally.

A more challenging task of composite texture inpainting is presented in Fig.2. From the zoom-in results in Fig.2 (c) and (d), we can see that there are obvious green plaques across the boundary between the rock and the grass layers by SE, while the rocks diffuse into the grass region by SR. So both algorithms have introduced unwanted objects due to their intrinsic greedy fashion. On the other side, our algorithm decreases this risk by exploiting the available information more fully and thus alleviates this negative effect. In Fig.3, this is more evidently demonstrated with regard to the mistakenly synthesized island region near the sea level by SE and SR.

5. CONCLUSIONS

This paper has proposed a novel algorithm for image inpainting based on Weighted Sparse Non-negative Matrix Factorization. The major contribution of this work is to cast the inpainting problem into the low-rank matrix recovery and completion framework. We no longer consider the single target patch as an incomplete signal, and try to fit it using the linear combination of several similar source patches under certain constraints as most of the existing techniques do; however, we integrated the target patch and the candidate patches as a higher level incomplete signal, and fit them simultaneously using the low-rank additive sparse linear combination of another self-adaptively constructed basis set in the transformed domain. Thus the information from these candidate patches is all combined. This approach is capable of inferring both structure and composite textures of large missing region with less greediness to prevent unwanted artifacts because of the more adequate exploitation of available information from multiple exemplars. It also achieves sharp inpainting results due to the introduction of sparseness prior on the combination coefficients. In the future, designing more effective mechanism for determining the filling order and applying suitable incremental NMF procedure to speed up the proposed approach will be conducted.

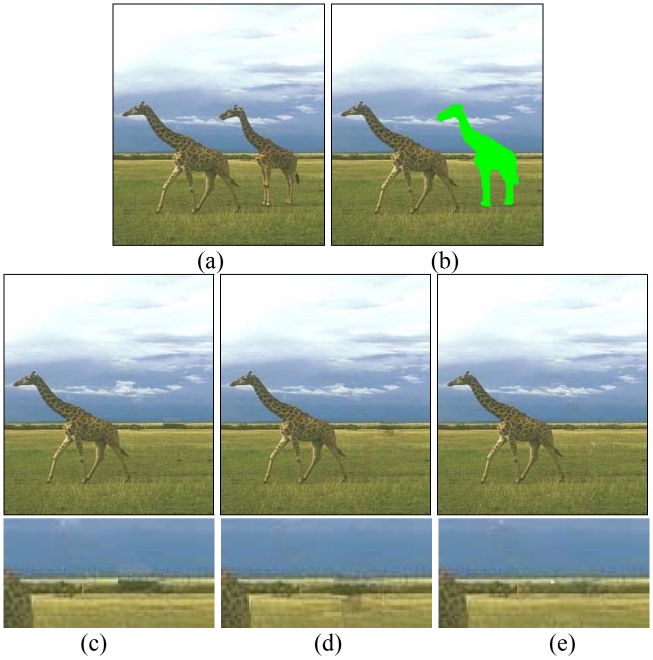

Fig.1. Structure and texture inpainting: (a) Original image. (b) Target region is marked in green. (c), (d), and (e) Inpainting results by SE, SR, and proposed algorithm, respectively. From top to bottom are the complete results and zoom-in detail comparisons.

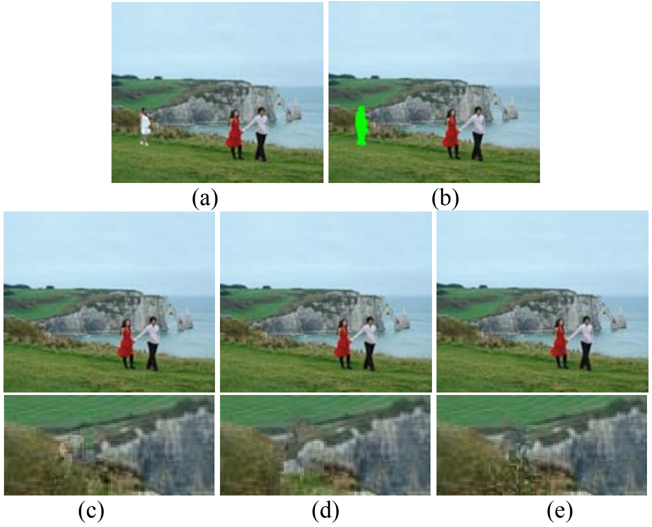

Fig.2. Composite texture inpainting: (a) Original image. (b) Target region is marked in green. (c), (d), and (e) Inpainting results by SE, SR, and proposed algorithm, respectively. From top to bottom are the complete results and zoom-in detail comparisons.

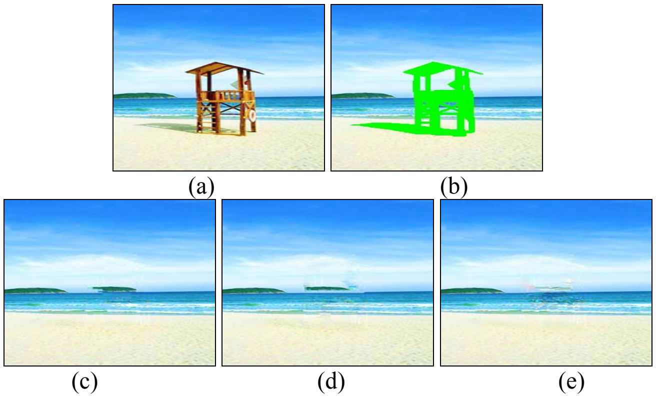

Fig.3. Unwanted artifact prevention: (a) Original image. (b) Target region is marked in green. (c), (d), and (e) Inpainting results by SE, SR, and proposed algorithm, respectively.

6. REFERENCES

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” SIGGRAPH, pp. 417–424, 2000.

[2] T. Chan and J. Shen, “Local inpainting models and TV inpainting,” SIAM J. Appl. Math., vol. 62, no. 3, pp. 1019-1043, 2001.

[3] A. Criminisi, P. Perez, and K. Toyama, “Object removal by exemplar-based image inpainting,” CVPR, pp. 721–728, 2003.

[4] A. Criminisi, P. Perez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” IEEE Trans. Image Processing, vol. 13, no. 9, pp. 1200–1212, 2004.

[5] P.O. Hoyer, “Non-negative sparse coding,” IEEE Workshop on Neural Networks for Signal Processing, pp. 557-565, 2002.

[6] Y.D. Kim and S. Choi, “Weighted nonnegative matrix factorization,” ICASSP, pp. 1541-1544, 2009.

[7] D.S. Lee and H.S. Seung, “Learning the parts of objects by non-negative matrix factorization,” Nature, vol. 401, pp. 788-791, 1999.

[8] L. Li and Y.J. Zhang, “SENSC: a stable and efficient algorithm for non-negative sparse coding,” Acta Automatica Sinica, vol. 35, no. 10, pp. 1257-1271, 2009.

[9] Y. Mao and L.K. Saul, “Modeling distances in large-scale networks by matrix factorization,” ACM SIGCOMM Conf. Internet Measurement, pp. 278-287, 2004.

[10] Y. Qin, L. Li, and Y.J. Zhang, “Evaluation of SENSC algorithm for image clustering,” ICIG, pp. 266-271, 2009.

[11] B. Shen, W. Hu, Y. Zhang, and Y.J. Zhang, “Image inpainting via sparse representation,” ICASSP, pp. 697-700, 2009.

[12] N. Srebro and T. Jaakkola, “Weighted low-rank approximations,” ICML, pp. 720-727, 2003.

[13] A. Wong and J. Orchard, “A nonlocal-means approach to exemplar-based inpainting,” ICIP, pp. 2600-2603, 2008.

[14] Z. Xu and J. Sun, “Image inpainting by patch propagation using patch sparsity,” IEEE Trans. Image Processing, vol. 19, no. 5, pp. 1153–1165, 2010.

[15] S. Zhang, W. Wang, J. Ford, and F. Makedon, “Learning from incomplete ratings using non-negative matrix factorization,” SIAM Int. Conf. Data Mining (SDM), pp. 549-553, 2006.