Supplementary Material
for
Illumination and Reflectance Spectra Separation of a Hyperspectral Image Meets Low-Rank Matrix Factorization

Yinqiang Zheng
National Institute of Informatics
yqzheng@nii.ac.jp

Imari Sato
National Institute of Informatics
imarik@nii.ac.jp

Yoichi Sato
The University of Tokyo
ysato@iis.u-tokyo.ac.jp

Abstract

In the supplementary material, we detail the similarities and discrepancies between the illumination and reflectance spectra separation (IRSS) problem and the well-known non-rigid structure-from-motion (NRSfM) problem.

1. The IRSS Problem

According to eq.(3) and eq.(4) in the main paper, the IRSS problem assumes the following low-rank matrix factorization formulation

$$\underbrace{\begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \cdots & \cdots & \cdots \\ d_{m1} & \cdots & d_{mn} \end{bmatrix}}_{D_{m\times n}} = \underbrace{\begin{bmatrix} l_1 & & \\ & \cdots & \\ & & l_m \end{bmatrix}}_{L_{m\times m}} \underbrace{\begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \cdots & \cdots & \cdots \\ r_{m1} & \cdots & r_{mn} \end{bmatrix}}_{R_{m\times n}}$$

$$= \underbrace{\begin{bmatrix} l_1 & & \\ & \cdots & \\ & & l_m \end{bmatrix}}_{L_{m\times m}} \underbrace{\begin{bmatrix} b_{11} & \cdots & b_{1s} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{ms} \end{bmatrix}}_{B_{m\times s}} \underbrace{\begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots \\ c_{s1} & \cdots & b_{sn} \end{bmatrix}}_{C_{s\times n}},$$

that is,

$$D_{m\times n} = L_{m\times m} R_{m\times n} = L_{m\times m} B_{m\times s} C_{s\times n}.$$

The illumination spectra matrix L is a $m \times m$ nonnegative diagonal matrix defined in the spectra domain. The reflectance spectra matrix R can be decomposed into the product of the spectral bases matrix B and the coefficient matrix C.

2. The NRSfM Probelm

The NRSfM problem under orthographic projection is a classical and noticeable problem in geometric vision [Ref.1,4,9,14]. Suppose that n feature points of a deforming object have been tracked across m image frames. Under orthographic projection, the 2D image coordinate $\mathbf{p}_{ij}$ of the j-th point in the i-th image relates to its 3D space coordinate $\mathbf{x}_{ij}$ by

$$\mathbf{p}_{ij} = R_i \mathbf{x}_{ij} + \mathbf{t}_i, 1 \leq i \leq m, 1 \leq j \leq n,$$

in which R_i is the 2×3 rotation of the i-th image such that $R_i R_i^T = I_2$. $\mathbf{t}_i$ is the translation, which can be simply eliminated by centralizing the image point coordinates.

By stacking all n points in m frames, the NRSfM problem can be formulated as

$$\underbrace{\begin{bmatrix} \mathbf{p}_{11} & \cdots & \mathbf{p}_{1n} \\ \cdots & \cdots & \cdots \\ \mathbf{p}_{m1} & \cdots & \mathbf{p}_{mn} \end{bmatrix}}_{\hat{P}_{2m\times n}} = \underbrace{\begin{bmatrix} R_1 & & \\ & \cdots & \\ & & R_m \end{bmatrix}}_{\hat{R}_{2m\times 3m}} \underbrace{\begin{bmatrix} \mathbf{x}_{11} & \cdots & \mathbf{x}_{1n} \\ \cdots & \cdots & \cdots \\ \mathbf{x}_{m1} & \cdots & \mathbf{x}_{mn} \end{bmatrix}}_{\hat{X}_{3m\times n}}.$$

By assuming that the deformation lies in a $\hat{s}$-dimensional linear subspace, the 3D structure matrix $\hat{X}$ can be expressed as the product of the trajectory bases matrix $\hat{B}$ and the coefficient matrix $\hat{C}$ such that

$$\hat{B}_{3m\times 3\hat{s}} = \begin{bmatrix} \hat{b}_{11} & 0 & 0 & \cdots & \hat{b}_{1\hat{s}} & 0 & 0 \\ 0 & \hat{b}_{11} & 0 & \cdots & 0 & \hat{b}_{1\hat{s}} & 0 \\ 0 & 0 & \hat{b}_{11} & \cdots & 0 & 0 & \hat{b}_{1\hat{s}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{b}_{m1} & 0 & 0 & \cdots & \hat{b}_{m\hat{s}} & 0 & 0 \\ 0 & \hat{b}_{m1} & 0 & \cdots & 0 & \hat{b}_{m\hat{s}} & 0 \\ 0 & 0 & \hat{b}_{m1} & \cdots & 0 & 0 & \hat{b}_{m\hat{s}} \end{bmatrix},$$

and,

$$\hat{C}_{3\hat{s}\times n} = \begin{bmatrix} \hat{c}_{11}^x & \hat{c}_{12}^x & \cdots & \hat{c}_{1n}^x \\ \hat{c}_{11}^y & \hat{c}_{12}^y & \cdots & \hat{c}_{1n}^y \\ \hat{c}_{11}^z & \hat{c}_{12}^z & \cdots & \hat{c}_{1n}^z \\ \cdots & \cdots & \cdots & \cdots \\ \hat{c}_{\hat{s}1}^x & \hat{c}_{\hat{s}2}^x & \cdots & \hat{c}_{\hat{s}n}^x \\ \hat{c}_{\hat{s}1}^y & \hat{c}_{\hat{s}2}^y & \cdots & \hat{c}_{\hat{s}n}^y \\ \hat{c}_{\hat{s}1}^z & \hat{c}_{\hat{s}2}^z & \cdots & \hat{c}_{\hat{s}n}^z \end{bmatrix}.$$

Table 1. Summarization of the connections and discrepancies between IRSS and NRSfM.

IRSS	$D = LR = LBC$				
	D	L	R	B	C
Meaning	Pixel Intensity	Illumination	Reflectance	Spectral Bases	Coefficients
				Coefficients	Image Bases
Constraints	Nonnegative	Nonnegative, Diagonal	Nonnegative	-	-
NRSfM	$\hat{P} = \hat{R}\hat{X} = \hat{R}\hat{B}\hat{C}$				
	$\hat{P}$	$\hat{R}$	$\hat{X}$	$\hat{B}$	$\hat{C}$
Meaning	Point Tracks	Motion	3D Structure	Trajectory Bases	Coefficients
				Coefficients	Shape Bases
Constraints	-	Orthonormal, Block Diagonal	-	Sub-Diagonal	-

To sum up, the NRSfM problem can be formulated as

$$\hat{P}_{2m\times n} = \hat{R}_{2m\times 3m}\hat{X}_{3m\times n} = \hat{R}_{2m\times 3m}\hat{B}_{3m\times 3\hat{s}}\hat{C}_{3\hat{s}\times n},$$

which is very similar to the low-rank factorization model of the IRSS problem. For example, the motion matrix $\hat{R}$ is block diagonal, which can be regarded as a time-domain counterpart of the diagonal illumination matrix L for IRSS defined in the spectral domain. We summarize the connections and discrepancies between IRSS and NRSfM in Table 1.

As shown in the main paper, when the spectral bases matrix B is unknown, it is ill-posed to factorize the observation matrix D into the illumination spectra matrix L and the reflectance spectra matrix R. Interestingly, for the NRSfM problem, when the trajectory bases matrix $\hat{B}$ is unknown, it is possible to uniquely determine the motion matrix $\hat{R}$ and the 3D structure matrix $\hat{X}$ up to an unknown similarity transformation in 3D. The reason is that, the motion matrix $\hat{R}$ is block diagonal, and its sub-blocks are row-wisely orthonormal. Unfortunately, for the IRSS problem, the illumination spectra matrix L is diagonal, and its diagonal elements are nonnegative. This nonnegativity constraint is too weak to better constrain the factorization.

It is our hope that the similarities and discrepancies summarized in Table 1 could inspire some collaborative effort in the photometric and geometric branches of computer vision, so as to better solve these two important problems.