APPENDIX A

LOCAL ILLUMINATION ESTIMATION

Using intrinsic representation of a texture image patch at location \( x \), we have:

\[
I = RL,
\]

where its equivalent representation in log space is:

\[
i = r + l,
\]

with lowercase letters indicating the log counterparts. It is apparent that in log space, the intrinsic decomposition is linear. Now, if we apply a low-pass filter \( \zeta \) to each texture patch in the original and log spaces, we have

\[
\zeta(I) = \zeta(RL), \quad \zeta(i) = \zeta(r) + \zeta(l).
\]

Since the log operator is non-linear, we cannot represent the linearized form in filtered log space in terms of filtered form in the original space. Thus, to tackle the problem, we need to linearize the log operator.

It is observed that while the log operator is significantly non-linear in the original image intensity range \([0, 1]\) (normalized), it is approximately linear in some other ranges. This can be made clear by the Taylor expansion of the log operator around a particular point \( x_0 \) as:

\[
\ln x = \ln x_0 + \frac{1}{x_0} (x - x_0) - \frac{1}{2x_0^2} (x - x_0)^2 + O \left[ (x - x_0)^3 \right].
\]

(4)

It is apparent when \( x_0 \) is relatively large, higher order terms with respect to \( x_0 \) can be discarded and \( \ln x \) can be linearized as

\[
\ln x \simeq \frac{x}{x_0} + \ln [x_0 - 1].
\]

(5)

This implies that when we map the input image range from \([0, 1]\) to \([I_0, I_0 + 1]\), the log space can then be linearized. In particular, if we set \( I_0 \) to be a relatively large value, the approximation is accurate enough (as can be verified by Eq. 4).

With such linearized approximation, we can modify the original intrinsic image representation such that the linearized log operator is applicable. To do this, we assume that the input image intensity \( I \) and the estimated illumination intensity \( L \) are mapped to the linearized range by specifying appropriate base values \( I_0 \) and \( L_0 \) such that

\[
I' = I_0 + I, \quad L' = L_0 + L,
\]

(6)

where \( I' \) and \( L' \) are mapped values in the linearized range for \( I \) and \( L \), respectively. By replacing the original \( I \) and \( L \) with \( I' \) and \( L' \) while keeping \( R \) unchanged, we have \( I' = RL' \). Applying the log operator on both sides of the above equation leads to:

\[
i' = r + l'.
\]

(7)

Since \( i' \) and \( l' \) are now in linearized range, we can rewrite \( i' \) and \( l' \) in terms of \( I \) and \( L \) as:

\[
i' = \frac{I}{I_0} + \ln I_0, \quad l = \frac{L}{L_0} + \ln L_0.
\]

(8)

If we define \( a = \frac{1}{\pi}, b = \ln I_0, a' = \frac{1}{\pi a}, \) and \( b' = \ln L_0 \), then the above equations can be re-written as:

\[
i' = aI + b, \quad l' = a'L + b'.
\]

(9)

Inserting these two representations into Eq. 7, we have:

\[
aI + b = r + a'L + b'.
\]

(10)

Now we have a linearized form which can be applied with the low-pass filter \( \zeta \) on both sides, and we use a bar to indicate such low-pass filtering:

\[
a\bar{l} + b = \bar{r} + a'\bar{L} + b'.
\]

(11)

Our statistical invariance experiments tell us that when such low-pass filter \( \zeta \) is a box filter, the statistical invariance is satisfied, and this is also true in log space. Suppose that given any intensity value \( I \), its probability to occur in a texture patch is \( p(I) \). Statistical invariance tells us that the first moment of \( p(I) \) (which is the mean value) is invariant, which leads to the result that

\[
S(x) = \int p[I(x)]I(x)dI(x)
\]

(12)

is invariant given any \( x \). This implies that \( p(I) \) is invariant. Thus, it is easy to verify that when replacing \( I \) with \( \ln I \), \( S(x) \) is still invariant. Therefore, \( \bar{r} \) can be replaced with a constant \( c \), and \( c = \ln C \), where \( C \) is the constant in the original space, which in fact quantifies the average reflectance intensity. As a result, we have

\[
a\bar{I} + b = c + a'\bar{L} + b',
\]

(13)

which indicates that the variation of the local color statistics is due to illumination.

To solve for \( L \), we need to approximate \( \bar{L} \) in terms of \( L \). If we write \( \bar{L} \) in integral form, we have:

\[
\bar{L}(x) = \frac{1}{A(\Omega)} \int_\Omega L(x)dx.
\]

(14)

By the mean value theorem of integral calculus, such integral can be approximated by the central value within the region \( \Omega \). More accurately, if we assume that the filter can be cancelled in space with respect to the center of the local patch, \( L(x) = \bar{L}(x) \) exactly. Such cancellation property actually requires that the local function within the patch region is odd-symmetric. Fig. 1 shows a local pattern of a given illumination field where the distribution of \( L(x) \) can be assumed odd-symmetric. In fact, the odd-symmetric assumption is not that strong. It can include many linear and non-linear cases since we only consider a local portion of the function. However, if the region is not
odd-symmetric, we can use $L$ as an estimate. Such estimate is accurate enough within a local region especially when the region is not large and illumination is smooth. This has been verified by our estimation results on a set of textured surfaces with different illuminations.

By replacing $\bar{L}$ with $L$, we have the following equation:

$$a\bar{I} + b = c + a'L + b' \quad (15)$$

with which we can solve $L$ by:

$$L = \frac{a}{a'}\bar{I} + \frac{(b - b' - c)}{a'} \quad (16)$$

If we define $k = \frac{a}{a'}$ and insert the corresponding representations for $a$, $b$, $a'$ and $b'$, we have:

$$L = k\bar{I} - I_0(\ln k + C) \quad (17)$$

To simplify, we enforce a constraint such that $I_0(\ln k + C) = 0$. This actually relates $c$ with $k$ as:

$$c = -\ln k \quad (18)$$

Since $c = \ln C$, we have:

$$k = \frac{1}{C} \quad (19)$$

This constraint leads to the following estimation equation for local illumination intensity at $x$:

$$L = \frac{k}{M} \sum_{y \in N(x)} I(y) = \frac{C^{-1}}{M} \sum_{y \in N(x)} I(y) \quad (20)$$

which is Eq. 1 in our main paper, where we can either specify $k$ or $C$ to compute the illumination. The reflectance can be directly computed by division from the intrinsic representation.

It is interesting to note that by enforcing the constraint, our final solution of the illumination does not rely on the mapping. The only parameter that is required to solve for the illumination is the output average reflectance, which ranges from 0 to 1. Also note that when $C = 0$, $k \to \infty$ and the illumination tends to infinity unless the original image is all black; when $C = 1$, $k = 1$, which forms the lower bound of $k$. Thus, $k \in [1, \infty)$. In practice, we choose $C \in [0.2, 0.8]$ which corresponds to $k \in [1.25, 5.0]$.

**APPENDIX B**

**LOCAL DEFORMATION ESTIMATION**

By applying eigen decomposition on $J$ in Eq.2 of Section 4.2, we can represent the mean structure tensor in patches around $x$ and $x'$ as $J$ and $J'$ in the following, respectively:

$$J = UDU^T, \quad J' = U'D'U'^T \quad (21)$$

Any linear deformation $T$ from $x$ to $x'$ will give:

$$J' = TJT^T \quad (22)$$

Suppose the deformation consists of only a rotation $R$ and an isotropic scaling $s$, then

$$T = sR \quad (23)$$

Thus we have:

$$J' = s^2RJR^T \quad (24)$$

By using eigen decomposition, we have:

$$U'D'U'^T = (RU)(s^2D)(RU)^T \quad (25)$$

Furthermore, since $R$ is orthonormal, it is obvious that we can obtain:

$$U' = RU, \quad D' = s^2D \quad (26)$$

Since $U^{-1} = U^T$, we obtain the rotation matrix from the patch around $x$ to the patch around $x'$ as

$$R(x') = U(x')U^T(x) \quad (27)$$

which is the Eq.3 in Section 4.2 of the main paper.

To estimate the scale change between the two patches, we can perform the least squares fitting on $D' = s^2D$, which minimizes $\|D' - s^2D\|^2$ w.r.t $s$. However, this may not be stable because the difference of gradient contrast between the two patches results in a change of $D'$ that may not be completely due to a scaling from $D$. 