Spatial dependencies between local luminance and contrast in natural images

Jussi T. Lindgren
Department of Computer Science and HIIT, University of Helsinki, Finland

Jarmo Hurri
Department of Computer Science and HIIT, University of Helsinki, Finland

Aapo Hyvärinen
Department of Computer Science and HIIT, University of Helsinki, Finland

Previous research has suggested only weak statistical dependencies between local luminance and contrast in natural images. Here we study luminance and contrast in natural images using established measures and show that when multiple measurements of these two local quantities are taken in different spatial locations across the visual field, strong dependencies are revealed that were not apparent in previous pointwise (single-site) analyses. We present a few simple experiments demonstrating this spatial dependency of luminance and contrast and show that the luminance measurements can be used to approximate the contrast measurements. We also show that relying on higher-order statistics, independent component analysis learns paired spatial features for luminance and contrast. These features are shown to share orientation and localization, with the filters corresponding to the features dependent in their outputs. Finally, we demonstrate that the found dependencies also exist in artificial images generated from a dead leaves model, implying that simple image phenomena may suffice to account for the dependencies. Our results indicate that local luminance and contrast computations do not recover independent information sources from the visual signal. Subsequently, our results predict spatial processing of local luminance and contrast to be non-separable in visual systems.

Keywords: luminance, contrast, independence, separate processing


Introduction

The visual world can be interpreted to consist of different cues such as color, motion, and shape. In vision research, the traditional view holds that these cues are recovered and processed from the retinal input by separable systems (for a review, see Casagrande & Xu, 2004). Such separation may appear in terms of different pathways (Ungerleider & Mishkin, 1982), cortical areas (Felleman & Van Essen, 1991), or subsystems (Ts’o & Roe, 1995), with varying degrees of interaction (DeYoe & Van Essen, 1988), possibly through feedback (Gilbert & Sigman, 2007). Why do such separable systems exist, and which visual cues or submodalities should we expect to have separable processing in visual systems? The recent research setting of natural image statistics (Simoncelli & Olshausen, 2001) would predict that the existing separable systems in visual processing are intimately connected to the statistical properties and dependencies in the visual environment, as relayed through the retina. The statistical analysis of the natural inputs can then be used to make predictions of the appropriate processing.

Here we investigate the statistical dependencies between luminance and contrast in natural images. These two quantities remain of considerable interest to vision research, both as modifiable parameters of experimental stimuli (Allard & Faubert, 2007; Badcock, Clifford, & Khoo, 2005; Geisler, Albrecht, & Crane, 2007; Henderson, Brockmole, & Gajewski, 2008; Sukumar & Waugh, 2007; Vladusich, Lucassen, & Cornelissen, 2006) and as properties of natural image data (Brady & Field, 2000; Frazor & Geisler, 2006; Mante, Frazor, Bonin, Geisler, & Carandini, 2005; Peli, 1990). Our current investigation is motivated by the results of Frazor and Geisler (2006) and Mante et al. (2005), suggesting local luminance and contrast to be independent or only weakly dependent in natural images. This immediately raises the question if local luminance and contrast computations allow recovering independent cues (or information sources) from the visual environment.

In this paper, we examine the computational utility of separated spatial processing of local luminance and contrast in natural images.
contrast and come to a negative conclusion from a statistical modelling perspective. In previous studies (Frazor & Geisler, 2006; Mante et al., 2005), local luminance and contrast were sampled and studied as scalar pairs, and as a consequence of this methodological choice, the previous analyses were oblivious to the spatial structures these quantities have in the input. However, if we consider local luminance and contrast as potential sources of visual information, the spatial, multivariate nature of the retinal input can no longer be neglected. Here we extend the previous analyses by taking multiple simultaneous measurements of local luminance and contrast across the visual field. We then apply a few experimental methodologies to analyze the cross-dependencies between these measurements. These experiments allow us to demonstrate strong spatial dependencies between local luminance and contrast in the data. Our results suggest that luminance and contrast computations do not recover independent sources of visual information, if multiple, spatially disjoint local measurements of these local quantities are taken simultaneously. We will also argue that although our empirical results are based on specific definitions of local luminance and contrast computation, this choice may be qualitatively unimportant, as long as the used measures correspond to intuitive definitions of local “base level” and local amount of “change.”

Finally, we ask that if weak dependencies of local luminance and contrast can be attained with artificial stimuli (e.g., Allard & Faubert, 2007) and in single-site analysis of natural images (Frazor & Geisler, 2006; Mante et al., 2005), where do the strong spatial dependencies in natural images arise from? To provide some tentative answers to this question, we show that similar dependencies exist in simple artificial images generated from a dead leaves model (e.g., Lee, Mumford, & Huang, 2001). Images from such a model are simple constructions made from randomly placed disks that can occlude each other. This suggests that relatively simple image properties may be sufficient to account for the found dependency structure between local luminance and contrast.

**General methods**

**Natural image data**

We used the natural image data set collected by van Hateren and van der Schaaf (1998). This data set contains over 4000 grayscale images representing natural scenes, each image having a resolution of 1024 × 1536 pixels (for a thorough description of the data, see van Hateren & van der Schaaf, 1998). Out of the two versions of the data set, we used the linear images (files named *.iml) to be able to compare our results to those of Frazor and Geisler (2006). We converted the intensity values to luminance (cd/m²) using the scale factors available on the data set Web page. Finally, we cropped 4 pixels from all sides in the images to avoid well-known border anomalies present in the data set images.

**Luminance and contrast measures**

To measure contrast, we use a local contrast computation closely resembling gain controlled estimation of standard deviation. This measure is becoming an accepted correlate of perceived contrast (Bex & Makous, 2002; Frazor & Geisler, 2006; Mante et al., 2005), a natural choice corresponding to first moment of a distribution (whereas contrast corresponds to the second). These measures are equal to those in (Frazor & Geisler, 2006; Mante et al., 2005), where weak dependencies were reported by studying single-variable luminances and contrasts of image patches. In the current paper we extend these measures across the visual field, i.e., we perform the measurements in each spatial location to get multivariate luminance and contrast response images instead of two scalars. We call these resulting response images channels, although the word “channel” is used only for illustrative purposes. We do not propose that the corresponding channels explicitly existed as neural representations.

We computed the local luminance and contrast channels as follows. First, to be in line with the setting used in (Frazor & Geisler, 2006; Mante et al., 2005), we used an isotropic, raised cosine taper window w to specify the localization,

$$w_{x,y}(i,j) = \frac{1}{2} \left( \cos \left( \frac{\pi}{p} \sqrt{(i - x)^2 + (j - y)^2} \right) + 1 \right), \quad (1)$$

where p is the mask radius (in pixels), (i, j) the location of the weighted pixel, and (x, y) the current center of the weighting mask. Due to the cosine-based window being cyclic, we modified Equation 1 by zeroing out the $w_{x,y}(i,j)$ for which $\|(i,j)-(x,y)\| > p$ (i.e., the weights that lie outside the radius).

Now, the pointwise local luminance and contrast measures used in (Frazor & Geisler, 2006; Mante et al., 2005) can be easily generalized into spatial versions. We denote local luminance, local contrast, and local luminance normalized contrast channels by LUM, RMS, and RMSL, respectively. We define them as

$$\text{LUM}(x,y) = \frac{1}{C} \sum_{i,j} w_{x,y}(i,j) I(i,j), \quad (2)$$
\[ \text{RMS}(x, y) = \sqrt{\frac{1}{C} \sum_{i,j} w_{x,y}(i,j) \left( I(i,j) - \text{LUM}(x,y) \right)^2}, \] (3)

\[ \text{RMSL}(x, y) = \frac{1}{\text{LUM}(x,y) + L_0} \text{RMS}(x, y), \] (4)

where \( I \) is an input image, \( L_0 \geq 0 \) a dark–light parameter, and \( C = \sum_i w_{x,y}(i,j) \). The normalization coefficient \( C \) is the same for all \((x, y)\). Note that the LUM, RMS, and RMSL channels are in spatial correspondence: for each location \((x, y)\), the measures are computed from exactly the same region of the input image \( I \) using the same weighting. Sampling paired scalar variables from the LUM and RMSL channels of our image set would then equal the pointwise setting of (Frazor & Geisler, 2006; Mante et al., 2005), providing that no weighting masks \( w \) between any two samples from the same image overlap.

The difference between Equations 3 and 4 is that in the latter, a model of divisive gain control is incorporated. We observed that this gain control can reduce the pointwise dependencies between local luminance and contrast, whereas comparing LUM and RMS without gain control revealed them to have pointwise correlations on our data (which would subsequently cause spatial dependencies between the two). Due to this finding, in the following we concentrate on LUM and RMSL. For RMSL, we noted that the dark–light parameter had some effect on the pointwise dependencies. In Mante et al. (2005), \( L_0 \) was equal to 0, and in Frazor and Geisler (2006) it was set to either 0 or 1 cd/m². In our preliminary experiments, we found that the value of \( L_0 \) required to achieve pointwise decorrelation depends slightly on the used radius \( p \), as shown in Figure 1. The decorrelating value of \( L_0 \) appears to be smaller for smaller radii \( p \), e.g., around 20 cd/m² for \( p = 4 \), and 30 cd/m² for \( p = 28 \). On the average over the tested radii, \( L_0 \) for decorrelation was 28 cd/m² with a standard deviation of 3 cd/m². As the obtained correlations between pointwise LUM and RMSL remain relatively small over the parameter ranges of \( L_0 \) and \( p \) in the plot of Figure 1, RMSL appears to be robust against the choices of \( L_0 \) and \( p \) over the whole data set. In the subsequent experiments, we used the value for \( L_0 \) that achieved decorrelation for the used mask radius. Setting the value of \( L_0 \) adversely in this way allows us to show that the gain control solution applied by RMSL is not sufficient to make RMSL spatially independent from LUM.

For computational efficiency, we implemented the LUM and RMSL computations using convolutions. Due to the sliding nature of Equations 2, 3, and 4, nearby in-channel pixels in the luminance (contrast) channel share some of their afferent pixels in the image \( I \). Later, we describe control experiments where no afferents were shared, verifying that the spatial dependencies are not due to the convolutive processing.

For illustration, a natural 768 \times 768 pixel image and its LUM and RMSL channels are shown in Figure 2 for \( p = 8 \) in Equation 1 and \( L_0 = 22 \) cd/m².

Figure 1. The correlation coefficient between pointwise LUM and RMSL when patches were sampled uniformly over the images for different values of the dark–light parameter \( L_0 \) of Equation 4 and weighting mask radius \( p \) of Equation 1. The black crosses mark the \( L_0 \) that attained zero correlation for the corresponding value of \( p \).
Experiments and results

Experiment 1: Pointwise dependency between luminance and contrast

First, we performed experiments similar to previous work (Frazor & Geisler, 2006; Mante et al., 2005), where dependencies between single points of local luminance and contrast were examined.

Methods

We sampled 10,000 LUM and RMSL pairs (radius $p = 4i, i \in \{1, \ldots, 16\}$) from all of the 4,000+ images.

Results

The results were already shown in Figure 1. Small pointwise correlations can be attained between LUM and RMSL. The amount of correlation depends somewhat on the values of the dark–light parameter $L_0$ and the radius $p$. This shows that these two parameters can slightly alter the statistical characteristics of the local luminance and contrast operator outputs.

Experiment 2: Simple experiment revealing spatial dependency

Do local luminance and contrast have spatial dependencies in natural images for Equations 2 and 4? To measure this, we first performed a simple experiment. This experiment verifies the intuition that when two distinct spatial locations have a large luminance difference, there should be a non-negligible contrast somewhere in the middle.

Methods

Let $I_1$, $I_2$, and $I_3$ be three horizontally adjacent 32 × 32 pixel image patches of the input image $I$, with no overlap between the patches. We sampled 10,000 such triplets from the used data set and then computed the (single-variable) local LUM for $I_1$ and $I_3$, and local RMSL for $I_2$, using $p = 16$ and $L_0 = 27$ cd/m² (for pointwise decorrelation). This gives a triplet of scalar variables, LUM($I_1$), LUM($I_3$), and RMSL($I_2$), with slight abuse of notation. Note that each of the three variables has been computed from disjoint pixels and none of the post-processings that we will use later for the ICA experiments have been performed. We computed a new variable $D = \left| \text{LUM}(I_1) - \text{LUM}(I_3) \right|$ and its median $M(D)$. If $D > M(D)$, we say the corresponding triplet has a high luminance difference (low, otherwise). Next we...
standardized the variable RMSL($I_2$) to unit variance and then split it to two sets, depending if the luminance difference of the triplet was high or low. If RMSL is independent of LUM, the two sets should be statistically equivalent.

**Results**

It turned out that the median for RMSL($I_2$) in the high difference set is 1.26, and for the low difference set only 0.57. According to a Kruskal–Wallis test, the null hypothesis that the medians are equal can be rejected with $p < 0.001$. We obtained similar results when the triplet was taken vertically or diagonally.

Note that we did not perform this analysis in hand-segmented regions (manual segmentation was used in Frazor & Geisler, 2006), and here the patches of the triplets can fall on different objects or textures. However, this experiment was only meant to illustrate in a simple manner that dependencies exist, not to quantify them. For pointwise analyses of dependencies in different image constituents, we refer the reader to (Frazor & Geisler, 2006). In the subsequent experiments, we will explore the nature of the found dependencies in more detail.

**Experiment 3: Predictability of RMSL from LUM**

Looking at Figure 2 subplots B and C suggests that luminance and contrast channel contents may have similar spatial organization on some level. Could this allow us to predict one channel from the other? We examined this question by examining if (in general) RMSL could be predictable from LUM. If the two channels were independent, it should not be possible to make even a weak correlate of RMSL using LUM as the only information source.

**Methods**

We first formed LUM and RMSL channels at different radii ($p$) in Equation 1 and then computed an estimate RMSL* based solely on the LUM channel (and not using the original image $I$) but now using a very small radius $p = 2$. The resulting approximate RMSL* image was then bilinearly interpolated to match the size of the true RMSL image. We performed this for 500 images sampled from our data set and computed the correlation coefficient between RMSL and RMSL* for each image and then averaged these obtained coefficients. Note that the choice of $p = 2$ was heuristic. Here this choice suffices to show that a strong spatial relation exists between LUM and RMSL.

**Results**

Figure 3A shows the obtained average correlation coefficients between RMSL and the LUM-based approximation RMSL* for different radii $p$. The obtained correlation is at the highest (mean $c = 0.88$) with radius $p = 4$ and decays with both larger radii and asymmetrically if the radii are different. With very disparate radii of the individual channel computations, the approximate

![Figure 3A](https://example.com/figure3a.png)

A) The average correlation coefficients between real RMSL and RMSL* approximated from LUM at different radii $p = 4i$, $i \in \{1, ..., 16\}$ of the original RMSL and LUM computation (see text). The coefficients indicate that when both channels are at the same scale, the luminance channel contents can be used to explain the contrast channel to varying degrees. The degree is stronger at smaller scales, and with growing difference between the scales, it gets smaller asymmetrically. (B) The corresponding standard deviations for the average correlations.

![Figure 3B](https://example.com/figure3b.png)
RMSL* is a poor correlate of the real RMSL, suggesting that the dependency between the two quantities may be weaker if the channels have a large difference in their scale of computation. But with similar scales, the strong correlations indicate that far from being independent, the LUM channel actually contains enough structure to approximate the RMSL channel to a noticeable degree. The corresponding standard deviations are displayed in Figure 3B, showing that as the analysis moves towards a more global scale, the predictability of contrast from luminance gets more unstable with the used method. This may reflect the scale of the content in the images themselves.

Note that although we showed that RMSL can be predicted from LUM to some accuracy, the opposite may not be trivially so. The reason is that RMSL cannot make a difference between dark and light surfaces, and it only indicates change but not its direction.

Experiment 4: Dependencies learned by independent component analysis

Finally, to obtain a better understanding of the nature of the dependencies between local luminance and contrast channels, we examined them by independent component analysis (ICA; see, e.g., Hyvärinen, Karhunen, & Oja, 2001). ICA has been successfully used previously in, e.g., learning natural image models that exhibit V1 simple cell receptive field characteristics (Bell & Sejnowski, 1997; van Hateren & van der Schaaf, 1998).

Methods

In its basic form, ICA assumes that the observed data \( x \), which is a random vector of dimension \( n \), has been generated by a linear generative model

\[
x = As, \text{ or equivalently,} \tag{5}
\]

\[
x_j = \sum_{i=1}^{n} A_{ji} s_i, \quad \text{for } j \in \{1, \ldots, n\}, \tag{6}
\]

where \( A \) is a constant mixing matrix (to be estimated) and \( s \) is a random vector of unknown independent and non-Gaussian components \( s_i \). The dimension of \( s \) is assumed to be equal to the dimension of \( x \), possibly after the observed data \( x \) have been reduced to a smaller dimension by principal component analysis (Hyvärinen et al., 2001). ICA tries to estimate both \( s \) and the parameter matrix \( A \) from the observed data \( x \); when the model holds, this can be done with very few assumptions (up to a scaling and ordering of the independent components and columns of matrix \( A \)). When the data does not follow the generative model of Equation 5, many ICA algorithms—including the one we use—can still be interpreted as minimization of higher-order statistical dependencies between the components of the estimated random vector \( s = Wx \equiv A^{-1}x \) (Hyvärinen et al., 2001). In addition, when statistical dependencies are reduced by searching for maximally non-Gaussian projections—as in our method—the method can be viewed as computing a sparse coding basis for the observed data.

The data for ICA estimation were obtained as follows. First, we computed the local LUM and RMSL channel of each image using Equations 2 and 4 with \( p = 8 \) in Equation 1. With \( p = 16 \), the results were qualitatively similar and are omitted. Next, we subsampled the resulting images by \( p/2 \). This was done to address the increased redundancy on the luminance channel due to the low-pass filtering of Equation 2. We did not attempt to make the subsampling optimal. Then, we took a natural logarithm of the luminance channel data since otherwise the dependencies revealed by ICA were less pronounced (logarithm or other non-linearity corresponding to retinal non-linearity is typically applied when ICA is used to model luminance; see e.g., van Hateren & van der Schaaf, 1998). From the resulting channels of the 4,000+ images, we sampled two sets of corresponding 200,000 small \( 21 \times 21 \) image patches at random. Each patch pair was obtained by sampling source image number and a point \((x, y)\) uniformly, so that a \( 21 \times 21 \) region starting from the sampled point would remain inside the channels. Then, this region was cropped from each channel to form the luminance/contrast patch pair (one such patch pair is illustrated in Figure 2 as a rectangle drawn on the two channels). We removed the DC component (mean) from each contrast and luminance patch separately. The obtained patches were then vectorized, and the two data sets corresponding to the two channels were separately normalized to have unit variance and then separately submitted to principal components analysis to remove noise (dimensionality reduction is a standard procedure in ICA to reduce the effect of noise). We kept 95% of the variance of each set, corresponding to 66/441 and 203/441 dimensions for luminance and contrast, respectively.

After reducing the dimension of both sets, we concatenated the two sets to form the observed 269-dimensional data for the ICA model estimation. Since the channels were already processed with PCA separately, no further dimensionality reduction was performed. For ICA, we used the FastICA algorithm (Hyvärinen, 1999) with tanh (\( \cdot \)) non-linearity and symmetric (simultaneous) estimation of all 269 components. After ICA estimation, the separate PCA matrices for luminance and contrast were used to transform the model \( A \) back to the original space for analysis.

Note that the postprocessing steps of the LUM and RMSL channels we described above are acceptable regarding the main results of the current paper: If two random vectors \( x \) and \( y \) are independent, so are \( f(x) \) and \( g(y) \)—postprocessing cannot make two independent
random vectors dependent (Papoulis, 1991). Here, these postprocessing steps were done to enhance the clarity of the ICA results. For experiments where ICA was not used, the postprocessings described in this subsection were not applied.

We use the following decompositions in describing our ICA results. Let \( x = [l; r] \) be a catenated data vector of luminance and contrast patches in correspondence. Now the weights of each column vector \( a \) of the estimated matrix \( A \) represent the learned visual features \([l^*; r^*] = a\) of that component. The corresponding row vectors \( q^T \in W \) represent the linear filters (receptive field models) computing each component response, i.e., \( s = q^T x = q^T[l; r] = q^T l + q^T r \). How ICA learns to represent catenated luminance and contrast image patches using these definitions is shown in Figure 4. In the subsequent analyses, we study the emergent feature pairs \([l^*; r^*]\), as well as dependencies between the filter responses \( q^T l \) and \( q^T r \). This was done through reshaping both the features and the filters back to the spatial form of \( 21 \times 21 \) pixels. Now, if according to ICA model, the two channels were independent, we would expect one of the filters \( q_l \) or \( q_r \) to be blank (with zero weights), with correspondingly blank feature as \( l^* \) or \( r^* \). Also, no dependencies should be observable between the filter outputs \( q^T l \) and \( q^T r \) in the condition of independence.

**Results**

ICA succeeds in revealing some intuitive dependencies between LUM and RMSL channels, providing that they have been suitably postprocessed as we described before. Results using RMS were qualitatively similar and have been omitted for brevity. Figure 5A shows the paired basis vectors \( l^* \) and \( r^* \) for the two channels. Most of the pairs turn out to have a similar pattern in both luminance and contrast channel. Some feature pairs and their corresponding filters have been magnified for convenience in Figures 5B and 5C. We will now turn to examine the learned dependencies more closely.

To better understand the regularities between the channels, we analyzed both the obtained model parameters \([l^*; r^*]\) for each component and the filter responses \( q^T l \) and \( q^T r \) on fresh data. We applied well-known Fourier techniques to the features \( l^* \) and \( r^* \) to demonstrate their regularities. This analysis would have been traditionally done to the filters, but in this case the filters turned out to be structurally less coherent than the features (compare Figures 5B and 5C). In Figure 6A, we display a histogram of correlation coefficients between the squared amplitudes of discrete 2D Fourier transforms of the feature pairs. The high correlation coefficients show that frequency and orientation are commonly shared by the pair. The other histogram in Figure 6A shows correlation coefficients of the spatial mask pairs, which is equal to computing the angle between the two vectors after they have been standardized to unit norm. A value of 1 indicates that the vectors are parallel (same phase), 0 that they are orthogonal, and -1 that they differ in the sign (opponent phase). It can be seen that most of the features lean toward phase agreement, but there is also a clear cluster of opponent phase features. Intuitively, since RMSL is invariant to stimulus phase, we would expect the filter pairs \( (q_l, q_r) \) and \((q_r, q_l)\) to be statistically identical in their responses, and hence that both phase-agreeing and phase-opponent features should occur in roughly equal proportions. However, as can be seen from Figure 6A, the phase-agreeing pairs dominate, and this tendency seemed to be stable across several bootstrapped experiments. We are uncertain whether this imbalance reflects some structure in the data or if it is a property of our modelling mechanism.

To analyze the response dependencies of the paired filters \( q_l \) and \( q_r \), we sampled 10,000 fresh patch pairs not used in the ICA estimation. In Figure 6B, we show a histogram of the correlation coefficients between \( q^T l \) and \( q^T r \) for all the filter pairs, computed over responses to the fresh patch pairs. We also show the correlation coefficients between the absolute values \(|q^T l| \) and \(|q^T r| \); these turned out to be 0.47 on the average. Together with the obtained shape pairs, this suggests that the same shapes tend to exist on both channels at the same time, but possibly with some phase difference. The lack of correlation in the signed outputs is likely because of linear filter output distributions on natural luminance images are typically symmetric, reflecting, e.g., that step edges are equally likely in both phases. However, regardless of luminance edge phase, the resulting contrast

\[
[l ; r] = s_1[l^*_1 ; r^*_1] + s_2[l^*_2 ; r^*_2] + \ldots + s_n[l^*_n ; r^*_n]
\]
luminance filter was paired with each of the obtained contrast filters. It turned out that for 86% of the filters, the correlation coefficient was highest when the luminance filter was paired with its ICA-designated contrast filter pair. This illustrates that the obtained pairings are not simply explained by the channels having isotropic energy correlation.

Finally, we computed the correlation coefficients of the absolute values of the filter responses when each obtained luminance filter was paired with each of the obtained contrast filters. It turned out that for 86% of the filters, the correlation coefficient was highest when the luminance filter was paired with its ICA-designated contrast filter pair. This illustrates that the obtained pairings are not simply explained by the channels having isotropic energy correlation.

shape is the same. Hence, the sign differences of the two filter responses lead to smaller signed correlation (similar arguments for dependencies in filter–rectify–filter setting can be found in Johnson & Baker, 2004).

Finally, we computed the correlation coefficients of the absolute values of the filter responses when each obtained luminance filter was paired with each of the obtained contrast filters. It turned out that for 86% of the filters, the correlation coefficient was highest when the luminance filter was paired with its ICA-designated contrast filter pair. This illustrates that the obtained pairings are not simply explained by the channels having isotropic energy correlation.

Figure 5. Results of learning dependencies between local luminance and contrast. (A) Arbitrarily selected subset of 152/269 learned basis vectors \([l^*; r^*] \) when ICA is applied to the concatenated luminance and contrast data. The rest of the basis vectors are similar. In each pair, the left-hand side corresponds to a luminance pattern and the right hand side to a contrast pattern. Any two non-zero patterns appearing together in a pair indicate a learned dependency between the two channels. (B) Three paired features magnified: an edge pair with agreeing phases, a bar pair with opponent phases, and a low frequency pair. (C) The filters \(q_l \) and \(q_r \) corresponding to the features in B. The left filter processes the luminance channel and the right filter the contrast channel, after which the two filter outputs are summed to produce the component response.

Figure 6. Some dependencies of the 269 shape pairs learned in the ICA experiment. (A) Similarities in the learned feature pairs. The histogram shows correlation coefficients of the squared amplitudes of discrete Fourier transforms of the paired stimuli, as well as correlation coefficients of the spatial masks. This confirms that the learned features on the two channels largely agree in their orientation and frequency (based on the Fourier analysis) and have tendency toward similar phases (spatial correlation coefficients near 1). (B) Between-pair correlation coefficients for the raw and absolute outputs of the corresponding filter pairs \((q_l, q_r) \) for luminance and contrast channels. The absolute correlations demonstrate clear redundancies between the two channels.
Control experiments and their results

We made several control experiments to verify that the obtained ICA results are not simple artifacts of the algorithm or postprocessing but actually reflect statistical properties of the input. These experiments were as follows:

1. We repeated our experiments so that all values inside a channel were computed from disjoint source image regions. This necessarily makes the obtained channels very low resolution (the image side length will be divided by the diameter 2p of the localization mask). Also, the contrast channel shapes become less continuous. However, qualitatively similar oriented features and feature pairings were obtained in this control condition, with the difference that the obtained contrast shapes were generally shorter, reflecting the discontinuity introduced to the contrast channel.

2. To verify that the generation of shape pairings for luminance and contrast are not a by-product of ICA, we made a new catenated data set of random pairings of luminance and contrast patches. In this case the two channels were independent by design and this was reflected in the ICA results: each component represented an oriented feature only in one of the channels, with the other channel blank.

3. We examined whether changing $L_0$, the weighting mask or the radius $p$ would remove the dependencies. Reasonable values of these parameters had visible effects on the learned features, but they still largely remained localized and oriented, with the corresponding filters bandpass. The qualitative dependencies between the channels appeared to be largely unaffected by these choices.

4. Are the presented results affected by performing the logarithmic non-linearity before the local luminance computation instead of using it as a postprocessing step? The results were similar in both conditions. It seems that the existence of some compression of the luminance signal is more important than whether it is compressed before or after the local luminance channel computation. Without compression, the results produced by FastICA were much closer to suggesting independence, with the correlation coefficients of the absolute filter pair responses close to 0.2 on the average.

5. We examined if the degree of PCA dimensionality reduction affects the dependencies. Although it did change the learned features and the filters in a similar fashion as PCA affects ICA when ICA is used for luminance modelling, PCA dimensionality reduction did not change the tendency of the ICA features to appear as pairs.

6. We repeated our ICA experiments with images of white (uncorrelated) noise that was sampled from the luminance distribution of the original image data. In this case, FastICA did not converge, which typically happens when the data is very far from the ICA model distribution or has more than one Gaussian direction (Hyvärinen et al., 2001).

Discussion

Relation to previous studies

A reason why our results differ from previous studies claiming weak or no dependencies between local luminance and contrast (Frazor & Geisler, 2006; Mante et al., 2005) is that instead of considering a large, unordered set of corresponding scalar pairs of LUM and RMSL, we consider taking multiple measurements of these quantities across the visual field. For clarity and simplicity, we took measurements at each pixel location (ignoring those boundary region pixels that might cause border effects due to the computation extending over the image area) and did not model, e.g., a higher foveal resolution. Our choice practically equals to using the local luminance and contrast operators as filters, and as a result, the filter responses form multivariate images or “channels” of luminance and contrast. Although the single scalar pairs may show only small dependencies (Frazor & Geisler, 2006), sampling such pairs from the images loses all information about the spatial structure of local luminance and contrast, and as we have shown, spatial structures in natural images result in dependencies between local luminance and contrast when multiple simultaneous measurements of these quantities are taken. From the viewpoint of probability theory, this result is an example of the fact that when more than two random variables are examined together, one can find dependencies which do not exist in purely pairwise analysis: pairwise independence of $k > 2$ variables does not guarantee full factorization of the $k$-dimensional joint distribution (Papoulis, 1991).

Structurally, our catenated ICA model in the ICA experiments very closely resembles a two-stream model (e.g., Johnson & Baker, 2004; Zhou & Baker, 1993; for a recent review see Landy & Graham, 2004). The main difference is that we specified the non-linear filtering stream as the isotropic RMSL representing a contrast channel. In addition, our functional form allows for summative integration of the two streams, where a single component output is $s = (q_l^Tl) + (q_r^Tr)$. Thus, our FastICA-based model estimation fits the weights for the two filters $q_l$ and $q_r$, simultaneously processing both channels. Due to statistical dependencies present in the data, luminance and contrast signals become mixed by the model,
provided that the luminance channel was compressed by a non-linearity. However, we do not propose that this emergent integration necessarily has any neural counterpart. Here the model was simply used as a tool to discover and characterize dependencies present in the data.

The dependencies we found agree with previous work on filter–rectify–filter (FRF) models, where dependencies between first- and second-order processing have been observed for Gabor filter models (Johnson & Baker, 2004). However, we studied RMSL contrast explicitly and not through its possible correlate from a hierarchical model of manually selected filters. In the ICA experiment, we assumed a linear basis for catenated luminance and contrast signals and attempted to computationally find a representation that minimizes dependencies between the resulting linear projection directions. The emerging filters we obtained for the individual channels are similar to those previously found by sparse coding of luminance in natural images (Bell & Sejnowski, 1997; Olshausen & Field, 1996; van Hateren & van der Schaaf, 1998), here illustrating that meaningful structure can also be learned in the contrast domain.

Measures of local luminance and contrast

How sensitive are the dependencies we report to the used definitions and their parameterizations? Although numerous contrast measures have been proposed in the literature (e.g., Bex & Makous, 2002; Frazor & Geisler, 2006; Mante et al., 2005; Moulden et al., 1990; Peli, 1990), we suspect that qualitatively similar spatial dependency structure would be expectable with all reasonable local measures. The reason is simply that if one local measure corresponds to base level (“luminance”) and the other to degree of change (“contrast”), then any variation in the multiple measurements of the base level necessarily reflects some degree of change in the original signal, and hence the base level measurements allow to predict activity in the degree of change measurements. A simple example of this is to think of a dark-to-light step edge, where one base level measurement is computed entirely on the dark side, and other one entirely on the light side, with a change measurement taken on top of the step edge itself. Now, the base level measurements will measure darkness and light, whereas the degree of change operator will react to the change at the step edge. But now notice that there is a difference in the two obtained measurements of the base level, and this difference can be used to predict that the degree of change operator responds in the middle.

Hence, we suppose that regardless of definitions, sensible local measures of luminance and contrast might not be able to decompose natural stimuli to independent sources of visual information.
each new disk was always painted behind the currently visible disks if there was overlap. This process was iterated until all untouched pixels were replaced by surface of some disk. The disks were placed randomly with random radius and random intensity. The disk placement was uniformly random and the disk radii sampled from $1/r^3$ distribution as in (Lee et al., 2001), with $r$ between 3 and 1024 pixels. The disk intensities were sampled from an exponential distribution with parameter $\mu = 1$. Finally, to avoid border effects, the central $1024 \times 1024$ pixel aperture of the canvas was taken as the final image.

We repeated our previous experiments with 250 images generated from the dead leaves model and got similar results as before. First, pointwise LUM and RMS are correlated, whereas for pointwise LUM and RMSL, the correlation can be made close to zero with a suitable choice of $L_0$. Second, contrast of a spatial location in the middle is dependent on the difference of luminances of two points that flank the middle location. Third, the RMSL channel can again be predicted from the LUM channel. Here the average correlation coefficient of the approximation to the actual channel was $c = 0.98$, when mask radius $p = 8$ was used. Finally, independent component analysis results in similar paired features on both channels where both features are share their spatial localization and orientation (not shown).

These results show that a fairly simple image creation process is sufficient for local luminance and contrast to have seemingly small pointwise correlations yet clear dependencies spatially. Note that there is no explicit complex phenomena such as texture, directed illumination, shadows, surface properties, other contrast modulation, or variable template shapes in 2D or 3D in these artificial images. Hence, such properties are not required to account for the dependencies we described. On the other hand, some properties of the dead leaves model seem necessary to attain the dependency structure described here. For example, Figure 7B shows an image from a two-tone dead leaves model where each disk can be either black or white (the actual intensities are irrelevant for the following, as long as the values are different). Although spatial dependencies between local luminance and contrast exist in these images as well, LUM, RMS, and RMSL are also all very strongly dependent pointwise. This seems natural, since with these two-tone images, midrange values on the luminance channel can be intuitively understood to indicate that two or more disks are causally responsible for the non-extremal luminance values, implying also contrast. If the luminance originated only from either black or white disk, its value would be one of two extremal values, and the corresponding contrast at that point would be zero. This difference between the two dead leaves models is illustrated in Figure 8, showing scatterplots of pointwise RMSL behavior as a function of LUM, for 1,000 samples, with radius $p = 8$ in Equation 1. In Figure 8A, source images were from the dead leaves model, and in Figure 8B, they were from the two-tone model. Clear dependency between pointwise LUM and RMSL can be seen in the two-toned case. To some extent, this same behavior would also exist in natural images thresholded to two levels, suggesting the approximately continuous scale of possible intensities to be important in making pointwise luminance and contrast appear independent.

Should there be statistically dominant complex phenomena in the data, the oriented dependencies between local luminance and contrast might be weaker. For example, it is possible to create artificial images where luminance and contrast filtering result in very different images (e.g., contrast modulation such as used in the artificial image in Figure 1 of Allard and Faubert, 2007, could be used for this effect). In natural images, luminance and contrast independence could exist in special conditions, such as camouflage, textures, or
contrast modulation by mist. It remains a possibility that our current data set might have subsets where local luminance and contrast are spatially less dependent. The result that the RMSL (at radius \( r = 8 \)) could be more easily predicted from LUM in the dead leaves images (mean \( c = 0.98 \)) than in natural images (mean \( c = 0.80 \), in Figure 3A) also hints at this direction. Yet, it is possible that some other correlate based on the LUM channel could match RMSL better in natural images than the simple application of RMSL we used. Our current results suffice to point out that simple image properties may play a statistically distinctive role in making local luminance and contrast dependent over natural images in general.

### Conclusion

We have shown that local luminance and local contrast have dependencies in natural images. The dependency structure found suggests that these two quantities are weakly dependent when examined pointwise but more strongly dependent spatially. Intuitively, when enough local measurements of the two quantities are sampled of the visual scene, corresponding shapes can be discovered to exist in both sets of measurements. Our results strengthened this intuition: We demonstrated that the contrast measurements can be closely approximated from the luminance measurements and also that independent component analysis learns paired visual features for the two quantities, where the paired features share orientation and localization. In terms of our experiments, images generated from a dead leaves model show a qualitatively similar dependency structure to natural images, implying that a relatively simple model of image formation is sufficient to account for the existence of these dependencies in natural images. To conclude, our results show a strong redundancy between local luminance and contrast and suggest that the corresponding computations do not recover two independent sources of visual information from natural visual input. Subsequently, our results predict that spatial processing mechanisms of local luminance and contrast are non-separable in visual systems.

### Acknowledgments

The authors wish to thank Mika Inki for helpful discussions, and the anonymous reviewers for helping to improve an earlier version of this manuscript. This work was supported in part by the project 116962 and grant 118653 (ALGODAN) of Academy of Finland, and in part by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778. This publication only reflects the authors’ views.

Commercial relationships: none.
Corresponding author: Jussi T. Lindgren.
Email: jtlindgr@iki.fi.
Address: Department of Computer Science and HIIT, University of Helsinki, Finland.

### References


