Measuring Perceptual Contrast in a Multi-level Framework

Gabriele Simone[1], Marius Pedersen[1], Jon Yngve Hardeberg[1], Alessandro Rizzi[2];
1: Gjøvik University College, Gjøvik, Norway. 2: University of Milano, Milano, Italy.

ABSTRACT

In this paper, we propose and discuss some approaches for measuring perceptual contrast in digital images. We start from previous algorithms by implementing different local measures of contrast and a parameterized way to recombine local contrast maps and color channels. We propose the idea of recombining the local contrast maps and the channels using particular measures taken from the image itself as weighting parameters. Exhaustive tests and results are presented and discussed, in particular we compare the performance of each algorithm in relation to perceived contrast by observers. Current results show an improvement in correlation between contrast measures and observers perceived contrast when the variance of the three color channels separately is used as weighting parameter for local contrast maps.

1. INTRODUCTION

Since the first studies on contrast in images, it has emerged how arduous it could be to give a definition of perceptual contrast and, moreover, how subjective and related to the observation task or observer experience this definition could turn out to be. For this reason, the first approaches to this topic have confined themselves to study the phenomenon at a global level, operating in controlled situations and under same constraints, the so-called "void conditions". After these very first experiments more complex measures have been devised, but a universal measure of contrast in images is still not clearly defined.[1–3]

After a brief introduction of some existing measure, we present a novel approach to measure perceptual contrast in digital images implemented as a modification of pre-exisitng algorithms.

2. BACKGROUND

Several contrast measures have been proposed so far. The classic approaches consist of global measures which result to be inadequate in measuring real visual configurations. In fact, the study of contrast in an image at a global level provides only a measure related on the maximum global difference in lightness and in some cases chromaticity. The response of the human visual system depends much less on the absolute luminance than on the relation of its local variations. Global measures have been mainly developed during the second half of the 20th century and luminance seems to play fundamental role in calculation. The first one was proposed earlier in 1927 by Michelson[4] and contrast is defined as follows: $C^M = \frac{L_{max}-L_{min}}{L_{max}+L_{min}}$. King-Smith and Kulikowski[5] (1975), Burkhardt[6] (1984) and Whittle[7] (1986) follow a similar concept replacing L_{max} or L_{min} with L_{avg}, which is the mean luminance in the image.

Even if they are useful applications on simple visual configurations, they are unsuitable for natural images. To overcome the limits of global measures, local measures have been developed later. We present here only a part of them, from which we started to implement the proposed approach.

Tadmor and Tolhurst[8] proposed in 1998 a measure based on the Difference Of Gaussian (D.O.G.) model. They propose the following criteria to measure the contrast in a *pixel (x,y)*, where x indicates the row and y the column:

$$c^{DOG}(x,y) = \frac{R_c(x,y) - R_s(x,y)}{R_c(x,y) + R_s(x,y)},$$

where R_c is the output of the so called central component and R_s is the output of the so called surround component.

The central and surround components are calculated as:

$$R_c(x,y) = \sum_i \sum_j Centre(i-x, j-y)\, I(i,j),$$

$$R_s(x,y) = \sum_i \sum_j Surround(i-x, j-y)\, I(i,j),$$

where *I(i,j)* is image pixel value at position *(i,j)*, while *Centre(x,y)* and *Surround(x,y)* are described by bi-dimensional Gaussian functions:

$$Centre(x,y) = \exp\left[-\left(\frac{x}{r_c}\right)^2 - \left(\frac{y}{r_c}\right)^2\right],$$

$$Surround(x,y) = 0.85\left(\frac{r_c}{r_s}\right)^2 \exp\left[-\left(\frac{x}{r_r}\right)^2 - \left(\frac{y}{r_r}\right)^2\right],$$

where r_c and r_s are their respective radiuses, parameters of this measure.

According to Tadmor and Tolhurst, the sums over x and y in the definition of R_c and R_s should extend over a square mask with a width equal to, respectively, $(2 \cdot 3 \cdot r_c) + 1$ and $(2 \cdot 3 \cdot r_s) + 1$. The simplest case, with $r_c = 1$ and $r_s = 1$, results in a 7×7 center mask and in a 13×13 surround mask.

In their experiments, using 256x256 images, the overall image contrast is calculated as the average local contrast of 1000 pixel locations taken randomly and anyhow that center mask and surround mask do not exceed the edges of the image:

$$C^{TT} = \frac{1}{1000}\sum_{i=1}^{1000} c_i^{DOG}$$

The number of pixels to consider in the calculation should change according to the image size.

In 2004, Rizzi et al.[3] proposed a measure, referred here as RAMMG (Figure 1 left), implemented as follows:

- It performs a pyramid subsampling of the image to various levels in the CIELAB color space
- For each level, it calculates the local contrast in each pixel by taking the average of absolute value difference between the lightness channel value of the pixel and the surrounding eight pixels, thus obtaining a contrast map of each level. The following weights are for the neighbor pixels:

$$M_{RAMMG} = \frac{1}{4+2\sqrt{2}} \begin{bmatrix} \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} \\ 1 & & 1 \\ \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} \end{bmatrix} \qquad (1)$$

- The final overall measure is a recombination of the average contrast for each level:

$$C^{RAMMG} = \frac{1}{N_l}\sum_l^{N_l} \bar{c}_l, \qquad (2)$$

  where N_l is the number of levels and $\bar{c}_l$ is the mean contrast in the level l.

The measure proposed by Rizzi et al.[1] in 2008, referred here as RSC, can be thought as the combination of RAMMG and Tadmor and Tolhurst measure (Figure 1 right). It works with the same pyramid subsampling as RAMMG but:

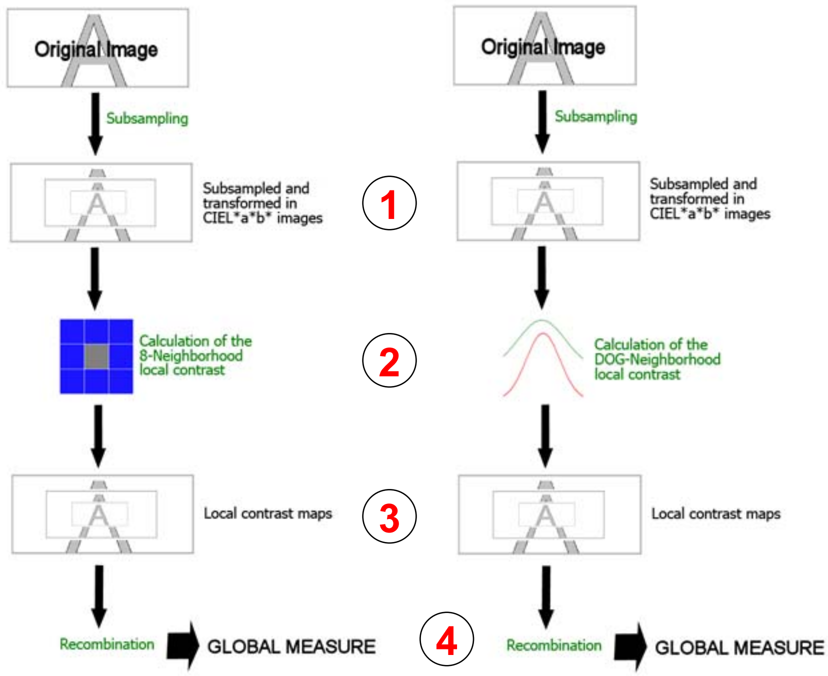

Figure 1. RAMMG and RSC algorithm steps comparison. The difference is found in the neighborhood calculation of local contrast, where RSC uses DOG-neighborhood while RAMMG uses a 8-neighborhood.

- It computes in each pixel of each level the DOG contrast instead of the 8-neighborhood local contrast
- It computes the DOG contrast separately for the lightness and the chromatic channels, instead of only for the lightness; the three measures are then combined with different weights

The final overall measure can be expressed by the formula:

$$C^{RSC} = \alpha \cdot C^{RSC}_{L*} + \beta \cdot C^{RSC}_{a*} + \gamma \cdot C^{RSC}_{b*}, \tag{3}$$

where α, β, γ represent the weighting of each channel.

3. THE PROPOSED APPROACH

As it can be seen in Figure 1, either for RAMMG or for RSC, we can distinguish four points where our analysis has been focused:

1. Pyramid construction
2. Neighborhood local contrast calculation
3. Local contrast maps
4. Global measure

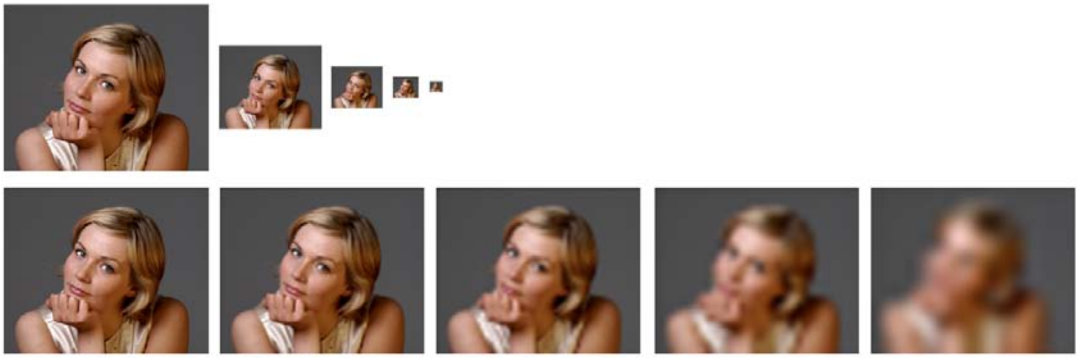

Figure 2. On the top a dimensional view is shown. On the bottom a detailed view is shown.

Figure 3. In red the missing frequency

3.1 Pyramid construction

The pyramid in RAMMG and RSC is build with a very simple not antialiased subsampling: the image is halved without prefiltering for each step (Figure 2 top). This approach is based on the principle that a multichannel analysis, not necessarily sophisticated from a frequential point of view, can be able to explain different mechanisms of perception (Figure 2 bottom).

Frankle and McCann[9] first and then Adelson et al.,[10] they proposed the use of multilevel as an important implementation feature to mimic HVS.

Adelson et al.[10] in their work on image compression, pattern recognition and image enhancement proposed a way of building the pyramid expressed by the series $P_1 = 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}...$ (Figure 3 up).

In our tests we have extended the number of pyramid levels defining the new series $P_2 = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}...$ (Figure 3 down). By preliminary tests, we found that the extended pyramid does not add substantial precision ending in a waste of computational time.

Differently from the previous works of Rizzi et al. 2004 and 2008, here we compute the pyramid using an antialiasing filtering at each level for avoiding artifacts at low resolutions.

3.2 Neighborhood local contrast calculation

In order to evaluate the avoidance by the authors of the simplest mask: $M_8 = \begin{bmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$, we have developed a very similar implementation of the RAMMG working with this mask from which we obtained equivalent results.

As mentioned before RSC is based on DOG model. In our test r_c has been set between one and three and r_s between two and four. Due to the computational cost and to a fast blurring of the image, we have decided to not go with larger Gaussians.

3.3 Local contrast maps and global measure

After building the pyramid and calculating the neighborhood local contrast, different local contrast maps are obtained at each level. For the RAMMG a recombination given by the mean of the averages for each level in lightness channel, results in the final overall measure described by Equation 2. The same recombination is done by the RSC but also on chromatic channels resulting in a final overall measure expressed by Equation 3. For a complete comparison with RSC, RAMMG has been extended to chromatic channels.

The overall contrast in one channel is given by the equation:

$$C_i = \frac{1}{N_l}\sum_{l}^{N_l}\overline{c}_l, \tag{4}$$

where N_l is the number of levels, $\overline{c}_l$ is the mean contrast in the level l and i indicates the applied channel.

As demonstrated by McCann,[9] Adelson et al.[10] and Peli[11] each level has a different contribution so we redefine the equation as follows:

$$C_i = \frac{1}{N_l}\sum_{l}^{N_l}\lambda_l \cdot \overline{c}_l, \tag{5}$$

where N_l is the number of levels, $\overline{c}_l$ is the mean contrast in the level l and i indicates the applied channel as before and the new parameter λ_l is the weight assigned to each level l.

In our tests for the parameter λ we have chosen:

1. 1
2. $\frac{1}{n}$ where n is the corresponding level of the pyramid
3. $\frac{1}{m}$ where m is the mean of pixel values in each channel separately at each level of the pyramid
4. τ as the variance of pixel values in each channel separately at each level of the pyramid

The overall final measure is given by equation:

$$C^{MLF} = \alpha \cdot C_1 + \beta \cdot C_2 + \gamma \cdot C_3, \tag{6}$$

where α, β, γ are the weights of each color channel and MLF is just the name of our proposed measure. Supposing for example to measure contrast in CIELAB color space, the contrast measure equation results in:

$$C^{MLF} = \alpha \cdot C_{L*} + \beta \cdot C_{a*} + \gamma \cdot C_{b*} \tag{7}$$

In our tests for the parameters α, β, γ we have chosen respectively:

1. 1, 0, 0
2. 0.333, 0.333, 0.333
3. 0.5, 0.25, 0.25
4. 0, 0.5, 0.5
5. $\frac{1}{v}$ where v is the mean of pixel values in each channel separately
6. ω as variance of pixel values in each channel separately

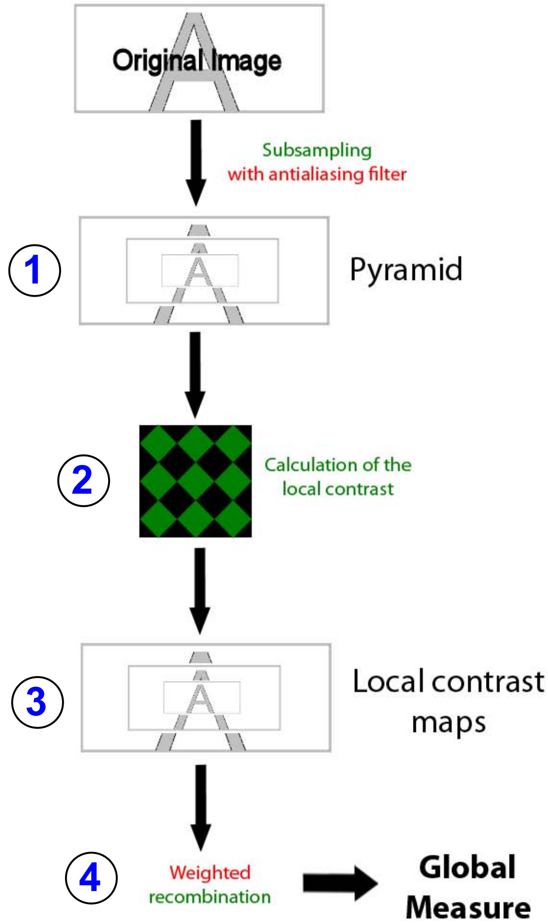

Figure 4. Proposed Measure

Of course settings 1, 3 and 4 have no sense for RGB and XYZ color space but possible future studies for hue predominance in contrast measure will consider unbalanced channel weights. In part it's already done by using as weights statistical measures taken from the image (settings 5-6). Pedersen et al[2] demonstrated that the Lab variance is a good indicator of users perceived contrast. For this reason we have focused on the choice to use variance as weighting paramater for local contrast maps and for color channels.

The general structure of our proposed measure can be seen in Figure 4 where in red are shown the most important novelties:

- Antialiasing filter in pyramid
- Weighted recombination of local contrast maps

In this framework RAMMG and RSC previously developed can be considered just special cases with uniform weighting levels and uniform weighting channels.

4. RESULTS AND COMPARISON WITH USER RATING

In order to evaluate the goodness of the algorithms in relation to perceived contrast a psychophysical experiment have been carried out.

The test set is the same presented by Pedersen et al.,[2] composed of 15 different images, representing different characteristics. 17 observers were asked to rate the contrast in the 15 images. Nine of the observers were experts, i.e. had experience in color science, image processing, photography or similar and eight non-experts had no or little experience in these fields. All observers were recruited from Gjøvik University College, both students and employees. Observers rated contrast from 1 to 100, where 1 was the lowest contrast and 100 maximum contrast. The observers were told to rate the contrast as they comprehended contrast, i.e. no definition of contrast was made by the researchers before starting the experiment. All observers had normal or corrected to normal vision. Each image was shown for 40 seconds with a surrounding black screen, and the observers stated the perceived contrast within this time-limit. The experiment was carried out on a calibrated CRT monitor, LaCIE electron 22 blue II, in a gray room. The observers were seated at approximately 80 cm from the monitor, and the lights were dimmed and measured to approximately 17 lux.

Around 10000 algorithm variations have been calculated so given the huge amount of data we show results with Pearson correleation greater than 0.75 for RAMMG and greater than 0.8 for RSC, in addition to default parameters.

Table 1. Yellow cells indicate the standard values. Grey cells indicate the highest Pearson and Spearman correlation. All numbers are round to 3 decimals. λ equal to τ indicates that each level of the pyramid is weighted with the variance of the three channels separately. $\frac{1}{v}$ indicates that the considered channel is weighted with the reciprocal of its mean while ω indicates that is weighted with its variance.

$C^{MLF(RAMMG)}$								
ColorSpace	λ	α	β	γ	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
CIELAB	1	1	0	0	0,409	0,130	0,397	0,143
CIELAB	1	0.333	0.333	0.333	0,537	0,039	0,472	0,076
RGB	τ	0.333	0.333	0.333	0,775	0,001	0,710	0,003
RGB	τ	ω	ω	ω	0,775	0,001	0,710	0,003
RGB	τ	$\frac{1}{v}$	$\frac{1}{v}$	$\frac{1}{v}$	0,778	0,001	0,693	0,004
CIELAB	τ	0.5	0.25	0.25	0,782	0,001	0,599	0,018
CIELAB	τ	ω	ω	ω	0,783	0,001	0,576	0,025
CIELAB	τ	0.333	0.333	0.333	0,783	0,001	0,576	0,025

Table 2. Yellow cells indicate the standard values. Grey cells indicate the highest Pearson and Spearman correlation. All numbers are round to 3 decimals. λ equal to τ indicates that each level of the pyramid is weighted with the variance of the three channels separately. $\frac{1}{v}$ indicates that the considered channel is weighted with the reciprocal of its mean while ω indicates that is weighted with its variance.

$C^{MLF(RSC)}$										
ColorSpace	rc	rs	λ	α	β	γ	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
CIELAB	1	2	1	0.333	0.333	0.333	-0,226	0,418	-0,214	0,443
CIELAB	1	2	1	0.5	0.25	0.25	-0,159	0,571	-0,197	0,483
CIELAB	1	2	1	1	0	0	0,016	0,954	-0,213	0,447
RGB	1	2	τ	0.333	0.333	0.333	0,818	0,000	0,683	0,005
RGB	2	4	τ	$\frac{1}{v}$	$\frac{1}{v}$	$\frac{1}{v}$	0,819	0,000	0,715	0,003
CIELAB	3	4	τ	0.5	0.25	0.25	0,822	0,000	0,719	0,003
RGB	2	3	τ	$\frac{1}{v}$	$\frac{1}{v}$	$\frac{1}{v}$	0,826	0,000	0,774	0,001
CIELAB	2	3	τ	0.5	0.25	0.25	0,829	0,000	0,738	0,002
CIELAB	2	4	τ	0.5	0.25	0.25	0,830	0,000	0,819	0,000
RGB	2	4	τ	ω	ω	ω	0,844	0,000	0,803	0,000
RGB	2	4	τ	0.333	0.333	0.333	0,845	0,000	0,785	0,001

As first we can notice from all the tables the absence of results with XYZ color space, which seems to be inappropriate for contrast measure. RGB and CIELAB appear to be equal as results show a difference not greater than ± 0.03.

What we can point out is that the parameter λ, giving a weight to each level of the pyramid, is the key of improving contrast measure especially if λ equal to τ or to be more precise equal to the variance of each channel separately (setting 4).

Furthermore, unlike mentioned in previous studies, α, β, γ loose their importance as we can see that results don't have so much difference.

For the RSC algorithm has to be underlined also that it works better with greater values of r_c and r_s compared to the standard values given by Tadmor and Tolhurst in their previous studies.

Same discussion can be done for Spearman correlation which is anyway always lower than Pearson correlation. With a p-value of 5% (or 0.05) there is only a 5% chance that results you are seeing would have come up in a random distribution, so you can say with a 95% probability of being correct that the variable is having some effect, assuming your model is specified correctly. For both measures of correlation a very low p-value confirm the correctness of our measures.

In conclusion we propose our measure of contrast, based on RSC neighborhood computation, that we call $C^{MLF(RSC)}$ and defined as follows:

$$C^{MLF(RSC)} = \omega_R \cdot C_R^{RSC} + \omega_G \cdot C_G^{RSC} + \omega_B \cdot C_B^{RSC} \tag{8}$$

while

$$\begin{aligned} C_R^{RSC} &= \tfrac{1}{N_l}\sum_l^{N_l} \tau_{R,l} \cdot \overline{c}_{R,l}, \\ C_G^{RSC} &= \tfrac{1}{N_l}\sum_l^{N_l} \tau_{G,l} \cdot \overline{c}_{G,l}, \\ C_B^{RSC} &= \tfrac{1}{N_l}\sum_l^{N_l} \tau_{B,l} \cdot \overline{c}_{B,l}, \end{aligned}$$

where N_l is the number of levels, $\overline{c}_{R,l}$, $\overline{c}_{G,l}$, $\overline{c}_{B,l}$ are respectively the mean contrast in the level l for R, G and B channel, $\tau_{R,l}$, $\tau_{G,l}$, $\tau_{B,l}$ are respectively the variance of pixel values in the level l for R, G and B channel, ω_R, ω_G, ω_B are respectively the variance of pixel values for R, G and B channel used for recombining the overall contrast of each channel C_R^{RSC}, C_G^{RSC} and C_B^{RSC}.

5. CONCLUSIONS

In this paper, we have proposed and discussed new approaches for contrast measures. We have improved previous algorithms by using different local measures of contrast and with parameterized way to recombine local contrast maps and color channels.

The idea to retrieve information from the image itself, for example the variance of the RGB channels, and use them as parameters seems to be promising for having a general contrast measure independent by many custom parameters.

The psychophysical experiment clearly shows a considerable improvement of the contrast measures especially when variance is used as parameter.

Future studies could consider an extended set of images, including also computer generated images, and the evaluation of the contrast measures in relation to user perceived contrast in uncontrolled environment. Besides the use of gaze information and saliency maps could be also an interesting way to explore.

REFERENCES

[1] Rizzi, A., Simone, G., and Cordone, R., "A modified algorithm for perceived contrast measure in digital images," in [*CGIV*], (2008).

[2] Pedersen, M., Rizzi, A., Hardeberg, J. Y., and Simone, G., "Evaluation of contrast measures in relation to observers perceived contrast," in [*CGIV 2008 - Fourth European Conference on Color in Graphics, Imaging and Vision*], 253–256, IS&T, Terrassa, Spain (Jun 2008).

[3] Rizzi, A., Algeri, T., Medeghini, G., and Marini, D., "A proposal for contrast measure in digital images," in [*CGIV 2004 – Second European Conference on Color in Graphics, Imaging and Vision*], (2004).

[4] Michelson, A., [*Studies in Optics*], University of Chicago Press (1927).

[5] King-Smith, P. E. and Kulikowski, J. J., "Pattern and flicker detection analysed by subthreshold summation.," *J Physiol* **249**, 519–548 (August 1975).

[6] Burkhardt, D. A., Gottesman, J., Kersten, D., and Legge, G. E., "Symmetry and constancy in the perception of negative and positive luminance contrast," *J. Opt. Soc. Am. A* **1**(3), 309 (1984).

[7] Whittle, P., "Increments and decrements: luminance discrimination," *Vision Research* (26), 1677–1691 (1986).

[8] Tadmor, Y. and Tolhurst, D., "Calculating the contrasts that retinal ganglion cells and lgn neurones encounter in natural scenes," *Vision Research* **40**, 3145–3157 (2000).

[9] Frankle, J. and McCann, J., "Method and apparatus for lightness imaging," (4) (1983).

[10] Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J., and Ogden, J. M., "Pyramid methods in image processing," *RCA Engineer* **29** (Nov. 1984).

[11] Peli, E., "Contrast in complex images," *Journal of the Optical Society of America* **7**, 2032–2040 (1990).