16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

WORD DESCRIPTORS OF IMAGE QUALITY BASED ON LOCAL DISPERSION-VERSUS-LOCATION DISTRIBUTIONS

Alfredo Restrepo (Palacios) and Giovanni Ramponi

IPL, DEEI, University of Trieste
Via Valerio 10 – 34127 Trieste, Italy
phone: + (39) 040.5587140, fax: + (39) 040.5583460, email: arestrepo, ramponi@units.it; arestrep@uniandes.edu.co
web: www.units.it/ipl

ABSTRACT

We consider an image descriptor that is derived from the scatter plot of range.vs.midrange values, taken uniformly across the image. The points in the scatter plot are classified as belonging to one of four main regions: *high-contrast/medium-luminance*, *low-contrast/high-luminance*, *medium-contrast/medium-luminance* and *low-contrast/low-luminance;* finally, a *word descriptor* is obtained by quantizing, normalizing and ordering the percentages of points in each of the four regions.

1. INTRODUCTION

We derive an image descriptor that measures the relationship between the local contrast and the local lightness in the image, measured respectively with the help of *estimators of location and dispersion,* that is, statistical estimators of distribution parameters such as the mean or center of mass, and the variance or width; examples of estimators of location are the average, the sample median and the midrange (the average of the min and the max), while the sample variance, the range (the max minus the min) and the quasiranges are examples of estimators of dispersion. Image quality measures defined along these lines have been previously considered, e. g. in [3]. One of the purposes of the paper is to shed light on what a relation between local luminance and local contrast should be in order to have different types of visually appealing images. Scatter plots of local dispersion versus location (*d-l plots,* for short) carry a fair amount of the information in the image; the information is further reduced to a *word descriptor* obtained by classifying the points as belonging to one of four main regions, counting them, normalizing and ordering the numbers. The four regions in question are labeled as A, B, C and D and correspond to *high-contrast/medium-luminance*, *low-contrast/low-luminance*, *medium-contrast/medium-luminance* and *low-contrast/high-luminance.* See Figs. 1 and 2.

Rather than using the sample standard deviation and the average, as estimators of dispersion and location as e.g. in [2], we use estimators based on *order statistics*; namely, the range and the midrange. The pair range-midrange behaves analogously to the pair deviation-average, with several advantages derived from the fact that the range and midrange are linear functions of the min and the max (order) statistics.

Techniques for image enhancement, such as *gamma correction,* have clear and predictable effects on d-l plots which in turn help explaining the effectiveness of the techniques. Let us consider the particular case of luminance correction of faded photographs, as implemented in [1]; an appropriate value g of gamma is estimated as a value that makes the correlation coefficient $r_{s,m}$ between the local standard deviation s and the local average m near 0. The correlation coefficient was found to depend in a monotonic nondecreasing fashion on g; for very light images, $r_{s,m}$ is negative while for very dark images, $r_{s,m}$ is positive. The application of gamma correction to an image determines a well defined flow of the points in d-l plots from which the behavior of $r_{s,m}$ can be inferred.

Figure 1: Image “Couple”

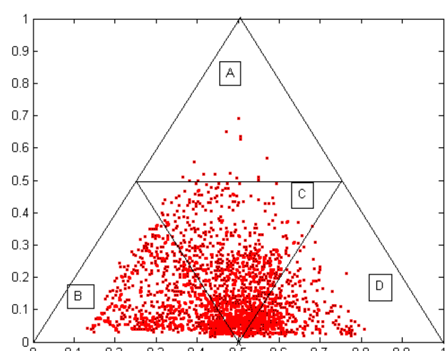

Figure 2: d-l distribution corresponding to Image “Couple” above

The percentages of the points of the scatter plot that fall in each of the regions A, B, C and D, allow for a compact description of the type of image. In Section IV, based on the relative weights of these regions for a given image, we derive a four-letter descriptor that characterizes the image in the joint dimensions of luminance and contrast.

It is not clear whether there is an ideal dependence between dispersion and location for good visual quality images; Oakley and Bu [3], for example, take the local standard deviation and the local average as estimators and assume that the local standard deviation should be proportional to the pixel average (a situation analogous to that in Weber's law) for certain types of images. This in fact is not true in general since, as mentioned, many good quality images have a correlation coefficient of local dispersion and location very near to zero. With a slighltly different definition of contrast, it is argued in [4] that there is an independence between luminance and contrast in natural scenes.

2. THE RANGE, MIDRANGE PAIR

We code the luminance component in the interval [0, 1]; also, unless otherwise stated, the shape of the window is 5×5. Both the range (the max minus the min) and the midrange (the average of the max and the min) are linear combinations of the minimum and the maximum. Besides any properties of robustness that the range and the midrange may have as estimators of dispersion and location, their use allows for an intuitive reading of d-l plots; for example, the line $range = 2 * midrange$ corresponds to $min = 0$ while the line $range = 2 - 2 * midrange$ corresponds to the line $max = 1$.

d-l plots, with dispersion d measured by the range and location l by the midrange, have points restricted to the triangle with vertices (0, 0), (0.5, 1) and (1, 0). Part of the usefulness of the (midrange, range) plane comes from the intuitiveness of the (min, max) plane. For example, since the max is larger than or equal to the min, the pair (min, max) lives on the triangle with vertices (0, 0), (1, 1) and (0, 1) of the max-min plane. There are also analytical advantages to the (midrange, range) pair; for example for independent and identically distributed data, the theoretical joint distribution of the max and the min (and therefore that of the range and the midrange) are easier to deduce than that of the sample variance and sample mean.

The pair (midrange, range) is obtained from the pair (min, max) with the help of the linear transformation $R^2 \rightarrow R^2$ with matrix

$$\begin{pmatrix} 1/2 & 1/2 \\ -1 & 1 \end{pmatrix}$$

In the (midrange, range) plane, the lines thru the origin with slope k correspond to the equation min = ((2-k)/(2+k))max while the lines thru the point (1, 0) with slope k' correspond to the equation max = ((2-k')/(2+k'))min + 2k'/(2+k'). For example, the line range=midrange corresponds to the line min=(1/3)max and the line midrange=1-range to the line max=(1/3)min+2/3. Likewise, the lines of slope k" thru the point (midrange, range) = (0.5, 0) correspond to the equation (k"-2)max + (k"+2)min = k"; for example, the lines of slopes 2 and -2 correspond to , min=0.5 and max= 0.5, respectively.

We use the lines of range=0.5, min=0.5 and max=0.5 as division lines of the (midrange, range) triangle, for the purpose of classifying the points in the scatter plots; we subdivide the triangle with vertices (0, 0), (0.5, 1), and (1, 0) of the (midrange, range) plane, into four triangles A, B, C and D, with vertexes
$A = \{(0.25, 0.5), (0.5, 1), (0.75, 0.5)\}$,
$B = \{(0, 0), (0.25, 0.5), (0.5, 0)\}$,
$C = \{(0.25, 0.5), (0.5, 0), (0.75, 0.5)\}$ and
$D = \{(0.5, 0), (0.75, 0.5), (1, 0)\}$.

3. ON D-L PLOTS

Dark images tend to have plots concentrated in the B region while light images tend to have plots concentrated in the D region. We argue that d-l plots give important information regarding image quality, in the sense that visual quality requires variety in the combination of local contrast and local intensity in the image. In particular, images that live only in the B and D regions are to be considered of poor quality.

Gamma correction is a technique that can be used for the enhancement/correction of images, e.g. of very dark and very light images. The visualization of the effects of gamma correction on the corresponding d-l plots of the images helps explain not only the change in luminance that the images undergo but also the changes in contrast and the relationships between the changes in contrast and the changes in luminance. As gamma increases (gamma larger than one) a flow of the points in the plane occurs; mainly, points in region D move to regions A and C, points in regions A and C move to region B, points in region B concentrate (towards the point of zero lightness and zero contrast) and remain in region B; conversely, for values of gamma smaller than one, points in Region B move to regions A and C, points in regions A and C move towards Region D and points in Region D remain in that region, concentrating to the point of largest intensity and zero contrast. Fig. 3 shows an extreme case of a light (faded) image while Fig. 7 shows a dark image; after gamma correction, both the d-l plots are more uniformly distributed and the visual quality of the images are improved.

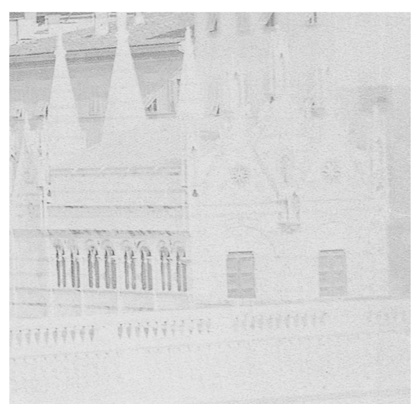

Figure 3: Image "Pisa", size: 512×512.

The effect on scatter d-l plots of (globally) raising to a power the value of the pixels in an image can be seen as a result of the the way the points (min, max) cahnge. Consider first the effect of raising to a power on the max and min. Since the function x^g (for positive g) is a monotonic nondecreasing function, the ordering of the points of the sample is respected; therefore the pair (min, max) becomes the pair

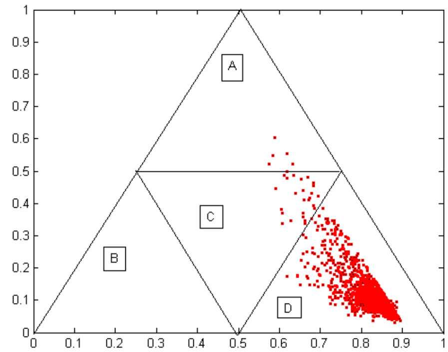

Figure 4: “Pisa”: d-l distribution.

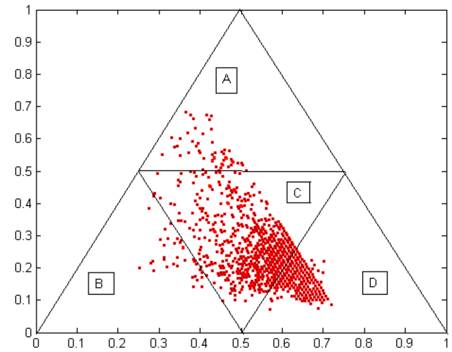

Figure 6: “Pisa”: gamma corrected, resulting d-l distribution.

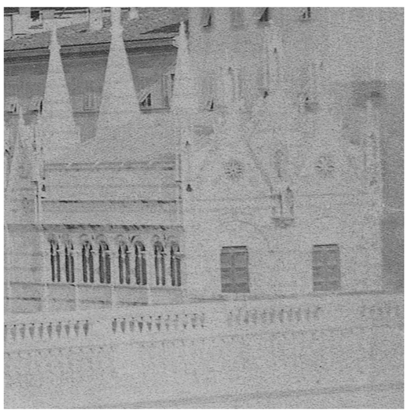

Figure 5: Image “Pisa”, gamma corrected, g=3.

Figure 7: Image “Girl” size: 256×256

($\min^g$, $\max^g$). On the other hand, the difference x^g - x for values of x in the interval [0, 1] is zero at 0 and 1 and maximal at x = $(1/g)^{\frac{1}{g-1}}$; for example, for g=0.5, the max of the difference in absolute value occurs at x = 0.707 while for g=2 the max of the abs of the difference occurs at x = 0.5. So, for intermediate values of the max and the min, the shift is maximal while for the extreme values the shift is null. The displacement of points in the (min, max) plane results from a vectorial composition so that the shift is in the NE direction for values of gamma smaller than one and in the SW direction for values of gamma larger than one. Translating this to the midrange-range plane, for gamma larger than one there is a movement towards the right for values of gamma smaller than one and towards the left for values of gamma larger than one. On one extreme, for very large values of g the points will accumulate near the vertex (midrange, range)= (0, 0) and on the other extreme, for values of gamma positive but near 0, will be near the vertex (midrange, range)= (1, 0). As a function of g, the range is low for very large and very small (positive) values and has a maximum at an intermediate value of gamma. The midrange on the other hand decreases monotonically with g.

Among the images that are strongly represented in the A region are texture images and white noise images, see Fig. 13; texture images tend to be strongly represented in the C region as well. The effects of i.i.d. noise on d-l plots can be studied theoretically with some distribution analysis. The joint distribution of the max and the min of a sample of n data from an underlying population with probability density function f and cumulative distribution function F is given by [6]

$$f_{mn,mx}(a,b) = n(n-1)[F(b)-F(a)]^{n-2} f(a) f(b)$$

For an underlying uniform distribution U[0,1], with F(t)=t and f(t)=1, the joint density becomes,

$$f_{mn,mx}(a,b) = n(n-1)[b-a]^{n-2}, a \in [0,1], a \leq b \leq 1$$

which has a maximum at (min, max) = (a, b) = (0, 1) and expectation (E[min], E[max]) = $(\frac{1}{n+1}, \frac{n-2}{n-1})$; thus, translating the result to the (midrange, range) plane, the largest likelihood will correspond to the point (0.5, 1) and the cloud of points will have a center of mass at (midrange, range) = $(\frac{n^2-3}{2(n+1)}, \frac{n^2-2n-1}{n^2-1})$. A strong A region may be an indication of noise or fine textures in the image. Consider for example the images in Figs. 11 and 12, and their corresponding d-l plots. Regions containing white noise will normally have a medium luminance, depending on their mean value, and a contrast that will depend on the strength and distribution of the noise: impulsive noise (with a heavy tailed distribution) will have a larger contrast than uniform noise; higher range values are probable for larger windows. Textured regions, because of their higher spatial organization, depend more on the size of the window.

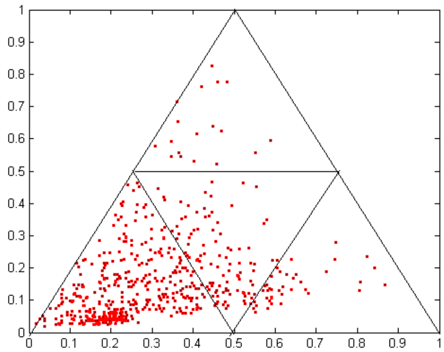

Figure 8: “Girl”

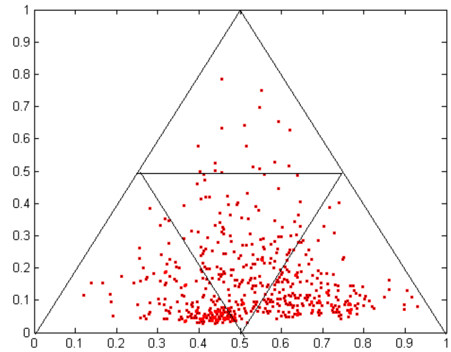

Figure 10: “Girl”: gamma corrected

Figure 9: Image “Girl”, gamma corrected, g=0.5

4. WORD DESCRIPTORS

For a given image, we compute the number of points in the scatter plot that fall on each of four triangles A, B, C and D that subdivide the basic triangle with vertices (0, 0), (0.5, 1), (1, 0) in the dispersion-location plane. Based on these distributions, we derive a four-letter word descriptor of the image as follows: we compute the percentages of points on each of the regions and order, in decreasing order, the names of the corresponding regions A, B, C and D; a descriptor of the image is obtained. It has 24 possible values as 4!=24; nevertheless, some are rather unlikely. In order to have an initial reference, based on an image set of 35 images we computed the min, max, average and median values of the percentages corresponding to each region A, B, C and D. The results are summarized in Table 1 below.

TABLE 1. Statistics corresponding to image set considered

region	min	max	median	average
A	0	16	1	3.04
B	0	89.3	16.9	27.03
C	0	64.09	17.7	22.59
D	1	98.4	40.71	43.94

We further refine the word descriptors by using respectively lower-case or upper-case letters, for each region, depending on whether or not the percentage is below the corresponding median, listed in the third column in Table 1. This says whether a given percentage on a given region is a relative large, or a small one.

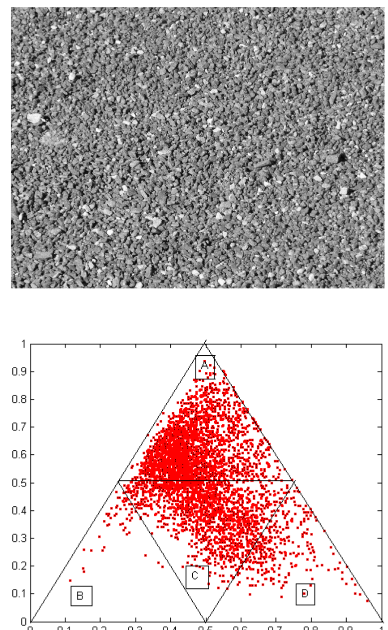

Figure 11: Image “Texture1” (McGill CCI database)

TABLE 2. Statistics corresponding to 8 images considered of good (4) and bad (4) qualities

	region	min	max	average
GOOD	A	2.4	16	8.20
	B	6.4	55.64	33.56
	C	21.35	38.7	31.92
	D	12.56	38.4	25.98
	region	min	max	average
BAD	A	0	55.81	16.76
	B	0	12.02	3.31
	C	0	60.58	24.50
	D	1	9.77	30.68

Table 2 gives statistics regarding a subset of 8 images subjectively classified as Good and Bad. The corresponding descriptors were CdAb, CBdA, BCdA and Bdca, for the good images and Dacb, Da’b’c’ (primed letters indicate abscence of points), ACdb and CdbA, for the bad images.

Under gamma correction, the image Pisa underwent a change in word descriptor from Dacb to DCab; its quality improved by a decrease of lightness and an increase of texture.

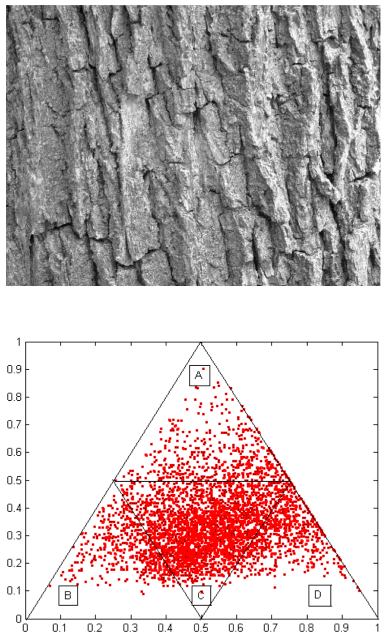

Figure 12: Image “Texture2” (McGill CCI database)

The image girl changed from BcdA to Bdca; it became less dark and got more detail as well. Texture1 (ACdb) has higher contrast than Texture2 (CdbA); in each case, local contrast occurs above minimal levels, also regions B and D are under-represented; this in opposition to good quality images which often include regions that are nearly constant.

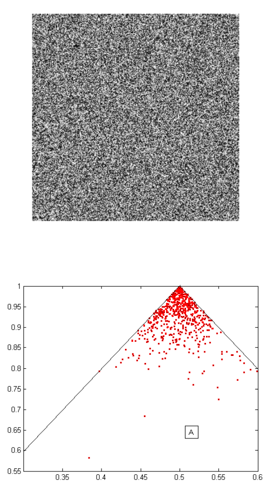

Figure 13: Image “Uniform noise”

5. CONCLUSION

Broadly, we conclude that the regions B and D should be balanced so that the image is neither too dark nor too light; also, the regions A and C should be at the end of the descriptor, otherwise the image is likely to have too much detail. Likewise, at least 20 percent of the points should fall on each of the regions B and D, otherwise the image is too light or too dark; a 5 percent minimum on the A and C regions is highly desired as well. Also, at least two large-case letters should be present; of course this will depend on the underlying data set used; in addition, a complete lack of points in any of the regions is an indicator of bad image quality. We considered the effects of changing the size of the window; the resulting d-l distributions remained quite unchanged, pointing to a robust image measure. For images with extreme resolutions, different window sizes will be convenient; also, as the window size increases there will be a greater chance for larger ranges and, obviously, for very large window sizes the overall midrange and range of the image will be obtained. For example, we considered window sizes 13 (in an aztec-cross shape) and 3×3 for the image GIRL; the general shape of the scatter plots remained unchanged while some points decreased their vertical coordinate (dispersion or range) by small amounts. Variations of the theme are possible along several directions. The proposed word descriptor surely has applications in image classification and retrieval, for queries of the type, *dark, light, low-contrast* etc. The use of letters allows for a compact descriptor and for speedy retrieval even if a large number of images may turn up. Different estimators of location and dispersion have different advantages; in particular, if the noise is an issue, it is known that the average, median and midrange are the maximum likelihood estimators of the mean for populations with Gaussian, Laplacian and uniform distributions, respectively. [5].

Regarding a desirable dependency of contrast on luminance, even though it seems to depend on the type of image, for natural scene images, there are important arguments that support an uncorrelatedness of the two dimensions [4].

6. ACKNOWLEDGEMENTS

Image Pisa was kindly provided by F.lli Alinari SpA. This work was partially supported by the FIRB project no. RBNE039LLC, 2005-08, and by a grant of the University of Trieste. [1]

REFERENCES

[1] G. Ramponi. “Adaptive contrast improvement for still images and video frames.” *IEEE- EURASIP Workshop on Nonlinear Signal and Image Processing, NSIP-07,* Bucharest, Romania, Sept. 10-12, 2007.

[2] A. Restrepo and G. Ramponi. “Filtering and luminance correction of aged photographs.” *IST/SPIE E.I. Sci. and Techn.* San Jose, Jan. 27-31, 2008.

[3] J.P. Oakley and H. Bu. “Correction of simple contrast loss in color images.” *IEEE trans. on IP,* vol. 16, no. 2, 2007.

[4] V. Mante, R.A. Frazor, V. Bonin, W. Geisler and M. Carandini. “Independence of luminance and contrast in natural scenes and in the early visual system.” *Nature Neuroscience,* vol. 8, no. 12, pp. 1690-1697, 2005.

[5] A. Restrepo and A.C. Bovik. “On the Statistical Optimality of Locally Monotonic Regression, ” in *IEEE Trans. on ASSP, vol. 42, N. 6, pp. 1548–1550* 1994.

[6] H.A. David. *Order Statistics, 3rd ed.* Wiley, 2003.

[1] A.Restrepo is on leave from the Dpt. of Ingenieria Electrica y Electronica of the Universidad de los Andes at Bogota, Colombia