Fast Surface Grading Using Color Statistics in the CIE Lab Space

Fernando López[1], José Miguel Valiente[1], Ramón Baldrich[2], and María Vanrell[2]

[1] Universidad Politécnica de Valencia, 46022 Valencia, Spain
flopez@disca.upv.es
[2] Universitat Autònoma de Barcelona, CVC, 08193 Cerdanyola del Vallès, Spain

Abstract. In this paper we approach the problem of fast surface grading of flat pieces decorated with random patterns. The proposed method is based on the use of global statistics of color computed in the CIE Lab space. Two other fast methods based on color histograms [1] and Centile-LBP features [8] are introduced for comparison purposes. We used CIE Lab in order to provide accuracy and perceptual approach in color difference computation. Experiments with RGB were also carried out to study CIE Lab reliability. The ground truth was provided through an image database of ceramic tiles. Nevertheless, the approach is suitable to be extended to other random decorated surfaces like marble, granite, wood or textile stuff. The experiments make us to conclude that a simple collection of global statistics of color in the CIE Lab space is powerful enough to well discriminate surface grades. The average success surpasses 95% in most of the tests, improving literature methods and achieving factory compliance. [3]

1 Introduction

The background problem is to solve the question of surface grading of flat pieces decorated with random patterns. These include surfaces from nature (wood, marble or granite) and artificial surfaces (ceramic tiles or textile stuff). The aim of surface grading is to split the production into different classes sorted by their global appearance, which is crucial to achieve competitive quality standards. Industries related with the manufacturing of these products rely the task of grading on human operators. This grading is subjective and often inconsistent between different graders [7]. Thus, automatic and reliable systems are needed. Also, real time compliance is important in order to make systems able to inspect the overall production at on-line rates.

In the last decade many approaches about surface grading were developed, mainly for the industrial sectors of ceramics, marble, granite and wood. Boukouvalas et al [1][2][3] proposed color histograms and dissimilarity measures of these distributions to grade ceramic tiles. No real time compliance was studied.

Other works were related with an specific type of ceramic tiles, the *polished porcelanic* tiles, which imitate granite appearance. These works included texture

[3] Partially supported by FEDER-CICYT(DPI2003-09173-C02-01) and Keraben S.A.

features. Baldrich et al [4] proposed a perceptual approximation based on the use of discriminant features defined by human classifiers at factory. These features were mainly related to grain distribution and size. The method included grain segmentation and features measurement. Lumbreras et al [5] joined color and texture through multiresolution decompositions on several color spaces. They tested combinations of multiresolution decomposition schemes (Mallat's, *àtrous* and wavelet packets), decomposition levels and color spaces (Grey, RGB, Ohta and Karhunen-Loève transform). Peñaranda et al [6] used the first and second histogram moments of each channel of the RGB space. This simple approximation, together with a deep studied inspection system, were able to comply time requirements for on-line inspection. In Baldrich and Lumbreras's works there are no study about time compliance.

On wood grading, Kauppinnen [7] developed a method based on the percentile features of histograms calculated for RGB channels. These features are also called Centiles. Kyllönen et al [8] made an approach using color and texture features. For color they chose the above mentioned Centiles, and LBP (Local Binary Pattern) histograms for texture description.

Lebrun and Macaire [9] described the surfaces of the Portuguese "Rosa Aurora" marble using the mean color of the background and mean color, absolute density and contrast of marble veins. They achieved good results but their approach is very dependent on the properties of this marble. Finally, Kukkonen et al [10] presented a system for the grading of ceramic tiles using spectral images. Spectral images have the inconvenient of producing great amounts of data.

	ground truth	**features**	**time study**	**accuracy %**
Boukouvalas	ceramic tiles	color	no	-
Baldrich	polished tiles	color/texture	no	92.0
Lumbreras	polished tiles	color/texture	no	93.3
Peñaranda	polished tiles	color/texture	yes	-
Kauppinen	wood	color	yes	80.0
Kyllönen	wood	color/texture	no	-
Lebrun	marble	color/texture	no	98.0
Kukkonen	ceramic tiles	color	no	80.0

Table 1. Summary of surface grading literature.

Many of these approaches were very specialized in a specific type of surface, others did not achieve good enough accuracy, and others did not take into account the time restrictions of a real inspection at factory. As a result of this, we think surface grading is still an open research field. In this paper we present a generic method suitable to be used in a wide range of random surfaces; ceramic tiles, marble, granite, wood, textile stuff, etc. The approach uses fast and simple statistics of color, achieving good results with a representative data set of ceramic tiles. Thus, the method is appropriate to be implemented on systems with real time requirements, typical in these contexts.

2 Lab Statistics

The presented method is simple, a set of statistical features describing color properties are collected. The features are computed in a perceptually uniform color space, the CIE Lab. These statistics form a feature vector used in the classification stage where the well known k-NN method [11] was chosen as classifier.

CIE Lab was designed to be perceptually uniform. The term 'perceptual' is refered to the way that humans perceive colors, and 'uniform' implies that the perceptual difference between two coordinates (two colors) will be related to a measure of distance, which commonly is the Euclidean distance. Thus, color differences can be measured in a way close to the human perception of colors.

The images of the data set were acquired originally in RGB, then, conversion to CIE Lab coordinates was needed. This conversion is made using the standard RGB to CIE Lab transformation [12] as follows.

RGB to XYZ:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

XYZ to CIE Lab:

$$L = 116(Y/Y_n)^{1/3} - 16$$
$$a = 500((X/X_n)^{1/3} - (Y/Y_n)^{1/3})$$
$$b = 200((Y/Y_n)^{1/3} - (Z/Z_n)^{1/3})$$

X_n, Y_n, and Z_n are the values of X, Y and Z for the illuminant (reference white point). We followed the ITU-R Recommendation BT.709, and used the illuminant D_{65}, where $[X_n\, Y_n\, Z_n] = [0.95045\ \ 1\ \ 1.088754]$.

We proposed several statistical features for describing surface appearance. For each channel we chose the mean, the standard deviation $\sigma(z)$ and the average deviation $ADev(z)$.

$$\sigma(z) = \sqrt{\frac{\sum_{i=1}^{L}(z_i - m)}{L-1}} \qquad\qquad ADev(z) = \frac{1}{L}\sum_{i=1}^{L}|z_i - m|$$

where z is the random variable, L size of the data set and m the mean value of z values.

Also, by computing the histogram of each channel, we are able to calculate histogram moments. We defined two blocks of histogram moments; one from 2nd to 5th and the other from 6th to 10th. The nth moment of z about the mean is defined as

$$\mu_n(z) = \sum_{i=1}^{L}(z_i - m)^n p(z_i)$$

where z is the random variable, $p(z_i)$, $i = 1, 2, \ldots, L$ the histogram, L the number of distinct variable values and m the mean value of z.

3 Literature methods

For comparison purposes we selected two methods from literature: color histograms [1] and Centile-LBP [8]. They are similar to ours, both are generic solutions with low computational costs. Color histograms are 3D histograms (one axis per space channel) which are compared using dissimilarity measures. In [1] they used the *chi square test* and the *linear correlation coefficient*.

$$\chi^2 = \sum_i \frac{(R_i - S_i)^2}{R_i + S_i} \qquad\qquad r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})}\sqrt{\sum_i (y_i - \bar{y})}}$$

When comparing two binned data sets with the same number of data points the *chi square* statistic (χ^2) is defined as above, where R_i is the number of events in bin i for the first data set, and S_i is the number of events in the same bin for the second data set. The *linear correlation coefficient* (r) measures the association between random variables for pairs of quantities (x_i,y_i), $i = 1,...,N$. The mean of the x_i values is $\bar{x}$ and $\bar{y}$ is the mean of the y_i values.

The Centiles, are calculated from a cumulative histogram $C_k(x)$, which is defined as a sum of all the values that are smaller than x or equal to x in the normalized histogram $P_k(x)$, corresponding to the color channel k. Finding a value for a percentile is finding the x when $C_k(x)$ is known, thus, requiring an inverse function of $C_k(x)$. Let $F_k(y)$ be the percentile feature, then $F_k(y) = C_k^{-1}(y) = x$, where y is a value of the cumulative histogram in the range [0%,100%].

The Local Binary Pattern (LBP) is a texture operator where the original 3x3 neighborhood is thresholded by the value of the center pixel (figure 1b). The values of the pixels in the thresholded neighbourhood are multiplied by the weights given to the corresponding pixels (figure 1c). Finally, the values of the eight pixels are summed to obtain the number of this texture unit. Using LBP there are 2^8 possible combinations of texture numbers, then a histogram collects the LBP texture description of an image.

6	5	3
7	5	2
9	3	7

(a)

1	1	0
1		0
1	0	1

(b)

1	2	4
8		16
32	64	128

(c)

1	2	0
8		0
32	0	128

(d)

LBP=1+2+8+32+128=171

Fig. 1. Computation of local binary pattern (LBP).

In [8] Centile and LBP features were combined in one measure of distance and then used the k-NN classifier. For Centile features they used the Euclidean distance in the feature space. For LBP they used a log-likelihood measure: $L(S,R) = -\sum_{n=0}^{N-1} S_n ln R_n$, where N is the number of bins. S_n and R_n are the sample and reference probabilities of bin n. The distances were joined by simply adding them. Previously both distances were normalized using the min and max values of all the distances found in the training set.

4 Experiments and Results

All the experiments were carried out using the same data set. The ground truth was formed by the digital RGB images of 492 tiles acquired from eight different models, each one with three different surface classes given by specialized graders at factory. For each model there were two close classes and one class far to them.

Models were chosen representing the extensive variety that factories can produce, a catalogue of 700 models is common. But, in spite of this great number of models, all of them imitate one of the following mineral textures; marble, granite or stone. Fixed pattern models are a subset of random pattern models.

	classes	**tiles/class**	**size (cm)**	**pattern**	**aspect**
Agata	13, 37, 38	16	33x33	fixed	marble
Berlin	2, 3, 11	24	16x16	random	granite
Firenze	9, 14, 16	20	20x25	random	stone
Lima	1, 4, 17	24	16x16	random	granite
Oslo	2, 3, 7	24	16x16	random	granite
Toscana	13, 18, 19	16	33x33	random	stone
Vega	30, 31, 37	20	20x25	fixed	marble
Venice	12, 17, 18	20	20x25	random	marble

Table 2. Ground truth of ceramic tiles.

Digital images of tiles were acquired using an illumination system spatially and temporally uniform. Spatial and temporal uniformity is important in surface grading [1,4,6] because variations on illumination can produce different shades for the same surface and then misclassifications. The illumination system was formed by two special high frequency fluorescent lamps with uniform illuminance along its length. For overcoming variations along time, the power supply is automatically regulated by a photoresistor located near fluorescents.

Two sets of experiments were made to demonstrate the feasibility of Lab statistics for solving the problem of surface grading. Firstly, experiments of statistics where carried out for the CIE Lab and RGB spaces. Classification was made using the half of the samples as training set and the remaining half as test set. Values of 1, 3, 5 and 7 were used for the k factor of the k-NN classifier.

The performance results of several statistics sets are shown in table 3. The error rates were computed as the average error ratios achieved over all models. More combinations of statistics were tested, but only the most prominent are presented. The last two columns corresponds to the averaged error rate and the 95% confidence intervals [11] respectively. The table is divided in two blocks, the first one corresponds with CIE Lab experiments. Here, the majority of sets have confidence intervals under the maximum error rate of 5% which is the factory requirement of performance. The best choice was to use the mean color plus the standard deviation. Histogram moments did not introduce any improvement. The second block collects the results of RGB which presents significant less discriminative power than CIE Lab.

mean	stddev	avedev	2-5th ms	6-10th ms	lab	rgb	error %	95% c.i
x					x		13.2	[10.3, 16.4]
x	x				x		**1.2**	**[0.33, 2.3]**
x		x			x		**3.0**	**[1.6, 4.7]**
x	x		x		x		**3.2**	**[1.7, 4.9]**
x	x		x	x	x		3.3	[1.9, 5.2]
x						x	13.4	[10.4, 16.6]
x	x					x	7.9	[5.7, 10.6]
x		x				x	7.3	[5.1, 9.9]
x	x		x			x	5.9	[4.0, 8.3]
x	x		x	x		x	6.7	[4.6, 9.2]

Table 3. Accuracy results of statistics sets in CIE Lab and RGB spaces.

In second place, experiments for color histograms and Centile-LBP were carried out. Once again, classification was made using the half of the samples for training and the remaining half for testing. In Centile-LBP experiments the original log-likelihood formula, the *chi square test* and the *linear correlation* $coefficient$ were used for measuring histograms differences.

The results of table 4 show that Centile-LBP achieves the best error rates when using RGB, but none of both methods achieves factory compliance because all of their confidence intervals surpass the max error rate of 5% required at factory. Comparing with table 3, Lab Statistics presents significant improvement in performance an also is the only method with confidence intervals complying the max factory error.

	Chi	Corr	Log	Lab	RGB	error %	95% c.i
Color Histograms	x			x		9.7	[7.2, 12.6]
Color Histograms		x		x		11.5	[8.8, 14.6]
Color Histograms	x				x	11.1	[8.5, 14.2]
Color Histograms		x			x	12.4	[9.5, 15.5]
Centile-LBP	x			x		5.6	[3.6, 7.8]
Centile-LBP		x		x		5.1	[3.3, 7.4]
Centile-LBP			x	x		8.7	[6.4, 11.5]
Centile-LBP	x				x	5.3	[3.5, 7.6]
Centile-LBP		x			x	4.6	[2.8, 6.6]
Centile-LBP			x		x	6.7	[4.6, 9.2]

Table 4. Accuracy results of Color Histograms and Centile-LBP.

Figure 2 shows the best performance response of each method itemized by models. Color histograms and Centile-LBP approaches, contrasted with Lab Statistics, present greater irregularity and more models are over the factory max error.

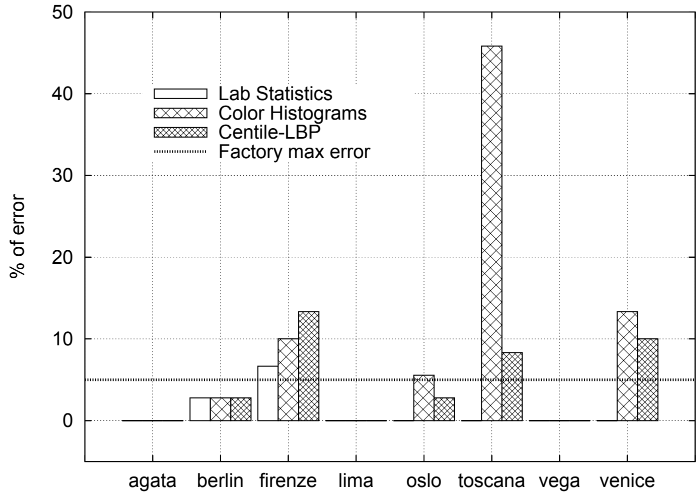

Fig. 2. Best accuracy results of Lab Statistics, Color Histograms, and Centile-LBP.

Finally, we measured the timing costs of the methods using a common PC (see figure 3). All the approaches have a theoretical computational cost of $\Theta(n) + C$, where n is the image size and C is a constant which varies depending on the approach. Lab statistics and color histograms were penalized because they need the conversion from RGB to CIE Lab. But, if we take away the RGB to CIE Lab conversion then Lab statistics achieves the best timing response.

5 Conclusions and further work

A fast method for the application of surface grading has been presented. The method uses simple statistics computed in a perceptually uniform color space, the CIE Lab. This approach performs well discriminating correctly surface grades among several types of surfaces representing a common catalogue of ceramic tiles. The benefit of using CIE Lab is demonstrated comparing with RGB results.

Other two methods coming from the literature were implemented and tested for comparison purposes. From the point of view of performance, color histograms achieved the worse results while Centile-LBP had intermediate results. The best accuracy response corresponded to Lab statistics, which also was the only method achieving factory compliance in performance with confidence intervals under the max error limit. From the point of view of timing costs, Centile-LBP had the best response, but Lab statistics was not too far away and timing can be easily improved transferring the RGB to CIE Lab conversion to hardware or using parallel processing systems.

Further work will extend the image database with more models and samples. Also, a deep study of real time compliance will be made simulating factory load and using parallel processing systems based on cluster and MPI technology.

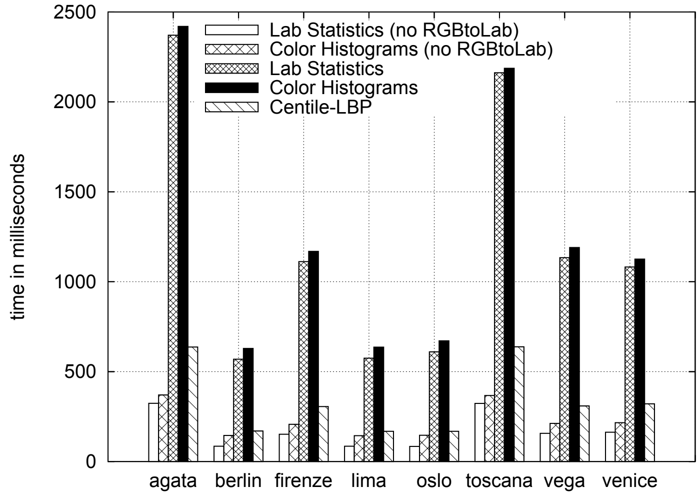

Fig. 3. Timing for the best accuracy results of each method.

References

[1] C. Boukouvalas et al. Color grading of randomly textured ceramic tiles using color histograms. IEEE Trans. Industrial Electronics. 46(1):219–226, 1999.

[2] C. Boukouvalas and M. Petrou. Perceptual correction for colour grading using sensor transformations and metameric Data. Machine Vision and Applications, 11:96-104, 1998.

[3] C. Boukouvalas and M. Petrou. Perceptual correction for colour grading of random textures. Machine Vision and Appl., 12:129-136, 2000.

[4] R. Baldrich, M. Vanrell and J. J. Villanueva. Texture-color features for tile classification. EUROPTO/SPIE Conf. on Color and Polarisation Techniques in Indust. Inspection, Germany, 1999.

[5] F. Lumbreras et al. Color texture recognition through multiresolution features. Int. Conf. on Quality Control by Artificial Vision. 1:114–121, France, 2001.

[6] J. A. Peñaranda, L. Briones and J. Florez. Color machine vision system for process control in ceramics industry. SPIE. 3101:182–192, 1997.

[7] H. Kauppinen. Development of a color machine vision method for wood surface inspection. Phd Thesis, Oulu University, 1999.

[8] J. Kyllönen and M. Pietikäinen Visual inspection of parquet slabs by combining color and texture. Proc. IAPR Workshop on Machine Vision Appl., Japan, 2000.

[9] V. Lebrun and L. Macaire. Aspect inspection of marble tiles by color line-scan camera. Int. Conf. on Quality Control by Artificial Vision, France, 2001.

[10] S. Kukkonen, H. Kvinen and J. Parkkinen. Color Features for Quality Control in Ceramic Tile Industry. Optical Engineering. 40(2):170–177, 2001.

[11] R.O. Duda and P.E. Hart. Pattern classification and scene analysis. John Wiley and Sons, New York, 1973.

[12] G. Wyszecki and W.S. Stiles. Color sciencie: concepts and methods, quantitative data and formulae. Wiley, 2nd Edition, New York, 1982.