**Periodicals of Engineering and Natural Sciences** ISSN 2303-4521
Vol. 8, No. 2, June 2020, pp.784-789

# Luminance pyramid for image generation and colorization

**Mariam A. Ali[1], Bushra KH. Alsaidi[2], Bara'a Ali Attea[3]**
[1] University of Baghdad/ College of Administration and Economics, Department of Financial and Banking Sciences
[2] University of Baghdad/ College of Administration and Economics, Department of industrial management
[3] University of Baghdad /College of Science, Computer Science Department

**Abstract**

Many image processing and machine learning applications require sufficient image feature selection and representation. This can be achieved by imitating human ability to process visual information. One such ability is that human eyes are much more sensitive to changes in the intensity (luminance) than the color information. In this paper, we present how to exploit luminance information, organized in a pyramid structure, to transfer properties between two images. Two applications are presented to demonstrate the results of using luminance channel in the similarity metric of two images. These are image generation; where a target image is to be generated from a source one, and image colorization; where color information is to be browsed from one colored image to another but grayscale image. The results demonstrate the suitability of luminance in achieving the two selected applications.

**Keywords**: Image generation, Image colorization, Luminance pyramid

***Corresponding Author:***

**Bushra KH. Alsaidi**
University of Baghdad, College of Administration and Economics
Department of industrial management
E-mail: d.bushra_alsaidi@coadec.uobaghdad.edu.iq

## 1. Introduction

Image feature selection and representation is a large open problem and active area of research in machine learning and image processing. Additionally, selection of the suitable feature(s) is the key starting point to building image processing and machine vision systems. In many ways, the success or failure of an algorithm depends greatly on an appropriately designed feature selection and representation [1]. In general, in the computer vision community, one can classify features as low level, intermediate level and high level. Low level deals with pixel level features such as chromatic and achromatic components of image pixels, pixel neighborhood statistics (e.g., mean, standard deviation, and Euclidian distance), and neighborhood shape and size. High level deals with abstract concepts (such as types and semantic meaning of image's objects), and intermediate level deals with something in between.

In pixel level, RGB channels of the image are used, traditionally, as the feature vector. However, this choice is insufficient for large area of image processing applications. For example, recently an active area of applications includes transferring properties between images. One such approach is that stated by Hertzmann et al [2], specifying transformations between images-image analogies (including different types of filters: blurring, embossing, artistic filters of oil, water, and pastel appearance, texture, and color transfer) can be provided in a very natural means and without specifying an explicit definition of the filters. Instead, it requires an appropriate definition of "*similarity*" to measure the relationship between source of property image and target image.

In this case, RGB channel alone may not contain enough data to match between the images as one may wish to process a target image giving a source image with completely different colors. An alternative way is to build machines that can mimic human's ability to process visual information. Even through much about the human vision system remains unknown, many biological vision theories exist. These theories could provide guidance to building practical engineering solutions to vision tasks. For example, we are much more sensitive to changes in the luminance (intensity) channels than in chromatic channels [3]. In other words, luminance vision is able to detect sharp edges and fine details of the patterns and textures in the image whereas color vision is left to "fill in" the color of the objects and forms [4].

The luminance can be computed in a number of ways. It requires converting the color space from correlated RGB space to another space which can de-correlate the intensity information from the chromatic information.

Examples of de-correlated color spaces are: YIQ and YUV [5], $l\alpha\beta$ [6], and CIECAM97 [7]. These transformations can be illustrated by the following equation:

- Between RGB and YUV

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.300 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \text{ and } \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0 & 0.0 & 1.140 \\ 1.0 & -0.394 & -0.581 \\ 1.0 & 2.028 & 0.0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix}$$

- Between RGB and YIQ

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \text{ and } \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.620 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.108 & 1.705 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

- Between RGB and $l\alpha\beta$

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Then convert the data to logarithmic space:

**L**=log $L$ , M**=log ,** **S**=log $S$

$$\begin{bmatrix} \ell \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{L} \\ \mathbf{M} \\ \mathbf{S} \end{bmatrix}, \text{ and } \begin{bmatrix} \mathbf{L} \\ \mathbf{M} \\ \mathbf{S} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \ell \\ \alpha \\ \beta \end{bmatrix}$$

$L = 10^{\mathbf{L}}$ , $\boldsymbol{M} = \mathbf{10}^{\mathrm{M}}$, $S = 10^{\mathrm{S}}$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}$$

- Between RGB and CIECAM97

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \begin{bmatrix} A \\ C1 \\ C2 \end{bmatrix} = \begin{bmatrix} 2.00 & 1.00 & 0.05 \\ 1.00 & -1.09 & 0.09 \\ 0.11 & 0.11 & -0.22 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}, \text{ and}$$

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3279 & 0.3216 & 0.2061 \\ 0.3279 & -0.6353 & -0.1854 \\ 0.3279 & -0.1569 & -4.5351 \end{bmatrix} \begin{bmatrix} A \\ C1 \\ C2 \end{bmatrix}, \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}$$

## 2. Work on luminance processing

This feature (i.e. luminance channel) has been used and found to be effective in many literatures. For example, Hertzmann et al in [2][8] use luminance feature as similarity metric to process images to approach some approximation to the style of an example target image. Their image analogies support many novel image processing operations. The heart of their algorithm is depended on the best neighborhood match of pure luminance values (in YIQ color space), blurred luminance, and/or luminance orientations represented as a multi-scale Gaussian pyramid. Their matching process proceeds from coarsest resolution of the pyramid to the finest one, synthesizing a multi-scale representation of output image, one level at a time. More specifically, image analogies synthesizes images drawn from the statistical distribution of square-shaped neighborhoods in another image.

A. Shah et al [9] developed pixel-based method to create textures through deterministic searching process. The feature used in searching involves luminance value of L-shape neighborhood for each synthesized pixel. To determine the pixel value of the output image, at a given location, it spatial neighborhood is compared against all possible neighborhoods from the input texture, and the input pixel with the most similar neighborhood is assigned to the current output pixel. Additionally, A.A. Efros and W.T. Freeman [10] present patch-based image quilting for texture synthesis and transfer. The constraints used in their work are luminance information including intensity, blurred value of intensity, and local image orientation angles. Similarly, M.Ashokhmin [11][12] uses luminance feature in his texture synthesis and transfer work.

Welsh et al [13] and Reinhard et al [7] also use luminance statistics to transfer color properties between a source color image and a target grayscale one, rather than choosing RGB colors from a palette to color individual components of the target image. Moreover, in their image colorization work, Vieira et al [14] index an image database depending on the luminance information. They use luminance vector ($l\alpha\beta$ color space) to compute each image signature. Then, the signature of a query (grayscale) image is compared with those stored in the database and the content of the database image with the most similar luminance vector is returned. Then, the algorithm adds to each scalar pixel (only luminance) of the query image, the chromatic components of an automatically chosen pixel of the returned image.

## 3. The luminance pyramid

In this section, we explore how to use image luminance feature (organized in a pyramid structure) in two applications: image generation (to generate a target image from a given source one) and image colorization (to color a target image from a given source colored one). The steps of how to construct luminance pyramid and how to use it to search for luminance similarity between the two images follow.

First, the luminance component of both source, S, and target, T, images are organized into a pyramid by successively down sampling, so that the coarsest level (or apex layer) of the pyramid has the luminance image at the lowest resolution while the finest level (or base of the pyramid) has the luminance image at full resolution. The down sampling of order *N* proceeds in the following way: for each un-overlapped luminance block of $N \times N$ pixels at coarse level $l_i$, calculate their average and set it to the value of the luminance of the sub-sampled version of the image at the next coarser level $l_{i+1}$ (i.e. sub-sampling results in 1/*N* as many pixels in each dimension of level l). Repeat this down sampling process until we reach the apex level which forms only $N \times N$ pixels. For implementation simplicity, the same of the image is set to the power of $N$ . Figure 1 depicts an image pyramid of three levels.

Figure 1. Image pyramid. The size of the image at the base level ($l_0$) is 256×256, where at $l_1$ is 64×64, and at the apex level ($l_2$) is 16×16. *N* was set to 4.the image [15]

Second, for both source and target images, go through the pyramid starting from the layer beneath the apex and search for the closest match between each target “patch of *N*×*N* pixels” in scan-line order and all source *N*×*N* patches. The best matching is based on the sum of pixel-wise luminance difference.
Once the best source patch is identified for each target patch, ascend through the image pyramid and repeat the search process. For each patch in a given resolution $l_i$, mask those patches in the layer directly beneath which can be projected back up to that current patch. Then search only the patches that have been masked in both source and target image pyramids, and repeat the process on up to the base of the pyramid.
In other words, the search routine proceeds from coarsest resolution to finest resolution, one level at a time, comparing luminance statistics pertaining to each patch in the target image against luminance statistics for every patch in the in the source image, and the “best” match is found.
Reaching the base level of both *S* and *T*, we can then use the luminance feature of the closest-matching patches as directives in two selected applications. As a proof of concept, we first tried generating the target image from the source one. It only requires to paste onto the resulting image the source patches that correspond to the best match with the target patches. On the other hand, in image colorization, the chromatic components of the best source patches are added to the target luminance patches to get a colorized target image. Given a pair of closest-match patches, one consists of achromatic and chromatic components and the second contains only achromatic information, add to each pixel in the latter patch the chromatic information of the closest-match pixel from the first patch.

## 4. Results

Figures 2 and 3 depict results obtained from utilizing luminance pyramid for image generation and colorization. Images used are of different sizes (even some images are depicted in reduced sizes). As can be carried out from the results, using luminance channel as a similarity metric between two different images is sufficient enough to produce acceptable and visual results. Moreover, instead of full searching for the best match for each target pixel, we organized luminance information of the two images, in a pyramid structure and in each layer; search for the best match between target patches (a square block of *N*×*N* pixels) and those source patches come from the best patch in layer beneath the current layer. As can be seen from the figures, the quality of image colorization results is acceptable and even comparable to those results of Welsh et al [13] and Hertzmann et al [2].

Figure 2. Results of image generation. Each result is depicted (from right to left) generated target image, target image source image, and the. *N* was set to 2, 4, and 5 respectively, while *l* was set to 2 or 3 according to the size of the image

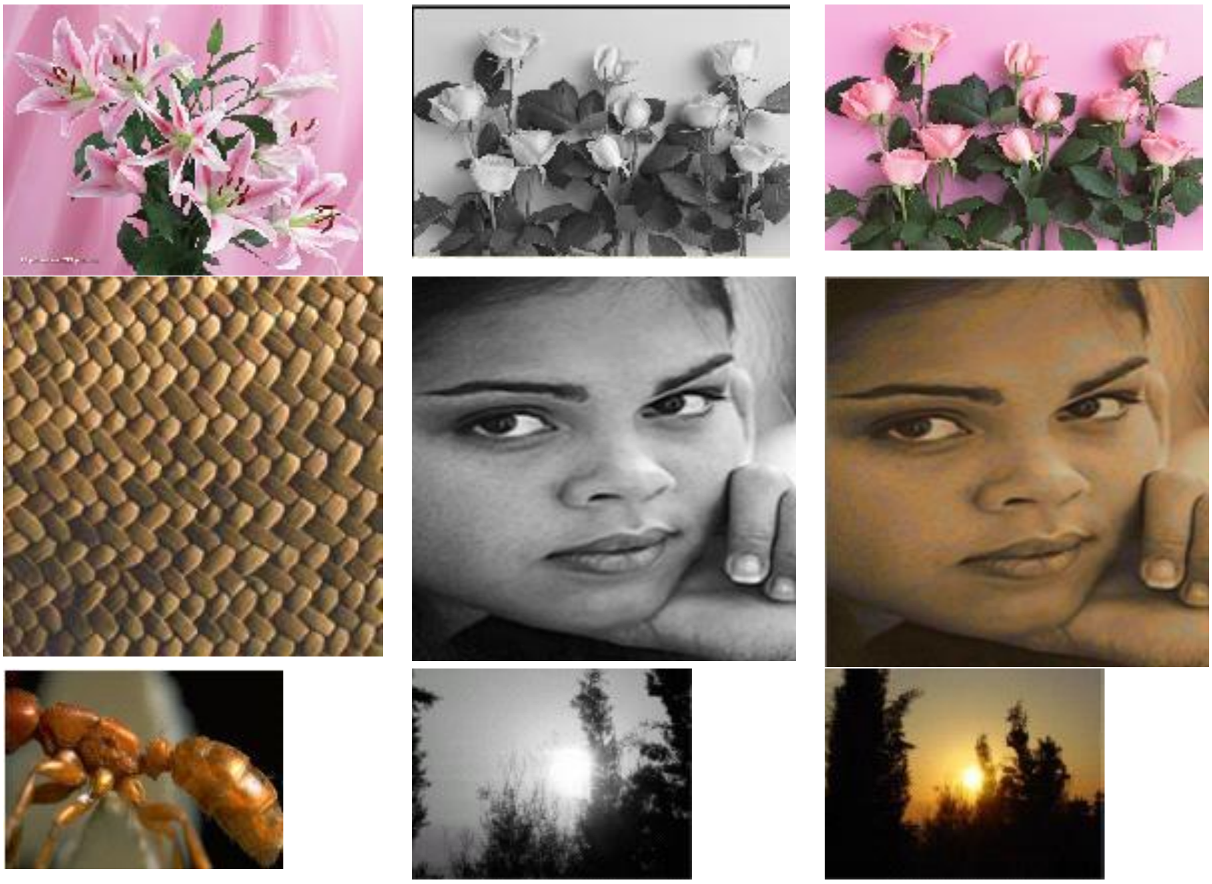

Figure 3. Results of image colorization. Each result is depicted (from right to left) the colorized target image, target image, and source image. *N* was set to 5 or 3, while *l* was set to 2

## 5. Conclusion

This paper presents a simple an effective usage of an important image feature (luminance), in pyramid structure, in two application image generation and colorization. We see that the pure value of intensity information is suitable enough to control the similarity metric between two different images so as to transfer either achromatic components from one image to another.

**References**

[1] B. Kh. AlSaidi B. J. Al-Khafaji S. Abed Al Wahab AlSaidi et al.: “Content Based Image Clustering Technique Using Statistical Features and Genetic Algorithm” Engineering, Technology & Applied Science Research Vol. 9, No. 2, 2019, 3892-3895 3892

[2] A. Hertzman, C.E. Jacobs, N. Oliver, B. Curless, and D.H. Salesin, “*Image analogies*”, In Proc. Pf ACM-SIGGRAPH 2001, pp.327-340, 2001.

[3] B. Wandell, Foundations of Vision. Sinauer Associatiates lnc., 1995.

[4] Guoping Qiu and Kin-Man Lam, “*Frequency Layered Color Indexing for Content-Based Image Retrieval*”, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 12, NO. 1, JANUARY 2003.

[5] Samruddhi Y. Kahu, Rajesh B. Raut, Kishor M. Bhurchandi, “*Review and evaluation of color spaces for image/video compression*”, https://doi.org/10.1002/col.22291, First published: 08 November 2018.

[6] Z. Pan, Z. Dong, and M. Zhang, “*A new algorithm for adding color to video or animation clips*”, Journal of WSCG pp. 2-3, 2004.

[7] E. Rrinhard, M. Ashikhmin, B. Gooch, and P. Shirley, “*Color transfer between images*”, IEEE Computer Graphics and Applications, pp. 34-41, 21(5), 2001.

[8] A. Hertzmann, “*Algorithms for rendering in artistic styles*”, Ph.D. dissertation, New York University, 2001.

[9] Ankit Shah, Parth Bhatt, Prof. Kirit J. Modi, “*Image Enhancement Techniques by Texture Synthesis*”, International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012).

[10] A.A. Efros, and W.T. Freeman, “*Image quilting for texture synthesis and transfer*”, SIGGRAPH 2001, pp. 1-6, 2001.

[11] M. Ashikhmin,” Synthesizing *natural textures*”, 2001 ACM Symposium on Interactive 3D Graphics, pp. 1-3, 2001.

[12] M. Ashikhmin, “*Fast texture transfer*”, IEEE Computer Graphics and Applications, pp. 1-4, 2003.

[13] T. Welsh, M.Ashikhmin, and K. Mueller, “*Transferring color to grayscale images*”, ACM Trans. On Graphics, 21, pp. 277-280, 2002.

[14] L. F. M. Vieira, R. D. Vilela, E. R. do Nascimento, F. A. Fernardes Jr., R. L. Carceroni, and A. A. Araujo, “*Automatically choosing source color images for coloring greyscale images*”, IEEE Computer Society, Pages 3-4, 2004.

[15] F. Petrie and J. Show “The *Big Book of Painting Nature in Watercolor”* 1990