

# ¿Es útil la armonía cromática en visión por computador?

Adrià Ciurana Lanau

**Resumen** – El proyecto trata un problema poco abordado en el campo de la visión por computador: el contenido armónico y su interpretación. El objetivo inicial era explorar la capacidad de los métodos computacionales para valorar la calidad cromática de las imágenes basándose en criterios de armonía de su contenido. El análisis de las métricas usadas dio paso a su aplicación en problemas más complejos como la generación de modelos computacionales, la corrección del contenido cromático, la interpretación semántica de la imagen o la búsqueda de imágenes en grandes BD. El trabajo realizado se ha basado inicialmente en los conceptos usados en el campo de diseño, adaptándolo a la VC. Se ha mejorado las métricas usadas hasta el momento proponiendo el uso de una medida más eficiente en términos de interpretación del contenido.

**Palabras claves** – calidad de imagen, corrección de color, descripción de emociones, recuperación de imágenes, información cromática, visión por computador.

**Abstract** – The project addresses an issue little discussed in the computer vision area: the harmonic content and its interpretation. The initial objective was to explore the ability of computational methods to value the chromatic quality of the pictures on the basis of the harmony criteria. The analysis of the metrics used led to its application to more complex problems such as the generation of computational models, correction of chromatic content, the semantic interpretation of the image or image search in large DB. The work was based initially on the concepts used in the field of design, adapting it to the computer vision. The metrics used until now have been improved suggesting the use of more efficient measurement in terms of the content interpretation.

**Index Terms** – image quality, colour correction, emotion description, image retrieval, chromatic information, computer vision.



## 1 OBJETIVOS

En la Visión por Computador, independientemente del ámbito en el que lo apliquemos, nuestro mayor problema es cómo y cuanta información podemos obtener a través de las imágenes. Las imágenes de por sí son una estructura matricial de datos, pero en su conjunto aportan muchísima información tanto de color, topológica y/o contextual. Debido a la dificultad de obtener información precisa y verídica siempre se buscan nuevas metodologías que ayuden en esta tarea.

Dicho esto podríamos distinguir entre tres niveles de información contenida en las imágenes según su complejidad de análisis:

- El color hace referencia a la información puntual que ofrece cada pixel.
- La topológica, a la unión y la información que es asociada a un conjunto de píxeles, ya sea por su estructura, geométrica, o asociación a un objeto.
- El contexto hace referencia aquello que los humanos asociamos a una imagen pero que no está implícita en ella (Por ejemplo emociones, deducciones, etcétera).

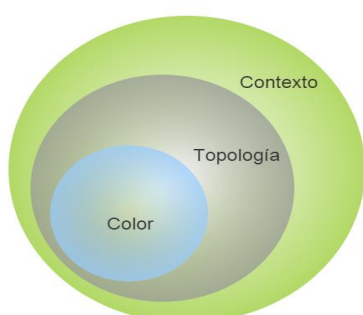


Figura 1 Diagrama de Venn sobre la información que ofrece una imagen.

Concretamente en este Trabajo de Fin de Grado, analizaremos e investigaremos sobre la información que nos es capaz de ofrecer la información cromática implícita en la imagen.

Para obtener la información cromática de la imagen, inicialmente usaremos herramientas ya existentes que podremos ver en el próximo apartado del Estado del Arte, y ver si somos capaces de mejorarlas. El análisis del color en imágenes es un campo vasto. En nuestro caso nos centraremos en las características armónicas de las imágenes.

Por otro lado este trabajo también tiene como fin, investigar hasta donde podemos llegar solamente con la información cromática de la imagen y generar modelos que permitan segmentar y explicar esa información.

Así, posteriormente se pueden usar métricas que estén relacionadas con la información cromática de la imagen y vincularlas a métodos que permitan inferir información

del contenido de la imagen y aplicables a problemas reales. No se pretende encontrar una solución universal al problema de la interpretación armónica de la imagen sino adentrarnos en su entendimiento y proponer metodologías para su uso computacional.

También es importante destacar que, por sí sola, la información cromática tiende a ser insuficiente para muchos problemas del ámbito de la Visión por Computador (Como analizaremos en este artículo) pero sí puede aportar cierta información que complementa herramientas de más alto nivel.

## 2 ESTADO DEL ARTE

El uso del término “armonía de color” ha sido y es un problema abierto sin una definición precisa. El principal campo de aplicación en las últimas décadas ha sido en el campo del diseño, del cual se alimentan las principales aproximaciones computacionales existentes.

Los primeros estudios en el campo de la armonía de color, están presentes en la Teoría del Color. Más concretamente la podemos encontrar en la rueda de Itten[1].

Ésta relaciona los colores discretamente y definiendo entre ellos armonías. La rueda de Itten, está muy presente en el campo del diseño donde muchas herramientas se basan en su uso para precisar colores a la hora de transmitir, una u otra información al observador.

Análogamente a la rueda de Itten, a nivel computacional, nos encontramos con la definición de los perfiles de Tokomaru[3] que redefinen los perfiles armónicos. El uso de un esquema unidimensional, obliga a buscar un espacio de color donde la separación cromática respecto las otras componentes esté muy diferenciada, el espacio de color seleccionado por cumplir dichas características es el HSV[2] (o HSL), donde solamente la componente cromática (Hue) es la que se tiene en cuenta en el análisis de los perfiles.

Una vez definidos los perfiles donde se contrastarán la información la imagen a analizar, se hará uso de métricas que ayuden a deducir que imágenes presentan más o menos parentesco a uno de ellos, lo que nos llevará a usar técnicas de comparación de señales, eliminación de ruido, etcétera.

## 3 TRABAJO REALIZADO

En este apartado del proyecto se explicará todo el contenido realizado, a lo largo del TFG. Es importante constatar que debido a la limitación de páginas del propio artículo, algunos términos no podrán ser explicados de forma extensa y serán citados únicamente.

El trabajo constará de dos bloques claros, pero que realmente se han realizado intercaladamente, aunque se explicarán uno tras otro. Estos bloques son los siguientes: Investigación y Aplicación.

## 3.1 INVESTIGACIÓN

Este bloque incluye el proceso de análisis de las alternativas existentes y sus posibles mejoras.

### 3.1.1 MODELOS ACTUALES

El proceso de investigación se inicia a continuación de los conceptos preconcebidos en el Estado del Arte, una vez conocemos que artículos y metodologías ya existentes, empezamos a investigar si estas pueden sernos útiles en nuestro proyecto.

Si es cierto que la información acerca del campo de la armonía de color, es poco abundante y principalmente nos basamos en tres artículos:

1. Color Harmonization de Cohen et al.[3].
2. Saliency-Guided Consistent Color Harmonization de Yoann Baveye et al.[4].
3. No-reference Harmony-guided Quality Assessment et al.[5].

#### 3.1.1.1 COLOR HARMONIZATION Y PERFILES DE TOKOMARU

El artículo que se presenta no solo hace referencia a una posible métrica de la armonía del color en las imágenes, sino que pone hincapié en como corregir el color, pero la segunda parte del artículo no tendrá relevancia para este proyecto.

Debemos partir de que todo el análisis se realiza en el histograma cromático<sup>1</sup> de la imagen, es decir, la imagen se convierte al espacio HSV, donde se realiza un histograma de la componente cromática (Hue).

La metodología usada en Color Harmonization, es la de comparar un conjunto de perfiles que definen armonías concretas con el histograma cromático de la imagen.

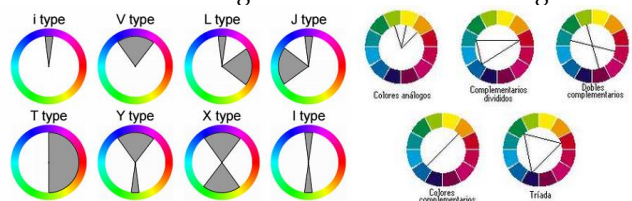


Figura 2 Perfiles de Tokomaru en el espacio unidimensional cromático Vs perfiles usados por los diseñadores en la rueda de Itten.

A continuación, se muestran los perfiles de Tokomaru junto a los perfiles que usan los diseñadores, donde cada uno corresponde a cada una de las armonías que podremos clasificar en un caso la información cromática, en el otro simplemente el color de forma visual.

El proceso de comparación entre el perfil de Tokomaru y el de la imagen, se realiza usando el perfil de Tokomaru como una máscara (Donde la región gris suma). Además es importante destacar que los perfiles son invariantes a

<sup>1</sup> El histograma cromático solo hace referencia a la componente Tono

(Hue) de la imagen y normalmente no se tiene en cuenta la Saturación.

traslación así que se realiza dicha comprobación para todos los ángulos posibles.

Curiosamente no solo se tiene en cuenta la componente hue, sino que también se tiene en cuenta la componente de saturación, dicha justificación viene dada por la pregunta ¿Qué sucede con las imágenes acromáticas? Al coger la componente de Saturación, podemos filtrar y dar menos pesos en la métrica a píxeles que tengan una pureza de color menor.

$$F(X, (m, \alpha)) = \sum_{p \in X} \|H(p) - E_{Tm(\alpha)}(p)\| \cdot S(p) \quad (1)$$

donde  $F(X, (m, \alpha))$ , es el valor de la comparación entre la imagen  $X$  y el perfil de Tokomaru  $m$  rotado  $\alpha$  grados,  $H(p)$  es el valor en el histograma cromático de un pixel  $p$ ,  $E_{Tm(\alpha)}(p)$  es el perfil de Tokomaru  $m$  rotado  $\alpha$  grados en el pixel  $p$ , y  $S(p)$  es el valor en el histograma de saturación en el pixel  $p$ .

La elección final del modelo óptimo viene dada por la búsqueda del ángulo con menor error (Eq. 2) y el modelo que mejor ajusta (Eq. 3).

$$M(X, T_m) = (m, \alpha_0), \alpha_0 = \underset{\alpha}{\operatorname{argmin}} F(X, (m, \alpha)), \quad (2)$$

$$B(X) = (m_0, \alpha_0), m_0 = \underset{m}{\operatorname{argmin}} F(X, M(X, T_m)) \quad (3)$$

### 3.1.1.2 NO-REFERENCE HARMONY-GUIDED QUALITY ASSESSMENT

En esencia, la finalidad es la misma que el Color Harmonization, pero incluyendo continuidad en los píxeles frontera de los perfiles de Tokomaru y también en la métrica que se usa a la hora de comparar los perfiles con el histograma cromático de la imagen.

La respuesta en este caso ya no es binaria, sino que los perfiles son una función de densidad suavizada en los márgenes. Veamos una comparativa:

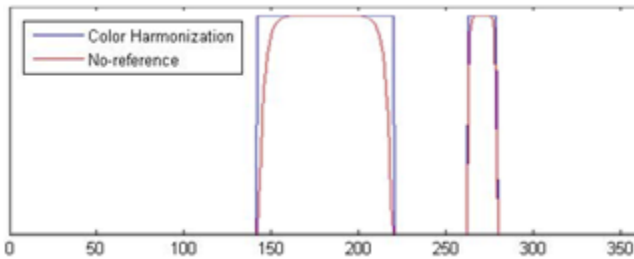


Figura 3 Comparativa perfiles en Color Harmonization y la suavización en No-Reference Harmony-Guided Quality Assessment.

La métrica en este caso es la divergencia KL[6] (Se sustituye por Eq. 1):

$$\hat{\alpha}_m = \underset{\alpha}{\operatorname{argmin}} \sum_h M(h) \cdot \ln \left( \frac{M(h)}{P_m(h - \alpha)} \right) \quad (4)$$

donde  $\hat{\alpha}_m$  es el ángulo óptimo para el perfil  $m$ .

Esta métrica obtiene resultados más precisos pero adolece de problemas en la interpretación de la distribución modelo cuando ésta es nula. Para obtener el mejor perfil y ángulo también hace uso de la Eq.2 y Eq.3.

### 3.1.1.3 DIVERGENCIA KL VS DISTANCIA JS

Estos problemas no tienen en cuenta la información cromática presente en la imagen fuera de los rangos que define cada uno de los perfiles de Tokomaru. Esto es debido a que la divergencia KL, no está definida por todos los valores, sino solo  $P_m(h - \alpha) = 0 \rightarrow M(h) = 0$ , sino el valor  $M(h)$  no se tiene en cuenta.

Para solventar el problema y tener en cuenta la totalidad de los puntos, se usa la distancia Jensen-Shannon[7], basada en la entropía de la información definida como (Sustituye Eq. 4):

$$\hat{\alpha}_m = \sum_h H \left( \frac{M(h) + P_m(h - \alpha)}{2} \right) + \frac{1}{2} \left( \sum_h H(M(h)) + \sum_h H(P_m(h - \alpha)) \right) \quad (5)$$

Para demostrar el erróneo funcionamiento de KL, hemos generado una imagen sintética simple que representa un tipo de armonía (Colores opuestos en el espacio HSV, concretamente el perfil I).

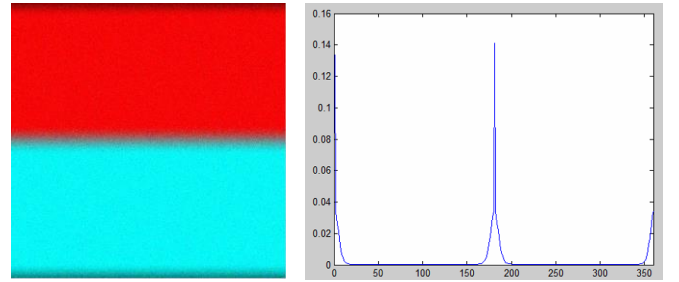


Figura 4 Imagen y su respectivo histograma cromático. Donde se observa dos picos espaciados 180° (Uno en 0° y el otro en el 180°).

Los resultados<sup>2</sup> son los siguientes:

Perfil	KL	JS
X-Type	1.7614	0.5022
I-Type	3.3732	0.1386

Podemos ver que la divergencia KL selecciona erróneamente el perfil X.

En posteriores apartados, podremos verificar el funcionamiento de estas metodologías.

<sup>2</sup> Los resultados completos se muestran en el apéndice (Ap. 1).

### 3.1.2 GENERACIÓN DE MODELOS

Inicialmente se ha evaluado los modelos ya existentes, pero ¿Pero podemos generar otros que sean más efímeros a la realidad? El proceso de investigación no solo se limitaba a usar los modelos ya existentes, sino que a su vez queríamos buscar otros que quizás explicaban mejor las armonías presentes en las imágenes estéticas.

Es por eso, que se usaron técnicas de clustering no-supervisadas para observar qué patrones tenían las imágenes y si podríamos obtener perfiles que nos ayudasen en un futuro a clasificar.

Cabe decir que para realizar dicho análisis, se ha usado una base de datos AVA basada en imágenes consideradas estéticas[8] e imágenes de DPChallenge. De estas se ha seleccionado un subconjunto filtrando por diversos factores como la puntuación, si la imagen es armónicamente variada y con suficiente información cromática<sup>3</sup>. Quedándonos finalmente con 30.836 imágenes con las que realizaremos el proceso de aprendizaje.

Para realizar todas las comparaciones entre imágenes, necesitábamos gran capacidad de cálculo, así que se usó el clúster del CVC<sup>4</sup>. Dicho uso queda justificado por el hecho de tener que calcular  $\frac{n \cdot (n-1)}{2}$ , distancias. Para  $n = 30.861$  son 475.414.030 cálculos.

Para ello se usó la similitud del coseno normalizada, creando la matriz de distancias para extraer las imágenes más parecidas entre ellas.

La distancia vendrá definida por la siguiente ecuación:

$$CS(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}, \quad (6)$$

$$xcorr(A, B) = 1 - \max_{\alpha} CS(A, B(-\alpha)), \quad (7)$$

$$M(i, j) = xcorr(H(i), H(j)) \forall i, j < \#Imgs \quad (8)$$

donde  $xcorr(A, B)$  será mejor distancia CS entre un histograma cromático A y otro B desplazado circularmente.  $M(i, j)$  corresponde a la celda  $i, j$ -ésima de la matriz de distancia y  $\#Imgs$  es el número total de imágenes.

Una vez procesados los datos, podemos realizar el proceso de clustering. A lo largo de este proyecto se han analizado dos metodologías para clusterizar las imágenes:

- Dendogramas.
- K-means.

#### 3.1.2.1 DENDOGRAMA

El dendograma[9], es un método cuya entrada es la matriz de distancias entre los puntos. Éste se basa en realizar uniones binarias, jerarquizando de modo que inicialmente se unen imágenes (hojas) y posteriormente se unen conjuntos de imágenes con distancias cercanas entre sí.

El proceso de clustering, se llama “linkage” y el resultado final es un árbol con un nodo raíz nulo (Incluye todas las imágenes). Una vez generado el árbol, este es podado a partir de una profundidad seleccionando los clústeres que necesitemos.

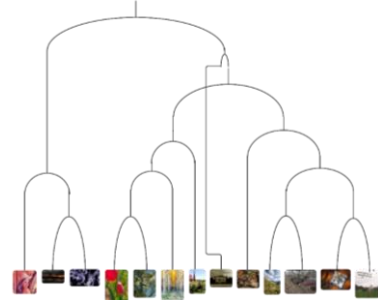


Figura 5 Ejemplo sintético de un dendograma

El clustering permite saber a qué conjunto pertenece cada imagen, ahora debemos obtener una representación de dicho conjunto.

La representación de cada clúster, se realiza mediante acumulación, donde todos los histogramas cromáticos se acumulan, estos además son desplazados para que la unión sea lo más correcta posible.

Sea  $\tilde{H}_i(j)$  el histograma resultado de acumular los  $j$  primeros histogramas correspondientes al cluster  $i$  (el orden es aleatorio). Buscamos encontrar  $\tilde{H}_i(n_i) \forall i$  donde  $n_i$  es el número de imágenes en el clúster  $i$ , definiéndolo de la siguiente manera:

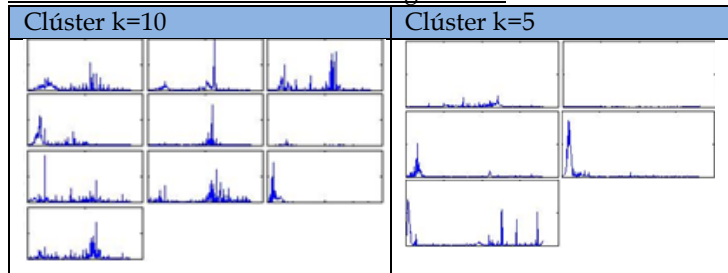
$$\tilde{H}_i(1) = H_{i,1}, \quad (9)$$

$$\tilde{H}_i(j) = \frac{j-1}{j} \tilde{H}_i(j-1) + \frac{1}{j} \text{Desp}(H_{i,j}, \tilde{H}_i(j-1)), \quad (10)$$

$$\text{Desp}(A, B) = \begin{cases} A(-\alpha_0), \\ \alpha_0 = \underset{\alpha}{\operatorname{argmax}} CS(A, B(-\alpha)) \end{cases} \quad (11)$$

donde  $H_{i,j}$  es el histograma de la  $j$ -ésima imagen perteneciente al clúster  $i$ .  $\text{Desp}(A, B)$  busca la mejor correlación entre los histogramas A y B.

Los resultados<sup>5</sup> obtenidos son los siguientes:



<sup>3</sup> La selección de imágenes se explica en el apéndice (Ap. 2).

<sup>4</sup> El proceso de paralelización se explica en el Informe Progreso I (Ap. 3.3)

<sup>5</sup> Por el reducido espacio en el que presentar los resultados, se ha obviado la importancia del histograma (Medida aproximada que indica la

importancia que tiene respecto al conjunto un clúster). Presente en el apéndice (Ap. 3).



No se observa en ninguno de ellos una tendencia concreta y definida específicamente como sucede en los perfiles de Tokomaru. Idealmente el resultado deberían ser perfiles cuasi-disjuntos como ocurre en el casos de perfiles de Tokomaru

En el clúster con  $k=10$ , algunas armonías son parecidas a los perfiles de Tokomaru, pero realmente el ruido entorpece la evaluación de los propios resultados. Y otras en cambio pueden llevar a deducir múltiples perfiles, debemos recordar que los perfiles armónicos deben ser simples.

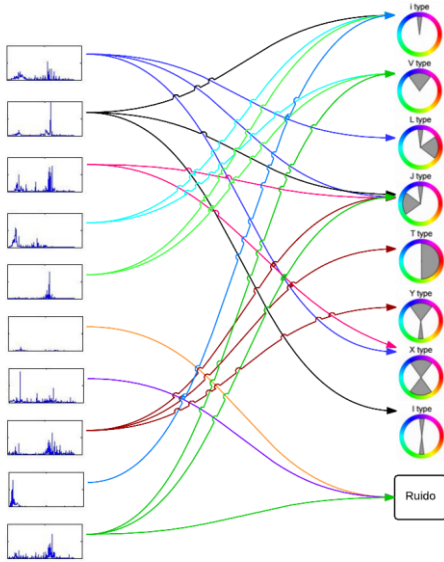


Figura 6 Comparativa clústeres obtenidos con  $n=10$ , respecto los perfiles de Tokomaru

La forma de interpretar los resultados, sería tener en cuenta que los perfiles de Tokomaru son realmente histogramas cromáticos proyectados en coordenadas polares, pero realmente pertenecen a un rango  $[0, 360)$ . Teniendo en cuenta esta premisa podemos realizar ciertas comparaciones con los perfiles de Tokomaru. Es muy difícil deducir si realmente los clúster obtenidos podrían ser un perfil armónico específico, o si el ruido los ha modificado considerablemente a lo largo del proceso. Este fenómeno se puede producir a causa del ruido acumulativo que se produce a cada iteración de la creación del histograma del clúster.

En el caso del clúster de  $n=5$ , podemos ver a simple vista como dos histogramas casi no aportan ninguna información sobre como son los datos y son los que más importancia tienen.

Por lo tanto no se considera una buena metodología para generar modelos, aunque se desconoce si realizando algún proceso de suavización y evitando el ruido acumulado a lo largo del proceso, se podrían obtener mejores resultados.

### 3.1.2.1 K-MEANS

Otra metodología que se probó para obtener modelos armónicos de los histogramas cromáticos, es usar el algoritmo de K-Means [10].

El problema se debe a que K-Means usa para realizar el clustering los vectores asociados a cada muestra, pero aquí no tenemos vector alguno, sino solamente las distancias comprendida entre todos los histogramas cromáticos (Matriz de distancias).

Para solventar este problema, usaremos el Multidimensional Scaling[11], una técnica que permite obtener vectores aleatorios comprendidos en un espacio vectorial de  $N$ -dimensiones que cumplan las restricciones de la matriz de distancias.

MATLAB incorpora en su librería, un algoritmo de M.S., y además devuelve el esfuerzo a cada dimensión (Análogo a la importancia que tiene cada dimensión respecto los datos) y éstas están ordenadas por importancia.

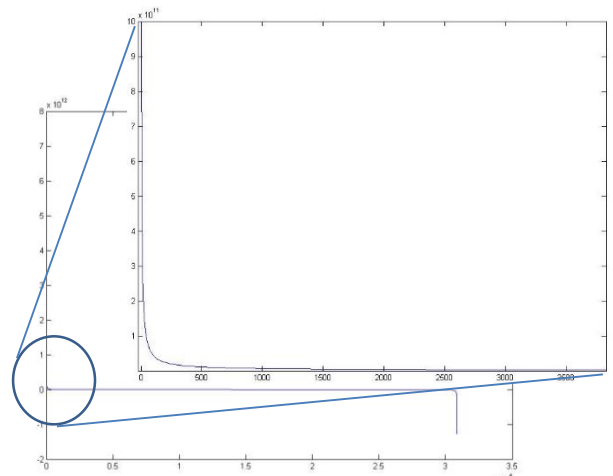


Figura 7 Importancia de cada dimensión (35000 dimensiones en total). Haciendo zoom a las 3500 primeras dimensiones.

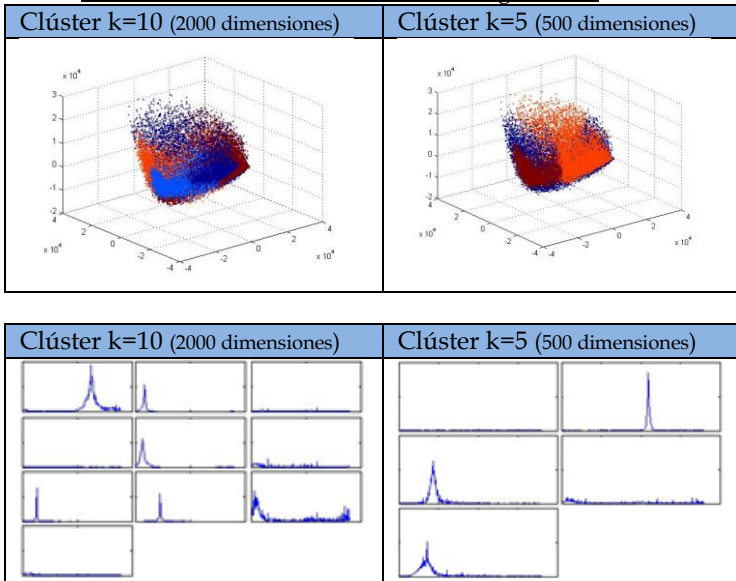
Como podemos observar en el grafico solamente las prácticamente 1500 dimensiones contienen toda la información (71%). Las demás dimensiones aportan muy poca información y complican considerablemente el proceso de clustering por lo que se descartan.

Una vez determinado el número de dimensiones, se aplica el algoritmo K-means para obtener los clústeres. Para reducir el ruido de los modelos resultantes se optó por seleccionar la distribución más cercana al centro como representante del clúster en lugar de usar el promedio.

$$H_{c_i} = H_{\arg \min d(\bar{p}_i, \bar{p}_j)}, \quad j \in c_i \quad (12)$$

donde,  $H_{c_i}$  es el histograma del cluster  $i$  y viene determinado por el histograma que esta mas cerca al cluster (Y pertenece a él) en el espacio vectorial generado.

Los resultados<sup>6</sup> en este caso son los siguientes:



Análogamente a que sucedía en los histogramas cromáticos de la metodología basada en dendogramas, los resultados no aportan mucha información.

Prácticamente solo se observa un tipo de armonía en los histogramas (armonía análoga o lo que sería equivalente a los perfiles i y X de Tokomaru) y otros muchos de ellos solamente son histogramas planos uniformes con ruido en algunos valores cromáticos (No permiten deducir armonía alguna). Es por ello que con estos resultados, junto a los anteriores no podemos concluir nada en claro.

### 3.2 APLICACIÓN

En el proceso de aplicación, se ha buscado y propuesto aplicaciones interesantes que podrían trabajar a partir de la información cromática desde el punto de vista de la armonía (Usando los conceptos anteriores).

Concretamente, se ha pretendido realizar cuatro aplicaciones distintas buscando una funcionalidad interesante y motivadora. Inicialmente todas ellas se han planteado en un entorno de programación MATLAB para simplificar su desarrollo y algunas de ellas una vez verificada su funcionalidad se han querido trasladar a Android.

Las aplicaciones<sup>7</sup> realizadas son las siguientes:

- Cualificador de belleza mediante los perfiles de Tokomaru.
- Armonía de color y emociones.
- Corrector de color basado en armonía.
- Comparador de imágenes basado en armonía.

#### 3.2.1 CUALIFICADOR DE ARMÓNIA CROMÁTICA BASADO EN PERFILES DE TOKOMARU

Este aplicativo es realmente sencillo, distinguimos 8 perfiles armónicos como anteriormente hemos explicado.

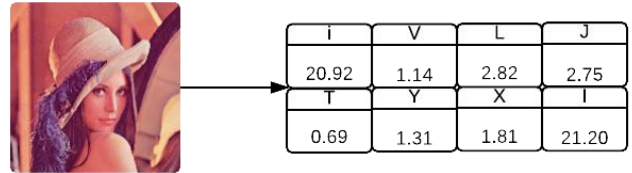


Figura 9 Valores por cada perfil de Tokomaru usando las métricas de No-reference Harmony-Guided Quality Assessment y la divergencia KL.

Es por ello, que queremos usar la información obtenida en los perfiles de Tokomaru para indicar cuanto de buena es una imagen.

Una vez obtenido tanto el mejor perfil armónico como su valor, se aplica una función sigmoide que indica usando un valor de  $[1, \dots, 0]$  la calidad de la imagen.

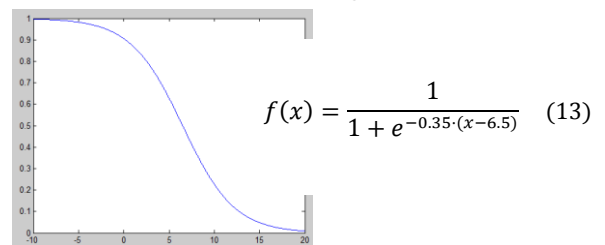
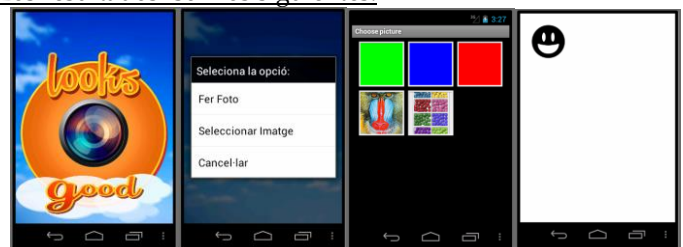


Figura 8 Función sigmoide usada para extraer un valor de  $[0, \dots, 1]$  que indicará la calidad de la imagen.

Debido a limitaciones de tiempo en la planificación, no se ha podido profundizar mucho en el uso de las sigmoides ni en los resultados obtenidos. Pero hubiese sido muy interesante realizar:

- Una sigmoide por cada tipo de armonía.
- Realizar un proceso de aprendizaje que determinase los mejores parámetros de cada sigmoide.
- No usar solamente la mejor armonía, sino hacer uso de todas en global para cualificar la imagen.

Los resultados<sup>8</sup> son los siguientes:



<sup>8</sup> Se muestra tanto el resultado como pipeline del aplicativo. Hubiese sido interesante poder realizar un análisis de los resultados obtenidos haciendo uso de diferentes metodologías para obtener la calidad de la imagen. Se usa el "emoticono" para representar la calidad de la imagen.

<sup>6</sup> En las imágenes de clustering solo se muestran las 2 dimensiones más significativas. Los resultados se pueden apreciar mejor en el apéndice (Ap. 4).

<sup>7</sup> En todas las aplicaciones se mostrará algún resultado, pero los restantes resultados se podrán ver en el apéndice.

### 3.2.2 ARMONÍA DE COLOR Y EMOCIONES

Las imágenes transmiten emociones, pero estas están compuestas por diversa información. ¿Tenemos suficiente con la información cromática para realizar un clasificador sencillo? En la respuesta a esta pregunta, se desarrolla todo este apartado, donde se intentará buscar y clarificar si la información únicamente cromática es capaz por si sola de determinar las emociones que surgen en el observador.

En una primera fase de documentación para llevar a cabo el experimento, nos encontramos con la “rueda de las emociones” de Robert Plutchik[12], esta representa 8 emociones básicas y otras 8 compuestas por la básicas que reciben el nombre de avanzadas.

La “rueda de las emociones” tiene un parentesco muy similar a los espacios de color, como pueden ser HSV u Oponente donde la distribución se puede considerar parecida.

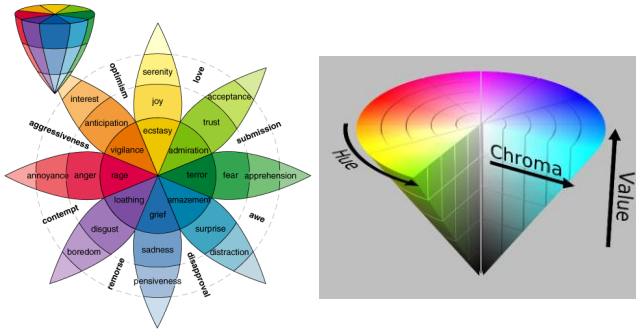


Figura 10 Comparativa de la rueda de las emociones con el espacio de color HSV.

Observando el gran parecido entre ambos, surgió la idea de que quizás podríamos ligar nuestros conceptos sobre la armonía de color con las emociones.

El algoritmo de armonía de color (Se basa en Eq.2, Eq.3 y Eq.5 para obtener los perfiles) devuelve 8 perfiles armónicos definidos por Tokomaru y cada una dispone de un peso ( $\overline{w}_i$ ) y ángulo ( $\overline{\alpha}_i$ ) asociado.

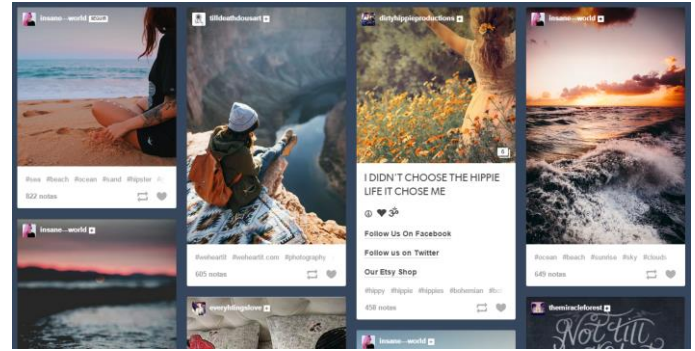
$$\overline{w}_i, \overline{\alpha}_i = \text{obtenerArmoniasJS}(hHue_i(\alpha)) \quad (14)$$

donde  $hHue_i(\alpha)$  es el histograma de cromático de la imagen  $I$ . En este proceso de aprendizaje, era necesario usar una Base de Datos, donde fuésemos capaces de aprender y en el transcurso de esa búsqueda es cuando vimos la inviabilidad de este aplicativo usando únicamente información cromática.

La base de datos que se uso fue “Tumblr” que nos permite buscar imágenes usando como tags las emociones que queríamos observar, el problema no era solo que muchas de ellas no corresponden realmente al tag que busca (ruido) sino que muchas de ellas son multinterpretables.

Estas son algunas de las búsquedas<sup>9</sup> que se realizaron:

Tag: Happiness (Alegria)



Realizando un análisis totalmente visual y basándonos en las emociones que nos transmiten a nosotros, como observador, podemos deducir lo siguiente:

- Prácticamente todas las imágenes tienen colores muy distintos y a primera apreciación no observamos ningún color primario (Como debería determinar la “rueda de las emociones”) o patrón armónico que permita relacionarlo con la emoción de “Alegria”.
- Muchas de ellas no dependen de la información cromática, ni tampoco de la topología (Aunque está claramente aporta más información), sino más bien de contexto en que son tomadas por el creador y percibidas por el observador. Véase un ejemplo, si observamos las dos primeras imágenes, el autor ha taggeado “Alegria”, porque quizás a nivel contextual pueden transmitir dicha emoción, pero obviamente esto no lo podemos obtener únicamente por la información implícita que hay en la imagen.

En conclusión, un análisis armónico y de información cromática no tiene suficiente peso para determinar las emociones que puede generar una imagen, y probablemente la información topología tampoco si no está unida a elementos contextuales que no están presentes en la información que la imagen ofrece, si no que es explícita a la misma.

### 3.2.3 CORRECTOR BASADO EN ARMONÍA

¿El posible usar la información armónica únicamente para corregir imágenes mejorando su estética? Se trata de usar la información cromática como fuente unívoca y mediante

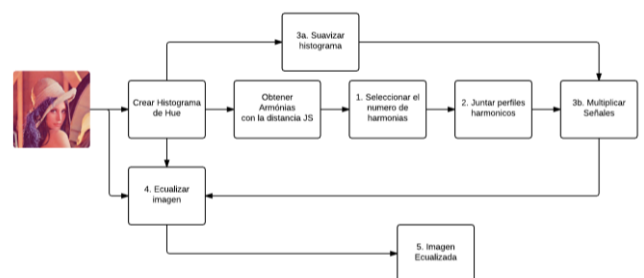


Figura 11 Pipeline de etapas en la corrección de una imagen. ella corregir algunos colores presentes en la imagen.

<sup>9</sup> Se incluye otro resultado en el apéndice (Ap. 5).



Inicialmente se crea el histograma cromático y se obtendrán las armonías usando la Eq.14.

### 1. Seleccionar el número de armonías:

#### a) Método A:

1. Primero de todo, invertiremos los valores, donde el 1 será el valor de dos histogramas iguales, usando Eq.14. Y se proyectarán a un kernel exponencial para separar mejor los datos.

$$\overline{vw}e_i = e^{k \cdot (1 - \overline{vw}e_i)}, \forall i, 0 < i < \# \overline{vw} \quad (15)$$

donde  $\overline{vw}e$  es el vector de pesos proyectado.

2. Se seleccionan las armonías con la media y la distribución estándar en un proceso iterativo.

$$selec_i = \left( \overline{vw}e_i - \bar{x}(\overline{vw}e_i) > \frac{\alpha}{j} \cdot \sigma(\overline{vw}e_i) + \beta \right), \quad \forall i, 0 < i < \# \overline{vw} \quad (16)$$

La Eq.16 se desarrolla iterativamente incrementando  $j$ , mientras el conjunto  $selec$  sea nulo (no se ha seleccionado ninguna aun) forzando a seleccionar un mínimo.  $\alpha$  y  $\beta$  son valores que se han elegido mediante el prueba y error ( $\alpha = \frac{3}{2}, \beta = 0$ ).

#### b) Método B:

Otro método se basa en fijar un umbral de aceptación respecto al mejor valor.

$$umbral = \alpha \cdot \max_i \overline{vw}e_i, \alpha \in [0.7, 0.95] \quad (17)$$

$$selec_i = \overline{vw}e_i \geq umbral$$

### 2. Juntar perfiles armónicos:

Los perfiles seleccionados anteriormente deben ser juntados en un único que usaremos para corregir la imagen.

$$\overline{vwn}_i = \frac{1 - \overline{vw}_i}{\sum_{j, selec_j=1} (1 - \overline{vw}_j)} \quad (18)$$

$$hHarmony(\alpha) = \sum_{i, selec_i=1} \left( \overline{vwn}_i \cdot \frac{hPerfil_i(\alpha - \overline{v\vec{a}}_i)}{\max(hPerfil_i)} \right) \quad (19)$$

La realizar la mezcla usaremos la ponderación de su peso original ( $\overline{vw}_i$ ) y el ángulo ( $\overline{v\vec{a}}_i$ ) para correlacionarlo correctamente entre ellos.

### 3. Obtener el histograma armónico:

Debido a la sensibilidad a la hora de modificar el "histograma cromático de la imagen", debemos realizar ciertas modificaciones al  $hHarmony(\alpha)$ , en esta etapa lo adaptaremos más a la imagen original para que las modificaciones no sean tan bruscas.

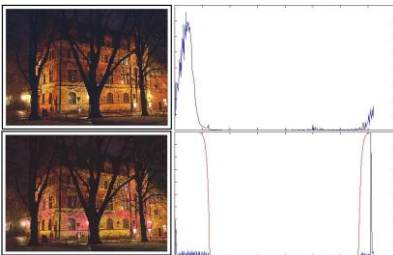


Figura 12 Imagen original y resultado sin aplicar un suavizado en el  $hHarmony$ .

Si se usa directamente el  $hHarmony(\alpha)$  en etapas posteriores, los cambios que se producen en la imagen son muy bruscos y se fuerza a la imagen a rellenar regiones con colores inexistentes, creando situaciones extrañas como los colores rosados de la Figura.12.

Principalmente se aplican dos pasos:

- a) **Suavización:** Se suaviza el histograma cromático de la imagen usando una gaussiana.  $hHue(\alpha)$  es el histograma cromático de la imagen original.

$$hSuav(\alpha) = \int_{-\frac{\mu}{2}}^{\frac{\mu}{2}} hHue(\alpha - x) \cdot \Phi_{\mu, \sigma^2}(x) dx \quad (19)$$

- b) **Producto de señales:** Se multiplica con el histograma suavizado con el obtenido en la anterior etapa ( $hHarmony(\alpha)$ ).

$$hHarmonyFix(\alpha) = hSuav(\alpha) \cdot hHarmony(\alpha) \quad (20)$$

### 4. Obtener transformación de ecualización

Se busca una función de LUT que permite transformar el histograma cromático de la imagen al histograma  $hHarmonyFix(\alpha)$ .

$$LUT = hEq(hHue, hHarmonyFix) \quad (21)$$

$$hHue_i(\alpha) \xrightarrow{LUT(i)} hHarmonyFix_i(\alpha) \quad (22)$$

### 5. Aplicar ecualización a la imagen

Usando la LUT obtenida en la anterior etapa, esta se usará para corregir solamente el canal de Hue de la imagen original.

$$hsv = rgb2hsv(I)$$

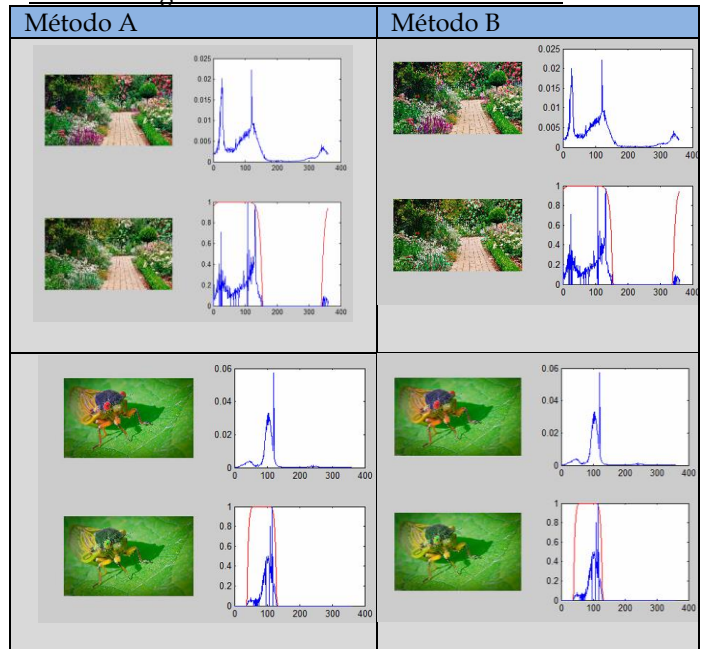
$$hue = canal(hsv, 1)$$

$$hueHarmony = LUT(hue(i, j)), \forall i \in [0, anchura(I)], \forall j \in [0, altura(I)]$$

$$hsvHarmony = mezclarCanales(hueHarmony, canal(hsv, 2...3))$$

$$IHarmony = hsv2rgb(I) \quad (23)$$

Estos son algunos de los resultados<sup>10</sup> obtenidos:



<sup>10</sup> Se muestran más resultados y una explicación adicional en el apéndice

(Ap. 6 y Ap. 11).



Los resultados que obtenemos tienden a ser antinaturales, posiblemente se mejore el histograma, pero se fuerza a que algunos píxeles modifiquen bruscamente su valor de hue, generando valores inesperados. Ejemplos:

- Flores del jardín obtienen un color inesperado (Imagen fila nº1, columna nº1).
- Ambos insecto pierden su color (Imágenes fila nº 1, columna nº 3, y fila nº 2, columna nº 2).

### 3.2.4 COMPARADOR DE IMÁGENES BASADO EN ARMONÍA

Se trata de comparar haciendo uso únicamente de la información cromática de cada imagen. Esta información es obtenida mediante Eq.14, inicialmente se hizo uso de la divergencia KL, pero demostrando la mejoría de la distancia JS se optó también por usarla.

Inicialmente este aplicativo se probó para comprobar el funcionamiento del “Color Harmonization”, pero se consideró interesante y se siguió desarrollando trayendo consigo las mejoras.

Para llevar a cabo el experimento, se seleccionaron 100 imágenes de consulta que se contrastarán contra otras 3000. Es decir, comprobaremos la afinidad cromática de 100 imágenes. Esto generará una matriz de 100x3000. Después seleccionaremos únicamente las 10 más parecidas.

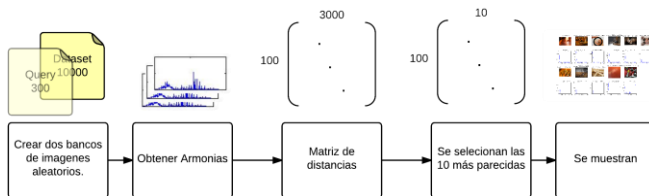


Figura 13 Pipeline del funcionamiento del comparador armónico.

#### 1. Se crean dos bancos de imágenes aleatorios:

Se crean dos bancos de imágenes aleatorios, donde ambos cumplen ciertas restricciones cromáticas (Para evitar imágenes con poca variedad cromática o imágenes puramente acromáticas). El banco “Query” serán las imágenes que buscaremos 10 imágenes afines, el banco “Dataset” serán donde buscaremos dichas imágenes similares. Los bancos un total de imágenes de 300 y 10000 respectivamente.

#### 2. Obtenemos las Armonías:

Aplicamos la Eq.14, para cada una de las imágenes de cada conjunto obteniendo para cada una: un vector de 8 posiciones que describe el peso en cada perfil  $\overline{vw}$  y otro vector de 8 posiciones con su ángulo  $\overline{v\alpha}$ .

#### 3. Matriz de distancias:

Para calcular la matriz de distancias, se debe calcular la distancia entre un histograma cromático y otro usando  $\overline{vw}$ . Se debe tener nuevamente un criterio de selección para que perfiles no importantes no añadan ruido innecesariamente. El criterio que se usó fue seleccionar para cada histograma los N perfiles más importantes. Los perfiles restantes se

anulan.

$$\overline{vws}_i = \begin{cases} \overline{vw}_i, & \text{si } \overline{vw}_i \text{ está en los } N \text{ más importantes} \\ 0, & \text{en otro caso} \end{cases} \quad (24)$$

Dependiendo del uso que queramos darle, se han realizado distintas distancias, sin tener en cuenta el ángulo o teniéndolo en cuenta.

#### a) Sin ángulo:

Únicamente la distribución cromática es comparada independientemente del ángulo en que se encuentre, es decir es invariante a traslación.



Figura 14 Ambas imágenes tienen las mismas relaciones armónicas pero distinta traslación. Si no tenemos en cuenta el ángulo ambas imágenes son equivalentes.

De forma que la distancia queda definida como:

$$\text{dist}(a, b) = \sqrt{\sum_i (\overline{vw}(a)_i - \overline{vw}(b)_i)^2} \quad (25)$$

donde  $\overline{vw}(a)$  es el vector de pesos de a y  $\overline{vw}(b)$  el de b.

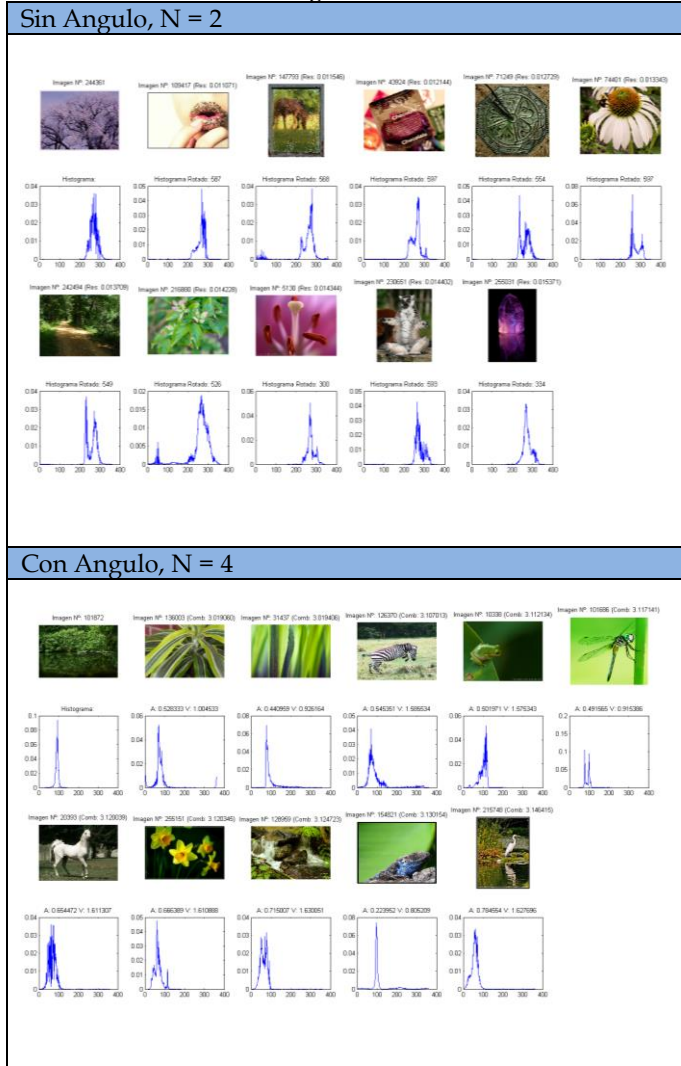
#### b) Con ángulo:

Si tenemos en cuenta el ángulo, la comparativa de la Fig.14 queda anulada y no tendrán parentesco, el ángulo nos fuerza a tener en cuenta el color de la imagen y no únicamente las relaciones cromáticas.

Para tener en cuenta el ángulo se genera una distancia que combina ambas informaciones.

$$\begin{aligned} \overline{vW}(a, b) &= \overline{vw}(a) - \overline{vw}(b) \\ \overline{vA}(a, b) &= \min \left( \left| \overline{v\alpha}(a) - \overline{v\alpha}(b) \right|, \left| \overline{v\alpha}(a) - (360 - \overline{v\alpha}(b)) \right| \right) \\ \text{dist}(a, b) &= \sqrt{\sum_i \left( \left( 1 + \overline{vW}(a, b)_i \right) \cdot \left( 1 + \frac{\overline{vA}(a, b)_i}{180} \right) \right)^2} \quad (26) \end{aligned}$$

Los resultados son los siguientes:



Los resultados en ambos casos son francamente satisfactorios, solamente usando la información cromática de la imagen y concretamente el tono. En otros ejemplos<sup>11</sup> se ven posibles errores debido a solo coger la componente hue, probablemente el algoritmo funcionaría mejor cogiendo más información sobre el color de la imagen (Haciendo uso de la componente Saturación, por ejemplo).

## 4 CONCLUSIONES

La conclusión general que obtenemos después del transcurso es que la información cromática es una información muy limitada, y que si más bien es cierto que nos da alguna información, no podemos basarnos solamente en ella para poder realizar aplicaciones con cierta complejidad sino que deberíamos usarla únicamente como soporte o complementaria para obtener mejores resultados.

Las relaciones armónicas son muy interesantes en ciertos entornos donde el color varía (cambios de luz por ejemplo), sería interesante usar estas técnicas para añadir información adicional y mejorar clasificadores que no deben ser

sensibles al color, con entornos que sí lo son.

Y finalmente que, aún obteniendo poca información con la información cromática, está normalmente presenta mucho ruido.

## 5 AGRADECIMIENTOS

Agradecer enormemente el apoyo a lo largo de todo el trabajo a Ramon Baldrich. He recibido un soporte ejemplar, tanto acerca de qué metodologías serían interesantes usar, problemas encontrados (Que han sido muchísimos), artículos interesantes y como gran motivador a la hora de obtener resultados no esperados en el trabajo. Considero que aunque no haya hecho un gran avance en la “Armonización del Color”, si he aprendido técnicas novedosas e interesantes para aplicar, herramientas para resolver problemas y mejorar mi capacidad de presentar artículos técnicos.

## 6 BIBLIOGRAFÍA

- [1] Johannes Itten. (2015, May 4). In Wikipedia, The Free Encyclopedia. Retrieved 18:53, Mayo 19, 2015, from [https://en.wikipedia.org/w/index.php?title=Johannes\\_Itten&oldid=660760362](https://en.wikipedia.org/w/index.php?title=Johannes_Itten&oldid=660760362)
- [2] Modelo de color HSV. (2014, 16 de agosto). Wikipedia, La enciclopedia libre. Fecha de consulta: 11:31, febrero 26, 2015 desde [http://es.wikipedia.org/w/index.php?title=Modelo\\_de\\_color\\_HSV&oldid=76398012](http://es.wikipedia.org/w/index.php?title=Modelo_de_color_HSV&oldid=76398012).
- [3] Cohen-Or, D et al. (2006, July). Color harmonization. In ACM Transactions on Graphics (TOG) (Vol. 25, No. 3, pp. 624-630). ACM.
- [4] Baveye, Y., Urbanet at. (2013). Saliency-guided consistent color harmonization. In Computational Color Imaging (pp. 105-118). Springer Berlin Heidelberg.
- [5] Chamaret, C., & Urban, F. (2013, June). No-reference harmony-guided quality assessment. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on (pp. 961-967). IEEE.
- [6] Kullback–Leibler divergence. (2015, April 19). In Wikipedia, The Free Encyclopedia. Retrieved 10:06, April 23, 2015, from [http://en.wikipedia.org/w/index.php?title=Kullback%E2%80%93Leibler\\_divergence&oldid=657166704](http://en.wikipedia.org/w/index.php?title=Kullback%E2%80%93Leibler_divergence&oldid=657166704)
- [7] Jensen–Shannon divergence. (2015, April 17). In Wikipedia, The Free Encyclopedia. Retrieved 10:35, May 30, 2015, from [http://en.wikipedia.org/w/index.php?title=Jensen%E2%80%93Shannon\\_divergence&oldid=656862833](http://en.wikipedia.org/w/index.php?title=Jensen%E2%80%93Shannon_divergence&oldid=656862833)
- [8] Murray, N. et al. AVA: A large-scale database for aesthetic visual analysis. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on (pp. 2408-2415). IEEE.
- [9] Hierarchical clustering. (2015, April 23). In Wikipedia, The Free Encyclopedia. Retrieved 10:37, April 23, 2015, from [http://en.wikipedia.org/w/index.php?title=Hierarchical\\_clustering&oldid=658784565](http://en.wikipedia.org/w/index.php?title=Hierarchical_clustering&oldid=658784565)
- [10] K-means clustering. (2015, June 19). In Wikipedia, The Free Encyclopedia. Retrieved 19:00, June 19, 2015, from [https://en.wikipedia.org/w/index.php?title=K-means\\_clustering&oldid=667635397](https://en.wikipedia.org/w/index.php?title=K-means_clustering&oldid=667635397)
- [11] Multidimensional scaling. (2015, May 17). In Wikipedia, The Free Encyclopedia. Retrieved 19:01, June 19, 2015, from [https://en.wikipedia.org/w/index.php?title=Multidimensional\\_scaling&oldid=662818125](https://en.wikipedia.org/w/index.php?title=Multidimensional_scaling&oldid=662818125)
- [12] Contrasting and categorization of emotions. (2015, May 10). In Wikipedia, The Free Encyclopedia. Retrieved 19:02, June 19, 2015, from [https://en.wikipedia.org/w/index.php?title=Contrasting\\_and\\_categorization\\_of\\_emotions&oldid=661730332](https://en.wikipedia.org/w/index.php?title=Contrasting_and_categorization_of_emotions&oldid=661730332)

<sup>11</sup> Se muestran en el apéndice (Ap. 7).

## APÉNDICE

En el apéndice se mostrarán tanto resultados, explicaciones y problemas citados en el propio artículo, como otros que no aparecen implícitamente.

### A.1 RESULTADOS EXPANDIDOS

#### 1. Resultados completos divergencia KL vs distancia JS:

Perfil	KL	JS
i-Type	19.1349	0.4058
V-Type	18.1681	0.6344
L-Type	18.1953	0.6423
J-Type	18.1950	0.6421
T-Type	18.4002	0.7290
Y-Type	2.6532	0.4335
X-Type	1.7614	0.5022
i-Type	3.3732	0.1386

#### 2. Selección de imágenes:

La selección de imágenes se realiza mediante unos prerrequisitos que se deben cumplir, estos aseguran variedad cromática, tags que no nos interesan, challenge que no nos interesen y una puntuación mínima que nos pueda asegurar que la imagen es estéticamente buena.

- No se tienen en cuenta imágenes que tengan challenge con el id: 910, 191, 529, 823, 1201, 51, 614, 1253, 108, 1192, 18, 278, 583, 1368, 924, 1137, 1246 o tags con el id 21.
- La puntuación mínima es de: 6.
- Se usan la siguiente restricción cromática en cada imagen:

$$F = \begin{cases} \text{canales}(I) \geq 3 \\ \|S(p) > 0.1 \wedge V(p) > 0.25\| > 0.25 \cdot \|I\|, p \in I \end{cases}$$

donde  $I$  es la imagen y  $\|I\|$  el número total de píxeles de la imagen.

#### 3. Ponderación de importancia por clúster:

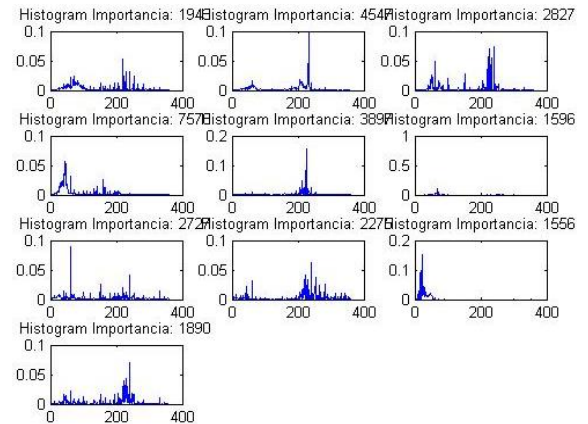
Aunque esta no se tuvo en cuenta a la hora de presentar los resultados debido a que los resultados no eran buenos y por consecuencia no hacía falta saber que clústeres tenían más relevancia que otros. Por otro lado, sí que se desarrolló una métrica que permitiese obtener un valor compacto. Esta métrica hace uso de todos los histogramas que caen dentro del clúster y a su vez la distancia mínima que hay entre histogramas.

$$Imp_i = \frac{(1 + \max_j w_j - w_i) \cdot \|c_i\|}{\sum_k (1 + \max_j w_j - w_k) \cdot \|c_k\|}$$

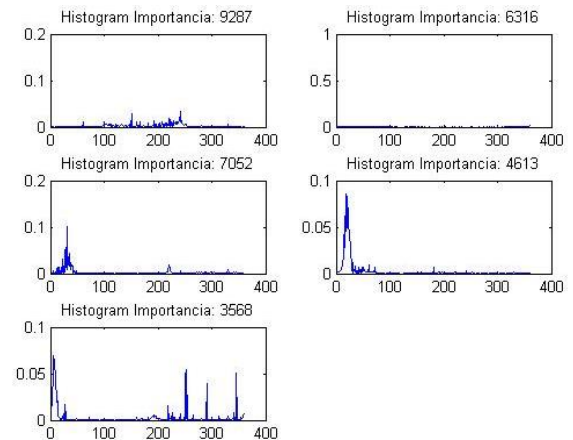
donde  $w_i$  es el peso del cluster  $i$ -ésimo del dendrograma, este es un peso que genera el método de clustering para evaluar la distancia intracluster.

#### 4. Resultados completos clustering K-means:

##### Clúster de 10

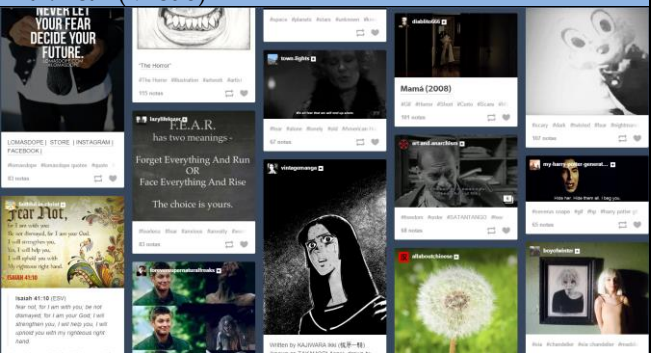


##### Clúster de 5



#### 5. Otro resultado de la búsqueda semántica de emociones:

##### Tal: Fear (Miedo)



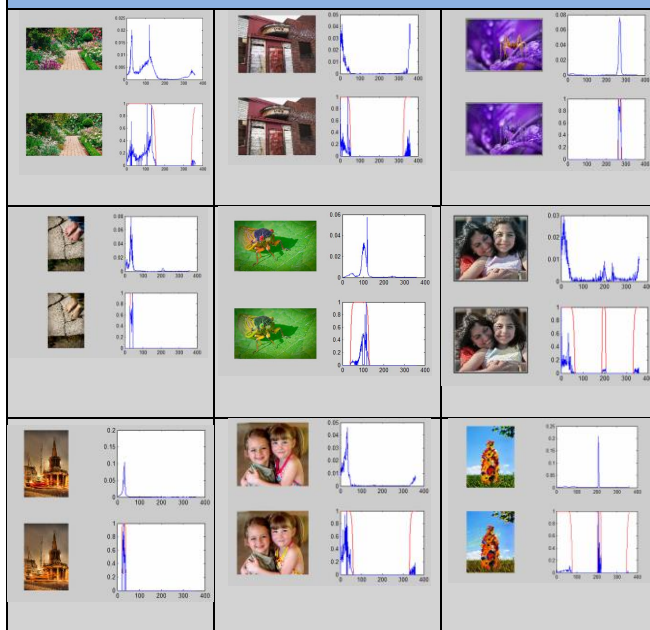
Si aplicamos la misma metodología para analizar estas imágenes que en el tan anterior, encontramos respuestas parecidas:

- Muchas de las imágenes taggeadas con "Miedo" son acromáticas.
- Algunas de ellas también usan elementos contextuales como podemos ver en la esquina superior-derecha. El uso de elementos infantiles, en entornos oscuros para transmitir miedo, pero este concepto es puramente contextual y cultural, y esa información no se encuentra en la imagen.
- Otras en cambio, aumentan el contraste y usan contextos más caóticos.

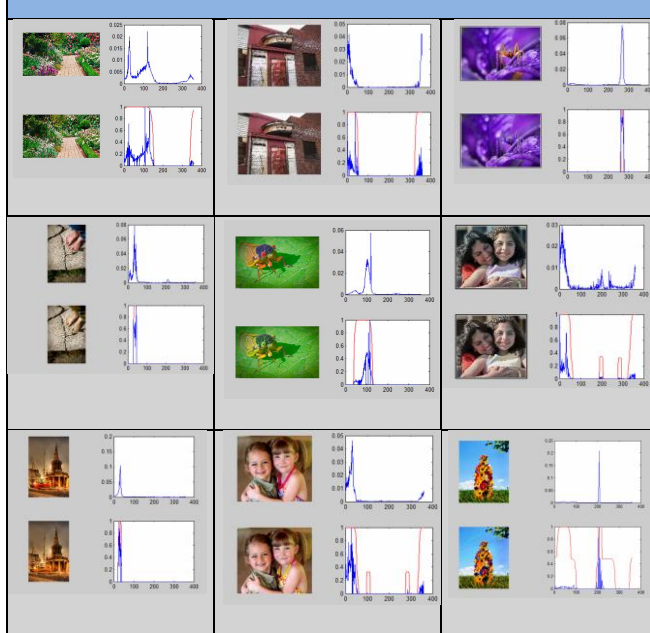


## 6. Resultados completos del corrector basado en armonía:

### Resultados método selección A:



### Resultados método selección B:

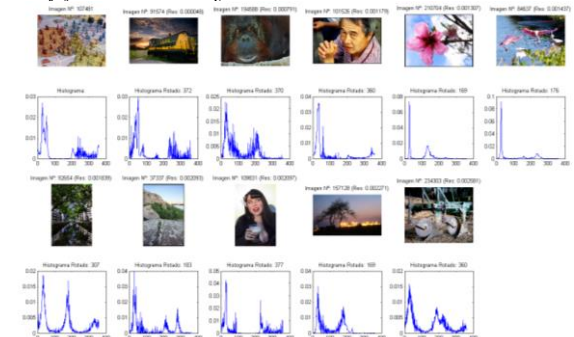


## 7. Resultados completos del comparador de imágenes:

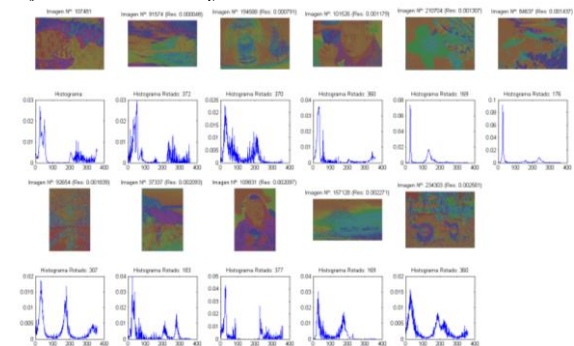
A continuación se mostrarán los resultados obtenidos, también se ha fijado en el caso “sin ángulo”, la saturación y el valor para que se pueda observar la similitud armónica.

### Sin ángulo y 1 perfil seleccionado

*Sin fijar saturación y valor*

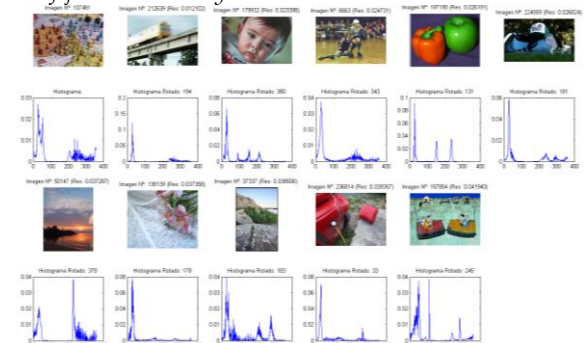


### Fijando saturación y valor

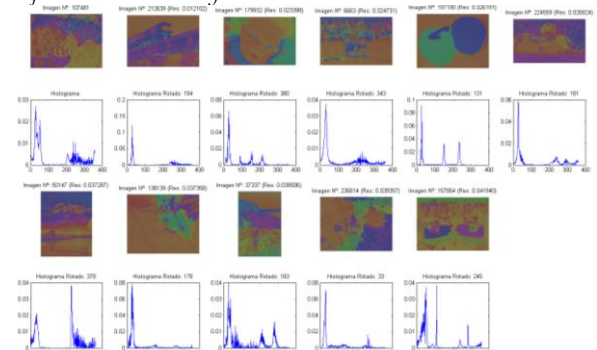


### Sin ángulo y 3 perfiles seleccionados

*Sin fijar saturación y valor*

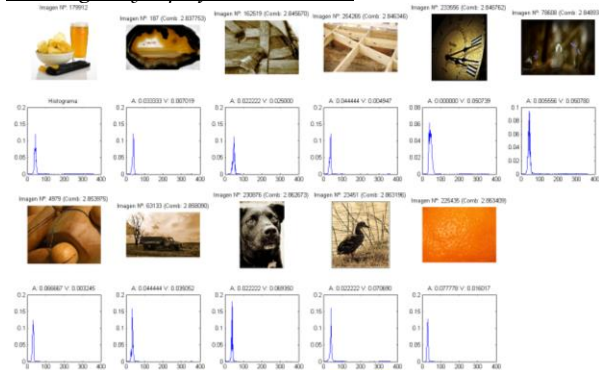


### Fijando saturación y valor

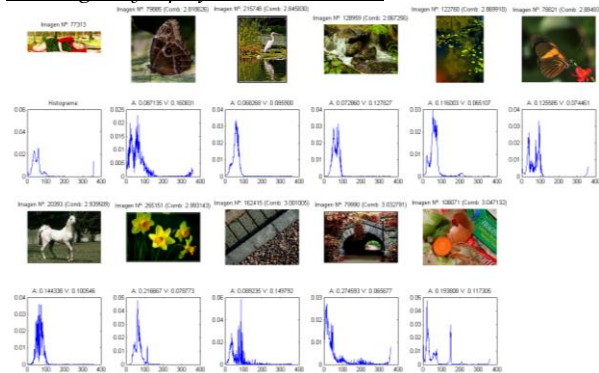




### Con ángulo y 1 perfil seleccionado



### Con ángulo y 3 perfiles seleccionados



Como podemos observar cuanta más perfiles seleccionemos, encontramos información normalmente más concreta.

## A.2 PROBLEMAS ENCONTRADOS

### 8. Divergencia KL:

La dificultad para implementar la divergencia KL reside en la facilidad de obtener infinitos e indeterminación. A nivel individual, si aplicamos la divergencia KL no tendríamos problemas. Pero en nuestra formula disponemos de un sumario, donde si un valor da infinito, entonces todo el resultado es infinito y lo mismo sucede con una indeterminación. Eso hace casi imposible que la medida permita comparar perfiles, así que se debieron tomar ciertas decisiones a la hora de la implementación para hacerla viable, incluyendo bias y evadiendo esos casos.

### 9. Ruido en el histograma acumulado (Generación de modelos por dendograma):

Uno de los problemas que encontramos al acumular histogramas, es que el ruido en cada uno de ellos también lo hacía, produciendo resultados inesperados y que el orden de aplicación fuese importante.

### 10. Aplicativo OpenCV:

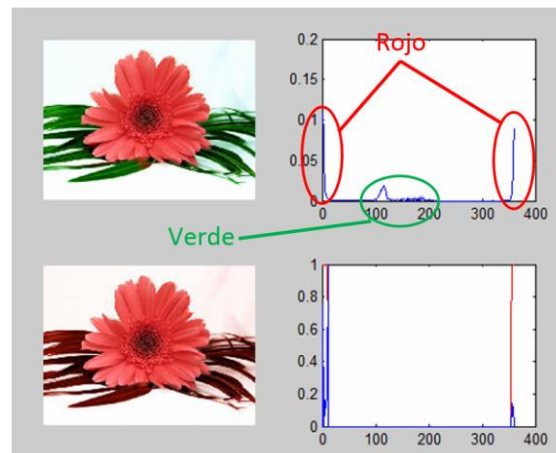
Para trasladar al móvil algunas de las aplicaciones, se hizo uso de la librería OpenCV, esta trabaja a nivel nativo de Android, es decir no se programa en Java, sino que se trabaja en C y se genera el bitstream que será llamado desde la aplicación Android cuando sea necesaria. Aunque al final, no se consiguieron los resultados desea-

dos en el lenguaje de alto nivel MATLAB, sí que se implementaron algunas librerías paralelamente indispensables para poder hacer uso de los algoritmos.

Y hubo muchísimas dificultades para que funcionasen. Por falta de tiempo y cambios en la planificación solo se ha podido probar su funcionamiento en el emulador de Android que incluye Eclipse.

### 11. Explicación complementaria del corrector basado en armonía:

Estos errores, también podrían deberse a que ciertas imágenes a nivel visual apreciamos colores en concreto, pero estos colores en el histograma de hue están esparcidos en un conjunto de ángulos de hue que no tienen suficiente peso como para contrarrestar picos de otros colores que no están tan esparcidos. Véase la siguiente figura:

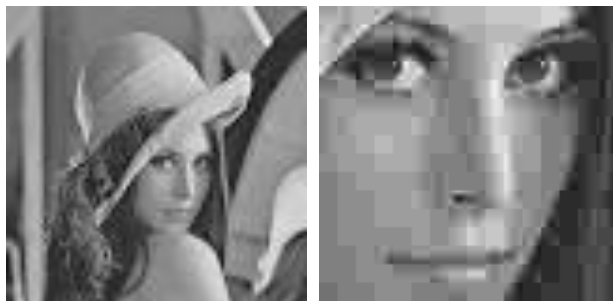


En este caso, la imagen está compuesta principalmente por los colore cromáticos: rojo y verde.

Como podemos observar en el "histograma de la imagen", el rojo está mucho más focalizado en un ángulo, mientras que el verde esta esparcido en un conjunto (Aunque visualmente vemos un color, esto se podría mejorar metodologías existentes como son las de: MacAdam ellipse, y homogeneizando las distancias de los distintos colores).

Como consecuencia, el rojo se considera un perfil armónico y el verde se elimina, cuando realmente esta imagen tiene un perfil asociado (En la realidad, directamente complementario).

Otro problema que puede suceder, se debe a como se guardan los datos en JPEG. Al realizar el proceso de compresión JPEG se aplica la transformada DCT y esta altera los colores en bloques de 8x8, haciendo que a nivel visual no apreciamos diferencias, pero que los valores si se hayan modificado por otros.



Si observamos la figura, a simple vista la cara podría tener tonalidades parecidas pero si hacemos zoom, vemos que realmente son bloques de tonos de grises bastante distintos (bloques de 8x8).