

Gamut Extension for Cinema

Syed Waqas Zamir, Javier Vazquez-Corral, and Marcelo Bertalmío

Abstract—Emerging display technologies are able to produce images with a much wider color gamut than those of conventional distribution gamuts for cinema and TV, creating an opportunity for the development of gamut extension algorithms (GEAs) that exploit the full color potential of these new systems. In this paper we present a novel GEA, implemented as a PDE-based optimization procedure related to visual perception models, that performs gamut extension by taking into account the analysis of distortions in hue, chroma and saturation. User studies performed using a digital cinema projector under cinematic (low ambient light, large screen) conditions show that the proposed algorithm outperforms the state of the art, producing gamut extended images that are perceptually more faithful to the wide-gamut ground truth, as well as free of color artifacts and hue shifts. We also show how currently available image quality metrics, when applied to the gamut extension problem, provide results that do not correlate with users’ choices.

Index Terms—Gamut mapping (GM), gamut mapping algorithm (GMA), gamut extension (GE), gamut extension algorithm (GEA), color contrast, color reproduction, variational methods.

I. INTRODUCTION

THE normal human eye contains three types of cone cells that respond to incident light to produce the sensation of color. Two lights that produce the same cone response triplet are perceived as having the same color, even if they have a different power spectrum. This allows to generate any perceivable color by a proper mixture of any three given lights (as long as they are colorimetrically independent, i.e. that the mixing of two of them does not produce the same color as the remaining one), in what is known as the trichromacy property. Therefore, given three light ‘primaries’, any color is characterized by the triplet of weights with which it can be generated as a mixture of the primaries. So colors can be represented as points in a three-dimensional space, although it is common to ignore light intensity and just represent the chromatic content of light as two-dimensional points on a plane: Fig. 1 shows the standard CIE xy chromaticity diagram, where the horseshoe-shaped region corresponds to the chromaticities of all the colors a standard observer can perceive. While the trichromacy property states that any color can be expressed as a linear combination of a given set of primaries, it is important to note that the weights for the linear combination can be negative. Since most display systems are based on mixing three primary lights by regulating the power contribution of each, the color set that a display can generate (its *color gamut*) is limited to the colors that can be obtained with linear combinations that use *positive* weights,

because negative weights imply a physically-unrealizable light of negative power. Given that each primary can be represented as a point in the CIE xy diagram, the color gamut of a display is therefore the triangle that has the primaries as its vertices, i.e. the region covered by linear combinations with positive weights, and as a consequence tri-chromatic displays are unable to reproduce many colors that we can perceive.

M. R. Pointer in [34] analyzed 4089 samples of frequently occurring real surface colors and derived what is commonly known as “Pointer’s gamut”, shown in Fig. 1. This figure also shows the standard gamuts DCI-P3 [3], used for digital cinema projection and based on the gamut of Xenon lamps, and BT.709 [11], with primaries close to the phosphor primaries of CRTs and used for cable and broadcast TV, DVD, Blu-Ray and streaming. Although both DCI-P3 and BT.709 cover a reasonable amount of Pointer’s gamut, many interesting real world colors fall outside these standard gamuts. To enhance the overall viewing experience both for cinema and TV, the media industry is continuously striving to improve image quality with higher frame rates, larger resolution, vivid colors, and greater contrast. Currently the emphasis is on high dynamic range (HDR) and wide color gamut (WCG) imaging. As pointed out in [15], HDR and WCG are independent image attributes (a picture may have a high dynamic range but a reduced color gamut such as BT.709, or the other way round) and current commercial systems do not yet support them, pending standardization of coding formats able to deal both with HDR and WCG and for which there have been a number of proposals, e.g. [15], [16], [35].

In 2012, the International Telecommunication Union-Radiocommunication (ITU-R) recommended a new standard gamut BT.2020 [10] for the next generation ultra-high definition TV that encompasses DCI-P3 and BT.709 and covers 99.9% of Pointer’s gamut. New laser projectors have monochromatic primaries with high color purity [1] and therefore they are able to cover the very wide BT.2020 gamut [23], [38], reproducing nearly every color found in nature and providing the audience with a more compelling color experience. But if the inputs are movies with DCI-P3 gamut, as virtually all professional movies currently are, the full color rendering potential of these new projectors can not be realized. The same issue happens presently when DCI-P3 projectors are used to display pictures that come in BT.709 (usually because the movie distributor wants to prevent issues with lower quality or older projectors). In both cases there is a pressing need to develop gamut extension (GE) techniques that automatically extend the gamut of the movie content, with the very challenging constraint that this has to be done in a way that the appearance of the gamut extended result preserves as much as possible the artistic intent of the content’s creator.

In the cinema industry, colorists at the post-production

Copyright (c) 2013 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Authors are with the Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08018, Barcelona, Spain. e-mail: {waqas.zamir, javier.vazquez, marcelo.bertalmio}@upf.edu

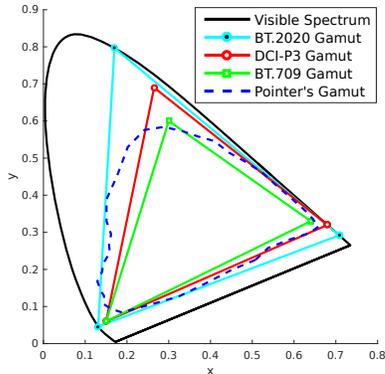


Fig. 1: Gamuts on CIE xy chromaticity diagram.

stage perform gamut modifications using three-dimensional look up tables (LUTs). These LUTs contain millions of entries and colorists only specify a few colors manually, while the rest are interpolated without taking care of their spatial or temporal context [8]. Subsequently, the resulting video may have false colors that were not present in the original material and intensive manual correction is usually necessary, commonly performed in a shot-by-shot, object-by-object basis. Nonetheless, these time consuming manual procedures by skilled technicians are preferred over automatic gamut mapping algorithms, because colorists always work following and respecting the vision of the movie creator [40], while automatic methods are prone to issues for which audiences are very sensitive and that therefore severely affect the quality of the results, like introducing color artifacts, hue shifts, and not preserving the integrity of memory colors (e.g. the blue of the sky, the green of grass) or skin tones [30].

The main contribution of this paper is to propose a gamut extension algorithm (GEA), based on perception models, that outperforms the state of the art in terms of faithfulness to the original material, as validated through a psychophysical evaluation done specifically for cinema using a digital cinema projector. The proposed GEA is capable of adapting itself according to the content of input images and aims to minimize distortions in hue, chroma and saturation. Another finding is that existing error metrics are not suitable for the GE problem, since their scores do not correlate with the results of the psychophysical experiments.

II. RELATED WORK

In general, GEAs are classified into two broad categories: global GEAs, that modify colors regardless of the pixel locations having those values, and local GEAs, which modify pixel values taking into account their local neighborhoods. Hoshino [19], [20] proposed pioneering global GEAs that extend colors from the limited gamut printed images to the color gamut of high-definition TV. Later, Kang et al. [21] and Anderson et al. [5] presented user assisted methods to deal with the problem of gamut extension. While all these aforementioned algorithms treat each color without analyzing the content of the input image, Pan and Daly [33], Casella

et al. [12], Heckaman and Sullivan [18] introduced methods that first classify the colors of the input image according to some criterion and then extend them. In particular, the work of [33] labels each color of a given image as skin or non-skin, [12] deals with objects of low chroma and high chroma differently, and [18] identifies certain memory colors such as green grass and blue sky, and renders them independently. Kim et al. [22] described a GEA with three type of extension strategies: chroma mapping, mapping along lines from the origin, and adaptive mapping that is a compromise between the first two strategies. Current state of the art global GEAs are those presented by Laird et al. [24] and they are explained in more detail in section IV-C as we use these algorithms to compare with the results of our GE algorithm.

Local GEAs are adaptive in nature and take into account the spatial color information of the original image during gamut expansion. This property certainly makes them flexible but at the same time far more complex and computationally expensive than global GEAs. Li et al. [26] presented a multilevel framework where first a non-linear hue-varying extension function is applied globally to expand the color gamut of the input image. Then in the second stage they make use of image dependent chroma smoothing to avoid over-enhancement of contrast and to retain the local information of the input image in the final reproduced image. Zamir et al. [41] proposed a spatial GEA that performs contrast enhancement using the perceptual variational color correction model [9] to expand the colors of an input image to a device gamut.

Gamut mapping algorithms can be evaluated either psychophysically or by using image quality metrics. The most common psychophysical method is the pair comparison, where two different gamut-mapped versions of an original image are shown to observers in isolation or alongside the original image. Observers are then asked to select the gamut-mapped image which exhibits more of the property (pleasantness, naturalness, or accuracy) being evaluated. In the case of GEAs, this procedure was used in Mujis et al. [31], and Laird et al. [24]. The latter described the psychophysical experiments to assess the performance of GEAs using simulated and actual wide-gamut displays. The alternative of the subjective experiment is to use objective quality metrics that are capable of finding specific distortions in reproduced images. There exists a vast variety of image quality metrics [6], [7], [13], [27], [28] in the literature that could in principle be used to quantify the results of GEAs; a few are perceptually-based. Hardeberg et al. [17] and Baranczuk et al. [6] presented psychophysical studies where they identify the best performing image quality metric for the gamut reduction problem. It is important to note that the ranking of color metrics for gamut reduction may not be consistent in the context of gamut extension if the metrics are not trained to predict well the distortions found in gamut extended images.

III. DISTORTION-BASED ITERATIVE GAMUT EXTENSION

A. Previous Approach: One-Shot Gamut Extension

In this section we briefly explain our previous work [42] on GE that we extend to develop the new GEA in this paper. In



Fig. 2: Gamut extension approach. (a) Gamuts on chromaticity diagram. (b) Gamut extension results. From left to right: input image ($\gamma = 0$), gamut extended image with $\gamma = 0.17$, $\gamma = 0.23$ and $\gamma = 0.35$. As the γ value increases the gamut becomes larger; notice the increment in saturation of helmets, socks, ski suits and shoes. Original image is courtesy of Kodak.

order to expand the colors from a smaller source gamut to a larger destination gamut, the algorithm of [42] adapts the image energy functional of Bertalmio et al. [9] that complies with some basic global and local perceptual properties of the human visual system and is closely related to the Retinex theory of color vision [25]. The authors present in [42] an adapted image energy functional

$$E(I) = \frac{\alpha}{2} \sum_x (I(x) - \mu)^2 + \frac{\beta}{2} \sum_x (I(x) - I_0(x))^2 - \frac{\gamma}{2} \sum_x \sum_y w(x, y) |I(x) - I(y)|, \quad (1)$$

whose minimization produces an image with extended gamut. In Eq. (1) α, β and γ are constant and positive weights, I is a color channel and has a range $[0, 1]$, μ is the average value of each channel of the original image I_0 , $w(x, y)$ is a normalized Gaussian kernel of standard deviation σ , and $I(x)$ and $I(y)$ are two intensity levels at pixel locations x and y respectively.

By minimizing the image energy in Eq. (1) the aim is to maximize the contrast (third term of the functional), while not departing too much from the original image (second term) and also preserving the gray world hypothesis (first term). It is formulated in [9] that the solution to the minimization of Eq. (1) can be found as the steady state of the evolution equation

$$I^{k+1}(x) = \frac{I^k(x) + \Delta t (\alpha\mu + \beta I_0(x) + \frac{\gamma}{2} R_{I^k}(x))}{1 + \Delta t (\alpha + \beta)} \quad (2)$$

where the initial condition is $I^{k=0}(x) = I_0(x)$. The function $R_{I^k}(x)$ indicates the contrast function:

$$R_{I^k}(x) = \frac{\sum_{y \in \mathcal{J}} w(x, y) s(I^k(x) - I^k(y))}{\sum_{y \in \mathcal{J}} w(x, y)} \quad (3)$$

where x is a fixed image pixel and y varies across the image. The slope function $s(\cdot)$ is a regularized approximation to the sign function, which appears as it is the derivative of the absolute value function in the third term of the functional; in [9] they choose for $s(\cdot)$ a polynomial of degree 7.

To perform gamut extension, Zamir et al. [42] first convert the RGB input image to the CIELAB color space, and then

maximize the contrast of only chromatic components ‘a’ and ‘b’ using Eq. (2). To show how the evolution Eq. (2) extends the color gamut, an example with several different gamuts (visible spectrum, source gamut, target gamut and reproduced gamut) on a chromaticity diagram is shown in Fig. 2a. It is important to recall from [42] that for each set of values for α , β , and γ the evolution Eq. (2) has a steady state. For example, it is shown in Fig. 2a that when $\beta = 1$, $\alpha = 0$, and $\gamma = 0$ we obtain the original image as the steady state of the evolution equation. Moreover, it can be seen in the same figure that as we increase γ the steady state of Eq. (2) has a gamut which is gradually larger. Fig. 2a also shows that the colors of the source gamut can be expanded to the destination gamut just by using a large enough value for γ ($\gamma = 0.35$ in this case). And to select an adequate γ value, the authors of [42] keep increasing the γ value and running evolution Eq. (2) to steady state until the gamut of the input image exceeds the target gamut up to a certain threshold T . This threshold T defines a stopping criteria according to which if $T\%$ pixels of the original image move out of the target gamut we should stop performing extension. Additionally, the threshold T controls the amount of saturation; a large value of T provides a higher level of saturation, whereas a small value of T produces a less saturated output. For each γ value, the corresponding reproductions are shown in Fig. 2b. After this, the colors that were placed outside the target gamut in previous iterations are mapped back inside using the gamut reduction algorithm [41]. Since the algorithm of [42] used a fixed value of threshold T for all the images, their results were often off from the ground truth. Moreover, the GEA of [42] causes hue shifts in some of the reproductions.

B. Proposed Method

In this section we present a new method that, unlike our previous GEA [42], works iteratively with added constraints to perform gamut extension in terms of the contrast coefficient γ . The general structure of our algorithm is as follows.

At each iteration, we run the evolution Eq. (2) for some particular α , β , and γ (in the first iteration the values are $\beta = 1$, $\alpha = 0$, and $\gamma = 0$) until we reach the steady state. For each

pixel of this steady state image we check that it simultaneously satisfies three constraints on saturation, hue and chroma: if it doesn't, then the pixel is marked as "done" and subsequent iterations will not modify its value (i.e. this pixel is now part of the final output, the gamut extended image). We move to the next iteration, by slightly increasing the value of γ and setting $\alpha = \frac{\gamma}{20}$. We run again evolution Eq. (2) until steady state and check whether any of the pixels violated any condition: those pixels are left unmodified for the following iterations. We keep repeating this process until the gamut of the original image exceeds the destination gamut up to a threshold T , at which point the iterations stop and all pixel values are final except for colors that now lie outside the destination gamut: these are mapped back inside using the gamut reduction algorithm [41].

In the following sections we describe in detail the constraints that are checked after each iteration and the way the destination gamut is constructed.

1) *Constraints*: These are the three constraints that pixels must satisfy in order to continue with the iterative process, otherwise we leave them untouched for the following iterations:

- *Saturation Constraint*: We define a condition to avoid pixels from getting desaturated, therefore each pixel must satisfy this constraint

$$S_{ae} - S_{be} > 0, \quad (4)$$

where S_{ae} and S_{be} denote saturation after extension and saturation before extension, respectively.

- *Hue Constraint*: The human visual system is highly sensitive to the changes in hue, and one major goal in GE is to produce images with unnoticeable hue shifts. In order to make our GEA robust against unpleasant color shifts, we make each pixel to comply with the hue constraint

$$|\theta_{ae} - \theta_{be}| < \epsilon_h, \quad (5)$$

where θ_{ae} , θ_{be} and ϵ_h represent hue angle after extension, hue angle before extension and hue threshold, respectively.

- *Chroma Constraint*: One of the main challenges in GE is that, after extension, colors of less saturated natural objects, skin tones, and shades of memory colors may appear too colorful and unrealistic. To deal with it, we manually segment a set of images that contain shades of sky and skin colors. We then define the crucial region on the chromaticity diagram in such a way that it encloses the chromaticity values of all the colors of these segments of images. The crucial region is shown in Fig. 6, and colors that fall inside this region should satisfy this chroma condition

$$|C_{ae} - C_{be}| < \epsilon_c, \quad (6)$$

where C_{ae} , C_{be} and ϵ_c indicate chroma after extension, chroma before extension and chroma threshold, respectively.

We want to mention that the values of the parameters ϵ_h and ϵ_c are of great importance; an example of an image with smooth color gradients is shown in Fig. 3 where it is evident that too small values of ϵ_h and ϵ_c produce an image with noticeable artifacts, whereas for slightly larger values of ϵ_h and ϵ_c the reproduction is free from false edges but may have

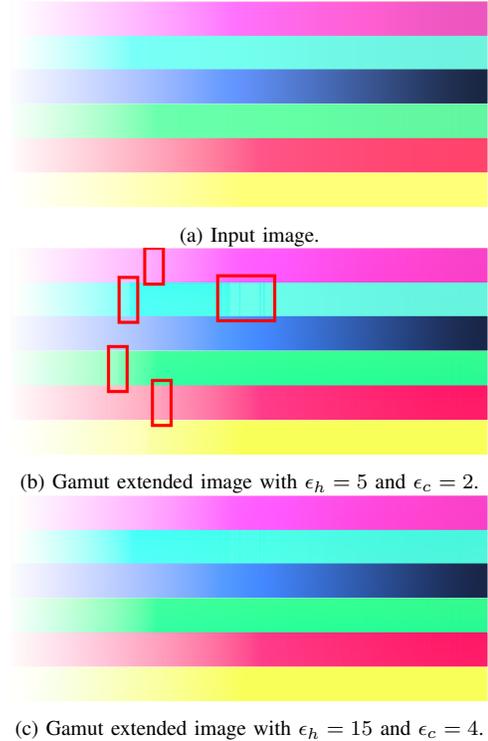


Fig. 3: Effect of parameters ϵ_h and ϵ_c . The red bounding boxes indicate the regions where artifacts (false edges) appear.

hue shifts. Therefore, these values should be adjusted in a way that the artifacts (false edges and hue shifts) stay below the visibility level. In the experiment section we propose values for these parameters that are suitable even for large differences between source and destination gamuts.

2) *Scaled Destination Gamut Computation*: One of the problems in gamut extension is that the majority of GEAs are either image dependent or perform well for a few combinations of source and destination gamuts [30]. To reduce this dependency issue, we compute a scaled version of the original image gamut and use it as the destination gamut.

Given two points p_0 and p_1 , the following parametric representation describes the line passing through them

$$p = p_0 + \eta(p_1 - p_0), \quad (7)$$

where η is a scaling factor and $\eta > 1$ will provide a point p on the line further from p_0 than p_1 .

To obtain the scaled target gamut, given an original image I_0 and a target gamut (TG), we first convert the image colors into luminance (Y) and chromaticity values (x and y) [37]. Next we calculate a reference point p_{r1} by taking the mean of those chromaticity points (x and y) that make the two-dimensional convex hull of the image I_0 . Both the reference point p_{r1} and the convex hull of I_0 are shown in Fig. 4a. We then define a set of lines (L) by substituting the reference point $p_0 = p_{r1}$ and each vertex of the convex hull of I_0 as p_1 in Eq. (7). Finally, we generate new points (one from each line in L) using a same value of scaling factor ($\eta_1 > 1$) that is chosen in such a way that none of the new points falls outside the TG and at least

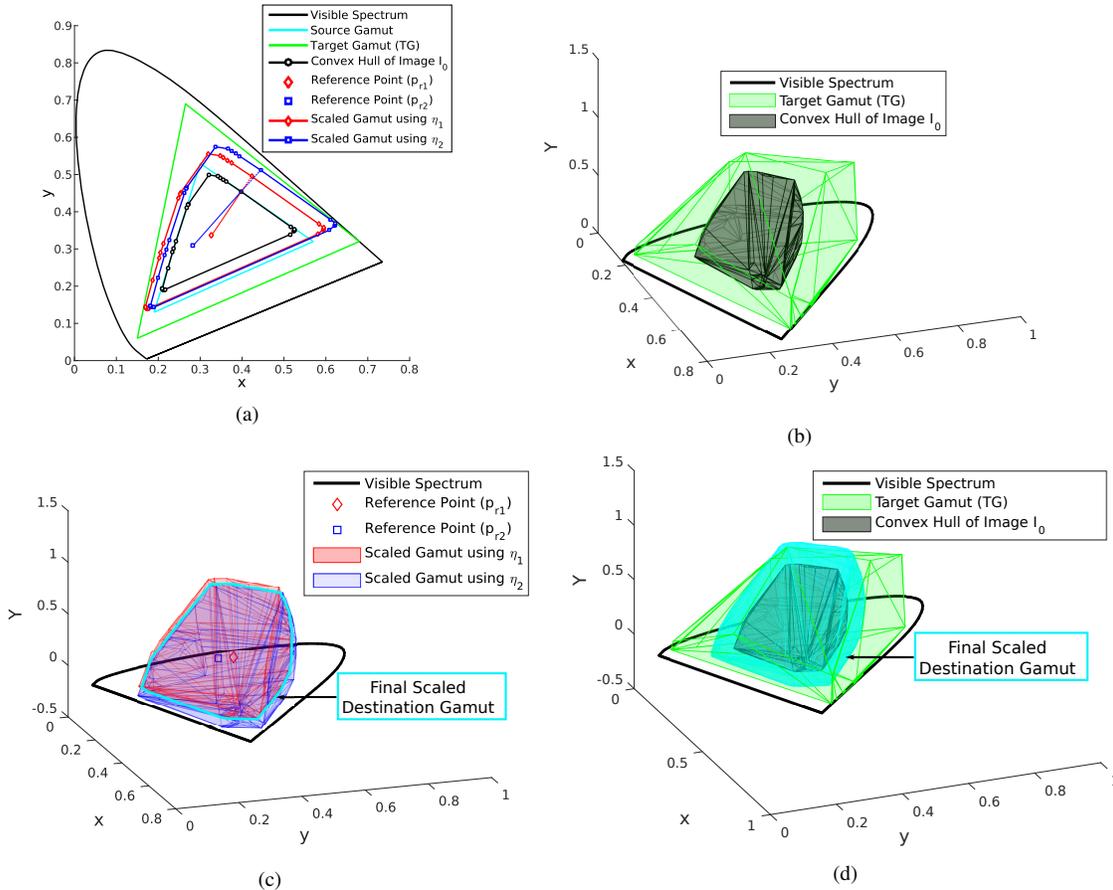


Fig. 4: Scaled destination gamut computation.

one of them touches the boundary of the TG as shown in Fig. 4a. Similarly, we calculate another scaling factor η_2 , but this time using the mean of all chromaticity values of the image as reference point p_{r2} and vertices of the convex hull of I_0 .

Once we have the scaling factors η_1 and η_2 , we apply them using Eq. (7) on the xyY triplets that make the three-dimensional (3D) convex hull of the original image I_0 (shown in Fig. 4b) to obtain two 3D scaled gamuts. Note that these two scaled gamuts, depicted in Fig. 4c, are obtained using two reference points (p_{r1} and p_{r2}); one of which is the mean of the xyY points that make the 3D convex hull of the input image, and the other is the mean of all xyY values of the image. Finally, we create the final scaled destination gamut by computing the intersection of both 3D scaled gamuts as illustrated in Fig. 4c. An example with all the relevant gamuts is shown Fig. 4d.

IV. QUALITATIVE EXPERIMENTS AND RESULTS

A. Methodology

One major goal of this paper is to investigate psychophysically the performance of GEAs in cinematic conditions using a digital cinema projector (Barco DP-1200 [2]). A GEA must respect as much as possible the artist's vision, and in this sense it is completely different from a process like colorization of black and white pictures. In general, and not only in the case

of legacy material, there never is a ground truth. The original material in a reduced gamut is all there is, and this material has to have its gamut extended while preserving as much as possible its original appearance. In our tests we actually have a ground truth because for evaluation purposes we start from a wide gamut material, reduce its gamut, extend it with a GEA and then ask users to compare the result with the wide gamut ground truth, but in general a wide-gamut ground truth is never available. Nonetheless, by showing that a GEA performs well as evaluated using ground truth data, we expect that it also performs well when ground truth data is not available, which is a most common approach in image processing (e.g. denoising algorithms are evaluated in terms of PSNR by comparing with a ground truth "clean" image that is never available in real scenarios, segmentation algorithms are evaluated by comparing their results to those of manual segmentations, etc.) One could question the choice of evaluation criteria: why ask users to choose the most accurate result instead of the one they find most pleasant? The reason is that, for a GE technique to be adopted by the movie industry, it must yield gamut extended results that preserve as much as possible the artistic intent of the content's creator. Designing GEAs for pleasantness does not guarantee this, usually quite the opposite: apart from general inter-subject variability in terms of preference, there is also a strong correlation between colourfulness and perceived image

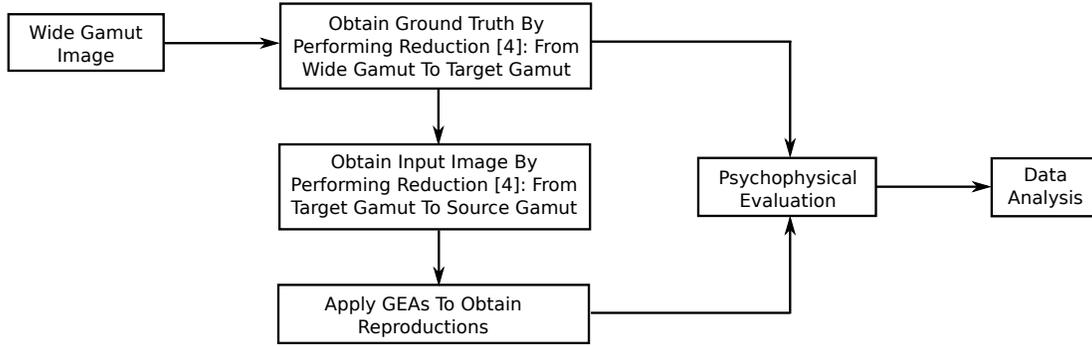


Fig. 5: A schematic of our framework.

quality [14], so user tests based on subject preference would rank higher those GEAs that increased color saturation even if that implied a departure from a subdued color palette chosen by the movie artist. User tests based on accuracy, on the other hand, are much less subjective (as users are simply asked to estimate differences with respect to the provided reference) and the preservation of the artistic intent is in a sense “built-in” in the procedure, since the ground truth acts as a stand-in for the content’s creator intent.

We show all the stages of our experimental framework in Fig. 5. The first task is to obtain both the wide-gamut ground truth images and the limited gamut input images from the given wide-gamut test images. The wide-gamut test images may have colors that fall outside the gamut of our cinema projector, therefore, to create the ground truth, we map the colors of the test images to the gamut of the projector by using the state of the art gamut reduction algorithm of Alsam and Farup [4]. All the gamuts that are used in this paper are shown in Fig. 6, and primaries are mentioned in Table I. It is important to note that in our experiments we make use of a cinema projector that has the same red and green primaries as those of the DCI-P3 standard but has a slightly different blue primary. Nonetheless, both the DCI-P3 gamut and the projector’s gamut cover almost the same amount of area on the chromaticity diagram as it can be seen in Fig. 6. Therefore, we refer to the projector’s gamut as DCI-P3 in the rest of this paper for the sake of simplicity. Next, to create the gamut-limited input images, we apply the gamut reduction method of [4]. Once we have the input images ready, we apply to them our proposed algorithm and four competing GEAs to generate reproductions for the following two different experimental setups:

- 1) *Mapping from small gamut to DCI-P3 gamut*: as laser displays with their extended gamut capabilities are becoming popular, in the near future the common case will be to have large color differences between the standard gamuts and displays’ gamuts. Therefore, we create this setup to investigate the behavior of GEAs when the difference between source and target gamut is large. To this end, we map the source images from the small ‘Toy’ gamut (slightly smaller than the BT.709 gamut) to the large DCI-P3 gamut. On the chromaticity diagram, the difference in gamuts for this setup is almost equal to the difference between DCI-P3 and BT.2020.

- 2) *Mapping from BT.709 to DCI-P3 gamut*: in this setup we mimic the practical situation where the source material has BT.709 gamut and we map the source colors to the colors of the DCI-P3 gamut.

Next we performed a subjective evaluation (for both experimental setups) with 15 observers of which ten were male and five female and their age was between 27 and 44 years, with average of 32 years. All observers were tested for normal color vision using the Ishihara color blindness test. To mimic the lighting conditions of a real cinema, we created a low light ambience where the illumination measured at the screen was around 750 lux and the ambient illuminance was 1 lux. During the experiment there was not any strong colored object in the field of view of the observers. The glare-free screen used in our experiments was 3 meters wide and 2 meters high. Each observer was instructed to sit approximately 5 meters away from the screen.

In the psychophysical experiments, we used a forced-choice pairwise comparison technique where observers were shown three images on the projection screen: the target gamut ground truth image (in the middle) and a pair of reproductions (one image on the left side and the other on the right side of the ground truth). For both setups, observers were asked to make selections according to the instructions: a) if there are any sort artifacts in one of the reproductions, choose the other, and b) if both of the reproductions have artifacts or are free from artifacts, choose the one which is perceptually more closer to the ground truth. Since there were 30 test images, 5 algorithms and 2 experimental setups, each participant had 600 comparisons to judge in total, but to avoid fatigue, we split these comparisons in four sessions that were conducted on four different days. Although there was no time restriction to make choices, each observer took approximately 30 minutes to complete one session. Finally, we analyze the psychophysical data using the work of Morovič [29], that is based on Thurstone’s law of comparative judgement [39].

Moreover, in order to validate the robustness of the competing GEAs explicitly, 9 experienced observers (who belong to the image processing community and participated in various psychophysical studies) were shown a pair of images on the projection screen; one of which was the ground truth and the other was the reproduction. We asked them to examine artifacts and hue shifts in the reproductions as compared with

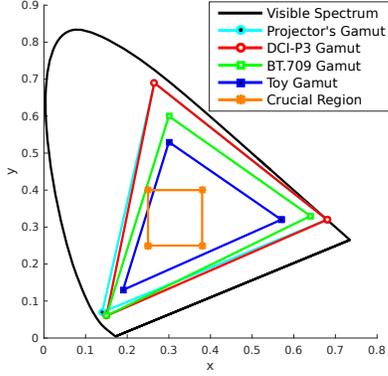


Fig. 6: Gamuts on chromaticity diagram.

TABLE I: Primaries of gamuts.

Gamuts	Red Primaries		Green Primaries		Blue Primaries	
	x	y	x	y	x	y
BT.2020	0.708	0.292	0.170	0.797	0.131	0.046
BT.709/sRGB	0.640	0.330	0.300	0.600	0.150	0.060
DCI-P3	0.680	0.320	0.265	0.690	0.150	0.060
Projector	0.680	0.320	0.265	0.690	0.140	0.070
Toy	0.570	0.320	0.300	0.530	0.190	0.130

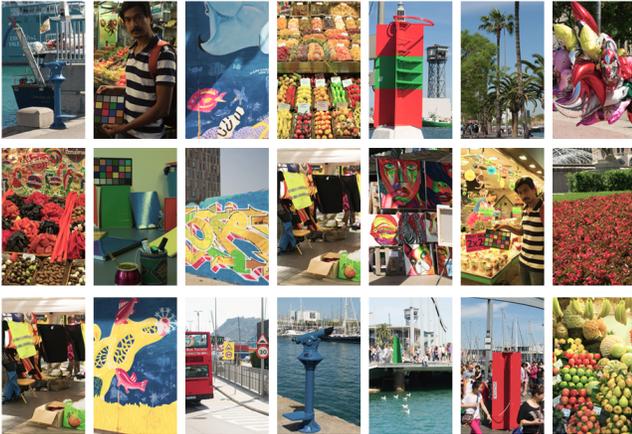


Fig. 7: Some of the wide-gamut images used in our tests. Note that only the central part of the images is shown.

the originals.

B. Creation of Wide Gamut Test Images

To the best of our knowledge, there is not any publicly available colorful wide-gamut dataset that can be used to assess the performance of GEAs. Therefore, we create wide-gamut images using a camera that is capable of capturing images in raw format. Once we have raw images, we associate with them the wide-gamut color space (ProPhoto RGB) to obtain true color images in Adobe Lightroom. Along with the 21 images shown in Fig. 7, we use 9 other test images that were from movies such as ‘Amazing Spider-Man’, ‘127 Hours’, ‘Tangled’, ‘Rio’, and ‘Requisitos para ser una persona normal’.

C. Competing GEAs

The four state of the art GEAs [24] that we compare with the proposed algorithm are briefly described as follows:

1) *Same Drive Signal (SDS)*: The most commonly used GE algorithm is the SDS [24], where the RGB primaries of the input material are linearly mapped to the RGB primaries of the display device enabling the SDS algorithm to make full use of the wide-gamut of the display.

2) *Hybrid Color Mapping (HCM)*: It is a combination of the SDS algorithm and the true-color algorithm. The true-color algorithm maintains the color information of the input image into the destination gamut without applying any sort of extension, meaning that the output of true-color is nothing but a representation of the input image in a wide-gamut color space:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{true-color} = M_{Destination}^{-1} M_{Source} \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{Source}, \quad (8)$$

where both M_{Source} and $M_{Destination}$ are 3×3 transformation matrices to convert the values of RGB to XYZ color space.

The HCM algorithm [24] linearly combines the output of the true-color and SDS algorithms based on the saturation of the input image:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{HCM} = (1 - \kappa) \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{true-color} + \kappa \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{SDS}, \quad (9)$$

where κ is a mixing factor that works as a function of saturation:

$$\kappa(S) = \begin{cases} 0, & \text{if } S \leq S_L \\ \frac{S - S_L}{S_H - S_L}, & \text{if } S_L < S < S_H \\ 1, & \text{if } S \geq S_H \end{cases} \quad (10)$$

S_L and S_H are constants to define the ranges of saturation for the mixing function κ , and their values that we used in our experiments are $S_L = 0.2$ and $S_H = 0.6$ for the setup 1, and $S_L = 0.8$ and $S_H = 1$ for the setup 2. .

The key idea behind the HCM algorithm is to preserve natural colors, leaving unmodified the low saturated colors such as skin tones, while mapping highly saturated colors using the SDS algorithm.

3) *Chroma Extension (CE)*: The SDS algorithm works by extending the input signal to the destination gamut. However, hue differences between source and destination gamut may lead SDS to produce an image that exhibits hue shifts. To tackle this problem, the chroma extension algorithm proposed in [24] maps colors of the source gamut to the reproduction gamut along lines of the chroma axis in the CIELCH color space, while keeping lightness and hue constant.

4) *Lightness Chroma Adaptive (LCA)*: The lightness-chroma adaptive algorithm [24] performs GE by altering both lightness and chroma while keeping the hue constant.

Both CE and LCA algorithms make use of the so called high chroma boost (HCB) function which smoothly maps colors of an input image in a manner that the high chroma objects get more boost in saturation than the low chroma ones. This

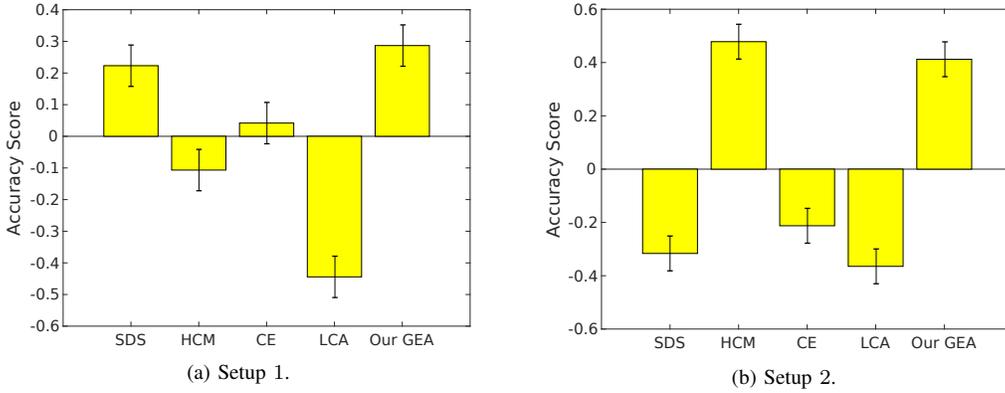


Fig. 8: Accuracy scores using 15 observers and 30 images.

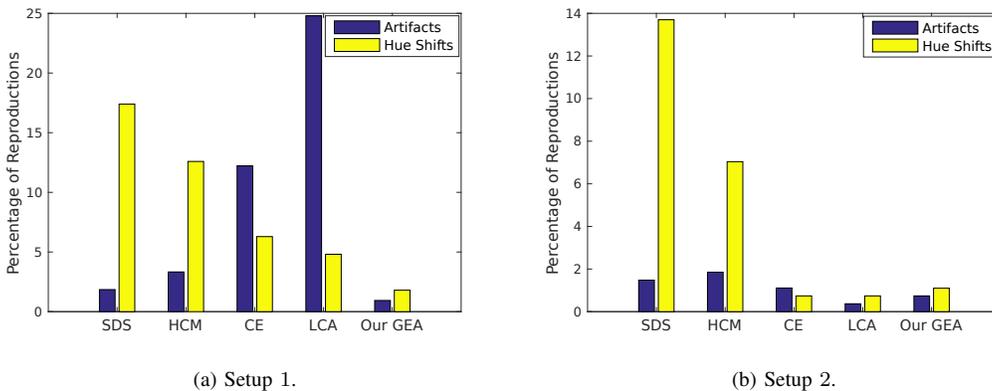


Fig. 9: Percentage of reproductions in which 9 experienced observers noticed visual distortions.

approach aims to preserve memory color objects as they often have less saturated values, and extends colors of artificial objects (such as plastics) that have high chroma.

D. Settings for Proposed Method

We work in CIELAB color space and the parameter values that we use in all our experiments are $\beta = 1$, $\epsilon_h = 15$, $\epsilon_c = 4$, $\Delta t = 0.10$, $T = 1\%$ of the total number of pixels, and the gamma increment of $\Delta\gamma = 0.01$ between two consecutive iterations. The proposed GEA in a non-optimized MATLAB implementation running on a machine with 8 cores 3.4-GHz CPU takes (on average) 4.5 minutes to process an image of resolution 656×1080 pixels.

E. Results

In order to compute the accuracy scores from the raw psychophysical data, we use the data analysis procedure presented in [29]. The analysis of psychophysical data of the setup 1 presented in Fig. 8a shows that, when there is a large difference among the source-target gamut pair, the proposed GEA produces images that are perceptually more faithful to the originals as compared with the other competing algorithms. The observers declared LCA [24] as the least accurate algorithm, whereas both the CE [24] and the HCM

[24] algorithms rank third and fourth, respectively. In Fig. 9, we report the percentage of reproductions for which 9 experienced observers, on average, noticed the visual distortions: artifacts or hue shifts. For the setup 1, the LCA and CE algorithms produce images with loss of texture due to over-saturation and it can be seen in Fig. 9a that the subjects noticed artifacts in 25% of the reproductions obtained using the LCA algorithm and in 12% of the images in the case of the CE algorithm. The observers confirmed that the proposed GEA produces images with very low error rate, around 2%. One such example is shown in Fig. 10 where it is clearly visible that the colors reproduced by our GEA are artifact free and perceptually more faithful to the ground truth than those of LCA and CE algorithms. Moreover, it can also be seen in the said figure that the LCA and CE algorithms turn subtle spatial color details of the woman's apron into noticeable unpleasant color gradients. Even though both the SDS [24] and HCM algorithms do not introduce much noticeable artifacts, their reproductions show strong hue shifts. We show an example in Fig. 11 where the hue shifts are evident on the floor in the results of the SDS and HCM methods. Note that the images depicted in Fig. 10 and Fig. 11 were in DCI-P3 format, originally. Since we are limited by sRGB standard (that has the same primaries as of BT.709) for the paper, in order to present results we show only those colors that are inside the sRGB gamut and mask (in green) the rest of the colors.

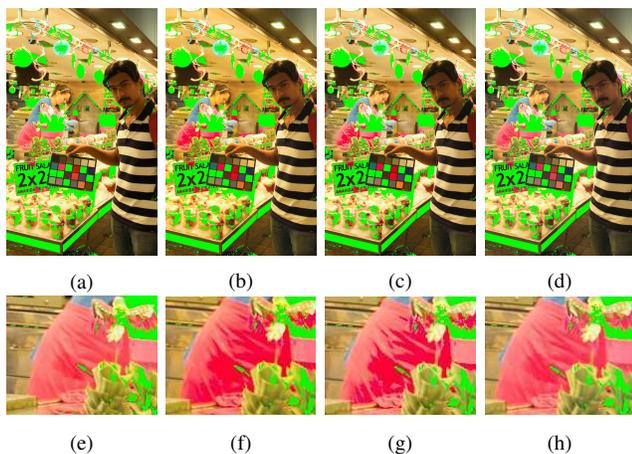


Fig. 10: Example of artifacts. (a) Ground Truth. (b) Output of CE algorithm [24]. (c) Output of LCA algorithm [24]. (d) Output of our GEA. (e)-(h) Zoomed-in view of the regions cropped from the top row. Note that these are wide-gamut images where out-of-sRGB pixels are masked green.

In Fig. 8b we show the results of setup 2 where it can be seen that, when the difference between source and target gamut is smaller, the ranking order of the GEAs changes dramatically. The HCM algorithm that ranked as the second least accurate in the previous setup becomes the most accurate method. Our GEA shows comparable performance with the HCM algorithm. Similar to the results of setup 1, the LCA algorithm produces gamut extended images that are least faithful with the original content. It is also evident from Fig. 8b that the SDS method, unlike in setup 1, shows poor performance under setup 2. Fig. 9b shows that the SDS and HCM algorithms produce gamut extended images with strong color shifts for 13.6% and 7% of the input images, respectively. It can be seen in the same figure that none of the competing algorithms produce images with distinct visual artifacts for setup 2, in which there are small color differences between source and target gamut.

Trends for both experimental setups show that the proposed GEA is the most consistent and reliable method for both small and large color differences among the source-target gamut pair.

1) *Temporal Consistency Test*: In order to examine the temporal consistency of GEAs, we conducted a psychophysical study with 9 experienced observers and two colorful image sequences with different levels of motion. The representative frames for both image sequences are shown in Fig. 12. In this experiment, the gamut extended videos obtained using different GEAs were shown in isolation (without any ground truth) to each observer and they were asked to inspect the following attributes: temporal color consistency (objects should retain the same hue, chroma and brightness), global flickering, local region flickering, and excessive noise. None of the observers noticed any temporal artifacts, which supports our choice to apply all competing GEAs on each frame independently. Finally, we want to stress that the quality of the input video is of high importance; if it contains any spatial artifacts due to compression or noise they may become prominent in the reproduced video.

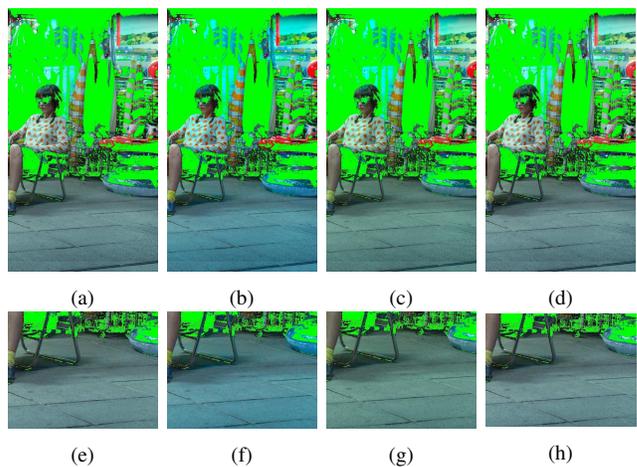


Fig. 11: Example of hue shifts. (a) Ground Truth. (b) Output of SDS algorithm [24]. (c) Output of HCM algorithm [24]. (d) Output of our GEA. (e)-(h) Zoomed-in view of the regions cropped from the top row. Note that these are wide-gamut images where out-of-sRGB pixels are masked green.

V. IDENTIFICATION OF ERROR METRIC SUITABLE FOR THE GE PROBLEM

We would like to find out whether or not there exists an error metric that is suitable for the gamut extension problem. First, we examine the prediction quality of nine error metrics: CIE ΔE , CIE ΔE_{94} [7], CIE ΔE_{2000} [28], color image difference (CID) [27], and the ones presented in [13] such as mean square error (MSE), structural content (SC), maximum difference (MD), Laplacian mean square error (LMSE) and normalized absolute error (NAE). We test the efficiency of the GEAs for setup 1 and setup 2 using these error metrics, that find distortions between ground truth and reproductions. We list the overall error for all the images in Table II. There it can be seen that for setup 1, only two out of nine error metrics selected the observers' preferred GEA (our proposed algorithm) as the best performing algorithm. Both the LCA and the HCM algorithms were chosen as the most accurate methods by three different error metrics, whereas none of the tested metrics picks SDS as a good performing method: all these results contradict the findings of the psychophysical experiment.

For setup 2, five out of the nine error metrics tested declare HCM as the most efficient algorithm, which is consistent with the results of the psychophysical experiment. However, the same image quality metrics rank other GEAs very differently as compared with the observers' choices. Another notable finding is that the CIE ΔE metric and its variants (CIE ΔE_{94} and CIE ΔE_{2000}) show a similar trend and pick LCA and HCM as the most efficient algorithms for setup 1 and setup 2, respectively.

Even though some error metrics were able to make right predictions, we check the correlation between the choices made by our observers in the psychophysical experiments and the predictions of the error metrics to confirm the reliability of these image metrics. To achieve this, we use the hit rate (h)



Fig. 12: Representative frames of image sequences with toy gamut.

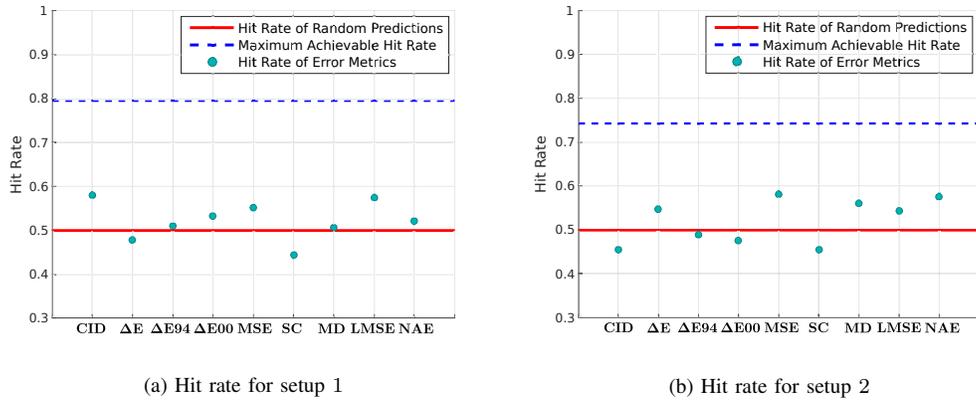


Fig. 13: Hit rates obtained by image quality metrics for experimental setups.

TABLE II: Predictions of image quality measures: error across all images.

Error Metrics	Setup 1: Toy to DCI-P3 gamut					Setup 2: BT.709 to DCI-P3 gamut				
	SDS	HCM	Chroma	LCA	Our GEA	SDS	HCM	Chroma	LCA	Our GEA
CID	0.0903	0.0858	0.0862	0.1013	0.0987	0.0645	0.0624	0.0658	0.0657	0.0681
ΔE	0.1926	0.1766	0.1636	0.1396	0.1796	0.1603	0.1401	0.1679	0.1680	0.1896
ΔE_{94}	0.2941	0.2684	0.2527	0.2203	0.2766	0.2203	0.1927	0.2141	0.2142	0.2194
ΔE_{00}	0.3058	0.2802	0.2489	0.2315	0.2670	0.2302	0.2037	0.2317	0.2301	0.2326
MSE	0.0060	0.0059	0.0109	0.0144	0.0116	0.2034	0.2192	0.1318	0.1317	0.1205
SC	0.6157	0.6052	0.5449	0.5212	0.5037	0.4885	0.4933	0.5268	0.5269	0.5209
MD	0.3442	0.3435	0.5096	0.5853	0.4393	0.3698	0.3172	0.3644	0.3651	0.2586
LMSE	0.3133	0.3154	0.2361	0.3124	0.2093	0.2034	0.2192	0.1318	0.1317	0.1205
NAE	0.1005	0.0967	0.0953	0.0977	0.1031	0.0636	0.0503	0.0748	0.0747	0.0640

approach that is defined as

$$h = \frac{c}{N} \quad (11)$$

where N is the total number of choices for any experimental setup and c is the number of correctly predicted choices. We operate on the raw visual data, and as we know the choice of the observer for each paired comparison in the psychophysical experiment, we consider a choice is correctly predicted if an image quality metric gives less error for the image selected by the observer. Any quality measure that makes random predictions is expected to achieve a hit rate of $h = 0.5$. Since each image pair is compared by several observers and considering that not all of them have made the same choice, we computed the *maximum achievable hit rate* to which an

error metric will approach if its predictions are in agreement with the majority of observers' choices for all image pairs.

Since our purpose is to find a metric that works for both small and large color differences between source and target gamuts, in Fig. 13 we present the hit rates to validate the image quality metrics for both setup 1 and setup 2. It can be seen that the error metrics made, most of the time, random predictions and their hit rates are significantly lower than the maximum achievable hit rate. Note that if an error metric has a very low hit rate (i.e. less than 0.2), one can make it a good image quality measure by inverting its predictions for all image pairs. However, none of these metrics achieves either a very low hit rate or a very high hit rate and therefore they are unreliable to be used for evaluating results in the gamut extension problem.

Without an efficient error metric, gamut extension cannot be posed as an optimization scheme as in [32], [36] where, given a reference image and its gamut mapped version, the idea is to alter the gamut mapped image to minimize the distortions as computed by an objective metric. Consequently, it leaves us with only one option for the moment to validate GEAs, that is to perform subjective studies which are cumbersome and time consuming.

VI. CONCLUSION

In this paper we have proposed a perception-based iterative GEA that adapts itself according to the content of the input image and produces gamut extended images that are free from visual distortions, natural in appearance and perceptually faithful to the original material. Moreover, we have presented a psychophysical study to investigate the quality of the results produced by the proposed algorithm and four other GEAs under cinematic conditions using a cinema projector. The analysis of psychophysical data shows that our algorithm outperforms the competing methods when the difference between source and destination gamuts is large (setup 1), whereas the results of the proposed GEA are comparable with the best performing algorithm (HCM) when the source-destination gamut pair has smaller differences in colors (setup 2), but our method produces much less artifacts than HCM. Although setup 2 is more common right now, setup 1 will become highly relevant with the popularization of new laser projectors. Finally, we tested the prediction quality of various image quality metrics in order to identify a suitable image measure for the GE problem. We concluded that none of the test metrics show any useful correlation with the psychophysical choices made by the observers.

VII. ACKNOWLEDGEMENTS

This work was supported by the European Research Council, Starting Grant ref. 306337, by the Spanish government and FEDER Fund, grant ref. TIN2015-71537-P (MINECO/FEDER,UE), and by the Icrea Academia Award. The authors are grateful to all the participants of the psychophysical experiments. Many thanks go to Stephan Cattan, from Deluxe (Spain), for providing us with wide gamut examples and for all his help and suggestions. Last but not least, the authors would like to express their gratitude to Barco N.V. for their invaluable help, without which this work would not have been possible.

REFERENCES

- [1] <http://spectrum.ieee.org/consumer-electronics/audiovideo/lasers-coming-to-a-theater-near-you>. [Online; accessed 13-June-2016].
- [2] <http://www.barco.com/en/Products-Solutions/Projectors/Digital-cinema-projectors/2K-Digital-cinema-projector-for-screens-up-to-12m-40ft.aspx>. [Online; accessed 13-June-2016].
- [3] SMPTE RP 431-2:2011. D-cinema quality – reference projector and environment, 2011.
- [4] A. Alsam and I. Farup. Spatial colour gamut mapping by orthogonal projection of gradients onto constant hue lines. In *Proc. of 8th International Symposium on Visual Computing*, pages 556–565, 2012.
- [5] H. Anderson, E. Garcia, and M. Gupta. Gamut expansion for video and image sets. In *Proc. of the 14th International Conference of Image Analysis and Processing - Workshops*, pages 188–191, 2007.
- [6] Z. Barańczuk, P. Zolliker, and J. Giesen. Image quality measures for evaluating gamut mapping. In *Color and Imaging Conference*, pages 21–26, 2009.
- [7] R. S. Berns. The mathematical development of CIE TC 1-29 proposed colour difference equation: CIELCH. In *Proc. of the Seventh Congress of International Colour Association, B, C19.1C19.4*, 1993.
- [8] M. Bertalmío. *Image Processing for Cinema*, volume 4. CRC Press, Taylor & Francis, 2014.
- [9] M. Bertalmío, V. Caselles, E. Provenzi, and A. Rizzi. Perceptual color correction through variational techniques. *IEEE Transactions on Image Processing*, 16(4):1058–1072, 2007.
- [10] ITU-R Recommendation BT.2020. Parameter values for ultra high definition television systems for production and international programme exchange, 2012.
- [11] ITU-R Recommendation BT.709-5. Parameter values for the HDTV standards for production and international programme exchange, 2002.
- [12] S. E. Casella, R. L. Heckaman, and M. D. Fairchild. Mapping standard image content to wide-gamut displays. In *Color and Imaging Conference*, pages 106–111, 2008.
- [13] A. M. Eskicioglu and P. S. Fisher. Image quality measures and their performance. *IEEE Transactions on Communications*, 43(12):2959–2965, Dec 1995.
- [14] E. A. Fedorovskaya, H. de Ridder, and F. JJ Blommaert. Chroma variations and perceived quality of color images of natural scenes. *Color research & application*, 22(2):96–110, 1997.
- [15] E. François, C. Fogg, Y. He, X. Li, A. Luthra, and C. Segall. High dynamic range and wide color gamut video coding in HEVC: Status and potential future enhancements. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(1):63–75, 2016.
- [16] J. Froehlich, T. Kunkel, R. Atkins, J. Pytlarz, S. Daly, A. Schilling, and B. Eberhardt. Encoding color difference signals for high dynamic range and wide gamut imagery. In *Color and Imaging Conference*, pages 240–247, 2015.
- [17] J. Y. Hardeberg, E. Bando, and M. Pedersen. Evaluating colour image difference metrics for gamut-mapped images. *Coloration Technology*, 124(4):243–253, 2008.
- [18] R. L. Heckaman and J. Sullivan. Rendering digital cinema and broadcast TV content to wide gamut display media. *SID Symposium Digest of Technical Papers*, 42(1):225–228, 2011.
- [19] T. Hoshino. A preferred color reproduction method for the HDTV digital still image system. In *Proc. of IS&T Symposium on Electronic Photography*, pages 27–32, 1991.
- [20] T. Hoshino. Color estimation method for expanding a color image for reproduction in a different color gamut, May 1994. US Patent 5,317,426.
- [21] B. H. Kang, J. Morovič, M. R. Luo, and M. S. Cho. Gamut compression and extension algorithms based on observer experimental data. *ETRI journal*, 25(3):156–170, 2003.
- [22] M. C. Kim, Y. C. Shin, Y. R. Song, S. J. Lee, and I. D. Kim. Wide gamut multi-primary display for HDTV. In *Proc. of 2nd European Conference on color Graphics, Imaging and Vision*, pages 248–253, 2004.
- [23] Y. Kusakabe, Y. Iwasaki, and Y. Nishida. Wide-color-gamut super hi-vision projector. In *Proc. ITE annual convention*, 2013 (in Japanese).
- [24] J. Laird, R. Muijs, and J. Kuang. Development and evaluation of gamut extension algorithms. *Color Research & Application*, 34(6):443–451, 2009.
- [25] E. H. Land. The retinex theory of color vision. *Scientific American*, 237(6):108–128, 1977.
- [26] Y. Li, G. Song, and H. Li. A multilevel gamut extension method for wide gamut displays. In *Proc. of International Conference on Electric Information and Control Engineering (ICEICE)*, pages 1035–1038, 2011.
- [27] I. Lissner, J. Preiss, P. Urban, M. S. Lichtenauer, and P. Zolliker. Image-difference prediction: From grayscale to color. *IEEE Transactions on Image Processing*, 22(2):435–446, 2013.
- [28] M. R. Luo, G. Cui, and B. Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application*, 26(5):340–350, 2001.
- [29] J. Morovič. *To Develop a Universal Gamut Mapping Algorithm*. PhD thesis, University of Derby, UK, 1998.
- [30] J. Morovič. *Color gamut mapping*, volume 10. Wiley, 2008.
- [31] R. Muijs, J. Laird, J. Kuang, and S. Swinkels. Subjective evaluation of gamut extension methods for wide-gamut displays. In *Proc. of the 13th International Display Workshop*, pages 1429–1432, 2006.
- [32] S. Nakauchi, S. Hatanaka, and S. Usui. Color gamut mapping based on a perceptual image difference measure. *Color Research & Application*, 24(4):280–291, 1999.

- [33] H. Pan and S. Daly. A gamut-mapping algorithm with separate skin and non-skin color preference controls for wide-color-gamut TV. *SID Symposium Digest of Technical Papers*, 39(1):1363–1366, 2008.
- [34] M. R. Pointer. The gamut of real surface colours. *Color Research & Application*, 5(3):145–155, 1980.
- [35] C. Poynton, J. Stessen, and R. Nijland. Deploying wide color gamut and high dynamic range in HD and UHD. *SMPTE Motion Imaging Journal*, 124(3):37–49, 2015.
- [36] J. Preiss, F. Fernandes, and P. Urban. Color-image quality assessment: from prediction to optimization. *IEEE Transactions on Image Processing*, 23(3):1366–1378, 2014.
- [37] G. Sharma. *Digital Color Imaging Handbook*. CRC Press, Inc., Boca Raton, FL, USA, 2002.
- [38] B. D. Silverstein, A. F. Kurtz, J. R. Bietry, and G. E. Nothhard. A laser-based digital cinema projector. *SID Symposium Digest of Technical Papers*, 42(1):326–329, 2011.
- [39] L. L. Thurstone. A law of comparative judgment. *Psychological Review*, 34(4):273–286, 1927.
- [40] A. Van Hurkman. *Color correction handbook: professional techniques for video and cinema*. Pearson Education, 2013.
- [41] S. W. Zamir, J. Vazquez-Corral, and M. Bertalmío. Gamut mapping in cinematography through perceptually-based contrast modification. *IEEE Journal of Selected Topics in Signal Processing*, 8(3):490–503, 2014.
- [42] S. W. Zamir, J. Vazquez-Corral, and M. Bertalmío. Gamut extension for cinema: Psychophysical evaluation of the state of the art, and a new algorithm. In *Proc. of IS&T/SPIE Human Vision and Electronic Imaging XX*, 2015.