RESTORATION OF IMAGE BURNOUT IN 3D-STEREOSCOPIC MEDIA USING INTER-VIEW GRADIENT INTERPOLATION

David Corrigan, François Pitié and Anil Kokaram

Department of Electrical and Electronic Engineering, Trinity College Dublin,
Dublin 2, Ireland.
email: {corrigad, fpitie, anil.kokaram}_at_tcd.ie
web: www.sigmedia.tv

ABSTRACT

Correcting colour imbalances between stereo views is a major task in the production of high quality 3D stereoscopic content. This paper addresses extreme cases of colour imbalance where the intensity range of one view exceeds the valid range causing image burnout in the degraded view. We propose a comprehensive framework for restoring the degraded view by using the other well-exposed view to interpolate both colour and texture information. The texture copy is done in the gradient domain by solving the Poisson equation and using the disparity map to guide the interpolation. We also propose a method of detecting burnout inspired by a technique for detecting poorly exposed regions in image stills. Our results show that it is possible to get a good restoration in stereo pairs with a significant degree of image burnout.

1. INTRODUCTION

Inconsistencies in the camera configuration of stereo rigs is the primary cause of visual discomfort when viewing stereo-3D video. Colour imbalances between stereo views are one of the most common problems associated with inconsistent camera configurations. These imbalances can cause one view for example to appear brighter or more washed out than the other. Typical causes of imbalances are differences in camera sensor characteristics as well as the choice of mirror rigs[1]. Colour balancing algorithms attempt to match the colour distributions of two images by either employing a linear mapping of the lower order moments of the colour histogram (typically mean and covariance) [13, 11] or adopting a non-linear histogram matching based approach [9, 12]. This paper deals with extreme cases of colour imbalances, where it is possible for the intensity range of the scene to exceed the dynamic range of the camera (Fig 1). Not only does one view appear brighter than the other, burnout occurs in the image highlights resulting in a loss of texture detail.

1.1 Related Work

Restoring image burnout in stereo pairs shares many similarities with the problem of estimating a well-exposed image from a series of bracketed-exposure stills of a static scene. The well-exposed image can be composed by estimating an intermediate high dynamic range (HDR) image (*e.g.* [5]) and applying a tone-mapping operator (*e.g.* [6]) to fit the HDR image to the range of the display device. An alternative direct approach [8] is to create the image by fusing together the well-exposed regions of each still. Similarly, the goal here is to restore the degraded view by fusing it with the relevant detail in the reference view. However, the viewpoints of the cameras are different and so the views need to be registered in order for fusion to take place. Since the intensity range in the reference view better fits the allowed range, it is also desirable to match the colour distribution of the degraded view to that of the reference view.

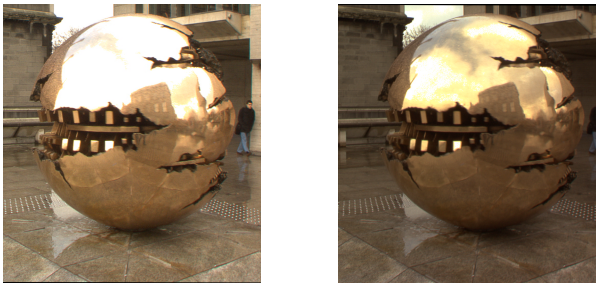

Figure 1: This stereo pair, taken from the Sigmedia Stereo Database [4], consists of one well-exposed view (right) and an over-exposed view (left) in which image burnout is present. Texture has been destroyed in the reflection on the sphere surface and in the sky behind the sphere. Our algorithms reconstructs the gradients inside the burnout regions by copying them from the well-exposed view and matches the colour distribution of the over-exposed view to well-exposed view.

A previous approach to the burnout problem was outlined in [4] which performed the texture transfer by copying the band-pass wavelet coefficients from the DT-CWT decompositions [7] of the degraded and reference views. The algorithm gave good results in examples where there were large non-elongated regions of burnout such as the example of the sphere shown in Fig. 1. However, since up to 9 levels of DTCWT band-pass coefficients were required to reconstruct the texture accurately, it did not perform well on elongated or small regions of burnout. Furthermore, it induced a spatially varying colour distortion caused by discontinuities in band-pass coefficients of the restored view. This acted to introduce a high frequency flicker to the restored sequence.

1.2 A New Approach

In this paper, we overcome the limitations of [4] by performing the texture copying in the gradient domain rather than the wavelet domain. By applying the Poisson partial differential equation (PDE) with Dirichlet boundary conditions it is possible to reconstruct the gradient field in the burnout region

This work was funded in part by grants from the Science Foundation Ireland PI award CAMP and the EU Fp7 i3dPost project.

[1] http://www.fxguide.com/quicktakes/nuke-stereo-master-class-plus-6-1-preview/

from the gradient field of the reference view in such away that minimises the visual discontinuity across the boundaries of the burnout regions. This technique is a well established method in the community for seamless merging of patches from different images [10, 1]. We also introduce a new method for the detection of burnout regions which is less reliant on morphological operations to get a complete mask than the method proposed in [4]. The final contribution presented in this work is a refined strategy for disparity estimation that gives more robust disparity estimates inside the burnout regions which allows for more robust texture transfer from the reference view.

The remaining sections of the paper are organised as follows. Section 2 introduces the restoration algorithm and describes its novel features. This is followed by the results section which shows some examples of restored frames and discusses the challenges of designing a practical algorithm to faithfully restore stereo-3D video sequences.

2. RESTORATION FRAMEWORK

The degraded view contains regions of image burnout and also contains a colour distortion in the non-burnout regions. The model of the the degraded view $G(\mathbf{x})$ is described in terms of the uncorrupted view $I(\mathbf{x})$ as

$$G(\mathbf{x}) = (1-\alpha(\mathbf{x}))f(I(\mathbf{x})) + \alpha(\mathbf{x})C_{\text{rep}}. \quad (1)$$

In this equation, the $\alpha(\mathbf{x})$ map specifies a binary mask that describes the degree of burnout in each region. In the burnout regions, where the value of $\alpha(\mathbf{x}) = 1$, the degraded image texture is replaced with a replacement colour C_{rep} which is assumed to be white in colour. $f(I(\mathbf{x}))$ describes the colour distortion of the hidden uncorrupted image.

The interpolation of burnout regions is guided by a stereo-view disparity model which states that

$$I(\mathbf{x}) = I_{\text{ref}}(\mathbf{x}+\mathbf{d}(\mathbf{x})) + e(\mathbf{x}) \quad (2)$$

where $\mathbf{d}(\mathbf{x})$ is the disparity estimate that maps pixels, $\mathbf{x}$, in the degraded view to pixels, $\mathbf{x}+\mathbf{d}(\mathbf{x})$, in the reference view and $e(\mathbf{x})$ represents the realisation of a random noise process. Consequently, a full disparity estimate is required to register the reference view. To restore the degraded view, it is necessary to correct the colour distortions in the non-burnout regions as well as to reconstruct the burnout regions. The restoration is described by

$$I_{\text{res}}(\mathbf{x}) = (1-\alpha(\mathbf{x}))f^{-1}(G(\mathbf{x})) + \alpha(\mathbf{x})I_{\text{ref}}(\mathbf{x}+\mathbf{d}(\mathbf{x})) \quad (3)$$

where $f^{-1}(.)$ describes the colour correction operation.

However, copying a region from the reference into the burnout region in image space will cause discontinuities across the boundary of the burnout regions. Our solution is based the gradient domain fusion approach [10, 1] used for seamless cloning of image patches. The remainder of this section describes our solutions to each of the problems of colour-correction (*i.e.* finding $f^{-1}(G(\mathbf{x}))$), burnout region detection (finding $\alpha(\mathbf{x})$), disparity estimation as well as describing the gradient domain data interpolation in the burnout regions.

2.1 Burnout Detection

Image burnout results in pixels that have a bright white colour. Hence, in the HSV colour space the burnout pixels have high value and low saturation components. As burnout regions also contain little texture they also exhibit low spatial contrast. In [8], Mertens *et al.* propose a method for measuring the exposure quality based on these observations that gives high weights to regions that are well-exposed and low weights to under or over-exposed regions. It proposes separate quality metrics based on the luminance, saturation and contrast values at each pixel and creates a joint exposure quality weight based on a geometric average of the individual metrics.

In this paper, we propose a method for detecting a binary burnout mask based on these metrics. This is achieved by estimating an intermediate burnout confidence measure based on the saturation and luminance metrics proposed in [8]. These metrics are calculated on the degraded view $G(\mathbf{x})$ defined in (1). The confidence measure at a given pixel, $\omega(\mathbf{x})$, is given by

$$\omega(\mathbf{x}) = (1-S(\mathbf{x})^{\omega_s})(1-L(\mathbf{x})^{\omega_l}) \quad (4)$$

where $S(\mathbf{x})$ and $L(\mathbf{x})$ are the exposure quality metrics based on saturation and luminance respectively and ω_s and ω_l are a set of tuning parameters.

The range of both $S(\mathbf{x})$ and $L(\mathbf{x})$ is between 0 and 1 where a value close to 0 indicates that the pixel is more likely to belong to a burnout region and is close to 1 if the pixel comes from a well exposed region of the image. The tuning parameters ω_s and ω_l adjust the sensitivity of $\omega(x)$ to $S(\mathbf{x})$ and $L(\mathbf{x})$ respectively. For example, as the value of ω_l decreases, for a given $S(\mathbf{x})$ the confidence measure $w(\mathbf{x})$ decreases. Thus, from (4) $\omega(\mathbf{x})$ will have a value approaching 1 in burnout regions. The expressions for $S(\mathbf{x})$ and $L(\mathbf{x})$ follow closely metrics outlined in [8] and are given by

$$S(\mathbf{x}) = \max\big(G_r(\mathbf{x}),\, G_g(\mathbf{x}),\, G_b(\mathbf{x})\big) - \min\big(G_r(\mathbf{x}),\, G_g(\mathbf{x}),\, G_b(\mathbf{x})\big) \quad (5)$$

$$L(\mathbf{x}) = \prod_{k=\{r,g,b\}} \begin{cases} \exp -\frac{(G_k(\mathbf{x})-0.5)^2}{2\times 0.3^2} & G_k(\mathbf{x}) > 0.5 \\ 1 & G_k(\mathbf{x}) < 0.5 \end{cases} \quad (6)$$

where $G_r(\mathbf{x}), G_g(\mathbf{x}), G_b(\mathbf{x})$ are red green and blue intensities (RGB) of the degraded view and where $G_k(\mathbf{x})$ corresponds to each of the RGB channels in turn. All colour intensities are scaled between 0 and 1. The most significant difference from [8] is that in our luminance metric $L(\mathbf{x})$ only applies to the top half of the intensity range (*i.e.* $L(\mathbf{x}) = 1$ if the intensities in each channel are less than 0.5) since we are interested in detecting over-exposed rather than under-exposed regions of the stereo frame. The burnout mask $\alpha(\mathbf{x})$ is estimated by applying a threshold, ω_t, to the value of $\omega(\mathbf{x})$ such that

$$\alpha(\mathbf{x}) = \begin{cases} 1 & \omega(\mathbf{x}) > 0.5 \\ 0 & \omega(\mathbf{x}) \leq 0.5. \end{cases} \quad (7)$$

Effectively, the detection rate for α is governed by the choice of the tuning parameters ω_s and ω_l. Our assessments show that values of $\omega_s = 1$ and $\omega_l = 0.5$ give good detection rates. Since the PDE-based interpolation will only affect the regions where $\alpha(\mathbf{x}) = 1$, $\alpha(\mathbf{x})$ is dilated slightly to reduce the missed detection rate around the borders of the burnout region. As long as accurate disparity estimates are possible, false alarms do not adversely affect the quality of the result. Examples of the saturation masks calculated using this technique are shown in Fig. 3.

2.2 Correcting Colour Distortions

The choice of colour correction technique is guided by the nature of the distortion which in turn is dictated by the choice of camera rig. For side-by-side rigs colour distortions tend to be spatially invariant or global in nature. In mirror rigs, additional spatially varying colour distortions are added which requires that separate corrections are performed on local patches. Disparity estimation is necessary to find the corresponding patches in each view.

The examples of burnout which were used to test the proposed algorithm are contained in the Sigmedia Stereo Database[2] and were captured using a side-by-side stereo rig. This implies a spatially invariant colour distortion exists between the views. We use the linear colour transfer algorithm proposed by Pitié *et al.* [11] to match the colour distribution of degraded view to the reference view. As stereo colour distortions tend to be manifested as differences in contrast and brightness levels corresponding to variations in the 1^{st} and 2^{nd} order moments of the colour distributions, a linear distortion model is appropriate. In theory the colour distortion should be estimated only on the regions of the image that are not burnout regions. This would require a disparity estimate to propagate the burnout mask $\alpha(\mathbf{x})$ to the reference view to exclude the corresponding burnout regions. However, from a visual inspection it was observed that estimating the colour distortion on the entire frame did not noticeably reduce the quality of the colour correction. This avoids the need to propagate $\alpha(\mathbf{x})$.

2.3 Disparity Estimation

Disparity estimation between images at different exposures is challenging for most estimation algorithms since changes in illumination tend to adversely affect the match of features. This is further compounded when burnout is present, since the texture content in the over-exposed and reference views is different. As the disparity estimator will then attempt to match texture in the reference view that simply does not exist in the degraded view, the estimation will fail in an unpredictable manner (See Fig. 2).

Colour correcting the degraded view (See Section 2.2) can resolve the illumination issue, however the problem of disparity estimation inside the burnout region remains. Our solution is to distort the colour distribution of the reference view to the degraded view by applying the colour matching algorithm described in [11] and clip the intensity of the reference view where it exceeds the permissible range. This results in burnout in the reference view that is similar to the burnout regions in the degraded view (See Fig. 2) and improves the performance of disparity estimation inside the burnout region.

The disparity estimation process itself consists of two stages. An initial disparity estimate is found using the simple tree method of Bleyer and Gelautz [2]. The estimate is further refined by applying the gradient based motion estimation algorithm of Brox *et al.* [3]. The idea is to smooth out quantisation artefacts of the simple tree method and gives more precisely defined disparity discontinuities around the boundary of foreground objects.

[2] www.sigmedia.tv/stereovideodatabase

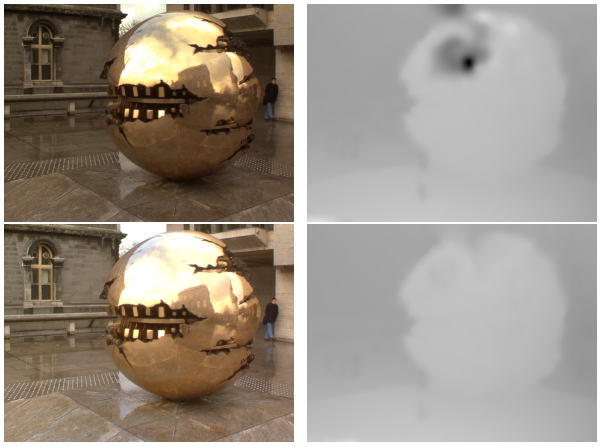

Figure 2: An illustration of the effect of burnout on disparity estimation robustness. The top row shows a portion of the estimated disparity (right) between the reference view (left) and the degraded frame shown on the left of Fig. 1. In the disparity map bright colours represent objects that are closer to the camera. In the disparity estimate (top right) there are distortions in the surface of the sphere as the disparity estimator attempts to match the texture in the reference view to the non-existent texture in the burnout regions of the degraded view. By artificially saturating the reference view (bottom left), the effects of these artefacts can be mitigated (bottom right).

2.4 Burnout Region Interpolation

The burnout regions are interpolated by reconstructing the gradient field of the restored image $\nabla I_{\text{res}}(\mathbf{x})$ inside the burnout regions (*i.e.* for all $\mathbf{x}$ where $\alpha(\mathbf{x}) = 1$). The optimum gradient field is the gradient field closest to the gradient field of the disparity estimate $\nabla I_{\text{ref}}(\mathbf{x} + \mathbf{d}(\mathbf{x}))$ in the least squares sense subject to the constraint that at all pixels outside the burnout region ($\alpha(\mathbf{x}) = 0$) $I_{\text{res}}(\mathbf{x}) = f^{-1}(G(\mathbf{x}))$, the colour corrected degraded view. This is equivalent to the solution of the discretised Poisson PDE with Dirichlet boundary conditions. The form of the equation for burnout detection is

$$\Delta I_{\text{res}}(\mathbf{x}) = \Delta I_{\text{ref}}(\mathbf{x} + \mathbf{d}(\mathbf{x})) \tag{8}$$

where Δ describes the discrete Laplacian operator. Perez *et al.* [10] describe in detail how a linear system of equations are constructed to solve the PDE and the details are not reproduced here. The system of equations is solved using Jacobi optimisation with a coarse-to-fine multigrid of 4 levels. An independent reconstruction is estimated for each of the colour channels.

Special consideration is paid to regions of the image around the border that are occluded in the reference view and results in data interpolation failure if burnout exists in these regions. To overcome this issue, a 2^{nd} interpolation pass is performed in the burnout patches that exist along these borders. An estimate of the affected regions can be made by finding all the pixels in the degraded view whose disparity vectors $\mathbf{d}$ point outside the valid bounds of the reference view. The intensities of the burnout pixels along these borders are computed by solving the Poisson PDE with described by

$$\Delta I_{\text{res}}(\mathbf{x}) = \Delta G(\mathbf{x}). \tag{9}$$

In other words, the gradient field of the restored view is matched to the gradient field of the degraded view. Although

Figure 3: This figure shows a restoration example from 3 different sequences. From left, the 1^{st} and 2^{nd} column shows the degraded and reference views. The 3^{rd} column shows the estimated burnout mask $\alpha(\mathbf{x})$ for the degraded view (burnout is marked white). The final restoration is shown in the last column.

Figure 4: A demonstration of the improved performance of the proposed algorithm (left) over [4] (right). The arrows in the right image right point to a bright halo around the hand caused by a discontinuity artefact introduced by the wavelet interpolation of [4]. This halo is more apparent when the restored sequence is played back as the artefact changes from frame to frame. There is no halo present when the proposed algorithm is applied (left). Note also how the proposed algorithm does not smear the finger at the top of the frame. This is due to the improved robustness of the disparity estimation employed here.

the missing texture in the burnout region is not recovered, the distortion in these regions is reduced.

3. RESULTS

Figure 3 shows 3 stereo pairs containing burnout restored using the proposed algorithm. It shows both the burnout detection matte and fully restored views for each pair. In the example on the first row, the texture on the surface of front of the building as well as the sky in the background is well reconstructed in the restored view (far right). The second example shows the restoration on the sphere example, and shows that not only are the reflections on the surface of the sphere restored but the algorithm can also restore the fringing artefacts on the tree branches located behind the sphere. The 3^{rd} example shows a pair taken from a sequence that is artificially brightened to induce burnout. The burnout regions in this example are elongated and the proposed method out-performs the method proposed in [4] (See Fig. 4).

Figure 5: This figure shows a closeup of the restored views from two consecutive frames of the sequence shown on the top row of Fig 3. An artefact is clearly visible in the right image along the horizontal edge above the column. This causes an popping effect when the restored sequence is played back.

Ultimately, restoring burnout is intended to improve the visual experience when viewing a stereo-video on a 3D display. An example of the visual impact of the restoration on one frame is shown in Fig. 6 which shows an anaglyph rendering of a stereo pair before and after restoration. The anaglyph of the degraded pair has a noticeable glare in the burnout regions compared to the restored anaglyph and results in greater eye strain when viewed for a significant period of time.

Temporal consistency is also crucial to the fidelity of the restoration. To test the temporal consistency of the restoration, we have restored 4 sequences from the Sigmedia Stereo Database[3]. The proposed gradient domain interpolation approach avoids the spatially varying colour distortions present

[3]The restored sequences can be viewed at `www.sigmedia.tv/research/Eusipco2011`

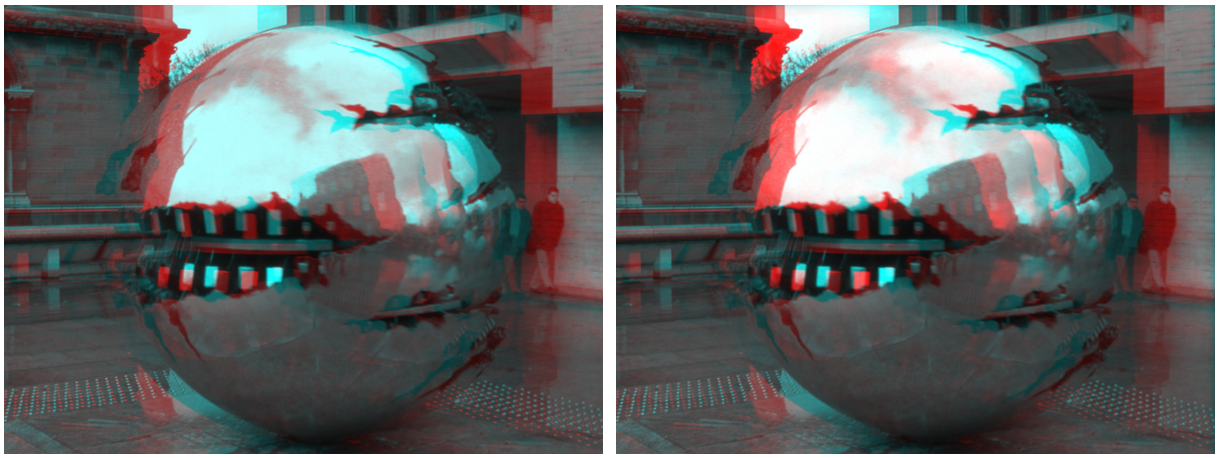

Figure 6: A rendered anaglyph of the stereo pair shown in the 2^{nd} row of Fig. 3 both before (left) and after (right) restoration. The burnout causes a glare in the left view when viewed with red/cyan anaglyph glasses. The restoration reduces this glare. (Note: This is best viewed in the electronic form.)

in [4]. Although the many of the restored frames are temporally consistent, there are some examples that are temporally inconsistent (See Fig. 5). The most visible instances of this are temporally unstable edges and noise like distortions in low texture regions. These failures occur due to temporal inconsistency in the disparity estimates. Therefore, a temporal smoothness constraint should be added to the disparity estimation process to boost temporal consistency. This is the subject of future research.

4. FINAL REMARKS

This paper has proposed a new algorithm to resolve the problem of image burnout in stereo pairs by interpolating the missing texture from the other well-exposed stereo view. A gradient domain interpolation based on solving the Poisson PDE is used to fill the burnout regions while avoiding visible seams across the region boundaries. The paper also proposed a new burnout detection method based on exposure-quality metrics. These advances, along with more robust disparity estimation, ensure that the proposed method outperforms the algorithm of [4], especially for smaller more-elongated burnout regions. Future work will investigate the gain in restoration temporal consistency by applying a temporal smoothness constraint to the disparity energy minimisation. We will also consider a joint iterative approach for disparity estimation and data interpolation.

REFERENCES

[1] A. Agarwala, M. Dontcheva, M. Agrawala, S. M. Drucker, A. Colburn, B. Curless, D. Salesin, and M. F. Cohen. Interactive digital photomontage. *ACM Transactions on Graphics*, 23(3):294–302, 2004.

[2] M. Bleyer and M. Gelautz. Simple but effective tree structures for dynamic programming-based stereo matching. In *International Conference on Computer Vision Theory and Applications (VISAPP) (2)*, pages 415–422, 2008.

[3] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *European Conference on Computer Vision (ECCV)*, pages 25–36, May 2004.

[4] D. Corrigan, F. Pitié, V. Morris, A. Rankin, M. Linnane, G. Kearney, M. Gorzel, M. O'Dea, C. Lee, and A. Kokaram. A Video Database for the Development of Stereo-3D Post-Production Algorithms. In *European Conference on Visual Media Production (CVMP '10)*, pages 64–73, London, UK, November 2010.

[5] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *ACM Siggraph*, pages 369–378, 1997.

[6] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. *ACM Transactions on Graphics*, 21(3):249–256, 2002.

[7] N. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Journal of Applied and Computational Harmonic Analysis*, 10:234–253, May 2001.

[8] T. Mertens, J. Kautz, and F. Van Reeth. Exposure fusion. In *Proceedings of the 15th Pacific Conference on Computer Graphics and Applications*, pages 382–390, Washington, DC, USA, 2007. IEEE Computer Society.

[9] J. Morovic and P.-L. Sun. Accurate 3d image colour histogram transformation. *Pattern Recognition Letters*, 24(11):1725–1735, 2003.

[10] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. *ACM Transactions on Graphics*, 22(3):313–318, 2003.

[11] F. Pitié and A. Kokaram. The Linear Monge-Kantorovitch Colour Mapping for Example-Based Colour Transfer. In *IET European Conference on Visual Media Production*, pages 1–9, December 2007.

[12] F. Pitié, A. Kokaram, and R. Dahyot. Automated colour grading using colour distribution transfer. *Journal of Computer Vision and Image Understanding*, 107:123–137, February 2007.

[13] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley. Color transfer between images. *IEEE Computer Graphics Applications*, 21:34–41, 2001.