EXTRACTION OF NON-BINARY BLOTCH MATTES

Mohamed A. Ahmed, François Pitié, Anil C. Kokaram

Sigmedia, Electronic and Electrical Engineering Department, Trinity College Dublin {ahmedelm@tcd.ie}

Fig. 1. *From left: Original patch, infrared (IR) scan and SDIp detected binary mask. Note how the IR scan shows the level of transparency, and that the SDIp detector overdetects the limit of the blotch.*

ABSTRACT

Automated blotch removal is important in film restoration and typically involves a detection/interpolation step. Current algorithms model the corruption as a binary mixture between the original, clean images and an opaque (dirt) field. This typically causes incomplete blotch removal that manifests as *blotch haloes* in reconstruction. This paper proposes a new approach by modeling the corruption as a continuous mixture between the two components and generating a solution using a Bayesian framework. We use novel priors, propose a computationally efficient scheme for implementation and our results show more complete blotch reconstruction.

Index Terms— Blotch, transparency, Bayesian matting, MRF, removal, opacity, non-binary, matte, frames, fusion.

1. INTRODUCTION

Audiovisual content represents a large part of our inherited culture. Most of this content is still stored on volatile media making them vulnerable to damage due to bad physical handling or normal storage degradations. Blotches are one of the most noticeable degradations. They are formed due to dust particles sticking to the recording medium, preventing light from passing through. As a result, they have random shapes and locations, occluding original data.

Automatic digital removal has become an important task in the postprocessing production workflow for some time now [1, 2, 3, 4]. Any blotch restoration algorithm consists of two main stages: blotch detection and blotch removal. In an ideal scenario, one would use infrared (IR) scans of storage medium for locating blotches. Infrared is transmitted through film but stopped by dust. As a result, the obtained scan is bright in original data regions and dark in dirt, ie. blotched regions. However, in many cases in film and print when an infrared scanner is not available, one is forced to use automatic blotch detection algorithms. Many of these algorithms have been designed for detecting a certain type of blotches known as opaque blotches [1, 2, 3, 4]. Opaque blotches block original data completely. Based on this property, a wide range of spatial and temporal detectors were introduced since 1984 using both heuristics and model based methods e.g. SDI, SDIp and 3D-AR [4]. Temporal detectors search nearby frames for discontinuities, while spatial detectors examine correlations in the current examined frame. It is also possible to combine both temporal and spatial detectors as presented in the work of Tilie et al. [1]. Generally, the output of such algorithms is a binary blotch matte where each pixel is classified either as a blotch or not.

In reality, most blotch pixels show a level of transparency between dirt and original data. Even if the blotch is opaque, the extremities of the blotch are typically transparent. Fig. 1 shows a typical image. Many authors have recognized this in the framework of water damage on still images [5, 6, 7]. Water damage is the result of water particles accumulating on recording medium leading to partially occluded regions known as semi-transparent water blotches. However these ideas have not been exploited in image sequences as yet.

We therefore propose to model an observed, corrupted image pel (blotch) $C(\mathbf{x})$, at site $\mathbf{x}$ as a linear mixture between a clean "background" image $B(\mathbf{x})$ and a fixed opaque colour $F(\mathbf{x})$ (0 for dark blotches and 255 for bright ones) as follows.

$$C(\mathbf{x}) = \alpha(\mathbf{x})F(\mathbf{x}) + (1 - \alpha(\mathbf{x}))B(\mathbf{x}) \qquad (1)$$

Here α is the mixing parameter (opacity matte) and $\alpha = 1$ implies total obliteration of the underlying data. This model is related to that used for digital matting by Chuang et al [8], but the application here is quite different. As the problem is image sequence based we are able to propose more effective priors and in the next section we develop the Bayesian approach to the solution. The novelty here is in the specification of priors and the combination of spatial and temporal solutions to maintain interpolation integrity despite motion errors.
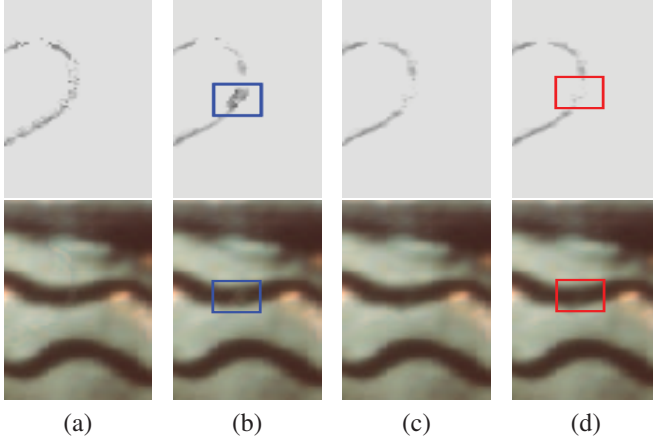
| (a) | (b) | (c) | (d) |

Fig. 2. *Top : Estimated mattes using from left: Bayesian Blotch matting, Method S, Method T, Fusion. Bottom: restorations. Method S reduces noise in Bayesian matting but color bleeding occurs across edges (blue rectangle). The fusion removes both artifacts, but it does not distinguish a dark blotch from a black background (red rectangle).*

2. BAYESIAN BLOTCH MATTING

We use a standard blotch detector (e.g. [4]) to overdetect the blotch regions. This results in an initial segmentation into two regions: 1) definite background and 2) suspected blotch region (see Fig. 1). Our observation model is the same of that proposed by Chuang et al. in equation 1 with the important exception that F is known. In this section we lay out the skeleton of Chuang's method and then build our modifications next. Our task at each pixel site is to refine our estimate for $B(\mathbf{x}), \alpha(\mathbf{x})$ given the observation model in Eq 1. This is an underconstrained problem since there is one equation and two unknowns $B(\mathbf{x}), \alpha(\mathbf{x})$ at each site. The solution is generated by incorporating colour priors (for B, F), and maximizing the conditional probability distribution $P(\alpha, B|C, F)$ at a site as follows (where we drop $\mathbf{x}$ for clarity).

$$\arg\max_{\alpha, B} P(\alpha, B|C, F) = \arg\max_{\alpha, B} P(C|\alpha, B, F)P(B)P(\alpha) \quad (2)$$

The likelihood $P(C|\alpha, B, F)$ is a normal distribution centered at $\hat{C} = \alpha F + (1-\alpha)B$ with variance σ_e^2 set to 1. The background distribution $P(B)$ is modeled as a multivariate normal distribution. The parameters for that distribution are estimated by using nearby definite background samples. This region ($\mathcal{R}$) is taken as a circular patch centered on $\mathbf{x}$ and large enough to contain at least 50 uncorrupted pixels. In Chuang's work, $P(\alpha)$ is constant hence does not effect the maximization process. Substituting background and error models in Eq. 2, taking log and maximizing w.r.t (α, B) as in [8], a set of 4 equations is obtained in the RGB space as follows.

$$\left[\sigma_e^2 R_B^{-1} + I(1-\alpha)^2\right][B] =$$
$$\left[\sigma_e^2 R_B^{-1}\overline{B} + (1-\alpha)C - (I(1-\alpha)\alpha)F\right] \quad (3)$$

$$\alpha = \frac{(C-B).(F-B)}{||F-B||^2} \quad (4)$$

where $(\overline{B}, R_B)$ is the background distribution mean and covariance respectively.

For each suspected blotch pixel, the corresponding (α, B) values are obtained by solving Eq. 3 and Eq. 4 iteratively using the mean of previously calculated opacities as a first estimate. In order to capture the richness of the original, clean data, the patch $\mathcal{R}$ is segmented into several clusters using [9]. A matting solution pair (α_j, B_j) is calculated for each background cluster j and the pair producing the highest conditional probability of Eq. 2 is selected.

Fig. 2a shows the extracted matte and restored data of Fig. 1 using this approach. Both matte and restored data are a bit noisy. This is mainly due to improper background cluster selection as choosing the solution pair (α_j, B_j) with maximum conditional probability of Eq. 2 is risky since different combinations of (α_j, B_j) could give similar probabilities. Increasing the number of background clusters give a better chance of choosing the right cluster, however, it often fails. Chuang et. al. tried to solve this problem by initializing α over the previously calculated values. However, this initialization depends on the scanning pattern and errors usually accumulate quickly. It is well acknowledged that these problems could be resolved by incorporating spatial smoothness into the framework [10]. This is discussed next.

3. SPATIAL PRIORS WITH MRFS

Priors for B, α can be made explicit in the framework using $P(\alpha, B|F, C, B^{\mathcal{N}}, \alpha^{\mathcal{N}}) \propto P(\alpha, B|C, F)P(B|B^{\mathcal{N}})P(\alpha|\alpha^{\mathcal{N}})$ Patch smoothness is imposed by modeling background and opacity values with MRFs on neighborhood $\mathcal{N}$. Assuming Gaussian observation noise and a Multivariate Normal distribution for the background colour B, yields

$$P(\alpha, B|C, F) \propto \exp-\left(\beta_l[||C - \hat{C}_j||^2/\sigma_e^2 + (B - \bar{B})^T R_B^{-1}(B - \bar{B})]\right)$$
$$P(B|B^{\mathcal{N}}) \propto \exp-\left(\beta_b \sum_{k \in \mathcal{N}} \lambda_k ||B_j - B_k||^2\right)$$
$$P(\alpha|\alpha^{\mathcal{N}}) \propto \exp-\left(\beta_a \sum_{k \in \mathcal{N}} \lambda_k |\alpha_j - \alpha_k|^2\right) \quad (5)$$

Recall that the background colour model is a multimodal Gaussian distribution and so the likelihood $P(\alpha, B|C, F)$ enforces the constraint that the combination of α, B should create an image close to the observed $\hat{C}$, as well as ensuring that B comes from one of the clusters in the background model. The priors are Gibbs Energy priors ensuring that B, α are both smooth within an 8-connected neighborhood $B^{\mathcal{N}}, \alpha^{\mathcal{N}}$. λ_k are the usual MRF hyperparameters and $(\beta_l, \beta_b, \beta_a)$ are weights that tradeoff the required level of smoothness. It is important to ensure that the smoothness constraint on B (β_b) is emphasized relative to the smoothness of α (β_a), a fact recognized in [10, 11]. This allows α to be more active than B, hence $\beta_a < \beta_b$. In addition, $\beta_l < \beta_b$ since the α, B pairs are not well discriminated by the likelihood alone. The values

for β's are set differently depending on whether a spatial or spatio-temporal solution (discussed below) is generated. Texture parameters λ_ks are set inversely proportional to gradients at the same locations in the motion compensated frame $\hat{I}_n$.

Solution: A solution to the problem of choosing α, B is created by using the standard Bayesian matting formulation (Eq. 3-4) to generate pairs of α, B candidates assuming that the background consists of 4 components and using the mean of the background cluster $\overline{B_j}$ as an initial guess for B. Hence there are 4 candidate pairs. The problem then is to select the optimal pair over the blotch in question. This is in effect a labeling problem and we employ QPBO graph cuts [12] to estimate the correct label configuration and so generate the solution for α and B respectively. Fig. 2b shows the extracted matte and restored data of Fig. 1 using the current approach (defined as Method S). The extracted matte and reconstructed data are now less noisy than Fig. 2a. However the smoothness constraints can produce bleeding artifacts as seen inside the blue rectangle in Fig. 2b.

4. SPATIO-TEMPORAL FUSION

Since blotches do not occur in the same place in consecutive frames (usually), then it is sensible to use the motion compensated previous or next frames (I'_{n-1}, I'_{n+1}) to generate a more accurate background model in the region of the blotch under consideration. Motion compensation is achieved here using a simple block matching search with a block size configured to include at least 100 uncorrupted pixels. To make this process more robust to blotch corruption, the α estimated using Method S in the current frame I_n are used as weights to weight out the effect of the blotch approximately. Background models $P(B|I_n, I_{n-1}, I_{n+1})$ are then estimated for each pixel at site $\mathbf{x}$ in the current frame n, from a 3x3 patch centered on $\mathbf{x}$ in the motion compensated frame $\hat{I}_n$. The small background patch size imposes further matte smoothness as it is expected for small data patches to vary smoothly from one pixel to another. Fig. 2c shows the extracted matte and restored data of Fig. 1 using this temporal approach, (defined as Method T). Extracted matte and reconstructed data are less noisy than Fig. 2a and nearby regions bleeding, as seen inside the blue rectangle in Fig. 2b, is reduced.

It is clear that when motion can be well estimated between frames Method T is a useful approach, but when there is pathological motion or severe corruption in the images, the spatial solution S would be an appropriate fallback. We propose then to generate an overall solution based on the fusion of Method S and T. Hence, for each pixel in the suspected blotch region, there are two possible solutions: (B_s, α_s) and (B_t, α_t) with negative log likelihoods $\mathcal{L}_s$ and $\mathcal{L}_t$ respectively. The task is to select one of these two solutions depending on the accuracy of the motion compensation. This accuracy can be related to the size of the displaced frame difference (DFD) between the corrupted and motion compensated frames. When

that DFD is large, the spatial solution should be favored, and when it is small, the temporal solution should be favored. We therefore modify the negative log likelihoods in Eq. 5 spatially and temporally to $\mathcal{L}'_s$ and $\mathcal{L}'_t$ respectively to account for this issue as follows.

$$\mathcal{L}'_s = ||C - \hat{C}_j||^2/\sigma_e^2 + (B_s - \bar{B}_s)^T R_{B_s}^{-1}(B_s - \bar{B}s)] + (Q - DFD)^2$$
$$\mathcal{L}'_t = ||C - \hat{C}_j||^2/\sigma_e^2 + (B_t - \bar{B}_t)^T R_{B_t}^{-1}(B_t - \bar{B}_t)] + DFD^2 \quad (6)$$

Q is a constant that can be related to a motion threshold and here we set it to 30 grey scale levels. When DFD is close to Q, $\mathcal{L}'_t$ is greater than $\mathcal{L}'_s$ and so the spatial solution is favored, while the opposite is true when DFD is small. Hence we have favored the motion based solution when the motion compensation has performed well. Using the same framework as in Sec: 3, the fused solution (B_x, α_x) is achieved by maximizing the global probability $P(\alpha, B|F, C, B^{\mathcal{N}}, \alpha^{\mathcal{N}}, \hat{I})$ using QPBO graph cuts as before. $(\beta_l, \beta_b, \beta_a)$ is set to $(0, 1, 0)$ in Method S and $(1, 1, 0)$ for spatio-temporal fusion to take into account the change in the relative importance of the energies between methods. Fig. 2d shows the extracted matte and restored data of Fig. 1 by fusing both Method S and T. Note how the fusion tends towards Method T due to good motion estimation and achieves an overall improved result.

5. RESULTS AND DISCUSSION

Fig. 3 compares results using Bayesian blotch matting and by fusing both Methods S and T. As a reference, temporal restoration results using a binary underdetected mask are displayed (overdetected masks can also lead to poor reconstruction). Our method does not produce visible artifacts and performs more believably than the reference binary blotch removal or Bayesian Matting. We have also found that the process is more robust to motion errors since the spatial method provides a nice fallback in case of pathological motion.

Note that the algorithm cannot distinguish a blotch from a dark background as it is assumed that F is black. As a result, a blotch lying on a black region gets classified as background ($\alpha = 0$). However, the restored data remains meaningful. Fig. 2d shows an example of this scenario where the fused matte represent complete transparency in the corresponding black region of the restored image (red rectangle), however, the reconstructed background is still consistent with the rest of the uncorrupted image which is acceptable for viewing.

Quantitative Evaluation: Ground truth mattes are derived from IR scans by gamma correction. Mean Squared Error (MSE) between this ground truth and the estimated mattes were measured for Method (C,F) and yielded for Fig. 3(a)-Fig. 3(e): $(0.0263, 0.0142)$, $(0.0159, 0.0161)$, $(0.0056, 0.0053)$, $(0.0088, 0.0051)$, $(0.0063, 0.0047)$. In most cases, Fusion was able to reduce MSE over Method C except for Fig. 3(b). In this case, the fused matte is mainly transparent as being lying on a dark background leading to incorrect matte extraction but correct background reconstruction. The average pro-

cessing time for a 720x576 frame is 25-30 seconds using 2.33 GHz Quad Core Processor and coding with MATLAB. Most of the time is consumed in the Bayesian Matting implementation while the rest is consumed in QPBO and a coarse motion estimation algorithm. QPBO is implemented efficiently by processing the small blotch masks only. This represents a coverage on average of much less than 1% of the picture data.

In conclusion then, the novelty of our algorithm is the exploration of blotch transparency. Our model for blotch transparency has been successfully combined with a practical initialization strategy (using SDIp) to generate useful restoration results. Our algorithm produces less noisy results than Bayesian blotch matting, maintains texture structure and detects blotch boundaries accurately. In the future we hope to explore other constraints on the underlying images e.g. using AR models for imposing smoothness.

6. REFERENCES

[1] S. Tilie, L. Laborelli, and I. Bloch, "Blotch detection for digital archives restoration based on the fusion of spatial and temporal detectors," in *9th International Conference on Information Fusion*, 2006, pp. 1–8.

[2] O. Buisson, B. Besserer, S. Boukir, and F. Helt, "Deterioration detection for digital film restoration," in *CVPR*, 1997, pp. 78–84.

[3] R. Storey, "Electronic detection and concealment of film dirt," *UK Patent Specification no. 2139039*, 1984.

[4] A.C. Kokaram, R.D. Morris, W.J. Fitzgerald, and P.J.W. Rayner, "Detection of missing data in image sequences," *IEEE trans. on Image Proc.*, vol. 4, no. 11, pp. 1496–1508, 1995.

[5] A. Greenblatt, K. Panetta, and S. Agaian, "Restoration of semi-transparent blotches in damaged texts, manuscripts, and images through localized, logarithmic image enhancement," *IS-CCSP 2008*, pp. 484–489, 2008.

[6] A.J. Crawford, V. Bruni, A.C. Kokaram, and D. Vitulano, "Multi-Scale Semi-Transparent Blotch Removal on Archived Photographs using Bayesian Matting Techniques and Visibility Laws," in *ICIP*, 2007, vol. 1, pp. 561–564.

[7] V.Bruni, A.Crawford, D.Vitulano, and F.Stanco, "Visibility based detection and removal of semi-transparent blotches on archived documents," in *VISAPP*, 2006.

[8] Yung-Yu Chuang, Brian Curless, David H. Salesin, and Richard Szeliski, "A Bayesian Approach to Digital Matting," in *IEEE CVPR*, December 2001, vol. 2, pp. 264–271.

[9] M.T. Orchard and C.A. Bouman, "Color quantization of images," *IEEE Trans. on Signal Processing*, vol. 39, no. 12, pp. 2677–2690, 1991.

[10] P.R. White, W.B. Collis, S. Robinson, and A. Kokaram, "Inference matting," in *CVMP*, 2005, pp. 168–172.

[11] A. Levin D. Lischinski and Y. Weiss, "A closed form solution to natural image matting," *IEEE Trans. on PAMI*, 2008.

[12] Endre Boros, P. L. Hammer, and X. Sun, "Network flows and minimization of quadratic pseudo-boolean functions," *Technical Report RRR 17-1991, RUTCOR*, 1991.

| (a) | (b) | (c) | (d) | (e) |

Fig. 3. *From Top: Blotched image, corresponding infrared scan, temporal restoration using underdetected SDIp mask, overdetected SDIp mask, extracted matte using Bayesian blotch matting, extracted matte by fusing Method S and T, Bayesian blotch matting restoration and our restoration. $(\beta_l, \beta_b, \beta_a)$ are set to $(0, 1, 0)$, $(1, 1, 0)$ for Method S and Fusion respectively, while Q set to 30 grey scale. Note how our approach is able to produce smoother results and maintains texture structure c.f Bayesian blotch matting.*