Image Colourization Using Artificial Intelligence

Abstract

Almost one and a half centuries of using black and white cameras has left a huge stock of images of historical figures, events and even nostalgic memories for many. Colourization of such images was usually done manually using digital image processing programs. However, this method is extremely time-consuming and requires experienced designers and artists who have a deep understanding of shading and colours. Owing to the recent advancements in artificial intelligence and computational technology, numerous studies have been published for automating this process. The very first of such algorithms were based on predefined methods but most of the research has shifted to data-driven machine learning models. The most successful models developed are semi-automated systems which need some simple inputs from the user, removing the need of a specialist. Research continues towards a fully automated system which can even guess the base colour of an object in the photograph. This paper starts by summarizing the challenges faced by the traditional algorithms and the necessary requirement of human inputs,followed by technical descriptions of methodologies used by a few key papers for automated colourizing and comparing the results and expectations. As accuracy is not a good measure of the colourizing ability, most studies used participants for rating the final result manually. Finally, we discuss the problems and challenges that researches have faced and some of the ideas proposed for improvements.

1. Introduction

After a century of black and white cameras, the first coloured cameras were invented and brought into public use around the 1950s, yet black and white cameras were prominent until the late 20th century. This has left us with a hundred and fifty years of greyscale photographs. Some of these photographs capture important historical events like The Solvay Conference of 1927. Colourization is also done for personal photographs to relive the memories of old time. This created a demand for specialized designers that colourized images manually in tools like gimp, for which designers have to be paid and significant time in order of hours is required, therefore it was not accessible to the common public economically. After computers became fast and accessible, many researchers have come up with automated solutions as opposed to manual colourization. The key difference between these procedures is that some information is obscured to a computer program in greyscale images themselves, the artists colourizing historical images do lots of research on the details to find the exact colours, also the weather conditions and season affect the tint colour. This information has to be provided in semi-automated solutions as well as manual colourization.

One of the first such systems were developed by Horiuchi (2004, p. 197) using probabilistic relaxation method which assumes Markov property of the input images, and Yatziv and Sapiro (2004, p.121) also published a non-learning algorithm based on luminance properties. Both of these are based on fixed mathematical model utilising concepts from probability theory, the latter being an improvement due to its less extensive computation. Welsh (2002, p. 277) proposed an innovative method of transferring the colour information from one image to another black and white image, this allows the algorithm to capture true colours of the environment, i.e. taking into account weather, season and location.

A solution to a simplified problem is given by Chang, Wei and Lee (1997, p. 165) and Sýkora et. al. (2004) which is colourizing cartoon images and videos, this process is much simpler than real images due to the lack of shading and not being constrained to colour choice. Machine learning algorithms have been the centre of focus in recent research, much of the work is done on variations of neural networks. Some of the most popular works using neural networks are Iizuka, Serra and Ishikawa (2016, p. 1) and Zhang, Isola and Efros (2016, p.649). The most successful of these neural networks are Convolutional Neural Networks (CNN) which have become the starting point of image processing in most of the machine learning models with images as input.

2. Traditional Non-Learning Algorithms for Colourization

These algorithms are primitive algorithms being proposed at the time machine learning was not popular due to computational power and data management available at that time, these algorithms work directly with the input without training on other photographs, thus the inner mechanism for these have better explainability than machine learning algorithms.

2.1 Probabilistic Relaxation

Horiuchi (2004, p. 197) is based on the Markov property assumption which states that the colour of a pixel is dependent only on its adjacent pixels. A combinatorial optimization problem is constructed that minimizes the pixel-wise differences in colours, and is dependent on some known colour values given beforehand. The performance of the model increases with the percentage of known colours. This property is exploited along with the knowledge of some known pixel values and the luminance extracted, the output is obtained by an optimization model, artefact free images were obtained not below than 10% prior colour information given.

2.2 Chrominance Blending

Yatziv and Sapiro (2004, p.121) introduced improvement over the algorithm given by Horiuchi (2004, p. 197) in terms of performance, this model can be used in both videos and images. It also solves the issue of artefacts for less prior information by using blending, this blending is derived from a weighted distance function efficiently computed from the luminance channel–a measure of light intensity at a point in the image. The algorithm can even be extended to re-colourization.

3. Variants of Image Colourization

Work has been done in problems similar to image colourization which has allowed advancements in the original problem of colourizing black and white photographs. Here we look at two such problems.

3.1 Colour Transfer Techniques

Colour transfer techniques in Reinhard et. al. (2013, p. 34) are widely used for recolouring a coloured image using a different coloured image as a reference. These techniques compute colour statistics in both the input and reference images and then establish mapping functions which map the colour distribution of a reference image to the input image. Welsh et al. (2002, p.277) proposed a general technique to colourize greyscale images by matching the luminance and texture information between images. This technique was improved by using a supervised classification scheme that analyzed low-level features Irony et al. (2005). Charpiat et al. (2008, p.126) proposed a global optimization framework that deals with multi-modality to predict the probability of possible colours at each pixel. Gupta et al. (2012) matched superpixels between the input image and the reference image using feature matching and space voting to perform the colourization.

3.2 Colourization in Cartoons

Sýkora, Buriánek, and Žára (2004) combined image segmentation, patch-based sampling and probabilistic reasoning for automated colourizing. This paper aims to colourize the old black and white images of cartoons, especially the digitized version of cartoons made from celluloid films using paper and cell shading. The whole colourization pipeline consists of five independent phases: image segmentation, foreground layer colour prediction, colour brightness modulation, background layer reconstruction and colourization and the final composition. However, the prediction performance of this method is greatly limited by the number of local structural matches between the example and target images. It may completely fail when the example and target images are widely dissimilar.

3.2.1 Image Segmentation

For the simplification of colour transfer into black and white cartoons, Sykora et al. (2004) suggested using unsupervised image segmentation. This method is suitable especially for cartoons which consist of two planar layers (background and foreground). The dynamic foreground layer contains homogeneous regions surrounded by visible outlines while the background layer is usually a more complicated textural image which remains static during the animation. This important property allows us to divide the original greyscale image into a set of regions using robust outline detectors and classify them roughly as foreground or background via region size thresholding.

4. Neural Networks and Machine Learning

A neural network is a series of algorithms that recognizes underlying relationships in a data set through a process that mimics the way the human brain operates. In this sense, a neural network refers to a system of neurons, either organic or artificial in nature.

Neural networks can adapt to changing input due to which the network generates good results without explicitly redesigning the output criteria. Such models are being used for classification tasks and low dimensional outputs that are being expanded into many tasks in recent times. They are applied to different tasks with images as an output like optical flow, superresolution, contour detection and semantic segmentation. Most of these are based on convolutional neural networks (CNNs) that can process images of any resolution and have less computational cost than even the simplest neural networks, artificial neural networks (ANNs), as a lesser number of parameters have to be trained.

Cheng et al. (2015) proposed a fully automatic approach in which various features are extracted and the different patches of the image are colourized using a small neural network. Joint bilateral filtering is used to improve the results. They used very little training data which greatly limits the type of images it is applicable to. Furthermore, they require a high-performance segmentation model to provide segmentation of the image.

4.1 CNN-Based Approach

Iizuka, Serra and Ishikawa (2016) used a novel CNN-based approach to automatically colourize greyscale images. It combined global-level features, e.g. day-night and indoor-outdoors information and the local features, e.g. the object boundaries and corners in an image, using CNNs. The deep network features a fusion layer that allows merging global information computed using the entire image with local information dependent on smaller image patches. The complete framework, including the global and local priors and the colourization model, is trained in an end-to-end manner. The performance of the results was determined by a user study with 92% natural results done on 1500 images. Apart from colourizing, the same model is shown to be able to perform style transfer as well as classification tasks. The solution works at realtime compatible costs, and can thus be used in videos as well.

4.2 GAN-Based Approach

Cao et. al (2017, p. 151) proposed a novel solution to automatically generate diverse colourization schemes for a greyscale image while maintaining the accuracy by exploiting conditional generative adversarial networks which not only solved the sepia-toned problem of other models but also enhanced the colourization diversity. The paper aims to colourize images without any prior data about the image to resolve the dependency on experts for providing colour scribbles. They introduced a novel generator architecture which consists of a fully convolutional non-stride structure with multi-layer noise to enhance diversity and multi-layer condition concatenation to maintain reality. The Turing test on humans resulted in a convincing 80% performance for an unsupervised algorithm.

5. Analysis and Discussion

In the statistical algorithms, the results were heavily dependent on prior user inputs, i.e. colour switches for a given photograph are necessary for accurate colouring. The neural network-based algorithms have performed better with lesser user interaction, the Cao et. al (2017, p. 151) model was completely unsupervised and still produced convincing results. Moreover, the CNN-based models are much faster in transforming images as the convolutional operation can be applied on the fly using graphical processing units (GPU) processors. The only advantage of the statistical algorithms by Horiuchi (2004, p. 197) and Yatziv and Sapiro (2004, p.121) is that a training set of images is not needed, however, initial input for some pixel values is still required.

The model given by Sýkora, Buriánek, and Žára (2004) is an unsupervised non-learning method based on segmentation but it is only limited to cartoon images with no shading. For most of the general-purpose uses, the newer algorithms are good enough, given that we are providing the colour data for a few points in the image. For mass conversion of images and colourizing videos, the unsupervised algorithms have decent results, but if desired, for important photographs, an expert would still produce much better results.

6. Limitations

A major limitation of computer-based image colourization methods is that a lot of extrinsic information that defines the choice of colour is not present in the black and white images as is. This fundamental problem is partially solved by making user-specific scribbles, but even that fails to consider the environmental factors. These factors can only be found by thoroughly investigating the historical texts and records of that time period, which is an expensive and time consuming process.

Neural networks based algorithms also require different kinds of training photos for different conditions, compiling such images is a task that still has to be done manually.

7. Conclusion

A realistic image colourization procedure is in popular demand for both personal use cases and viewing historical photographs from a new perspective. In this review article, we have discussed the different innovations made in image colourization, and compared the performances which were determined by human participants through Turing tests. We also noted the limitations of computational requirements and the need for prior input.

References

An, X. and Pellacini, F., 2008. AppProp: all-pairs appearance-space edit propagation. In ACM SIGGRAPH 2008 papers (pp. 1-9).

Cao, Y., Zhou, Z., Zhang, W. and Yu, Y., 2017, September. Unsupervised diverse colorization via generative adversarial networks. In Joint European conference on machine learning and knowledge discovery in databases (pp. 151-166). Springer, Cham.

Chang, C. W., and Lee, S. Y. 1997. Automatic cel painting in computer assisted cartoon production using similarity recognition. The Journal of Visualization and Computer Animation 8, 165–185.

Charpiat, G., Hofmann, M. and Schölkopf, B., 2008, October. Automatic image colorization via multimodal predictions. In European conference on computer vision (pp. 126-139). Springer, Berlin, Heidelberg.

Cheng, Z., Yang, Q. and Sheng, B., 2015. Deep colorization. In Proceedings of the IEEE International Conference on Computer Vision (pp. 415-423).

Chia, A.Y.S., Zhuo, S., Gupta, R.K., Tai, Y.W., Cho, S.Y., Tan, P. and Lin, S., 2011. Semantic colorization with internet images. ACM Transactions on Graphics (TOG), 30(6), pp.1-8.

Gupta, R.K., Chia, A.Y.S., Rajan, D., Ng, E.S. and Zhiyong, H., 2012, October. Image colorization using similar images. In Proceedings of the 20th ACM international conference on Multimedia (pp. 369-378).

IEEE transactions on image processing, 15(5), pp.1120-1129

Iizuka, S., Simo-Serra, E. and Ishikawa, H., 2016. Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. ACM Transactions on Graphics (ToG), 35(4), pp.1-11

Ironi, R., Cohen-Or, D. and Lischinski, D., 2005, June. Colorization by Example. In Rendering Techniques (pp. 201-210).

Levin, A., Lischinski, D. and Weiss, Y., 2004. Colorization using optimization. In ACM SIGGRAPH 2004 Papers (pp. 689-694).

Pan, Z., Dong, Z., and Zhang, M. 2004. A new algorithm for adding color to video or animation clips. In Proceedings of WSCG – International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, 515–519.

Reinhard, E., Adhikhmin, M., Gooch, B. and Shirley, P., 2001. Color transfer between images. IEEE Computer graphics and applications, 21(5), pp.34-41.

Sýkora, D., Buriánek, J. and Žára, J., 2004, June. Unsupervised colorization of black-and-white cartoons. In Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering (pp. 121-127).

Welsh, T., Ashikhmin, M. and Mueller, K., 2002, July. Transferring color to greyscale images. In Proceedings of the 29th annual conference on Computer graphics and interactive techniques (pp. 277-280).

Yatziv, L. and Sapiro, G., 2006. Fast image and video colorization using chrominance blending. Horiuchi, T., 2004. Colorization algorithm using probabilistic relaxation. Image and Vision Computing, 22(3), pp.197-202.

Zhang, R., Isola, P. and Efros, A.A., 2016, October. Colorful image colorization. In European conference on computer vision (pp. 649-666). Springer, Cham. [Online]