

A Decision Tree Based Image Enhancement Instruction System for Producing Contemporary Style Images

Meng-Luen Wu^(✉) and Chin-Shyurng Fahn

Department of Computer Science and Information Engineering,
National Taiwan University of Science and Technology,
Taipei 10607, Taiwan, ROC
{dl0015015, csfahn}@mail.ntust.edu.tw

Abstract. In this paper, we have proposed an image enhancement method by contemporary aesthetics criteria, which enables computers to produce visually favorable images automatically. The contemporary aesthetics criteria is obtained through data mining algorithms such as decision tree, support vector machine, and neural networks. In order to make computers adjust the images automatically to make them match the contemporary aesthetics criteria, the tree-based classification method is proposed for enhancement instructions. Our proposed system finds the reasons in the tree why an input image is not perceptually favorable and give improvement instructions accordingly. The training features are based on enhancement instructions, such as color component, saturation, sharpness, and so on. Preprocessing methods are also proposed for a more efficient labeling and better accuracy for image classification. The training samples are from both contemporary style high aesthetics quality images and those are not, which more than 15,000 training samples are used. The accuracy of our proposed method is above 95 %. The experimental result shows our system can give users or computers appropriate enhancement instruction efficiently.

Keywords: Professional photographing · Computational aesthetics · Autonomous photographing · Data mining · Image retouching · Image enhancement

1 Introduction

Users who upgrade their camera to high-end ones because they want more appealing pictures. Although high-end cameras provide more functionalities than low-end ones, they are not guaranteed to take better photos. A good image depends on user's handling, and the environmental light sources, as well as the content in the field of view. Automatic image enhancement in real-time can help user take better pictures without any manual operations. In existing camera products, the automatic adjustments focus on image quality enhancement, such as brightness, sharpness, and so on. The existing adjustments help producing clear pictures but will not necessarily produce appealing ones.

Figure 1 shows three images of the same scene. The three images are original raw image from image sensor, image enhanced by camera, and image enhanced by a



Fig. 1. Three images of the same scene: (a) original raw image from image sensor; (b) image enhanced by camera; (c) image enhanced by software darkroom.

software darkroom, and the difference among them is noticeable. The software darkroom is a key for popular and perceptually favorable images, which the software darkroom enables user to modify the contrast, sharpness, color component, brightness, and so on. The digital darkroom is usually operated manually. The digital darkroom is used for post-enhancement so it has limitations. For example, when an image is underexposure that some pixel values are zero, increasing the brightness of the image is useless for restoring the contents of these pixels. In this case, we know that to produce better images instructing the camera users to handle the cameras well is better than using the software darkroom, and the instructions must run in real-time.

If we deem the software darkroom as a post-enhancement for a popular image, the good handling of professional cameras is a pre-enhancement. Compared to post-enhancement, the pre-enhancement has more possibilities because the environmental conditions are still adjustable. Therefore, our proposed methods focus on autonomous handling instructions, which is the pre-enhancement method. In this paper, we proposed a real-time instruction system for users to adjust their camera parameters appropriately to get better and more appealing pictures. The adjustments can also be performed by machines automatically. The main challenge for this topic is that the concept of a good images differs from time to time, so we have to find out the criterions for popular contemporary style images, which is performed by data mining algorithms.

2 Feature Extraction

Choosing the right image aesthetics features is essential for both clustering and classification of professional and non-professional images, which is used to predict whether an image is possibly favorable [1].

In this paper, we focus on efficient image enhancement instructions that helps an amateur picture become professional. The features include color, brightness, contrast, sharpness, and so on, which are proposed as follows:

The color component is extracted by the following equation:

$$f_{colorcomponent} = \sum_{x=1}^{width} \sum_{y=1}^{height} \frac{D(c_l, c(x, y))}{width \times height} \quad (1)$$

where the width and height in the equation are for the image, and D is the Euclidean distance between two colors in CIELab color space, c_l is the color component to be

extracted in the CIE Lab color space. $c(x, y)$ is the color value of coordinate (x, y) in the same color space.

The achromatic degree can be obtained by:

$$f_{achromatic} = \sum_{x=1}^{width} \sum_{y=1}^{height} \frac{|\{(x, y) | Ch_R(x, y) = Ch_G(x, y) = Ch_B(x, y)\}|}{width \times height} \quad (2)$$

where Ch_R , Ch_G , Ch_B are for the intensity of red, green, and blue channel of a pixel respectively.

Six features according to this formula, which are red, green, blue, magenta, cyan, and yellow components.

Professional images are usually higher in contrast, which is defined as:

$$f_{contrast} = \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1 - d(i, j)) \frac{D(i, j)}{A_i A_j} \quad (3)$$

where $d(i, j)$ is the spatial distance between the centroids of two segmentations i and j ; $D(i, j)$ is the color distance between the two segmentations in the CIE Lab color space.

The color saturation feature is defined as:

$$f_{saturation} = \sum_{x=1}^{width} \sum_{y=1}^{height} \frac{s(x, y)}{width \times height} \quad (4)$$

where s is for the saturation of a pixel in HSV color space.

The degree of sharpness and blur of an image is stated as follows:

$$f_{sharpness} = \frac{|\{(u, v) | |F(u, v)| > \xi\}|}{width \times height} \propto \frac{1}{\sigma} \quad (5-1)$$

$$f_{blur} \propto \frac{1}{f_{sharpness}} \quad (5-2)$$

where $I_{blur} = G_\sigma * I$ is the blurred image derived through convolving the original image I with a Gaussian filter G_σ , and $F(u, v) = FFT(I_{blur}(x, y))$ is the blurred image transformed into the frequency domain via the fast Fourier transform. Here, ξ is set to 5.

In [2], the formula of the simplicity feature yields:

$$f_{simplicity} = \left(\frac{|\{l | k(c_l) \geq \gamma k_{max}\}|}{4096} \right) \times 10 \quad (6)$$

where $k(c_l)$ is the color count for color c_l , k_{max} is the maximum color count, and γ is set to 0.001. In this formula, the number of colors in the image is reduced to 4096; that is, the color counts of R, G, and B are all reduced to 16.

For an input image, the global brightness can be obtained by the following equation:

$$f_{brightness} = \frac{\sum_{x=1}^{width} \sum_{y=1}^{height} I(x, y)}{width \times height} \quad (7)$$

where $I(x, y)$ is the intensity of a pixel at (x, y) .

The chosen image features are as follows, $f_1 \sim f_6$ are color components, which are red, green, blue, cyan, magenta, and yellow. f_7 is the degree of achromatic. f_8 is the color contrast. f_9 is color saturation. f_{10} is sharpness. f_{11} is simplicity. f_{12} is the brightness.

3 Data Preprocessing

In this paper, the criterions of popular contemporary images are found via data mining. For better data mining results, the training data should be preprocessed properly. In this section, the feature selection and data clustering methods used in our system and are described as follows:

3.1 Feature Selection

The feature selection chooses a useful subset of features from the full subset of the training data. The benefit of feature selection is that less features need to be extracted. In addition, not every features are useful for analysis, which some of them would even influence the accuracy. This is because some feature values are redundant which are useless for analysis according to the information theory. What's worse, data belong to some features can be noisy and not correct. Therefore, feature selection is necessary for better analysis results.

The preliminary goal for feature selection is to find decisive features for data classification. The correlation feature selection (CFS) [3] selects the best feature subsets that have the best contribution to the training set. The CFS assumes that good feature subsets contain features that are highly correlated with the classification, and uncorrelated to others.

3.2 Data Clustering

In our proposed method, data clustering before classification is the major key for the success of training. The training samples are clustered into many groups for labeling. The images in the same group share similar features values, while images belong to different groups are perceptually different. Label the groups which are preferred by professional photographers as "favorable," and the other are labeled as "not favorable." Two preferred images may not be in the same group and they may perceptually different. For example, images under day light and in night scenes are visually different

and they will be clustered into different groups, but some groups of both of them are preferred by professionals as they have other subset features in common. To be brief, the purpose of data clustering is to group images with similar features together, and each groups can be labeled as “favorable” or “not favorable.”

We choose K-means algorithm for clustering. K-means algorithm is simple yet the number of clusters is controllable. When there are too much groups, the labeling is too detailed and does not give good classification result. On the other hand, when clustered into few groups, the labeling is too rough and still the classifier does not give good accuracy. Empirically, 20 groups is appropriate for labeling.

When the clusters are appropriately labelled, data belong to all labeled groups are set as inputs for the classifier, and the classifier will build separation boundaries for all clusters, which is a means of hybrid learning. The concept is shown in Fig. 2. In next section, the decision tree algorithm is used as one of the classifiers which generates readable rules for classification with the labeling.

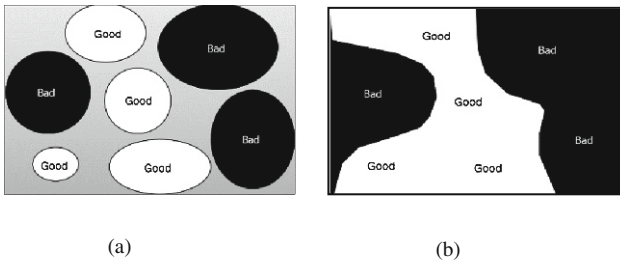


Fig. 2. Hybrid learning: (a) clustered data, which two clusters are distributed in the feature space; (b) a possible separation for two classes of data in the feature space with their respective labels. (Good = favorable by experts; Bad = not favorable by experts)

4 Decision Tree Based Instruction System

In this section, the method for autonomous photographing instruction is proposed. The tree based classification method is the core algorithm. In addition to our proposed methods, the limitations are also explained in this section.

4.1 Decision Tree

Decision tree [4] is a popular supervised learning approach because the decision process walks through the path of the tree, and each path can be written as a rule. In a decision tree, the nodes except the leaves represent a feature, and its child edges are predicates for the feature, such as “value larger than,” or “value less than.” A node without any children is a leaf node. The class labels are put in the leaf nodes of the decision tree.

Decision tree algorithms are based on information theory [5], which the main idea is to calculate the entropy of the classes of data when data are separated using specified feature and branching value.

A major problem for decision tree is over-fitting, [6] which can be solved by pruning. If the degree of the tree is too high, some unnecessary nodes are produced near the leaf, and lowers the accuracy of the decision tree. Therefore, pruning is needed for the tree. When a node is pruned by replacing it with a class leaf, and the accuracy of the tree is better after the replacement, the pruning is accepted, otherwise keep the original sub-tree.

4.2 Instruction Algorithm

When the decision tree is a binary tree, if the value of a node’s feature is greater than a specific value, it will be branched to a child, otherwise it will be branched to the other.

The features are abbreviated as $f_1, f_2, f_3, \dots, f_n$ respectively. For example, in Fig. 3, The red edges is decision path in the tree.

Given a path shown in red in Fig. 3(a), when the value of f_4 is less than or equal to 57.46, the image is classified as “bad.” (not favorable) If all features are conditionally independent in this decision tree, and f_4 has exactly two children, the image can be classified to “good” (favorable) simply by making the value of f_4 greater than 57.46, as shown in Fig. 3(b).

Through the decision process, the conditions that cause an photo not favorable can be known by user, and is improved by increasing or decreasing a feature in the image.

The decision path’s leaf node should have one leaf node sibling. For example, in Fig. 4(a), an image is classified as “bad.” (not favorable) if we adjust the value of f_2 to be greater than 41.91, the path will lead to another child of the f_2 , which is a sub-tree shown as orange in Fig. 4(b). The sub-tree does not guarantee that the adjusted image will be classified as “good,” (favorable). Therefore the instruction does not necessarily give expected result in this situation.

A feature in the decision path cannot have multiple occurrences, otherwise the adjustments may lead the path to another sub-tree. For example, in Fig. 5(a), an image is

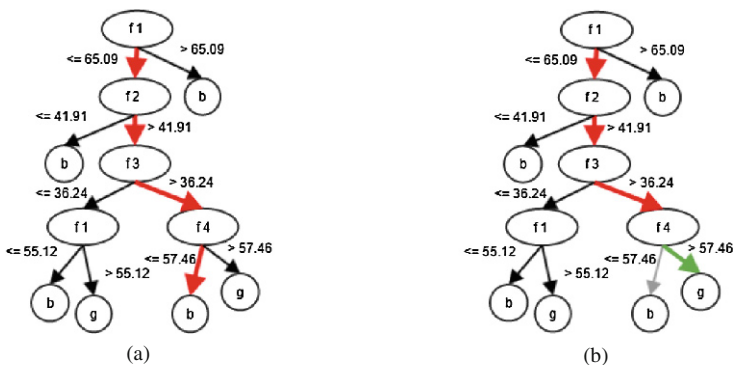


Fig. 3. The decision process of a decision tree and the instruction given method: (a) an image is classified as bad (not favored by expert) with the red path; (b) an instruction is given that user changes the environmental condition which makes the last edge in the decision process go to the green edge, and finally the image can be classified as good (favored by experts) 4.3 Limitations. (Color figure online)

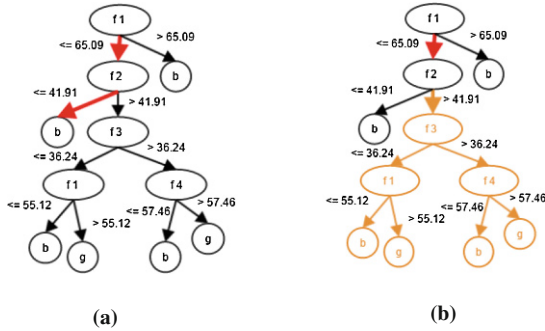


Fig. 4. Limitations for the instruction system: (a) an image is classified as bad (not favored by experts) with the red decision path; (b) altering the value of f2 is not effective because the decision path will lead to a subtree with two classes at leaves. (Color figure online)



Fig. 5. A limitation that the suggestion cannot be given: (a) an image is classified as bad (not favored by experts) with the red decision path; (b) altering the value of f1 near the leaf node will also change the path of the root node in the tree, which is also branched by f1. (Color figure online)

classified as “bad,” (not favorable) and the decision path have two “f1”s in it. If the value of f1 is adjusted to 60, fortunately, it will be classified as “good” (favorable) with the green edge in Fig. 5(b). However, if the value of f1 is adjusted to be greater than 65.09, it will be classified as “bad” (not favorable) with the red edge shown in Fig. 5(b).

Finally, it is recommended that two features in the decision tree must be conditionally independent, which means an adjustment of one feature must not influence another. For example, if “f1” is the feature of “red component” and “f2” is the feature of “green,” the enhancement of “red component” also improves “green” at times. Therefore, when training the decision tree, all selected features for training must be conditionally independent. For example, the feature of sharpness does not always influence those of colors.

5 Experimental Result

In this section, the experimental results performed on our proposed system is shown. First, we will introduce the experimental system setup. Second, the accuracy of image classifiers is described.

5.1 Experimental Setup

The experimental setup is as follows: The image sensor is 1/2.3", 20.7 MP image sensor with 27 mm lens and the aperture value is f/2.0. The resolution of the image in the software is 1920×1080 when capture and is downscaled to 320×240 for processing.

In the experiment, we choose 12 prominent features for training the decision tree, which are red, blue, green, cyan, yellow, magenta, achromatic, contrast, saturation, sharpness, simplicity, brightness. We choose 10,000 images for training, which is first clustered into 20 groups, and only 6 of them are labeled as positive (favored by experts) samples, while the other groups are labeled as bad (not favored by experts).

5.2 Aesthetics Value Prediction

In order to examine the feasibility of our proposed methods, we have verified the accuracy using different classification algorithms other than decision tree, including Naïve Bayesian, Support Vector Machine, Multi-layered Perceptron, Radial Basis Network, Adaptive Boosting, and Random Forest.

The parameters for the classifiers are as follows: For the support vector machine, the RBF function is chosen as the kernel function, the cost is set to 1. For the multi-layered perceptron, the number of hidden layers is set to half the sum of feature numbers and number of classes, which is 7 here. The learning rate is 0.3, and the training iteration is 500. For the Radial basis network, the minimum standard deviation is set to 0.1, clustering seed is 1, number of clusters is 2, and the ridge is 10^{-8} . For the Adaptive Boosting, the weak classifiers are decision stump, the number of iterations is 10, the seed is set to 1, and the weight of threshold is set to 100. For the J48 decision tree, the confidence factor is set to 0.25. For the Random forest, the number of trees is set to 100. Each of the algorithms are verified using 10-fold verification.

5.3 Evaluation of the Instruction System

In this section, the result for the instruction system is shown. Both the input image, the predicted result, and the decision process for the instructions are shown.

The following are good aesthetics value prediction results for the natural images. The images are predicted as accepted. The prediction process is also shown beside the images to be predicted. For cleaner explanation, the "g" in the leaf node means favorable, and "b" represents "not favorable."



Fig. 6. A flower image which is classified as “favorable”: (a) input image; (b) the decision path of the decision tree.



Fig. 7. The improvement of an image after applying the instruction from our proposed system: (a) the original image; (b) the improved image after applying the instruction.

In the Fig. 6, the image is classified as “favorable.” The decision process are also shown in the figures.

In the figures of the decision path, the black path is the actual path traversed by the decision process, and the gray path in the tree are not traversed. The “...” represents a sub-tree.

The Fig. 7 shows an improvement case after applying the instruction from our system. The instruction asks to improve sharpness, and we use an image editor to enhance the sharpness of the image. In this case, the sharpness is adjusted violently, and the increase of the sharpness also influences other features in the image. The image is classified as “favorable” after applying the instruction. However, the result in this case can be a coincidence because the decision path is not the same as the one before improvement. Fortunately, nodes in four degrees near the root are the same in two paths of the decision tree, which means pruning may be a solution to prevent this kind of problem.

The Figs. 8, 9 and 10 are successful examples that the images are improved by using the instructions from our system. In the three images, the features from the instructions are adjusted carefully that they would not change the decision path of the tree.

In Fig. 23, the feature f4 (cyan degree) is not enough in the image. The image is adjusted to the one which meets the conditions of a contemporary image after increasing the cyan degree of the image. Although the change in the image is subtle, the result is noticeable.

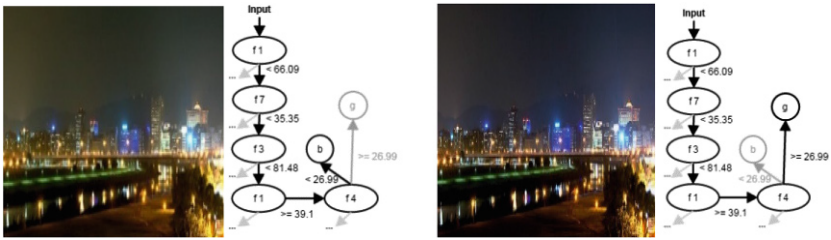


Fig. 8. The improvement of an image after applying the instruction from our proposed system: (a) the original image; (b) the improved image after applying the instruction.



Fig. 9. The improvement of an image after applying the instruction from our proposed system: (a) the original image; (b) the improved image after applying the instruction.



Fig. 10. The improvement of an image after applying the instruction from our proposed system: (a) the original image; (b) the improved image after applying the instruction.

In Fig. 9, the feature f1 (red degree) is not enough in the image. The image is adjusted to the one which meets the conditions of a contemporary image after increasing the red degree of the image.

In Fig. 10, the feature f12 (brightness degree) is too high in the image. The image is adjusted to the one which meets the conditions of a contemporary image after decreasing the brightness degree of the image.

6 Conclusion and Future Works

Image enhancement by expert retouching is an effective technique for making contemporary style images which is preferred by professional photographers. Traditionally, the enhancement is performed manually by experience or “trial and error.” A good handling of camera can reduce the time of image enhancement. Our proposed methods

can effectively tell users directly which features should be enhanced to make a favorable image with single instruction.

There are two issues to be solved for the system. First, when the sibling's leaf node is a sub-tree, the instruction cannot be given in the current method. Second, when changing a feature of the image violently, other features are possibly altered, which also changes the decision path. The solution of these issues would help improve the application and eliminate limitations of the proposed system.

Acknowledgement. The authors thank the National Science Council of Taiwan (R.O.C.) for supporting this work in part under Grant MOST 104-2221-E-011-032-MY3.

References

1. Fahn, C.-S., Wu, M.-L.: An autonomous aesthetics-driven photographing instructor system with personality prediction. In: Proceedings of International Conference on Computer Graphics, Visualization, Computer Vision, and Game Technology, Jakarta, Indonesia, pp. 13–19 (2013)
2. Yeh, C.-H., Ho, Y.-C., Barsky, B.A., Ouhyoung, M.: Personalized photograph ranking and selection system. In: Proceedings of International Conference on Multimedia, Firenze, Italy, pp. 211–220 (2010)
3. Christopher, J.C.: A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **2**(2), 121–167 (1998)
4. Apté, C.V., Weiss, S.: Data mining with decision trees and decision rules. *Future Gener. Comput. Syst.* **13**(2), 197–210 (1997)
5. Apté, C.V., Hong, S.J., Natarajan, R., Pednault, E.P., Tipu, F.A., Weiss, S.M.: Data-intensive analytics for predictive modeling. *IBM J. Res. Dev.* **4**(1), 17–23 (2003)
6. Quinlan, J.R.: *C4.5: Programs for Machine Learning*. Elsevier, Amsterdam (2014)