Collaborative Personalization of Image Enhancement

Juan C. Caicedo
Universidad Nacional
Bogotá, Colombia
jccaicedoru@unal.edu.co

Ashish Kapoor
Microsoft Research
Redmond, WA. USA
akapoor@microsoft.com

Sing Bing Kang
Microsoft Research
Redmond, WA. USA
sbkang@microsoft.com

Abstract

While most existing enhancement tools for photographs have universal auto-enhancement functionality, recent research [8] shows that users can have personalized preferences. In this paper, we explore whether such personalized preferences in image enhancement tend to cluster and whether users can be grouped according to such preferences. To this end, we analyze a comprehensive data set of image enhancements collected from 336 users via Amazon Mechanical Turk. We find that such clusters do exist and can be used to derive methods to learn statistical preference models from a group of users. We also present a probabilistic framework that exploits the ideas behind collaborative filtering to automatically enhance novel images for new users. Experiments show that inferring clusters in image enhancement preferences results in better prediction of image enhancement preferences and outperforms generic auto-correction tools.

1. Introduction

Despite advances in digital photography that improve camera functionality taking good quality pictures remain a challenge for casual photographers. Problems include incorrect camera settings and poor lighting conditions. Photos may be improved using image enhancement tools, but manually retouching every single photograph is infeasible. Auto-enhancement tools are available (e.g., from Picasa or Windows Live Photo Gallery), but as Kang *et al*. [8] have shown, users can have significantly different preferences that are not reflected in such tools.

We use Kang *et al*.'s observations to explore a new approach for image enhancement. We explore if groups of users have similar preferences. Our hypothesis is that such clusters do exist and that we can utilize that information in order to collaboratively enhance images. Intuitively, we expect that enhancement preferences should be similar for users with similar taste on similar images and consequently we can derive methods that would let us harness the enhancement efforts of multiple users to build predictive models. Such adaptation of the image enhancement problem to the collaborative setting is further strengthened by the ever growing use of web-based photo-sharing systems such as Facebook and Flickr. Adding a personalized image enhancement component to those sites would allow the system to suggest appropriate enhancement parameters for new images by utilizing prior enhancements made by other users with similar preferences.

Our work has three key contributions. First, we describe a comprehensive user study through Amazon Mechanical Turk in order to collect image enhancements from a large number of users. Second, we propose a probabilistic model that explicitly encodes clusterings of user preferences and derive an efficient inference method to collaboratively enhance unseen images for novel users. Finally, we empirically show that explicitly modeling the clustering of user enhancement preferences leads to better predictions of enhancement parameters than the existing one touch-button commercial auto-enhance functionalities.

2. Previous Work

The basic idea of image-derived priors is to exploit information in a predefined image set to automatically determine the most likely parameters for enhancing an image. It has been used for a wide range of tasks including denoising [3], color correction [13], and image restoration [2]. These approaches use generic data sets with a large number of images that do not provide any information about user identity or user preferences. Grabler *et al*. [4] use sample image manipulations to learn macros to be applied to images with similar content. Also, Joshi *et al*. [7] narrowed the domain of prior images to a person's favorite photographs to develop a personal image enhancement framework. In principle, these methods are user-dependent, but no studies were done to establish if their variance across users is significant.

The work closest to ours is that of Kang *et al*. [8], who proposed a model for personalization of image enhancement that learns user preferences by observing her choices in a training set. They conducted a user study that showed

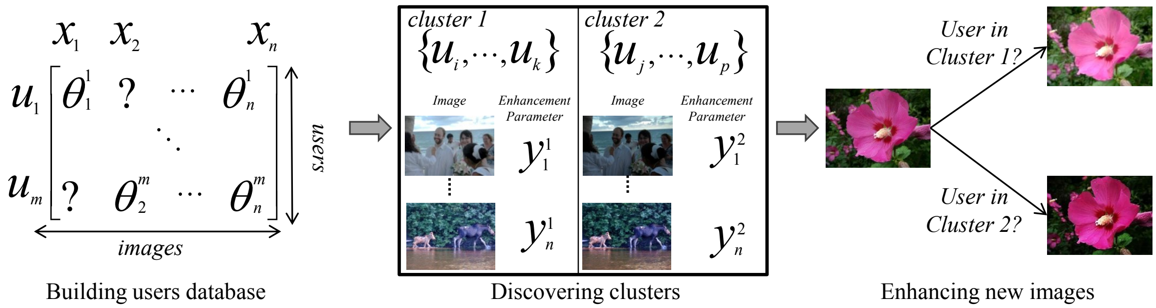

Figure 1. Overview of collaborative personalization of image enhancement. A database of users is built, with θ_i^u the observed enhancement vector for user u and image $\mathbf{x}_i$. Enhancements are not observed for every combination of images and users, which is indicated by the symbol '?' in the matrix. An algorithm is used to discover cluster membership for users, and to infer enhancement vectors $\mathbf{y}_i^c$ associated with cluster c for each i^{th} image. New images are enhanced based on similarities to existing images and cluster to which the user belongs.

that different users do have different preferences in image enhancement, and demonstrated that personalization is an important component to improve the subjective quality of images. Our method builds upon both personalization of image enhancement and collaborative filtering techniques. For personalization of image enhancement, our approach asks users to enhance a set of training images using an intuitive interface to collect preference information. In this portion of our work, we implemented methods similar to those in Kang *et al.* [8]. However, our work substantially extends that personalization approach to consider the contribution of other users with similar preferences.

Collaborative filtering is an approach to build recommender systems, which analyzes patterns of user interest in products to provide personalized recommendations [9]. Intuitively, collaborative filtering works by building a database of preferences for items by users, and matching new users against this database to find neighbors with similar preferences [11]. Most existing collaborative filtering methods attempt to predict a *scalar* quantity, and thus are hard to adapt to our scenario where our goal is to predict a vector of enhancement parameters. This is the first work that we know of that proposes the idea of using the *wisdom of crowds* to enhance images via collaborative filtering.

3. Overview

Figure 1 summarizes the flow of tasks towards building a system for collaborative image enhancement. We first acquire a database of images that are enhanced by multiple users, and represent the database as a table (leftmost of Figure 1) whose rows are individual users and columns are images. Each entry in the table corresponds to enhancement parameters associated with a different user u and image $\mathbf{x}_i$, and are represented by vector θ_i^u. The images used for training are a fixed set, and are reasonably representative of images that need enhancement. Note that this set of representative images is large; hence, it is unreasonable to expect that each user will provide enhancement parameters for all the images. Instead, every user enhances only a small subset of images, and as such, there are quite a few entries in the table that remain unobserved (denoted as '?').

Once this database is acquired, we analyze it to learn a statistical model that explicitly encodes clustering of users. Specifically, we recover groupings of users and estimate the preference parameters of each individual group for all the representative images (middle of Figure 1). Once such a statistical model is learned, we can enhance a new (i.e., unseen) image for any cluster by considering its similarity to the images in the representative set.

4. Building Blocks

Prior to running our user study, we need to pick the knobs required to perform image enhancement and select our representative set of training images. Also in order to effectively perform collaborative filtering, we require a distance metric between images that reflects differences in the respective enhancement operations required on those images.

Enhancement Operations: We considered two global enhancement operations: contrast manipulation (S-curve and Gamma correction) and color correction (tint and temperature), which as with Kang *et al.* [8], are represented by 5 enhancement parameters (two parameters for S-curve, one each for gamma-correction, tint and temperature).

Learning a Distance Metric between Images: One of the key ingredients in our approach is distance (or similarity) between images. In particular, given two images, we want to estimate the extent to which these images require similar enhancement parameters. We approach this as a metric learning problem, in which a distance between visual features is adapted to include information of enhancement parameters. For every image in the photo collection, we extract 6 color histograms from the RGB channels and the HSV components, which are used to build a unique image descriptor. Formally, let $g(\mathbf{x}_i)$ and $g(\mathbf{x}_j)$ be visual feature vectors for two images $\mathbf{x}_i$ and $\mathbf{x}_j$. We use the Mahalanobis distance in the visual feature space:

$$d_A(\mathbf{x}_i, \mathbf{x}_j) = (g(\mathbf{x}_i) - g(\mathbf{x}_j))^T A(g(\mathbf{x}_i) - g(\mathbf{x}_j)) \quad (1)$$

which is parameterized by a square matrix A to encode weights associated with different features and correlations amongst them in order to reflect the disparity in the enhancement parameter space. We learn the matrix A using an online algorithm following the methodology proposed by Jain *et al*. [6]. The target distance between any two images is considered to be the distance between their auto-enhancement parameters. Specifically, we generate constraints of the form $(g(\mathbf{x}_i), g(\mathbf{x}_j), \delta_{ij})$, where δ is the distance between the enhancement parameters (obtained from an auto-enhance utility) of the corresponding images. This specific online algorithm is well suited for our task, because our image collection has approximately 100 million constraints, which makes it non-trivial for other distance metric learning methods to handle.

Representative Image Selection: Selecting a good set of representative images is critical. We downloaded a set of real photographs from Flickr, following a "sampling through time" strategy, in which the parameter "date taken" is randomly changed [1, 5]. The list of keywords includes "landscape," "nature," "people," and "sports," among others, to cover a diverse range of photographs without focusing on particular objects. We filtered out images that are too small (images smaller than 800×800 are removed), black-and-white, and have very little intensity variation (measured by histogram entropy). In the end, our downloaded set consisted of more than 300,000 photos.

Given the set of images and the distance metric between images, we need to identify a subset of different images that represent the original set. This problem was approached as a sensor placement problem, in which a budget of sensors is available to be located in a space, deciding which positions to chose to maximally cover spatial monitoring. In our problem, the sensors are the available images in the given set where the similarity (covariance function) is defined by the kernel $K_{ij} = \exp(-d_A(\mathbf{x}_i, \mathbf{x}_j)/\text{mean}(d_A(:)))$. We used the algorithm described by Krause *et al*. [10], in which the elements that maximally reduce the entropy in the original set are iteratively selected.

In the Flickr snapshot, we also observed that users tend to upload many similar images. We applied the sensor placement algorithm to images from the same user, to keep only the three most "different" images from each individual. We then combined these filtered images and ran the algorithm again to select a final photo collection for the user study. We found that with 200 images, we kept about 84% of the information from a set of approximately 25,000 candidates. Figure 2 shows how the information gain reduces as long as informative images are being selected. These

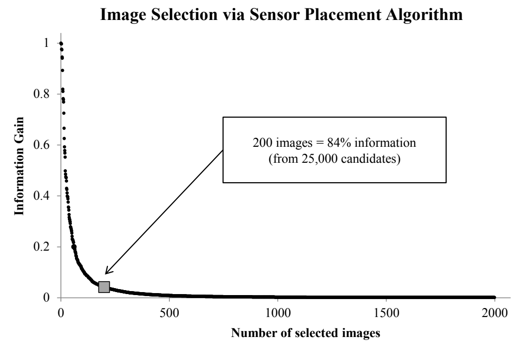

Figure 2. The image selection algorithm picks images that provide maximum increase in mutual information. This results in a decreasing information gain in the selected image set.

200 images sufficiently represent the visual feature space defined by the metric that we have learned.

5. User Study

The goal of this study was to build the user-image matrix of enhancement parameters. We opted for Amazon Mechanical Turk (MT) to engage people in our image enhancement task. Each person had to enhance a small number of images by selecting the most preferred enhanced version via an interactive web system. We managed to collect enhancement parameters from 336 valid users.

Two main criteria guided the design of the system for image enhancement used in this study: (1) easy of use and (2) web-oriented. For the first criterion, the user interface does not require parameter tweaking to apply enhancements. We implemented a user interface similar to the one proposed by Kang *et al*. [8], in which the system presents a 3×3 matrix, with an enhanced version of the image in each cell. In the first iteration, the system presents the original image in the center cell surrounded by 8 possible initial enhancements. When the user clicks one of these possible enhancements, new similar enhancements are computed to replace the current ones. When the user has identified the most preferred enhanced version, the system records the enhancement parameters that produce that image and presents another one. Figure 3 shows a screenshot of this user interface.

For the second design criterion, the application is required to be a web-oriented system, to allow people to enhance images online through MT. We opted to deploy our application using cloud computing services, which allows our system to scale up as needed to compute image enhancements online. In our user study, we asked each participant to enhance 20 images, which are randomly assigned from the collection of 200 photos. The assignment of images to users attempts to have approximately the same number of users enhancing each image. The maximum time allowed to complete this task was 1 hour and the reward was

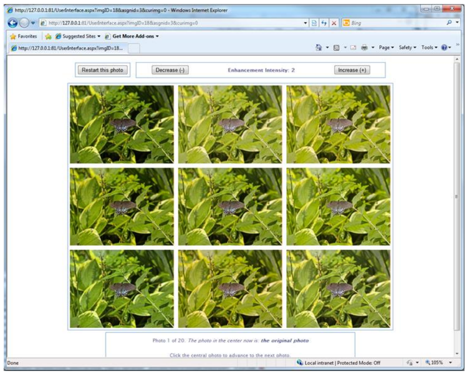

Figure 3. Web-based user interface for image enhancement.

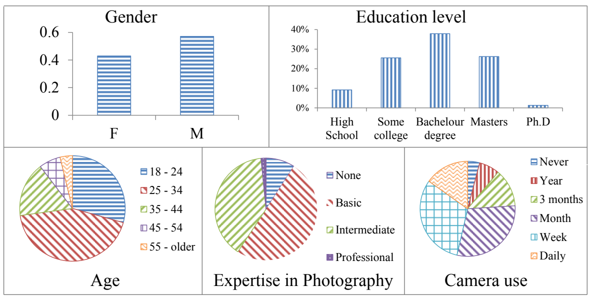

Figure 4. Demographics of subjects involved in the MT study.

US$1.50. The click count, click rate, and time stamps were recorded to identify spammers.

Right before starting the task, participants were asked to fill a questionnaire aimed to collect demographic information, including gender, age bracket, country, and education level. They also have to indicate their level of expertise in photography and frequency of camera use. Figure 4 shows the proportions of these attributes for 336 valid subjects. Most of the subjects are from USA (43%) and India (41%), with the rest from more than 20 other countries.

6. Enhancing Images Collaboratively

Let us denote the collection of all the n images as $\mathbf{X} = \{\mathbf{x}_i\}$. The user study provides us with sets of enhancement parameters chosen by all the m users. Let θ_i^u be the enhancement parameters chosen by user u for image $\mathbf{x}_i$. Because of resource and time constraints, every user enhances only a small subset of images (20 images out of 200) and consequently there are a lot of images for which we do not directly observe the enhancement parameters corresponding to each user. Our goal is to derive a framework that can (1) infer these missing enhancement parameters for every user in the study, and more importantly, (2) determine enhancement parameters for new images and users.

Our model is motivated by the methods in collaborative filtering. In particular, the first key underlying assumption here is that similar users should have similar enhancement parameters for similar images. Assumptions like this are also at the heart of many collaborative filtering methodologies; we could in fact adapt those methods to infer missing enhancement parameters for all the users in the study. However, most of the work in collaborative filtering focuses on predicting a scalar quantity while our goal is to model a vector of enhancement preferences. While off-the-shelf approaches can be individually adapted to each enhancement parameter, such simplistic scheme ignores the structure of relationships across parameters and will be sub-optimal.

In this work, we extend the collaborative filtering to a setting of structured prediction where we jointly predict all the components of the enhancement preferences. Specifically, our model not only encodes similarity across images and users but also models relationships between different components of the enhancement space. The notion of users similarity in our system corresponds to grouping of users into clusters. Thus, users are clustered if they have similar enhancement parameters for all the images. In the next subsections, we describe (1) a probabilistic graphical model for jointly predicting enhancement preferences that explicitly encodes similarity across images and groups users into clusters, (2) an efficient inference algorithm, and (3) extensions to do predictions on unseen images and users that were not part of the user study.

6.1. Probabilistic Model for Enhancements

We propose a model that encodes the dependence of enhancement parameters on image content as well as user preferences. Specifically, given the collection of images and clustering of users, we assume that there are latent enhancement preference vectors $\mathbf{y}_i^c$ which correspond to cluster c and image $\mathbf{x}_i$. Further, the enhancement parameter vectors we observe in the MT study are simply noisy versions of these latent true enhancement preferences. Figure 5 illustrates the factor graph corresponding to the proposed model. The observed enhancement preferences θ_i^u from different users are denoted as circles and we introduce a discrete random variable (squares) $h_u, u \in \{1, .., m\}$ for each user that indicates the cluster the user belongs to. We also use $\mathbf{Y}_i$ to denote collection of all true enhancement preferences for the i^{th} image across all the clusterings. The shaded nodes correspond to random variables that are observed. For example, in Figure 5, the enhancement preferences for user 1 are known for images 2 and n, while user m did not enhance image 2. This is consistent with our data set where users enhance only a subset of all training images.

The model also imposes smoothness constraints using a GP prior [12] in order to account for the image content. We use a Gaussian Process prior to enforce the assumption that "similar" images should have similar preference parameters. In particular, for each cluster c we induce the GP prior (denoted as $GP(c)$), where each of the five com-

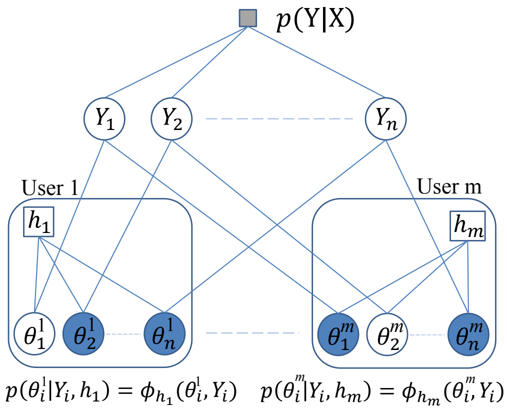

Figure 5. Factor graph depicting the proposed model. Shaded nodes correspond to observed preferences. Not all observations are available for all the images and users. Square boxes depict latent random variables corresponding to cluster membership.

ponents of the latent variables $\mathbf{y}_i^c$ and $\mathbf{y}_j^c$ are assumed to be jointly Gaussian with zero mean and the covariance specified using a kernel function applied to $\mathbf{x}_i$ and $\mathbf{x}_j$. Formally, $GP(c) \sim \prod_{p=1}^{5} \mathcal{N}(\mathbf{y}^c(p); \mathbf{0}, \mathbf{K})$. Here $\mathbf{y}^c(p)$ is the column vector of p^{th} component of enhancement preference for all images corresponding to the cluster c and $\mathbf{K}$ is a kernel matrix[1] with $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ and encodes similarity between pairs of images. In this paper, we use $K_{ij} = \exp(-d_A(\mathbf{x}_i, \mathbf{x}_j)/\text{mean}(d_A(:)))$. Note, that the GP prior above does not explicitly encode the relationship between the components of the parameters and we can assume that the components of $\mathbf{y}_i^c$ have been whitened (zero mean, with unit covariance) beforehand[2]. Also, note that all the dimensions of the latent variable $\mathbf{Y}$ are coupled in the model and performing inference will preserve the relationships between different components.

Let $\mathbf{\Theta}$ be all the enhancement preferences from all the users and for all the images. Our proposed model induces a conditional probability distribution $p(\mathbf{\Theta}, \mathbf{Y}, \mathbf{h}|\mathbf{X})$ using the GP prior $p(\mathbf{Y}|\mathbf{X})$, prior probabilities on the cluster membership $p(\mathbf{h})$, and the potential terms $p(\theta_i^u|\mathbf{x}_i, \mathbf{Y}_i, h_u)$ that link the latent image preferences to the ones that are observed. Thus, the conditional distribution induced by our model can be written as

$$p(\mathbf{\Theta}, \mathbf{Y}, \mathbf{h}|\mathbf{X}) = \frac{1}{Z} p(\mathbf{Y}|\mathbf{X}) p(\mathbf{h}) p(\mathbf{\Theta}|\mathbf{X}, \mathbf{Y}, \mathbf{h})$$
$$= \frac{1}{Z} \prod_{c=1}^{k} GP(c) \prod_{u=1}^{m} p(h_u) \prod_{i=1}^{n} \phi_{h_u}(\theta_i^u, \mathbf{Y}_i),$$

[1]This kernel matrix is a positive semidefinite matrix and is akin to the kernel matrix used in classifiers such as SVMs.

[2]In this work, we whiten the enhancement preferences before applying the model and re-project back to orginal space eventulally which results in preservation of the structure of the parameter space.

where Z is the partition function (normalization term) and the potential $\phi_{h_u}(\theta_i^u, \mathbf{Y}_i)$ corresponding to a user u and image $\mathbf{x}_i$ takes the following form:

$$\phi_{h_u}(\theta_i^u, \mathbf{Y}_i) \propto e^{-\frac{||\mathbf{y}_i^{h_u} - \theta_i^u||^2}{2\sigma^2}}. \quad (2)$$

Here, $\mathbf{y}_i^{h_u}$ are the hidden random variable for the same cluster as the cluster indicated by h_u and the image $\mathbf{x}_i$, and σ^2 is the noise parameter that determines how tight the relation between the smoothness constraint and the final label is. By changing the value of σ we can emphasize or de-emphasize the effect of the GP prior. In summary, the model provides a powerful framework for encoding dependence of enhancement parameters on image content (via the GP prior) as well as the clustering of users and allows us to combine the prior assumptions with the data that is observed in the MT study.

6.2. Inference in the Model

Given the observations $\mathbf{\Theta}_o$ from the MT study the key task is to infer the posterior distribution $p(\mathbf{Y}, \mathbf{h}|\mathbf{X}, \mathbf{\Theta}_o)$ over latent true enhancement preferences and the clustering membership for all the users. Performing exact inference is prohibitive as the joint distribution is a product of a Gaussian (GP prior and the $\phi(\cdot)$ potentials) and non-Gaussian terms (cluster membership). We resort to approximate inference techniques in order to get around this problem. In particular, we perform an approximate inference by maximizing the variational lower bound with the assumption that the posterior over the unobserved random variable $\mathbf{Y}$ and $\mathbf{h}$ can be factorized:

$$F = \int_{\mathbf{Y},\mathbf{h}} q(\mathbf{Y}) q(\mathbf{h}) \log \frac{p(\mathbf{Y}, \mathbf{h}|\mathbf{X}, \mathbf{\Theta}_o)}{q(\mathbf{Y}) q(\mathbf{h})}$$
$$\leq \log \int_{\mathbf{Y},\mathbf{h}} p(\mathbf{Y}, \mathbf{h}|\mathbf{X}, \mathbf{\Theta}_o),$$

where $q(\mathbf{Y}) = \prod_{c=1}^{k} \prod_{p=1}^{5} q(\mathbf{y}^c(p))$ is assumed to be a Gaussian distribution and $q(\mathbf{h}) = \prod_{u=1}^{m} q(h_u)$ is a discrete joint distribution over the unobserved labels. The approximate inference algorithm aims to compute good approximations $q(\mathbf{Y})$ and $q(\mathbf{h})$ to the real posteriors by iteratively optimizing the above described variational bound. Specifically, given the approximations $q^t(\mathbf{Y})$ and $q^t(\mathbf{h})$ from the t^{th} iteration and assuming uniform prior over $p(\mathbf{h})$ the update rules are:

$$q^{t+1}(\mathbf{y}^c) \propto GP(c) \prod_{\theta_i^u \in \mathbf{\Theta}_o} [\phi_c(\theta_i^u, \mathbf{Y}_i)]^{q^t(h_u=c)}$$
$$q^{t+1}(h_u = c) \propto \prod_{\theta_i^u \in \mathbf{\Theta}_o} \phi_c(\theta_i^u, \text{mean}(q^{t+1}(\mathbf{Y}_i))).$$

Intuitively, the update of image enhancement preferences considers the cluster membership from the previous iteration and uses it to decide if a data term should be included in update for each cluster. Similarly, the update for the posterior over the cluster membership considers mean enhancement preferences from the previous iteration. Thus, starting from a random initialization the parameters and posterior of the cluster memberships are iteratively updated untill convergence. Upon convergence, we obtain posterior distribution of cluster membership for each user ($q(\mathbf{h})$) as well the distribution over the true image enhancement preferences ($q(\mathbf{Y})$) approximated as a Gaussian distribution.

6.3. Handling New Images and Users

Besides inferring about the images in the database, we are also interested in predicting enhancement preferences for an image $\mathbf{x}_{test}$ that was not in the original training set as well as for a user who was not part of the user study.

The more straightforward case is where we infer enhancement preferences for a user u who was a part of the user study and had enhanced training images for us. Here, we simply need to run the inference algorithm where we augment the random variable $\mathbf{Y}$ with $\mathbf{Y}_{test}$, the collection of random variables that represent true enhancement parameters for the new test image across all the clusterings. Thus, the new image can be enhanced using the mean of inferred enhancement preferences corresponding to the cluster the user u belongs to. Note that if we have already performed variational inference on the training data, inference after the augmentation only requires one iteration of update equations. This is because the new image does not introduce any new information about clusterings of the user or enhancement parameters, hence, making it fairly efficient.

It is trickier if we need to enhance images for a user who is not part of the study. However, once we know the cluster membership of the new user, we can simply use the scheme mentioned above to estimate the enhancement preferences. Thus, the problem really reduces down to establishing the cluster membership of the user. To this end, we can employ several methods. For example, the new user can enhance some of the images from the training set, which in turn will give us evidence about membership. The user now can be considered a part of the available corpus and inference can be run as before. Alternatively, we can also engage the user in an introductory dialogue where training images are used; the resulting enhancements will indicate membership. In this work, we use the former approach of asking users to enhance a subset of training images to test the extension of the system to new users and images.

7. Experiments and Results

We first perform experiments to determine the correct number of clusters using the data collected in the MT study. Train-test splits were created by randomly choosing 10% of the observed enhancement vectors in the user-image matrix as test examples. The other 90% is used to run the inference algorithm described in Section 6.2 with σ^2 fixed to 10^{-3}. Evaluating the quality of user adjustments is difficult, since there is no universal agreement on the perceptual metric. We resorted to the objective measure of intensity difference looking at average Root Mean Square Error (RMSE) in parameter space (enhancement vectors) as well as image intensity space (resulting enhanced images).

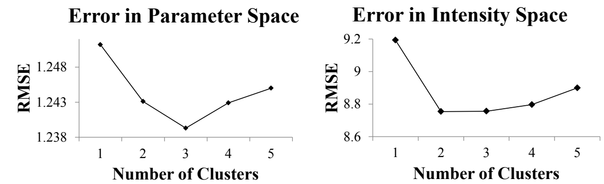

Figure 6. Average estimation error on training for different number of clusters.

Figure 6 highlights the variation of RMSE as we change the number of clusters in the model. Note, that considering only one cluster is equivalent to assuming that all users follow the same preferences for enhancing images. We observe that error is highest in both parameter and intensity spaces when the number of clusters is fixed to one and reduces significantly as we start incorporating additional number of clusters. This result is consistent with the observations reported by Kang et al. [8], where it was shown that different users do have significanlty different preferences. We also find that the error was minimum in the parameter space for 3 clusters and pretty close to minimum in the intensity space. Thus, we perform rest of the analysis and experiments by fixing the number of clusters to 3.

The proposed method can be thought of as a clustering algorithm with constraints imposed by similarity in the image appearance space. Consequently, it is in principle the same as an EM algorithm for learning a Gaussian mixture model, and it does not prefer that most users have their own distinct cluster. We think that the increase in error due to further addition of clusters might be due to the small number of enhancements operations allowed and we can expect to see different clusterings if additional operations are permitted.

Figure 7 shows example images enhanced according to the preferences of users in each of the three clusters, giving a general indication of the preferred appearance in each cluster. While the corrected images in cluster 2 are slightly more saturated than the others, images in cluster 3 seem to have more contrast. A more quantitative description of the preferred enhancements in each cluster can be seen in Figure 8, which presents the distribution of the five parameters as box plots. The box plots are over the means of the inferred distribution over $\mathbf{Y}$ (all 200 images for the 3 clusters), where the red line in each column denotes the median. Note that the largest difference across clusters is for the Power Curve and the S-Curve Inflection Point, suggest-

Figure 7. Example images enhanced according to the preferences discovered for each cluster.

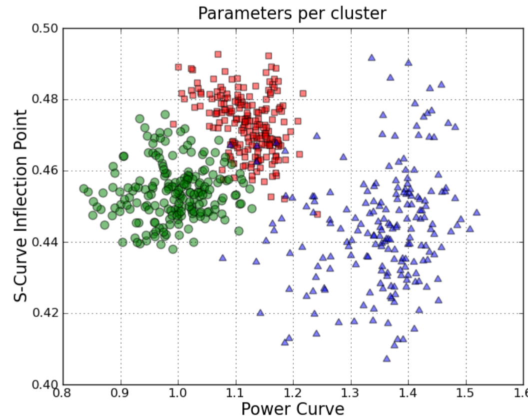

Figure 9. Scatter plot of all images according to parameters Power-Curve and S-Curve Inflection Point (which partially control global image contrast). Cluster 1 in red squares, Cluster 2 in green circles and Cluster 3 in blue triangles.

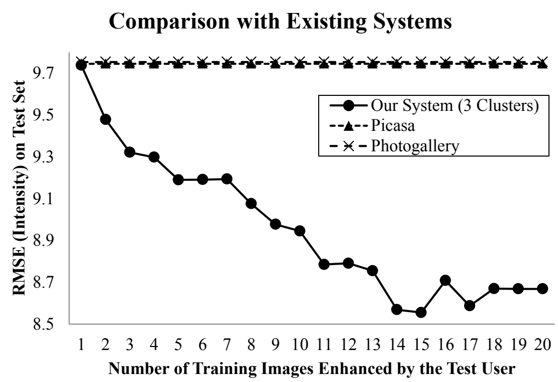

Figure 10. Average error of predicting personalized enhancement parameters. Auto-enhancement tools do not learn user preferences. Our approach predicts personalized enhancements with increasing accuracy as users provide new examples.

ing that variations in contrast is a dominating factor across the clusters. There are some differences across the rest of the three parameters as well suggesting a difference in parameter choice amongst the users.

In the scatter plot shown in 9, each point corresponds to an enhancement preference color-coded by the cluster to which it was assigned. Note that the points are clustered fairly evenly in this 2D projection of the space, further emphasizing the latent structure which is successfully discovered by the inference algorithm. We also analyzed statistics of demographic data across the clusters but the statistics were fairly similar across the clusters.

Next, we evaluate the ability of the proposed model to predict enhancement preferences for a new user on unseen images in a collaborative environment. In this test, 10% of all users were randomly chosen as test subjects and inference was performed using only the data observed for rest of the users. We simulate the real-life situation where every test user provides evidence about its cluster membership by enhancing some images. Since this work is about collaboratively using information about image enhancement available from many different users, we show results on scenarios where a test user has provided such little information that personalized model of [8] cannot learn. We look at the RMS error in intensity space as each user enhances one image at a time (chosen randomly) and compare them with results of auto-enhancement from Picassa and Windows Live Photogallery.

Figure 10 presents the plot of RMSE obtained by the system with increasing number of enhancements available from the user. We find a consistent reduction in estimation errors as more enhancements are added to the personal profile. This is due to the fact that additional enhanced images provides more information about the cluster membership of the user, enabling the inference procedure to estimate a better recommendation. We also observe that the performance of the collaborative approach is far better than the two auto-enhancement tools. While all the three approaches perform similarly when there is no evidence about the cluster membership of the user, we see that the performance of collaborative strategy greatly improves as user starts enhancing images. Also, note that the gains start showing up with as few as 2 images indicating that strong gains can be obtained with the collaborative enhancement of images.

Finally, Figure 11 shows several image examples with the corresponding enhanced versions produced by Photo-

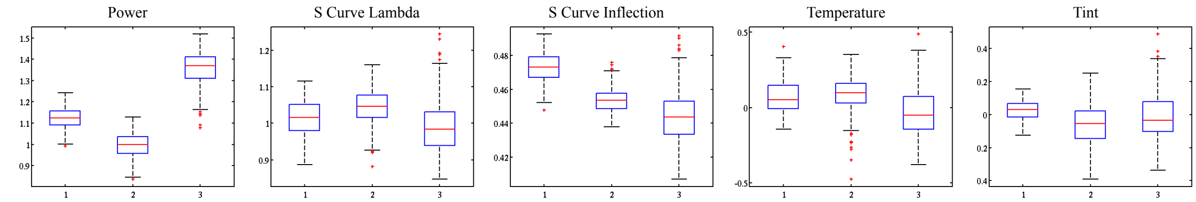

Figure 8. Box plots of values for each parameter (first three control global contrast, last two control color correction). The x-axis indicates cluster number and y-axis corresponds to range of values. Red lines denote the median, end lines are at the lower and upper quartile values, and crosses beyond the ends of the whiskers are outliers.

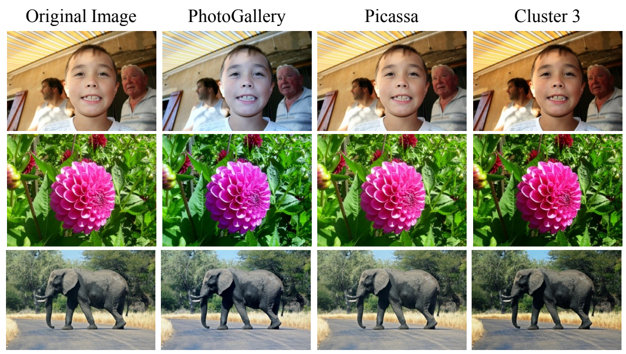

Figure 11. Example images with the corresponding enhanced versions generated by auto-enhancement tools and the proposed approach for a subject that belongs to cluster 3.

Gallery and Picassa. Also, the enhanced version of our method, for a subject of the cluster 3 is presented. In general, each tool produces a different enhancement resulting in a different final look and feel. Notice that the suggestion made by our method has been inferred using the subjective preferences learned from people in cluster 3, thus it is more likely to be preferred by a user belonging to that population.

8. Conclusion and Future Work

We present a novel approach for image enhancement that follows a collaborative framework to recommend corrections based on user preferences. Instead of building a system that is trained to enhance images for a specific user, our technique is a principled and practical way for collaboration at a web scale, and can encompass alternative parameterizations and error criterion. The main idea is we prediscover the clusters of personalized enhancement parameters through collaborative learning on large image collections. Then, for a new user, we need only figure out which cluster he/she belongs to.

Results of the cluster analysis showed the existence of three main groups, mainly characterized by differences in contrast preferences. Experimental results indicate that the collaborative enhancement strategy significantly helps in making better predictions of enhancement parameters than existing one touch-button commercial auto-enhance tools. We do not claim that our results extend to a much larger scale with many more users and a much richer set of enhancement knobs; more experiments are required to substantiate such a claim. Future work includes *active* selection of images that would enable better estimation of cluster membership with fewest image enhancements by the user.

References

[1] T. S. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y. Zheng. Nus-wide: a real-world web image database from national university of singapore. In *CIVR '09*. ACM, 2009.

[2] K. Dale, M. K. Johnson, K. Sunkavalli, W. Matusik, and H. Pfister. Image restoration using online photo collections. In *ICCV*. IEEE, 2009.

[3] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graph.*, 25(3):787–794, 2006.

[4] F. Grabler, M. Agrawala, W. Li, M. Dontcheva, and T. Igarashi. Generating photo manipulation tutorials by demonstration. *ACM Trans. Graph.*, 28(3):1–9, 2009.

[5] J. Hays and A. A. Efros. Scene completion using millions of photographs. *Commun. ACM*, 51(10):87–94, 2008.

[6] P. Jain, B. Kulis, I. Dhillon, and K. Grauman. Online metric learning and fast similarity search. In *NIPS*, 2008.

[7] N. Joshi, W. Matusik, E. H. Adelson, and D. J. Kriegman. Personal photo enhancement using example images. *ACM Trans. Graph.*, 29(2):1–15, 2010.

[8] S. B. Kang, A. Kapoor, and D. Lischinski. Personalization of image enhancement. *Computer Vision and Pattern Recognition Conference, CVPR*, 2010.

[9] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *IEEE Computer Journal*, pages 30–37, 2009.

[10] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.*, 9:235–284, 2008.

[11] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. *ACM, WWW*, 2001.

[12] M. Seeger. Gaussian Processes for machine learning. *International Journal of Neural Systems*, 14(2), 2004.

[13] R. Stanikunas. Investigation of color constancy with a neural network. *Neural Networks*, 17(3):327–337, April 2004.