Architecture of Feelings
Machine Learning Project 2

Carlos Megías, Jorge Sánchez and Kiarash Farivar
École polytechnique fédérale de Lausanne, Switzerland

***Abstract*—The objective of this project consists in predicting the feelings of people when they are immersed in an interior with a certain context, shapes of the facades and illumination. In this project, we attempt to determine the reactions that an scene characterized by this properties would incur in a person entering it. This prediction could be useful for a designer who is looking for an interior which causes a certain kind of feelings.**

I. Introduction

The project is based on an experiment realized using different scenes over people from Greece and Switzerland, using Virtual Reality to simulate these environments. The main goal is to know if a scene will provoke the people some of these reactions: pleasant, interesting, exciting, calming, bright, complex and spacious. We were asked to focus in excitation and also in calmness; and we have been focused mostly in the excitation feelings. At the end we have extrapolated our work to build good models for the other reactions too, as we show in the final section. After the stay of a person in the scene during some minutes of exposure, he/she is asked about his/her feelings, having to correspond them to integer qualification from 0 to 10. The scenes consist in two possible types of context: social and work; three sky types of illumination: clear sky with high sun angle, clear sky with low sun angle and overcast sky; and six patterns for the windows: horizontal bars, vertical bars, bamboo bars, squared, random bars and circular. The following images represent the patterns and an example of a scene:

Fig. 1: Patterns

Fig. 2: Example of scenario: social, circular and overcast

A. Our first approach

Initially we started working on the idea of pretending to know how a particular person entering a scene will feel rather than trying to know which feelings a scenario will provoke over people generally. This approach leads us to some problems: we only have the information about the gender and the country of each particular person, we had identical entries (people) with different outputs (feelings)... We tried to overcome this problems applying different methods such as Random Forests, Dense Neural Networks, and finally Matrix Factorization, getting a classifier for the kind of people (depending just in Gender and Country). After great discussions with Martin Jaggi and our supervisor, Kynthia Chamilothori, among all we decided to change our approach due to the mentioned problems and the real study application from the architecture point of view; since from the designer perspective, when you design a scene you do not usually know who or how is the people that is going to enter it.

B. Modeling the problem properly

In the new approach we cover, which is the main one in this report, we are predicting the feelings that an scene tends to provoke on the people who entered it based on its characteristics. At the end we want to have a global idea of the feelings that a particular interior may incur on the people. This study is more valuable for a designer who a priori does not know any information about the people who is going to see this scene. Being able to predict this kind of information based on images would be very worthy for an architect to have a previous knowledge what may help in the his/her decisions about the design. To proceed with this new approach we work with percentages of the people who felt (a specific feeling) over a threshold number in the scale from 0 to 10, doing this for every different scene (36 in total) and differentiating also by country (accordingly with our supervisor, a designer could usually know the country of origin of the people).

This report has been divided in seven sections. Section II presents what the provided data set contains and how it is initially distributed, to finally manage it to prepare everything for the next steps. Section III focuses on trying to understand and clean the available data. Finally, Sections IV and V are dedicated in processing the data and the selection of the model. The Section VI extrapolates the results obtained in Section V. It finishes with a summary and a discussion of the results in Section VII.

II. The data set

We received a set of images (the VR images and its parts or cubes) and two data sets. The first one contains 1590 rows (entries) and 18 columns where we can see the original samples of the experiment. Each column is explained below:

- *ID*: Identifier of the subject of the experiment (of one person).
- *Country*: (Greece 138 participants, Switzerland 127 participants)
- *Stimulus_SkyType*: 3 x sky types
- *Stimulus_Context*: 2 x context scenarios
- *Pattern*: 6 x patterns (See Figure 1)
- Outputs, what people felt in every scene (from 0 to 10): *pleasant, interesting, exciting, calming, complex, bright, view, spacious.*
- Scene metrics for describing the interior (the kind of room, light, ...) quantitatively: *contrast_mean_mSC5, contrast_mean_Michelson, contrast_mean_RMS, complexity_mean_JPEGtoBMP*. Obtained by the scenes images.

Note: The metrics in this data set are applied to the whole virtual reality image (in every interior).

The second one contains a table with 36 rows and 31 columns. This time we see the features of each one of the 36 scenes (the kind of furniture, light, ...) and several metrics for describing quantitatively the interior situation. Each column is explained below:

- Pattern, Context, SkyType columns same as explained in the other data set.
- *filename*: File where we have the cube map projections associated to every scene.
- Means and medians of the outputs (of what people have answered in their respective survey) in every scene: *mean_pleasant, median_pleasant,...*
- Scene metrics: same as in the other one but now we have also these metrics applied to the different parts of the image (specifically to cube3 part, referring to what you see on your front site when you are doing the experiment with VR, and to cube123 parts, referring to the left, front and right sites).

III. Exploratory Data Analysis

A. Cleaning phase

Exploring the data we first checked that there are no missing data. We also checked that the ratings are all between 0 and 10.

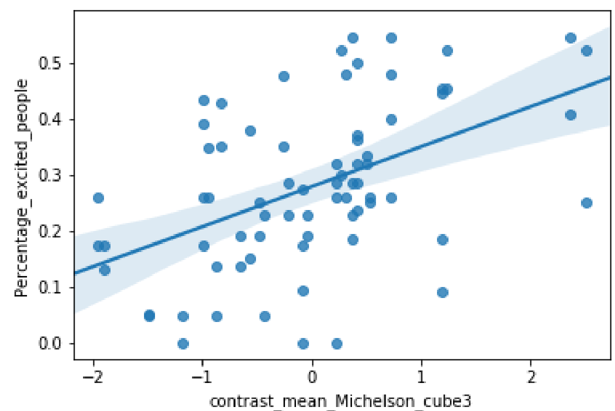
Having done this, we joined the two data sets in just one with all the experiment samples (1590) with all the scenes features both categorical(Pattern, Context, SkyType) and numerical (the metrics) features; removing *Gender* and the mean and medians of the feelings.

Then we computed for every scene (36 in total) the percentage (equivalently the proportion) of people who have been excited over a limit value, i.e the percentage of people who have answer more than the limit on this feeling in the survey (we have fix this limit in $>= 7$). (We finally have computed this percentage for every feeling; see last section).

And so our final raw data set is formed by 36 rows with the different scenes, its metrics and the percentages of excited people (computed for both countries separately)

B. Correlations

The correlations between the given labels and the label we have focus on, "exciting" are not very low but not high. We also drew the scatter plot for our features and they confirmed the relationship between current features and the output variable. To make our visualization complete we plotted the regression plot vs our output "exciting" (plots in the Appendix).



IV. Model selection with initial features

For trying to find the best model we have considered some of the most common and powerful methods of regression: Ridge Regression, Lasso, Random Forest and XGBoost (in their regression version). In this section we use the initial given features to build our model and how it works in terms of the error. The dataset are splitted into two different sets each one, using for training the 80% and the rest 20% for the test. We have executed a Grid Search with Cross-Validation for every predictor (Ridge, Lasso,...) using a pipeline for doing transformations depending on the specific method and indicating the parameters over which iterate to finally get the best model possible. More precisely, with the Ridge and Lasso regression we have used the polynomial expansion transformation; and so iterating in the both the ridge/lasso regularization hyperparameter (λ) and the degree of the polynomial expansion. On the other hand, with the Random Forest and the XGBoost regressors we have iterated over the number of estimators and the maximum depth parameters. Following this grid searches we found that the minimum error is obtained by the ridge regression, so in the FIGURE 3 we can see the cross-validation errors (the mean of them) of the Grid Search for each evaluated configuration of the hyper-parameters:

To evaluate the validity of the results, our Grid Search is minimizing the mean absolute error (MAE) over the mean of the test data of every iteration of the Cross Validation.

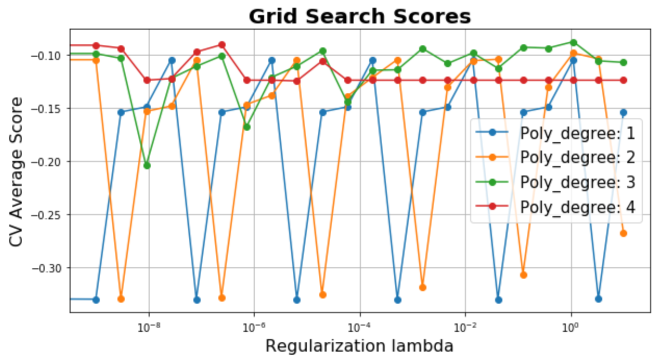


Fig. 3: Grid Search with Lasso regression

After this we compute the MAE over our separated test set obtaining the differences between the predicted test entries and the original test ones (and taking the mean). This allow us to know how close we are from the real percentages and let us compare which method leads us to better results.

V. Principle Component Analysis (PCA)

PCA is an algorithm that linearly transforms data to a new coordinate where our data has the most variance in direction of these coordinates.

The main reason we have used PCA was that we wanted to use the image data we had for each type of room. This would result in better performance in our algorithms since we let the algorithm choose what principle features of the image work best for performing the regression. So this method could have the potential of improvement compared to using some predefined features.

A. Down-scaling of the images

Before performing PCA we read our images and saw that the number of pixels of each one was over 15 million. Our computers obviously could not handle the PCA with such amount of features (even it was not possible to load a data frame).

So we first remove the color of the images (taking then in gray-scale), and then we explore the ways to resize or downscale the images; using finally a downscale function, which reduce the number of pixels in each image by taking the average of pixels over a block of pixels and replacing all the pixels in the block with a single pixel. The size of the block used is of (6,12) pixels. In this way this make possible for our computers to handle the data properly without losing too much performance over the images (e.g. clearly differentiable the shadow patterns). The Figure 4 shows this compression.

After that, we reshape our images into 1-dimension vectors and we joined them with the label of the ”percentage excited” and countries. Before looking for the best model, we standardize these new features (every vector position as new features).

B. Model selection with PCA features

Having the PCA features prepared we are ready to try the different methods we have already discussed in Section IV,

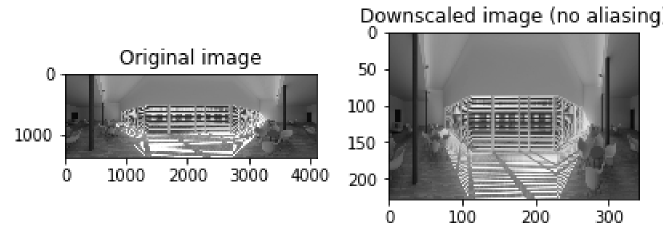


Fig. 4: Downscale of the original images

that is: Ridge, Lasso, Random Forest and XGBoost, for the *exciting percentage*. We first separate all the entries by country, because at the end, we are going to build one model for the Greek people and another one for the Swiss people (note that now we are not using dummy variable for distinguishing the country).

To find the best model we have tried all the four different methods mentioned above by doing a Grid Search with Cross Validation over the number of PCA components and some of the hyperparameter that these methods have: regularization for Ridge and Lasso, number of estimators and maximum depth for Random Forest and XGBoost. The number of PCA components is limited to the minimum between the number of features and the entries of the data set, so in our case, it is the last one which limits it to 27.

The best results obtained by every method are found with the number of PCA components and the values for their hyperparameters in TABLE I.

Method	N PCA	Regularization λ	N Estimators	Depth	Error
Ridge Regression	10	1.0	-	-	0.0517
Lasso Regression	22	0.047	-	-	0.0709
Random Forest	22	-	26	4	0.0834
XGBoost	18	-	40	1	0.0933

TABLE I: Best models for each method and best result: Exciting

Looking at the obtained results we find that the best results are obtained with the Ridge method, so for our final model we will use Ridge Regression. In Section VI we will see how having selected this method of regression works for the rest of feelings, computing the same Grid Search for each label.

C. Interpreting the results for the PCA and ridge regression

One advantage of the PCA is that the results can be somehow interpreted.

Our images (pixel by pixel) have the most variance (in grayscale) in the direction of these eigen-images, i.e. these images capture the most varying differences between all the images. What matters about the eigen-images is their direction, not their values. This intuitively translates into the contrast between the pixels. Even more intuitively, the patterns in the images is what matters not the individual values of each pixel. These patterns help us capture the most variance between our images.

However, PCA does not take into account the predictive power of each of the principal components regarding excitement. The weights of our ridge regression model tell us how much each of these eigen-images are important. Since the ridge regression is a linear classifier the higher the absolute value of the weight the more the impact of that eigen-image in predicting the excitement of that room.

We have computed the relative importance of the eigen-images by getting the sum of absolute values of weights and dividing each weight by it (including the sign of the weight). The absolute value of this measure tells us the magnitude of effect of the measure on the outcome. The following is the list of the top 4 positive eigen-images for the Greece data set.

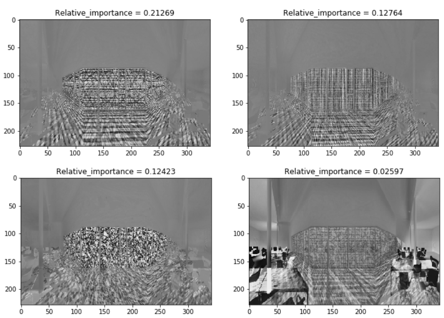


Fig. 5: PCA: Top 4 positive eigen-images

VI. Extrapolating results to every feeling

Having obtained a great model for exciting we finally can extrapolate this methodology to find the a good model for each one of the rest of feelings. The only thing we have to do is to repeat the process described in Section V Subsections C and D. As we got the best result with the 'exciting' label with Ridge Regression before the PCA, we have chosen this methodology for doing a grid search for each label and see how it is working with the others labels. So we get a different models for every feeling and show the test error of the best model obtained for each one. The models obtained for each label is given by the parameters in TABLE II and the FIGURE 6 shows the error obtained for every feeling using their particular best model.

Method	**N PCA**	**Regularization λ**	**Error**
Pleasant	18	13996	0.0983
Interesting	8	19351	0.1079
Exciting	10	1	0.0517
Calming	20	13996	0.0927
Complex	18	13996	0.0999
Bright	8	113896	0.1294
View	16	13996	0.0858
Spacious	9	113895	0.0744

TABLE II: Best Ridge Regression models for every feeling

(Note that all the errors are over the proportions; not percentages). As we see the error obtained for the other feelings is not very different from the obtained for exciting; however we clearly see that the best one is obtained with this label. This may lead us to consider that our work may be extrapolated reasonably to the rest of the labels and improved for each of them separately; so you could follow a similar procedure and you will find a good solution for a specific

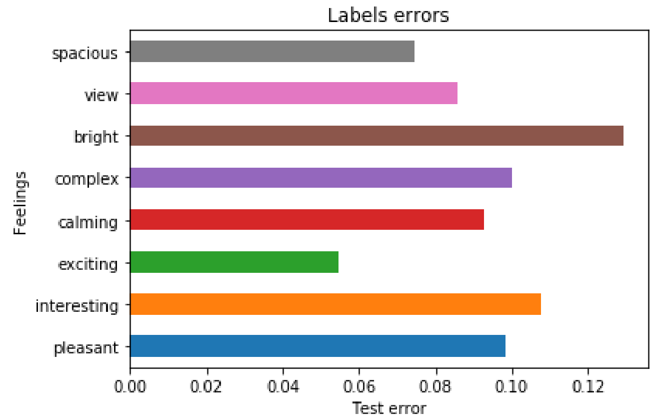


Fig. 6: Error for each label with best found model (feeling)

feeling. (Note that the best model for another label could be given by a regressor different to the Ridge's).

VII. Summary and Results

We have developed a study based on regression to find the way to predict if a scenario will provoke a particular feeling in people (in general) immersed in it. Initially we used some features from the general interpretation of the image to build our model. In order to try to improve the results we used PCA, getting new ways to get information about every scene of the data set and using it to generate new valuable features. In FIGURE 7 we show the comparison of the final error achieve (with the excitation) using the first features and the new ones.

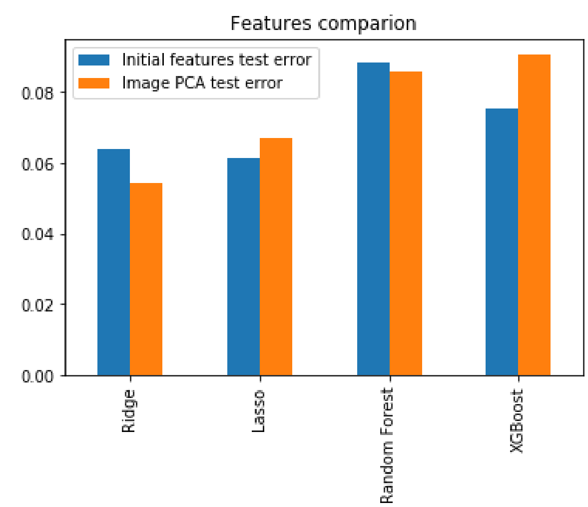


Fig. 7: Error for each label with best found model (feeling)

As we can see, depending on the regressor we get best results with some features or with the others. We finally observe that the final error obtained using the PCA features with Ridge Regression is the smallest, what means that we may have found the way to build a more accurate and valuable model with the new features. This last model will be our final model for predicting future scenarios.