Predicting positions using Independence Graphs

Volker Strobel

June 9, 2016

Abstract

In this report, we use variants of linear regression, independence graphs, and vine copulas, to analyze the distribution and predictive power of image features–color and edge histograms. We compare the predictive power of the used models using the mean absolute error of the predictions on a hold out test set. We study the associated uncertainty in the predictions and give indications, how the proposed model can be used in the physical world.

1 Introduction & Problem Statement

The goal of computer vision is to make machines see—for example, for the purpose of object detection, image restoration, or *localization*. In this report, we consider the following problem:

Based on a given patch of an image, we would like to predict, where the patch was taken in a larger image—the map image. While existing approaches extract key features of the current patch and the map image, these approaches are usually computationally complex and do not allow for in-depth analysis for the problem.

image patches from a given map image are generated. These images simulate camera images that could be obtained during a flight with a UAV. Therefore, we like to address the problem based on two approaches: (i) modify the environment (i.e., the map image) to make it better suited for the used approach, and (ii) modify the approach to make it better suited to the given map image.

The goal is to learn a model and predict the location of unseen image patches. Figure 1 shows the process of data set generation. To this end, we will construct a generative model for the distribution.

$$p(x, y, R, G, B) \tag{1}$$

The regression will be performed by deriving a point estimate from the discrminative model by using the *mode* of the distribution. We decided to create the generative model, since it allows to construct an image from the trained model.

(a) Map image (b) Random image patch

Figure 1: Figure 1a Shows the map image that is used for generating the image patches. Figure 1b shows one sample of the $N = 1000$ image patches that were generated for the generation of the data set.

$$p(x, y \mid R, G, B) \tag{2}$$

.

By construction, we know that x, y are independent, since the positions were sampled from independent uniform distributions.

This leads to the data set that can be found in the appendix.

In this setting, the availability of the data is not a problem and new data can be easily generated. What we would like to have is a dependency model that is able to capture the multivariate distribution of the sample data. This approach is based on Sklar's theorem that states that multivariate distribution can be described by marginal distributions and the dependence structure—-the copulas. Using bivariate copulas as building blocks, more complex interaction structures can be achieved by building *vine copulas*–nested sets of connected trees.

2 Methods

In our first model, we use three features: the average red, green, and blue color per image patch. The data set based on $N = 1000$ image patches can be found in the appendix.

The low values of explained variation for x (R^2(x; rest))and y (R^2(x; rest)) show that the used approach is not promising. This is not surprising, given the fact that the colors are non-linearly distributed in the image (see Figure**??** for the analysis of the colors). Two alternatives could be used to improve the model's performance. Here, we investigate both.

- Alternative 1: Change the map image such that higher R^2 are obtained
- Alternative 2: Use a more powerful model

red	3558				
green	1578	3354			
blue	1219	2749	4016		
x	-1199	-427	-343	22900	
y	-871	-389	616	467	12204
means	180	154	145	244	192
	red	green	blue	x	y

Table 1: The sample variance matrix of the data set

red	1.00				
green	0.46	1.00			
blue	0.32	0.75	1.00		
x	-0.13	-0.05	-0.04	1.00	
y	-0.13	-0.06	0.09	0.03	1.00
	red	green	blue	x	y

Table 2: The sample variance matrix of the data set

2.1 Different map image

In this case, the image was chosen by manual analysis of the features. A more interesting case would be to generate the image based on desired properties of the inverse correlation matrix.

An ideal inverse correlation matrix would look like:

One possible solution would be to set all correlations to a high value.

2.2 The backward approach

Knowing the criteria for computing R^2, we could construct optimal images for the given approach. These images will be gradient images. Since images consist of three channels, we could encode the information in two channels, and use arbitrary information in the third channel—and therefore even keep our initial

red	1.30				
green	-0.59	2.64			
blue	0.02	-1.81	2.37		
x	0.14	-0.02	0.01	1.02	
y	0.13	0.24	-0.32	-0.01	1.06
R^2	0.23	0.62	0.58	0.02	0.06
	red	green	blue	x	y

Table 3: Inverse correlation matrix

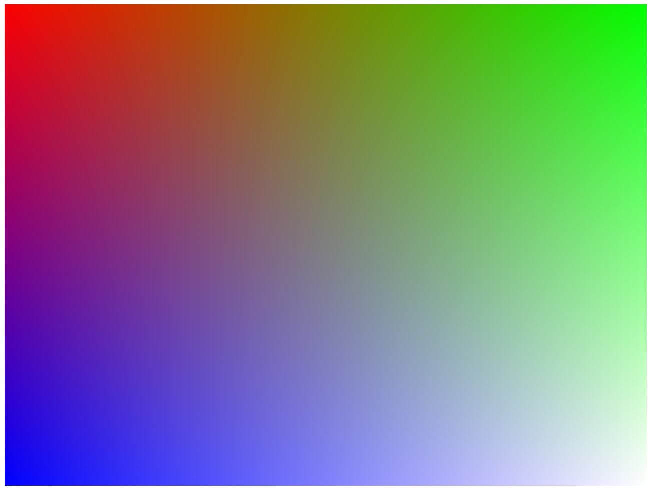

Figure 2: A different map image that makes use of the used features.

red	1.30				
green	-0.59	2.64			
blue	0.02	-1.81	2.37		
x	0.14	-0.02	0.01	1.02	
y	0.13	0.24	-0.32	-0.01	1.06
R^2	0.23	0.62	0.58	0.02	0.06
	red	green	blue	x	y

Table 4: Inverse correlation matrix

image to a great extent.

2.3 Vine Copula

In our final approach, we try to capture the dependence structure without modifications of the original map image. In order to do, we need a more powerful model. The use of *copulas* allows to model marginal distributions and dependence structure independently, allowing for convenient and powerful representation of joint probability distributions.

3 Discussion

The presented results indicate the suitability of our approach. In a real-world setting, run-time and computational complexity will be crucial which we left out here.

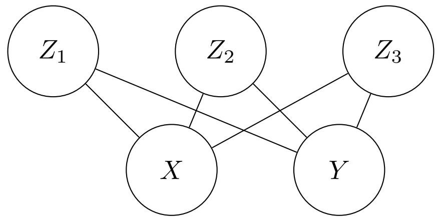

Figure 3: The graphical model used for generating ideal images.

4 Conclusion

In this report, the suitability of independence graphs for computer vision-based localization was shown. While the straight-forward application of linear regression was not able to capture the structure in the generated dataset, a modification of the image underlying the data set allowed to generate a simple and effective model, which lead to low mean squared errors. Since the modification of the underlying image might be undesired or not possible, two more powerful approaches have been presented. The Gaussian

Future research will include the modeling of time dependence: in a real-world scenario, successive images will be, leading a a Bayesian network, as presented by Daphne.

A Supplementary Material

B Dataset with average red, green, and blue values

red	green	blue	x	y
199.15	184.01	148.09	173.21	312.11
161.74	104.92	98.682	471.52	156.44
152.76	115.75	91.875	61.441	70.699
219.71	120.5	92.252	304.45	61.207
221.39	199.51	169.75	164.92	40.368
228.69	224.08	196.25	185.25	259.92
181.83	131.69	147.57	456.16	218.93
195.77	71.636	60.897	32.196	107.07
109.1	193.62	143.95	432.03	95.567
255	255	255	168.08	182.98
234.87	192.93	147.54	473.74	37.964
255	255	255	194.94	135.98
246.92	165.26	213.07	-9.5128	346.47
...	...	...	...	...