FINDING CORRELATION BETWEEN CUSTOMER TYPOLOGY AND SALES RESULTS IN ASSISTED RETAIL USING COMPUTER VISION

[a]JAKUB MACHALICKÝ, [b]LUBOŠ JURÁNEK

Mendel University in Brno, Zemědělská 1, Brno, Czech Republic
email: [a]jakub.machalicky@gmail.com,
[b]xjuranek@node.mendelu.cz

Abstract: We use computer vision capabilities to detect several basic attributes reliably recognizable from an image. We also analyze sales data to compare the connection between these data. The results are evaluated on a small sample by a human, and then everything is applied to a large set of data. The purpose of this comparison is to find a correlation with better (worse) sales results with collected data. This can be used to improve the planning of sales capacities, sales plans and other important indicators for retail store management. The article is narrowly focused on assisted retail in telecommunication business. For age and gender detection, the Cortana Analytics Suite from the Microsoft Azure platform is used. For other image recognition problems, custom algorithms have been created. All this has lead to the creation of a tool usable for anyone who wants to understand better who his customers in retail are.

Keywords: computer vision, retail, sales, footfall, azure

1 Introduction

The goal of computer vision is not only to see the image, but rather to get information from it (Szeliski, 2010). Such information, in the field of assisted retail (all customers are served by a human), is mainly the number of customers who have entered the store. This information is the basic pillar for calculating the main performance indicator of a given store – the Conversion Rate (Marketos and Theodoridis, 2006). In addition to this basic indicator, other information such as gender, age, color of clothing and entrance as a group ("family"), can be easily read. There are a lot of other parameters that can affect customer buying behavior, such as customer mood and walking speed. We are looking for methods of getting information about incoming customers and looking for correlations between the above attributes and sales data.

As Deshwal (2016) writes, there is high correlation in buying willingness differentiated by gender, age and other factors that cannot be detected by computer vision such as family income and education level. This research was done on companies that provide similar service to telecommunications services, with a high focus on maintaining a long-term relationship with customers. According to Roy Dholakia (1999) and Bakewell (2006), motivation for going shopping is connected with customers' mood, in particular cases. There is a difference between have-to-go shopping and want-to-go shopping. This mood can be represented with body poses, facial expressions and colors of clothing. We have neutral colors, negative colors and positive colors (Alanazi, 2018). Which color is positive and negative differs by region and culture. Green is connected with nature and fertility in Western Europe, while in Indonesia it is a forbidden color and in China it can mean cheating one's spouse in a sexual way (Scott-Kemmis, 2018). In this article, we want to find a correlation with sales in a manner independent of positive or negative colors. The hypothesis of this article is that this computer vision can be used to optimize sales planning in retail with the abovementioned factors influencing sales results.

The acquired image and structured data are anonymous and are provided with the consent of the store management. Image data are available for performing recognition. A camera placed above the store entrance takes this data. We also have data on customer requirements / sales and their time of those requests. These requirements are of two types. These are service transactions, such as when a customer claims goods or services or requests changes in services, etc. These are requirements that will not make any profit for companies. We also have sales transactions. The customer purchases a new product, a service, or changes the setting of existing services, but generally, such a transaction brings profit to companies. The number of customers does not match the number of transactions executed. Customers can go to the store only to get information or view the exhibited products. Vice versa, it is possible to observe more than one transaction per customer.

Data from the image and sales information cannot be linked – due to a lack of a common identifier. Therefore, data is divided into time blocks and only numbers in these blocks are compared.

The data obtained from the image is verified in a small section by the human eye, in order to determine the error rate of the methods used. Even this sort of verification is not unambiguous; it is difficult to determine the frontier classification in the age category.

2 Materials and Methods

2.1 Group Size Recognition

The input image for group size recognition is a snapshot from the camera system above the queue management system kiosk, created a few seconds after sending a request for a ticket. Detection is performed by comparing the difference between an image without customers and an input image (one example of group detection is in Figure 1, and an example of one customer detection is in Figure 2) (Piccardi, 2004). Image differencing is a common image processing technique used to determine changes between images. The basic prerequisite for a correct evaluation is that near this kiosk, queues of unrelated customers do not form; however, if more people are detected at a small distance from each other, it is really a group of customers who have come together.

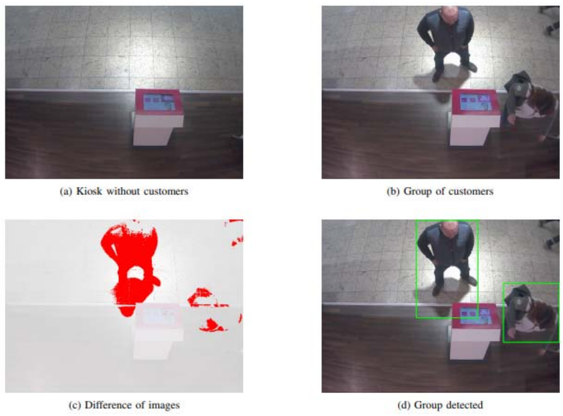

Fig. 1: Example of group detection during group size recognition

The field of view of the camera is not large enough for a greater group of people, and therefore, in the next evaluation of the collected data, we only consider the group flag where the True value corresponds to a group of two or more people, and the False value means only one customer.

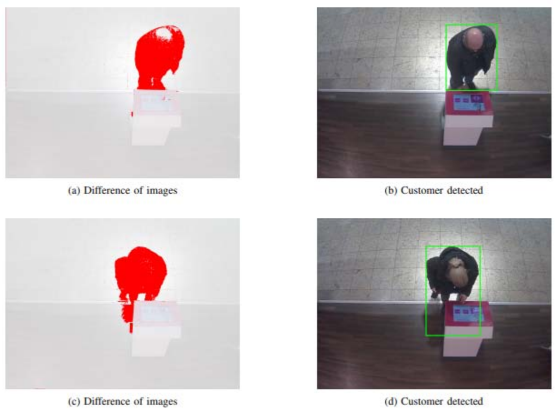

Fig. 2: Example of customer detection during group size recognition

2.2 Gender and Age Recognition

The input image for gender and age recognition is a snapshot from the camera system which is pointing directly towards the entrance to a retail outlet. Face detection and subsequent gender and age estimation are done through machine learning, specifically by using the Cortana Analytics Suite from the Microsoft Azure cloud computing platform (URL: https://azure.microsoft.com/cs-cz/).

The Face API of Microsoft Azure can extract facial landmarks from photographs, such as the locations of the pupils, eye corners, eyebrow edges, nose tip and lip boundaries. By comparing those points with AI learning models, this cloud service can estimate age based on data about how facial landmarks tend to shift and change with aging and gender based on data about what differences of facial landmarks there are between male and female.

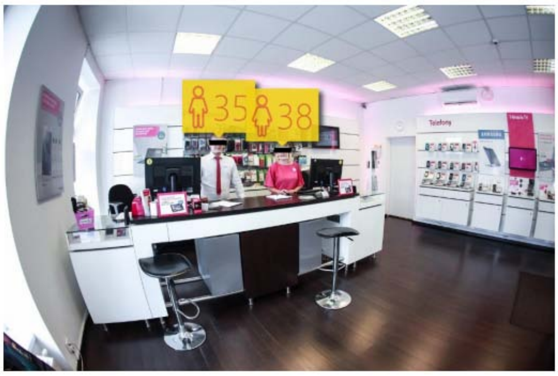

Fig. 3: Example of gender and age detection

Due to the anonymization of customers, the example in Figure 3 is only meant for illustration, being the image output of the HowOldRobot application (URL: https://www.how-old.net/), which serves as a demonstration of the Cortana Analytics Suite capabilities. The output of the Cortana Analytics Suite itself contains the following JSON data:

["face": {"age": 34.5, "gender": "Male"}, "face": {"age": 38.0, "gender": "Female"}]

In this particular case, the age estimation error is a few years compared to reality. Because we assume a fairly high error rate for all the input images, in the further evaluation of the collected data, we consider grouping the estimated age into five categories: 0–17 years, 18–26 years, 27–40 years, 41–60 years and 61+ years.

2.3 Color Recognition of Clothes

The input image for the color recognition of clothes is a snapshot from a camera system, observing the horizontal movement towards a retail outlet. In order to detect the color of the clothes in a way that isn't affected by the hair or the shadow of the customer, it is necessary to crop the input image from the bottom and top by a certain percentage of the height of the detected figure, for example, by 20 %. We need to extract, from the resulting image, a palette of predefined colors according to their similarity to the colors in the image (Comaniciu and Meer, 1997). Because the floor of the outlet is gray, we also have to neglect this color if it isn't the dominant color on the image. In this case, there is a high probability that the customer has clothing colored similarly to the floor color. The pair of examples is in Figure 4.

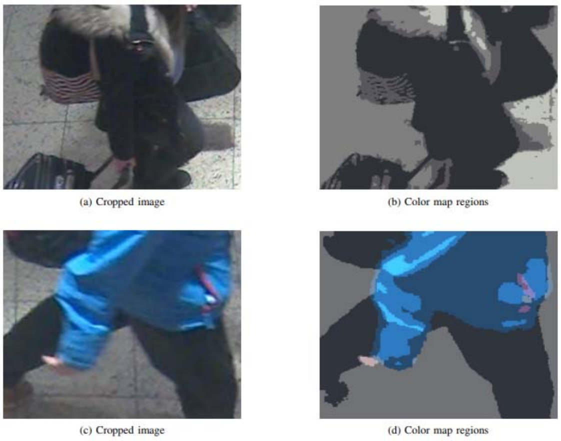

Fig. 4: Example of color detection of an incoming customer's clothes

In some cases, it is possible to subtract the background of the image, which is the floor, by algorithms for background-subtracting; for example, using frame differencing (Singla, 2014), and then not considering the background in color recognition. This method has only a 27 % success rate in the image set due to inappropriate lighting conditions. The pair of examples with successful background subtraction is in Figure 5.

Fig. 5: Example of color detection of incoming customer's clothes

The output of this detection is the set of colors and the corresponding percentage of their representation in the color map of the processed image. In the case of the last two examples, it's the following data in the JSON format:

["clothes": {"Grey": 71.8, "Blue": 28.2}, "clothes": { "Black": 85.1, "Red": 9.8, "Grey": 5.1}]

2.4 Output Information about Customer

After performing the above-described image recognition operations on camera images, each customer is provided with a list of properties in the following JSON format:

["customer": {"group": Boolean, "face": Array.of(Object), "clothes": Array.of(Object)}]

3 Results

The data was collected at 3 stores in one month, only on working days in the opening hours, between 9am and 9pm.

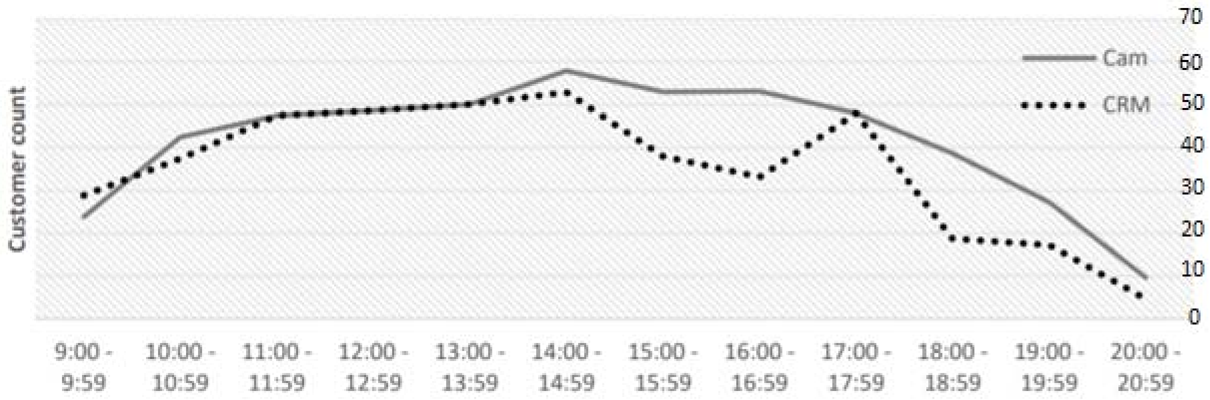

Fig. 6: Correlation between customer count gained by cameras and CRM data

The number of customers counted using cameras has been compared with the number of customers served by CRM (customer relationship management) systems. The number of customers is a correlation confirming the correctness of the methods used. The mathematical definition of correlation says that it is the degree to which two or more quantities are linearly associated. In a two-dimensional plot, the degree of correlation between the values on the two axes is quantified by the so-called correlation coefficient (Spiegel, 1992). For the purposes of this article, it is simplified to two directional curves. One is for CRM data and the second is for data extracted from the image. When the direction of the first curve is the same as the direction of the second curve, we call it correlation (increasing maximization vs. decreasing minimization). These curves are separated into hour-long blocks of time in which we look for those correlations. The deviations visible in the graph (Figure 6) are due to the arrival of groups, when only one person in the group is actually served in the CRM systems.

The data is grouped into hour-long blocks within which the maxima and minima (curve directions) are found, which can be used to determine that at a given time, more persons meeting those parameters have visited the store, and further notice, whether or not sales have been diverted at that time, up or down. Table I shows the distribution of the conversion rate over time and the distribution of traffic over time by the total number of units on the one-hour clock. This is the baseline for each of the further searches.

Tab. I.: Conversion rate and footfall according to time

Hour	Conversion rate	Footfall
9:00 – 9:59	0.08	0.05
10:00 – 10:59	0.02	0.09
11:00 – 11:59	0.03	0.09
12:00 – 12:59	0.05	0.09
13:00 – 13:59	0.03	0.10
14:00 – 14:59	0.22	0.13
15:00 – 15:59	0.17	0.10
16:00 – 16:59	0.06	0.11
17:00 – 17:59	0.15	0.09
18:00 – 18:59	0.07	0.08
19:00 – 19:59	0.07	0.06
20:00 – 20:59	0.22	0.01

Another output is the number of customers by gender. No data is available to verify the sex of the incoming customer other than the cameras. The data was verified by a brief observation of a man in parallel with the computer data, and the resulting deviation was assessed as negligible. The ratio of incoming customers by gender is shown in Table II. The percentages shown are from the total number of incoming customers.

Tab. II.: Incoming customers by gender and groups

Hour	Solo	Group	Female	Male
9:00 – 9:59	61%	39%	64%	36%
10:00 – 10:59	62%	38%	59%	41%
11:00 – 11:59	62%	38%	60%	40%
12:00 – 12:59	62%	38%	58%	42%
13:00 – 13:59	61%	39%	60%	40%
14:00 – 14:59	62%	38%	58%	42%
15:00 – 15:59	62%	38%	60%	40%
16:00 – 16:59	62%	38%	58%	42%
17:00 – 17:59	63%	37%	59%	41%
18:00 – 18:59	61%	39%	60%	40%
19:00 – 19:59	61%	39%	59%	41%
20:00 – 20:59	67%	33%	57%	43%

Table II also shows the arrival of customers in groups. The methods used were equally validated by observation. Unfortunately, this error rate is significantly greater. Groups do not always come to the shop together, but sometimes at different hours. However, it is true that a larger proportion of customers arrive alone and the resulting ratio, despite the inaccurate input data, is accurate.

The methods used make it possible to determine the age as an exact number. This information has been generalized and customers have been grouped. Customer arrivals by age are shown in Table III.

Tab. III.: Incoming customers by age

Hour	0–17	18–26	27–40	41–60	61+
9:00 – 9:59	11%	39%	28%	13%	8%
10:00 – 10:59	9%	39%	31%	12%	9%
11:00 – 11:59	10%	40%	30%	12%	7%
12:00 – 12:59	10%	39%	27%	14%	10%
13:00 – 13:59	10%	40%	29%	13%	9%
14:00 – 14:59	11%	39%	29%	13%	8%
15:00 – 15:59	12%	41%	28%	11%	8%
16:00 – 16:59	9%	42%	28%	12%	8%
17:00 – 17:59	11%	40%	28%	12%	9%
18:00 – 18:59	9%	42%	29%	12%	8%
19:00 – 19:59	11%	41%	27%	11%	9%
20:00 – 20:59	14%	39%	30%	6%	11%

The breakdown of customers in Table III by age was verified by a brief observation. Deviations due to grouping are negligible and these data are accurate.

The methods used provide information about all the customers' clothing colors. Only one dominant color was used for the evaluation. The color distribution is shown in Table IV. Colors other than black, brown, blue, grey and pink are omitted because it is impossible to group these colors together.

Tab. IV: Incoming customers by dominant color of clothes

Hour	Black	Brown	Blue	Grey	Pink	Other
9:00 – 9:59	26%	8%	12%	33%	3%	18%
10:00 – 10:59	31%	8%	11%	39%	6%	5%
11:00 – 11:59	25%	8%	12%	34%	5%	16%
12:00 – 12:59	31%	8%	12%	36%	5%	8%
13:00 – 13:59	27%	8%	13%	38%	5%	9%
14:00 – 14:59	25%	8%	12%	34%	5%	16%
15:00 – 15:59	26%	9%	12%	36%	4%	13%
16:00 – 16:59	26%	7%	11%	35%	4%	17%
17:00 – 17:59	29%	9%	12%	39%	5%	6%
18:00 – 18:59	31%	8%	12%	39%	5%	5%
19:00 – 19:59	25%	9%	13%	36%	3%	14%
20:00 – 20:59	32%	8%	8%	34%	4%	14%

3.1 Correlations

In addition, the maximum and minimum is searched for in each group. If a maximum or minimum is found, it is then compared with the conversion rate. If the conversion rate is smaller at that point or larger than the average, this is considered a correlation. These are divided into two types. The first is the correlation where the entire group only increases or decreases over all time blocks. The second is partial; here a time block is found where it decreases within the group, but a different time block with an increase is also found. All these correlations are shown in Table V.

Tab. V.: Correlations list

Group	Value	Correlation	Business	Hour	
Solo/group	solo	partial	decrease	19:00:00	19:59
Solo/group	solo	partial	increase	20:00:00	20:59
Gender	male	partial	decrease	9:00:00	9:59
Gender	male	partial	increase	20:00:00	20:59
Age	0–17	partial	decrease	18:00:00	18:59
Age	0–17	partial	increase	20:00:00	20:59
Age	18–26	full	decrease	-	
Age	27–40	full	decrease	-	
Age	61+	partial	decrease	11:00:00	11:59
Age	61+	partial	increase	20:00:00	20:59
Color	Black	partial	decrease	19:00:00	19:59
Color	Black	partial	increase	20:00:00	20:59
Color	Brown	full	decrease	-	
Color	Grey	full	decrease	-	
Color	Pink	partial	increase	14:00:00	14:59
Color	Pink	partial	decrease	19:00:00	19:59

4 Discussion and Conclusion

From the collected data, conclusions can be made. The data was collected over a period of twenty working days within one month (11/2018). As can be seen in Table VI, the total number of records is nearly thirty thousand. Records containing unsold transactions were neglected.

Tab. VI.: Data summary

Type	Count of type
cam	18828
sale	1739
care	8275
total	28842

Most of the correlations found are only partial, i.e., they are valid in a specific part of the day but are not general. For this reason, it is advisable to generalize the time divisions with a granularity of three. This is by four hours (morning, afternoon, evening). Then we can make conclusions in these blocks. Another factor that undoubtedly influences the conclusions is the seasons in two areas. It can be hypothesized that the color composition will be more distinguished in warmer months, among other things, because people have more exposed skin, the color of the skin would be neglected in the results, and then the impact of the remaining parts of the clothing may increase. Another hypothesis takes into account the division into time blocks where, for example, due to daylight and darkness, these shares can be different. These hypotheses will be part of further experiments.

It has to be said that these data were collected in the narrowly specific category of retail, telecommunications. They may vary considerably in another sector. At the same time, the period (autumn) in which the data was collected may vary considerably across the year.

In general, the following can be inferred from the results:

- Stores that are visited by more people aged 18-40 have lower potential
- Stores that are visited by more people dressed in brown and gray have lower potential
- Stores visited by more people over the age of 61 in the morning have lower potential
- Stores visited by more people over the age of 61 in the evening have higher potential
- Stores that are visited by more people dressed mostly in pink in the afternoon have a higher potential
- Stores that are visited by more people dressed mostly in pink in the evening have a lower potential

Partial results within one four-hour block are not conclusive enough.

Results show that factors such as age, gender and color of clothes affect sales results and customer willingness to buy. We can confirm that results from India (Deschwal, 2016) and from the USA (Roy Dholakia, 1999) are valid also in the Czech Republic in the field of telecommunications. We can develop an application which will use computer vision and update these statistics over time. With valid information, it is possible to achieve better sales results or lower costs. With information that one of our stores is visited mostly by customers aged 18-40 (which has lower potential, according to our results), we set our sales goals to be more motivational and get better performance from our sales representatives. This is the main benefit of this research. The chosen methodology has proved to be applicable for further research and commercial use.

Literature:

1. ALANAZI, Fatimah Mohammed 2018. Communicating emotions using cloth colors, body poses, and facial expressions on static virtual characters. Carbondale. Thesis. Southern Illinois University.
2. BAKEWELL, C., MITCHELL V. 2006. Male versus female consumer decision making styles. Journal of Business Research, 59(12), pp. 1297-1300. DOI: 10.1016/j.jbusres.2006.09.008. ISSN 01482963.
3. COMANICIU, D., and MEER, P. 1997. Robust analysis of feature spaces: color image segmentation. In Proceedings of IEEE computer society conference on computer vision and pattern recognition pp. 750-755. IEEE.
4. DESHWAL, P. 2016. Customer experience quality and demographic variables (age, gender, education level, and family income) in retail stores, International Journal of Retail & Distribution Management, 44(9), pp. 940-955. DOI: 10.1108/IJRDM-03-2016-0031.
5. MARKETOS, G., and THEODORIDIS, Y. 2006. Measuring performance in the retail industry (position paper). In International Conference on Business Process Management (pp. 129-140). Springer, Berlin, Heidelberg.

6. PICCARDI, M. 2004. Background subtraction techniques: a review. In 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583) (Vol. 4, pp. 3099-3104). IEEE.
7. ROY DHOLAKIA, R. 1999. Going shopping: key determinants of shopping behaviors and motivations, International Journal of Retail & Distribution Management, 27(4), pp. 154-165. DOI: 10.1108/09590559910268499.
8. SCOTT-KEMMIS, J. 2018. Empower yourself with color psychology [online]. https://www.empower-yourself-with-color-psychology.com/cultural-color.html
9. SINGLA, N., 2014. Motion detection based on frame difference method. International Journal of Information & Computation Technology, 4(15), pp. 1559-1565.
10. SPIEGEL, M. R. 1992. Correlation Theory. Ch. 14 in Theory and Problems of Probability and Statistics, 2nd ed., New York: McGraw-Hill, pp. 294-323.
11. SZELISKI, R. 2010. Algorithms and Applications. Computer Vision, Springer-Verlag New York, Inc., pp. 201-13.

Primary Paper Section: I

Secondary Paper Section: IN