Retail Store Location:

An application of signal processing to the detection of location opportunities for bakeries in France

Jerome BARAY - CREREG - Rennes University - 11, rue Jean Macé - 35000 Rennes

Gerard CLIQUET- CREREG - Rennes University - 11, rue Jean Macé - 35000 Rennes

Retail Store Location: an application of signal processing

to the detection of location opportunities for bakeries in France

Introduction

Selecting a convenient and profitable site for the establishment of a commercial activity is a risky complicated task, which necessitates to gather huge database of consumers if the location analysis needs to be accurate. In fact, the common ways exposed in the literature to find optimal locations take as first assumption that stores are to be as closed as possible to consumers, which leads logically to this high complexity degree problem. An other manner to tackle this question is to consider the presence of competitors. We assume that a high density of competitors on some wide areas of a global territory constitutes a clue, which indicates there a strong demand rate for the category of products distributed. The objective of the present paper is thus to describe a fast, demonstrative and easy method to detect geographical commercial opportunities in considering that there must be in these high demand areas, some less exploited clusters in which the level of competition is weaker althought the consumer potential is approximately the same. We will try to detect these gaps of commercial offer using signal processing and statistical analysis having always in mind the competitor's presence.

Theory and practice of commercial location choice

The major methods whose aim is to look for optimal commercial locations usually take as base of analysis the repartition of potential customers in space. Among the theoretical methods, the gravitation models i.e. law of Reilly [1], the MCI model or Interactive Model of Competition [2], seek to locate the outlets as nearest as possible to the customers so as to attract their greatest part whereas location-allocation models [3] try to find the closest sites to customers thanks to mathematical algorithms. In the last category, the objective of the p-median model [4] [5] [6] is to find the locations for a number of p activities having to provide some services or products to N customers in such a way that the sum of the distances separating each customer to the closest activity is minimal [7]. Only the

---

[1] REILLY W. J. (1931) *The Law of Retail Gravitation*, W. Reilly ed, 285 Madison Ave, New York, NY.

[2] NAKANISHI M. et COOPER L.G. (1974) Parameter Estimates for Multiplicative Competitive Interaction Models: Least Square Approach, *Journal of Marketing Research* 11: 303-311.

[3] GHOSH A., RUSHTON G. (1987) *Spatial Analysis and Location-Allocation Models*, Van Nostrand, Rheinhold, 1987.

[4] WEBER A. (1909) *Über den Standort der Industrien*, Tübingen, Traduction Anglaise de Friedrich (1929) *Theory of the Location of Industries*, University of Chicago Press, Chicago.

[5] HAKIMI S.L. (1965) Optimum Distribution of Switching Centers in a Communication Network and Some Related Graph Theoretic Problems, *Operations Research* 13, 462-475.

[6] REVELLE C. et SWAIN R. (1970) Central Facilities Location, *Geographical Analysis*, 2, 30-40.

[7] BEAUMONT J.R. (1987) Location-Allocation Models and Central Place Theory, *Spatial Analysis and Location-Allocation Models*, ed. Ghosh Avijit, Rushton Gerard, Van Nostrand Reinhold.

central places theory of Christaller [8], a space interaction model, takes into account the repartition of the competitors alone : indeed, Christaller 's theory assumes that geographical space understands a uniform distribution of identical customers: the optimal establishment is then at the center of a hexagon whose apices are occupied by six elementary stores.

In practice, professionals use rather their intuition to locate a new store [9] [10] and if not, some simple and quick statistical methods with socio-economic data like the multiple regression model and the analog method [11] [12]. They balk at adopting heavy methods or expert systems developed by researchers. This hesitation comes also from the fact is also that it is difficult to get some database of potential customers to feed such models. On the other hand, it is easier to get records of stores addresses extracted for example from the yellow pages. In the distribution of basic products, stores tend to avoid competition in establishing in unexploited commercial places. This leads us to propose a new method based on signal processing and statistical analysis in which the positioning of existing stores could be the basic information used to find some new optimal commercial locations taking also into account the type of customers patronizing this retail. Signal processing has already been used to outline trade areas and to find optimal locations in combination with a p-median model but formerly with databases of potential customers. This time, this powerful processing method applied on a database of stores will enable to detect free interesting commercial places and to outline the establishment areas of stores belonging to a specific commercial activity.

A new approach for the search of optimal locations

Our method for location search is divided in 3 stages, a mapping process, an outlining of the commercial establishment areas and a detection step inside the commercial areas to find the gaps related to the commercial existing offer.

More precisely, **stage 1** consists in mapping the density of stores on a global territory (a country or a region).

Stage 2 enables to segment the territory of analysis in the establishment areas of stores. These areas underlying at least a certain density of stores are outlined and featured thanks signal processing techniques.

---

[8] CHRISTALLER W. (1935) *Die Zentralen Orte in Süddeutschland*, G.Fischer, Germany, Jena.

[9] CLARKE I., HORITA M. et MACKANESS W. (2002) Intuition et Evaluation des Sites Commerciaux : Appréhender la Connaissance des Commerçants, , *Stratégies de localisation des Entreprises Commerciales et Industrielles : De Nouvelles Perspectives*, p.107, Gérard Cliquet et Jean-Michel Josselin éditeurs, De Boeck Université, à paraître.

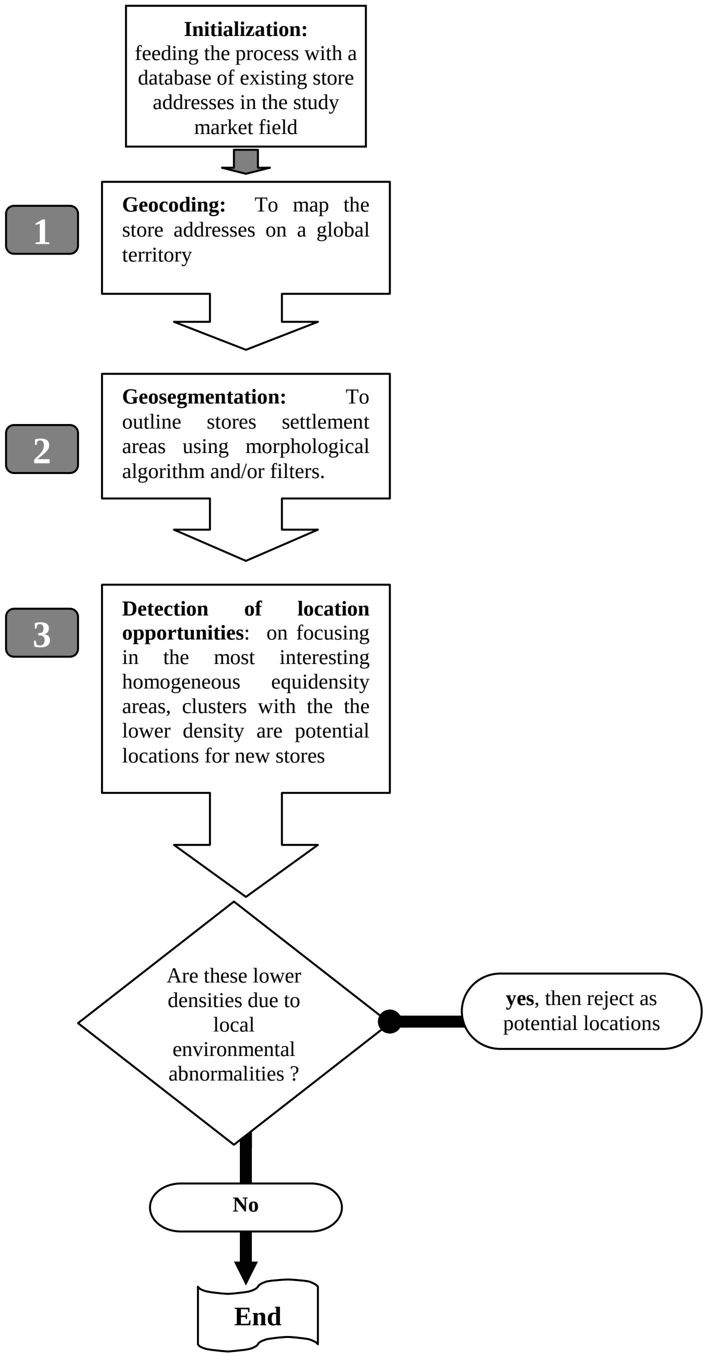
[10] HERNANDEZ J.A. (1998) *The Role of Geographical Information Systems Within Retail Location Decision Making*, PhD Thesis, The Manchester Metropolitan University Manchester.

[11] APPLEBAUM W. (1968) The Analog Method for Estimating Potential Store Sales, *Guide to Store Location Research*, Addison-Wesley, Reading, Mass.

[12] ROGERS D.S. (1980) 5 ways to Evaluate a Store Location, *Store Location*, 42-48.

In **stage 3**, the process focuses on the most interesting areas and reveals the sub-areas in which the density of stores is lower than normal. These clusters can constitute opportunities for the location of new stores if these sub-areas do not show lower densities due to atypical environment parameters decreasing the local commercial demand.

Flow chart of the new method of location search suggested



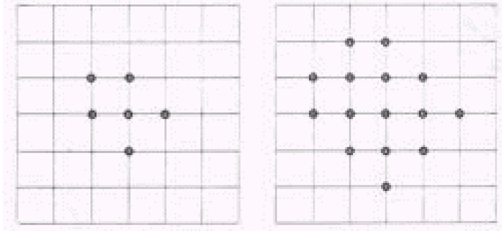
Thus, the method consists to find the borders of the settlement areas of stores on a territory. In a second stage, one concentrates the study on some of this regions to detect some possible gap in term of store facilities and densities, that means places where stores are expected to be found but are not there, according to the population or a statistical parameter linked to the commercial demand for the concerned activity. We will now study each phase more in detail.

Segmenting the territory in settlement areas the stores by signal processing

The objective of this stage is to specify the areas in which there is similar density of stores. The mapping of stores density must have been done accurately beforehand with a traditional mapping software i.e. MapPoint, MapInfo, and ArcView. These existing stores addresses extracted for example from the yellow pages can belong to different commercial fields in the retail sector (bakeries, jeweller's) or from the services (insurance, banks, leisure activities, campings,...) depending of course on which kind of commercial activity needs to be located. The map will show the density of stores concerning a socio-economical parameter or a combination of parameters in direct relation with the commercial demand evolving in the geographical space. It can be for example in the case of tobacconist location search, the number of tobacconists per thousand smokers or to locate telecom agencies, the number of agencies per thousand phone subscribers or more simply a ratio linked to the population or to its purchase power when considering the basic products or services distribution.

The boundaries of the stores major establishment areas will then be determined in using morphological functions directly on the preceding binarised electronical map made of pixels. The basic morphological functions are the *dilatation* and the *erosion*. The dilatation function (Fig. 1) consists in adding pixels to the edges of objects in binary images. A pixel is added (set to black) if four or more of its eight neighbours are black. Generally speaking, dilation connects discontinuous objects and fills in holes. On the contrary, the erosion removes pixels from the edges of objects in a binary image, where contiguous black areas in the image are considered objects, and background is assumed to be white. A pixel is removed (set to white) if four or more of its eight neighbours are white. Erosion separates objects that are touching and removes isolated pixels. A more elaborated function is the *opening*, which performs an erosion operation, followed by dilation: this process smoothes objects and removes isolated pixels.

Fig. 1 : The original shape The dilated shape



The opening function makes it possible to distinguish the shape of the stores settlement areas.

Their edges are simply drawn with a signal processing function named the Sobel edge detection operation. Two convolutions are done using the kernels shown below, generating vertical and horizontal derivatives. The results are then combined by using the square root of the sum of the squares of the two derivatives.

$$Sx = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad ; \quad Sy = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$$

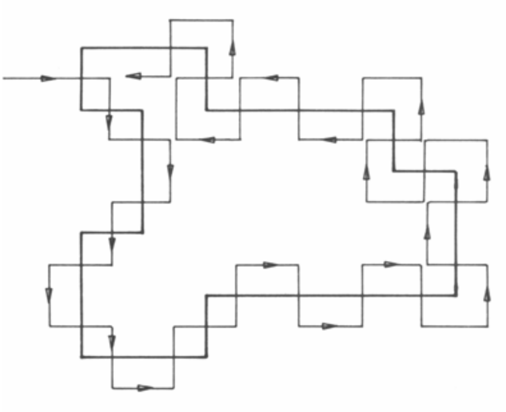
Horizontal detection matrix ; *Vertical detection matrix*

We must recall that the convolution between two matrices, an image fi,j (the map) and a filter Hi,j (the Sobel operators) is given by the following mathematical formula:

$$g_{i,j} = f_{i,j} * S_{i,j} = \sum_{\alpha=i-k}^{i+k} \sum_{\beta=j-k}^{j+k} f_{\alpha,\beta} \, S_{i-\alpha,j-\beta}$$

To specify the importance of each store settlement region, an edge detection sub-algorithm (Fig. 2) is used to scan their edge and calculate the corresponding area.

Fig. 2 : Example of a trade area border follow-up using the edge detection sub-algorithm



If the border is described in terms of Northern, Southern, Western and Eastern orientations, the surface of the zone is given by the following process:

1) u = 0 and t =0

2) **For** i = 1 **To** n **Do**

A) **If** a_i **=** East	**Then**	t =	t + 1	**Else Go To** B)
B) **If** a_i **=** South	**Then**	u =	u + t	**Else Go To** C)
C) **If** a_i **=** West	**Then**	t =	t - 1	**Else Go To** D)
D) **If** a_i **=** North	**Then**	u =	u - t	

3) S_g = u x S

where the value of the parameter U at the end of the procedure is the number of pixels contained in the trade area zone, T a counter and S the geographical unit surface of a pixel.

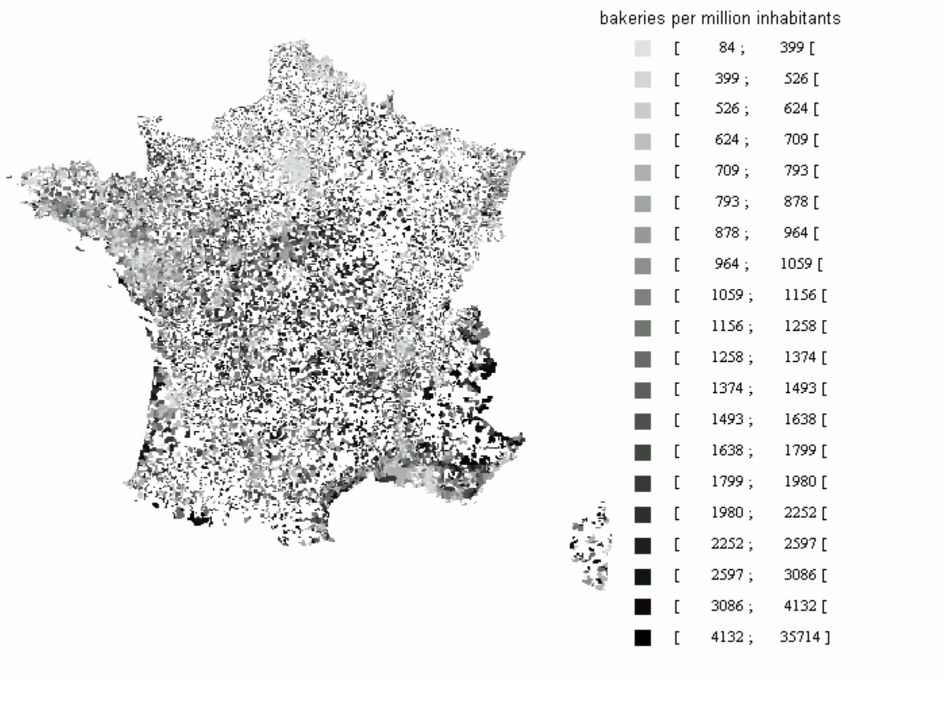
The detection of location opportunities on local scale

After having specified the settlement areas of stores, we will then focus on specific regions and work at a thinner scale. The cumulative histogram of the number of stores in function of the density can be plotted in the local regions to study the density repartition and detect at which level this density is lower than normal. A density of stores under a normal level in the area can indicate an opportunity for a new store establishment if this density has been properly calculated according to the right population or socio-economic parameter representing the flow of potential customers. Nevertheless, as it is likely not possible to include all parameters influencing the store patronage, it is recommended at this stage to review all variables playing a role in the demand to make sure a lower density is not due to environmental local differences. To illustrate this process, we propose to use our method to find potential locations for bakeries in France.

An application of signal processing for finding some optimal locations for bakeries in France

We have used a database of bakeries and a mapping software (Cartes & Données) to first plot the density of stores in French cities. The following map (Fig. 3) shows the number of bakeries per million inhabitants according to a statistical repartition in quantiles in which all grey shading underlines the same number of elements. The administrative cell in this map is the city. 20 classes have been defined with 654 cities in each. The average number of bakeries in a town is 2,54 with a maximum of 1102 bakeries for Paris city.

Fig. 3 : The density of bakeries in French cities according to the population



The areas in grey correspond to the settlement places of bakeries and on the contrary, the white areas where no bakeries can be found likely because the potential demand is too weak (the number of inhabitants in these regions is too small). Then, this map has been transformed in a binary image (Fig. 4), which shows only the presence (in black) or this absence (in white) of bakeries. The preceding greyscale image has been converted to binary by setting pixels that have been highlighted by thresholding to black (255 value), and all other pixels to white (0 value) thanks to the Scion image processing software. To segment the territory of analysis in the establishment areas of stores, we then performed an opening morphological transform: 4 openings (Fig. 5) were necessary to smooth the shapes, remove isolated presence of bakeries and to distinguish the settlement areas (a fifth opening would not have changed the general shapes of areas). A Sobel filter applied with a convolution product then showed the edges of the major settlement regions of bakeries in France (Fig. 6). These high-density regions should correspond to places with the highest demand for bread and relative products.

Fig. 4: The map of density converted in binary image

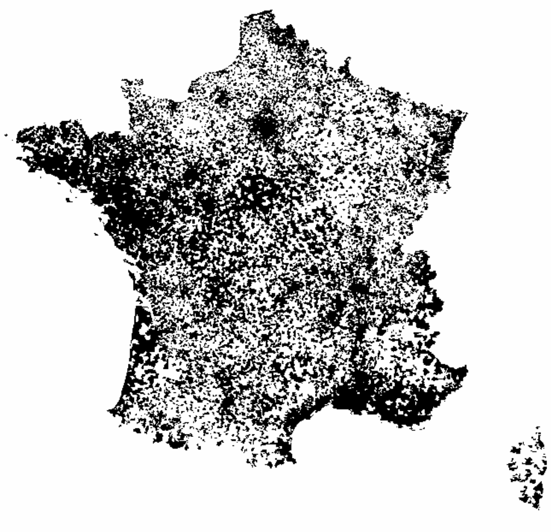


Fig. 5 : After 4 openings transforms

Fig. 6 : The edges drawn by a Sobel filter

Thus, the major areas where bakeries are present in great number are in reference to figure n°7 : 4 (Brittany-Pays de Loire), 11 (Mediterranean region - Rhone valley), 13 (Aquitaine), 7 (Lyons), 3 (Alsace), 2 (Paris), 1 (North), 15 (Tarn), 8 (Haute-Savoie), 12 (Grenoble), 9 (Puy de Dôme), 14 (Gorges du Tarn), 5 (Yonne), 16 (Castres), 6 (Nièvre), 10 (North Correze). The software gave us the area and the coordinates of these areas (Table 1) with the edge detection sub-algorithm. The origin corresponds to the upper left corner of the image.

Fig. 7 : The labelled regions of bakeries

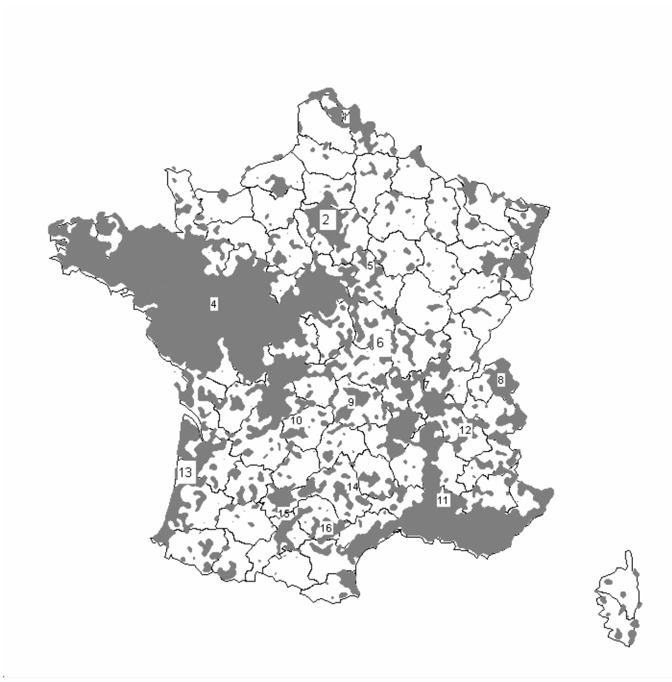


Table 1 : Area and coordinates of the regions

	Area	X	Y
1	1398	358	116
2	1536	338	225
3	1892	535	251
4	25121	221	311
5	577	384	271
6	543	396	355
7	2076	442	396
8	933	520	390
9	601	363	414
10	530	307	433
11	9589	459	516
12	615	483	443
13	2789	192	490
14	596	365	501
15	1296	294	530
16	569	337	545

Having determined the main settlement regions likely corresponding to a higher demand for the products distributed by bakeries, we can the focus on some of these regions to specify some interesting places to locate new bakeries. Region n°4 is the biggest area gathering a big amount of bakeries. If our objective were to locate some stores in the western part of France in such a region, then we would examine at a thinner scale where there could be some free spaces for new commercial activity in this field. The fact that bakeries are typically stores of proximity must also be kept in mind. We will then not choose to locate our stores nearby other bakeries but on the contrary in empty commercial clusters to avoid competition. Brittany (region 4, the most western part of France) is the biggest area having a high density of bakeries. The histogram representing the number of cities according to the number of inhabitants per bakery (in % for the upper diagram (Fig.9) and in number for the lower (Fig. 10)).

Fig. 9 : Number of cities in % according to the number of inhabitants per bakery

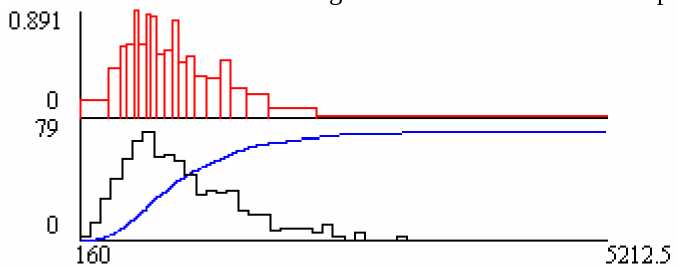
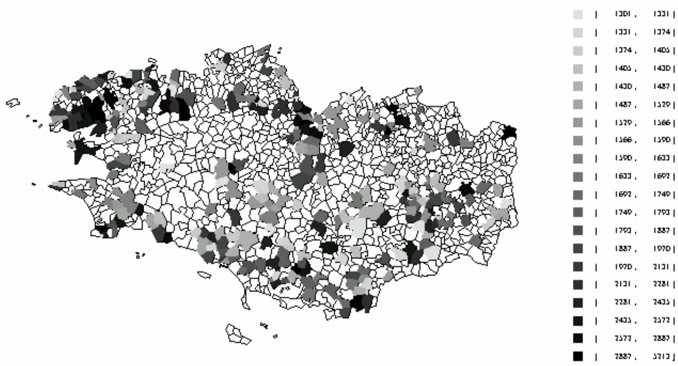


Fig. 10 : Number of cities according to the number of inhabitants per bakery (histogram) and cumulative curve

The upper bound of the histogram corresponds to 800 inhabitants for 1 bakery. Under this level, it is possible to say that the bakery facilities rate is high and above that it is low. We notice that there is an inflection for a lower presence of bakeries corresponding to 1300 inhabitants for 1 bakery. Above this rate, the cities are particularly and abnormally poor in bakeries. The following map (Fig.11) shows these cities with a very low rate of bakeries in a region where this activity is flourishing on the contrary.

Fig 11 : Cities with an abnormal rate of bakeries



The image processing software enables to label (Fig. 12) and to calculate the coordinates and areas of the major clusters of cities having this feature (Table 2).

Fig. 12 : the main clusters of cities having a very low rate of bakeries

Table 2 : the corresponding area and coordinates

	Area	X	Y
1	804	147	104
2	585	58	93
3	547	210	105
4	1951	73	121
5	1730	268	147
6	511	345	124
7	653	195	176
8	2381	377	214
9	924	232	200
10	1128	319	210
11	893	93	208
12	1348	189	239
13	1531	249	262
14	1049	311	283

These very low density clusters are : Plourin-les-Morlais (1), Plougerneau (2), Brest and surroundings (4), Plemy (5), Plouer-sur-Rance (6), Rostrenen (7), Rennes (8), Sourn (9), Paimpont (10), Plomelin (11), Caudan (12), Plumergat (13), Nivillac/St. Doley (14).

These clusters of cities can then feature some opportunities for the creation of a bakery for the level of offer in this commercial field is low compared to the bakeries density in the surroundings and in the region of Brittany on the contrary rather high. It would have been possible to try to locate a bakery in the areas with a weak density of stores outside the 16 French regions outlined by signal processing. But as the importance and turnover of bakeries are comparable and not so different, it would mean to find a site in areas where the demand for bread and related products is not so high for various cultural and socio-economic reasons. On first sigh, it appears but must be confirmed by further study that French areas with a high density of bakeries corresponds to populated regions having at once a rural, a tourist and a gastronomic reputation e.g. Brittany, Pays de Loire, Haute-Savoie, Mediterranean region, Alsace, Gorges du Tarn, Lyons. Their high density could be explained by the various types of customers bakeries can locally attract.

Limits and Perspectives

The method of the present paper based on a general outlining of commercial establishment areas with signal processing and a local statistical analysis shows a certain number of limits. First, although its advantage is to use mainly a database of existing stores and general statistics only, it has not the precision reached by a direct examination of potential customers. Of course, as said before, an exhaustive database of customers is harder to obtain and often impossible to build. Secondly, the procedure needs to have a good knowledge of the clients in the studied field to calculate the appropriate densities. In our study, we have used the simplest density based on population importance because we assumed that the number of bakeries in each city is correlated by this factor. To refine the results, it would be better to take into account a socio-economic variable or a combination of variables better representing the mass of consumers patronizing baker's shops that could drive to realize a survey among customers and a marketing segmentation. As a consequence of this last remark, this method is mainly valid in the case of establishing activities, which can be compared to those of the general database of stores. Starting up a new concept aiming to appeal a specific and different category of consumers would not lead to the same analysis. A promising perspective of research is precisely to locate such activities centred on a particular segment. The procedure would combine the mapping of statistic factors, signal processing and a location-allocation model a bit as done in a previous research but with an exhaustive database of consumers. We would then take into account and map not the densities of existing stores but directly the statistics representing the consumer target on a global territory. After outlining the areas gathering the main mass of consumers e.g. the trade area, a location-allocation model can be used to locate precisely the future stores.

Conclusion

The present paper exposes a fast and demonstrative method for outlining the main areas concentrating a specific commercial activity and detecting geographical establishment opportunities. Based on an image and a statistical process, it is composed of a mapping of existing stores, an outlining stage with a smoothing morphological function and a delimitating filter and finally a statistical analysis to detect facilities gaps in local areas showing the highest demand rates. This frame can be significant for the manager of a commercial network or for the businessman who looks for investment opportunities in unexploited areas. The method can also lead to make a complete list of these valuable free clusters in different fields which would be of a high value for chambers of commerce, consulting firms or real estate agents having a role of advisors for retail stores purchase or creation. Among its main benefits is the possibility to use the current database of stores easily available rather than unaffordable potential customer records. A new orientation of this research will be to introduce local statistics on consumers instead of a commercial database to refine the location search on a specific segment.

Summary

Retail Store Location: an application of signal processing to the detection of location opportunities for bakeries in France

The present paper describes a new, fast and demonstrative method to outline establishment areas of stores and to detect promising commercial locations. Based on signal processing and statistical analysis, the procedure begins by the precise mapping of stores density. Then, the major areas of competitors are delimitated thanks to an opening morphological function and a Sobel filter. Gaps of retail presence constituting possible interesting potential locations for new stores are highlighted in examining the repartition of densities at a local scale in the establishment areas having the higher densities. The method has been successfully used to find some location opportunities for bakeries in France.

Keywords: *retail location, signal processing, stores density, bakery*