Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2003 Proceedings

International Conference on Information Systems (ICIS)

December 2003

Visualizing Globalization: A Self-Organizing Maps Approach to Customer Profiling

Arnulfo Azcarraga
De La Salle University

Ming-Huei Hsieh
Yuan Ze University

Rudy Setiono
National University of Singapore

VISUALIZING GLOBALIZATION: A SELF-ORGANIZING MAP APPROACH TO CUSTOMER PROFILING

Arnulfo Azcarraga
College of Computer Studies
De La Salle University
Manila, Philippines
azcarragaa@dlsu.edu.ph

Ming Hsieh
College of Management
Yuan Ze University
Jung-Li, Taiwan

Rudy Setiono
School of Computing
National University of Singapore
Singapore
rudys@comp.nus.edu.sg

Abstract

This research demonstrates the usefulness of self-organizing maps (SOM) as an intuitive visual rendering of a globalization phenomenon. We propose a systematic neural-network-based segmentation scheme for identifying and subsequently profiling transnational segments based on consumers' desired benefits. In the study, SOMs are used in grouping survey respondents from 16 countries in the Asia-Pacific region, Europe, South America, and North America on the basis of their expressed preference toward certain car features such as styling, sportiness, fuel economy, and safety in accidents. These car features had been shown to form four major groupings: symbolic, utilitarian, sensory, and economic. The SOM-based clustering of the data yielded these same groupings of car features, but the economic and utilitarian clusters have been further subdivided into more specific benefits clusters. These benefits clusters have been used to identify a mixture of cultural and geographic factors that would segment the world market in such a way that countries within a market segment are homogeneous in terms of distribution of benefits sought. These market segments are subsequently analyzed for their socio-demographic profile. The paper concludes that SOM is not only an effective clustering method, it also provides an insightful visual depiction of the interrelationships of the clusters by positioning them in such a way that clusters that are spatially near each other resemble each other more.

Keywords: Self-organizing maps, neural networks, market segmentation, customer profiling

Introduction

Understanding customers has always been crucial to the success of any enterprise. With the trend toward globalization, the task of understanding customers and compiling customer profiles at the level of distinct transnational market segments has become even more crucial and difficult. In the past, most marketers clustered their customers according to demographic backgrounds. With ever-increasing competition and the growing trend of globalization, marketers today have no option but to specifically develop and position their brands within particular global or pan-regional segments that cross national boundaries in order to maximize profit. The general consensus is that a firm needs to target its products at specific segments, i.e., the region, country, or individual consumer groups across nations or within a nation, with homogeneous features, who are likely to exhibit similar purchasing patterns and behavior. Effective marketing strategies cannot be developed, however, unless firms comprehensively identify the needs of each specific market segment.

While the need for precise market segmentation in the global market has been recognized, and although research on market segmentation has a long history in marketing (Wind, 1978), there have been few systematic studies on international segmentation. Empirical studies in this area are limited by the sparse coverage of nations and the lack of adequate analysis methods that integrate micro- and macro-segmentation in a single framework.

This study proposes a systematic neural-network-based segmentation scheme for identifying and subsequently profiling transnational segments based on consumers' desired benefits. The segmentation technique is based on self-organizing maps (SOMs), which are effective not only for finding *clusters* (i.e., groupings of desired benefits) in a large volume of data but also for *visualizing* the intercluster disparities. Using a SOM, consumer segments are identified in individual national markets and those with similar distributions of benefit segments are aggregated into blocs of countries. Subsequently, differentiated benefit segments are identified across homogeneous countries and the underlying socio-demographic characteristics are portrayed. Once the socio-demographic profiles are extracted, they are used as supplementary labels for the SOM, which then allows for the visualization of the global market.

In general, the bases for segmenting international markets could be distinguished between macro- and micro-factors. The macro-level variables cover a broad range of geographic, macroeconomic, and culture factors. For instance, countries are usually clustered into four groups on the basis of trading blocs: (1)- North American countries, (2) South American countries, (3) member countries of the European Union (EU), and (4) Asia-Pacific countries, or three groups on the basis of level of economic development: (1) developed, (2) developing, and (3) less-developed countries. In addition to the trading affiliation and level of economic development, the cultural heterogeneity between countries is also used for segmenting countries (Hofstede 1980).

As opposed to macro-level country segmentation, consumer segmentation focuses on micro-level factors such as socio-demographics, psychographics and behavior variables. One important behavioral variable for market segmentation is the product benefit that each segment is seeking (Haley 1968). Empirically, benefit segments have been found to differ in terms of consumers' brand choices as well as their effective responses to marketing actions (Haley 1968; Wedel and Steenakamp 1991). Benefit segmentation, where the goal is to define an unfilled need in international markets, could be one of the effective segmentation bases for obtaining the transnational consumer segments (Hassan and Katsanis 1994).

The paper is organized as follows. After the introduction, the SOM methodology is discussed and the procedures to identify transnational market segments are outlined. In the same section, the SOM methodology is applied to an empirical study of a worldwide survey of automobile purchasing behavior. The resultant transnational benefit segments are then presented. The socio-demographic profiles of the benefits-sought clusters are then analyzed. Finally, conclusions, limitations, and perspectives for further research are identified.

Using Self-Organizing Maps to Identify Benefit Segments

The basic idea of clustering is to systematically find a set of criteria that would group data points in a certain number of clusters in such a way that each point of a given cluster tends to be more similar to points within the same cluster than to points in another cluster (Everitt,1974; Hartigan 1975; Spath 1980). Self-organizing maps can be used as effective clustering tools for large volumes of data. In addition to obtaining clusters, SOMs also provide a visual rendering of how the clusters are interrelated by *organizing* the clusters in a regular two-dimensional grid. Clusters that are spatially near each other in the map resemble each other more than clusters that are located relatively farther apart.

The SOM methodology dates back to the early 1980s (Kohonen 1982) and has been applied to a wide variety of applications (Kohonen 1990), which includes data mining (Kiang and Kumar 2001), marketing (Mazanec 2001), text organization and retrieval (Kohonen et al. 2000; Merkl 1998), and finance engineering (DeBoeck and Kohonen 1998). In this paper, we will only outline the basic ideas about the SOM methodology. In a SOM system, a map is usually a rectangular grid of nodes (some SOMs use hexagonal grid structures). All input units are connected to each node in the map, and the connection from each input unit to a node is represented by a connection weight. Training of the map consists of successively presenting input patterns through the input units and of adapting the various connection weights in the map. At each training cycle t, one training sample $\boldsymbol{x}^t$ is selected at random. Each node then computes its distance to the current input, using some distance measure like the Euclidean distance. The weights w_{ij}^t in the neighborhood of the node with the smallest distance (the winning node) are then updated using the following learning rule (Clark and Ravishankar 1990):

$$w_{ij}^{t+1} = w_{ij}^{t} + \lambda(t)\,(x_j^{t} - w_{ij}^{t}) \tag{1}$$

The gain parameter λ(t) and the size of the neighborhood decrease with the number of cycles, according to some parameter adjustment function (Kohonen 1995; Ritter et al 1992).

We trained a SOM using data collected by MORPACE International in an international survey. The data set covers the top 20 automobile markets consisting of 4,320 eligible new vehicle buyers during the period September-October 1997. Only respondents who have purchased or intended to purchase a passenger car are selected for analysis. Furthermore, survey samples from China, Russia, Turkey, and India were removed from the dataset due to the relatively modest qualified sample sizes in those countries. Consequently, a total of 2,385 respondent records from 16 countries were included in the study.

In the study, product benefit sought is measured by asking respondents to choose up to three benefits (out of 15) that they considered as most important when purchasing a new car. The list of benefits includes fun to drive, good acceleration and speed, good dealer services, good fuel economy, good styling, level of technology, luxury features, made to last, prestige, reliability, safety in accidents, sportiness, high quality, passenger space, and cargo/luggage space. According to a recent study, the dimensionality of the benefits listed above corresponds approximately to the brand concepts[1] proposed by Park et al. (1986). The four dimensions extracted are (1) the *symbolic* dimension including prestige, luxury features, styling and quality, (2) the *sensory* dimension including good acceleration and speed, fun to drive, and sportiness, (3) the *utilitarian* dimension including reliability, durability and safety in accidents, and (4) the *economic* dimension consisting of fuel economy and dealer service (Hsieh 2002).

A 16×16 SOM was trained using the converted binary data from the global samples. In a trained SOM, each node has a reference vector that represents the centroid of all data points whose distance to the given node is smaller than to any other node in the map. Once trained, the map effectively partitions the entire input space by assigning an input subregion to each node. After doing k-means clustering (k = 6) on the reference vectors, the trained SOM of Figure 1 is obtained. In the figure, cluster numbers are assigned to each node. By virtue of SOM's well-studied characteristics, it could be surmised that clusters 4 and 5 are somewhat related because the nodes that constitute these clusters are positioned spatially close together in the map. On the other hand, clusters 0 and 3 are positioned at opposite ends of the map, indicating that responses to the survey vary more significantly between these two clusters than between other pairs of clusters in the map. Note the clear delineation of clusters of SOM nodes except for cluster 2, which is treated as a special cluster. Cluster 2 was further split into clusters 2-a, 2-b, and 2-c. There is a separate SOM for these special clusters.

Each respondent record in our dataset is matched to the reference vector in the map with the smallest (Euclidean) distance. The cluster number of the nearest reference vector is associated with the respondent record accordingly. Thus, the entire dataset is now subdivided into subsets of respondent records for each cluster number. The number of respondent records assigned to each cluster varies. Cluster 0 has 248 respondents (10.40%), cluster 1 has 170 (7.13%), cluster 3 has 201 (8.43%), cluster 4 has 152 (6.37%), and cluster 5 has 280 (11.74%), and a significant number (1,334 or 55.93%) of the respondents are assigned to cluster 2. Since cluster 2 is a significantly sized cluster, we proceed to probe it further by doing a further k-means clustering on just the nodes associated to cluster 2. At k = 3, we are able to break-up cluster 2 into three clusters. Cluster 2-a retains 704 respondents (29.51%), 2-b has 139 respondents (5.82%) and the remaining 491 respondents (20.59%) belong to cluster 2-c.

As it is, however, a trained SOM is not useful unless it is labeled so that each node of the map can be understood to signify some concrete notion in the application domain. Typically, the labeling of SOM-trained maps is achieved by comparing the trained vector of each node with a predetermined list of known patterns, and then assigning labels to nodes depending on which of the known patterns are closest (most similar) to the given node (Haykin 1998; Kohonen 1995, 1999). In the case of the marketing data that we have here, there are no such known patterns. So instead, we label the nodes by further clustering them into groupings of similar nodes (based on reference vectors) and then labeling each node grouping according to the benefits-sought that it represents. The grouping of nodes is also done using a standard k-means clustering algorithm (Everitt 1974; Hartigan 1975).

---

[1]Brand concepts are defined as brand-unique abstract meanings that typically originate from a particular configuration of product features and a firm's efforts to create meaning from these arrangements (Park et al. 1991).

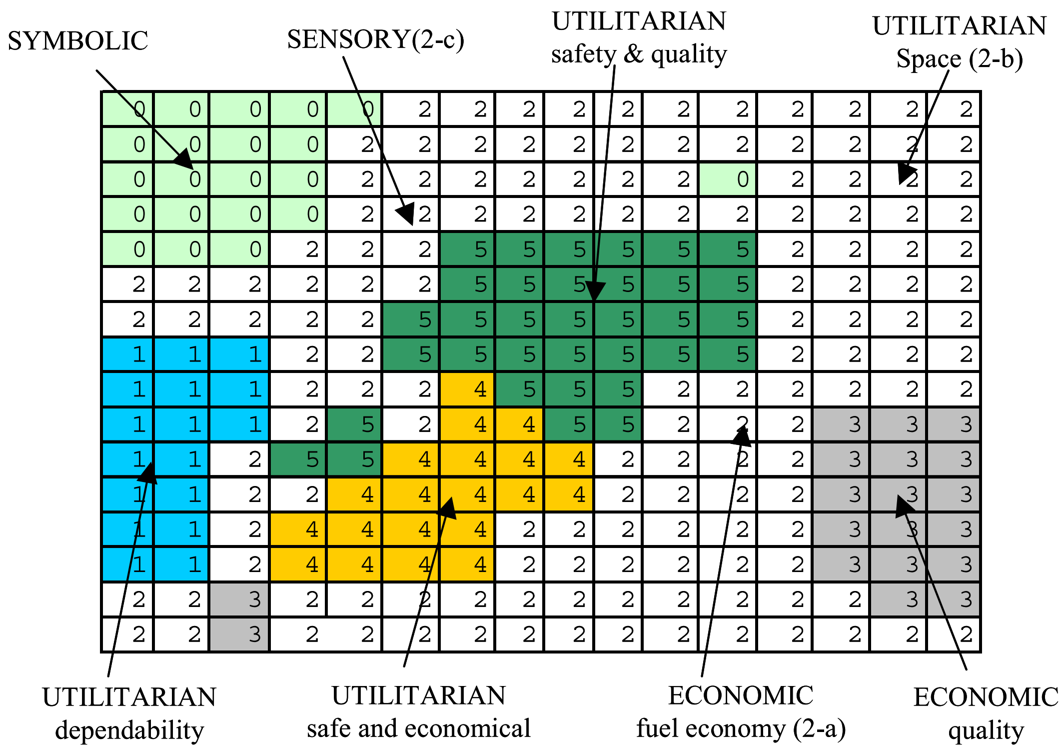

Figure 1. Trained 16×16 SOM with Nodes Labeled as Clusters 0, 1, 2, 3, 4, 5
(Cluster 2 is further subdivided into clusters 2-a, 2-b, 2c)

We inspect the respondent records of each cluster to generate the profile of benefits sought in each of the clusters. The distribution of the frequencies of each benefit-sought in the entire survey set is first computed. The frequency distributions in every cluster are then computed. If the benefits sought in each cluster were randomly distributed, then the distributions of the benefits-sought across clusters would be similar. Obviously, since they are not random, we expect the distributions to be different, and this would be the basis for a second level market segmentation. The best way to study the difference between the clusters is to compute the net deviations from the mean for each of the benefits sought as depicted in Table 1. The net-deviations are computed as follows:

$$\textit{net-deviations}\ (b,k) = [\textit{count}\%(b,k) - \textit{total}\%(b)]\ /\ \textit{average}(\textit{total}\%(b), k) \qquad (2)$$

where total%(b) is the percentage of survey respondents who chose benefit b; count%(b,k) is the percentage of respondents grouped under cluster k who chose benefit b; and average(total%(b), k) is just the average across all benefits b for a given cluster k.

On the basis of the net deviations, we identify the *primary benefit(s)* and the *secondary benefit(s)* sought. The primary benefits sought are those benefits whose net-deviations are more than one standard deviation from the mean, which are computed at a per cluster level. The rest of the benefits sought are considered secondary if there is a positive deviation from the mean of at least 5 percent. On the basis of these primary and secondary benefits sought, a *label* is given to each cluster. The primary and secondary benefits sought that were identified for each cluster plus the cluster labels are summarized in Table 2.

Table 1. Net Deviations for Each Cluster from Mean Percentage Over All Clusters.
(Figures in bold are those deviations that are more than 1 standard deviation from the mean deviation of each cluster)

Cluster	fun to drive	acceleration/speed	dealer service	fuel economy	styling	level of technology	luxury features	made to last	prestige	reliability	safety in accidents	sportiness	quality	passenger space	cargo/luggage space
0	-0.05	0.44	-0.54	-0.87	-0.37	0.13	**4.92**	-0.76	0.43	-0.91	-2.12	0.81	-0.73	-0.17	-0.26
1	-0.12	-0.87	-0.41	-0.07	-0.30	-0.42	-0.88	**3.55**	-0.21	**4.24**	-1.18	-0.23	-0.50	-0.66	-0.56
2-a	-0.20	0.27	-0.05	**1.84**	0.38	-0.23	-0.63	-0.48	-0.15	0.01	0.60	-0.13	-0.90	-0.10	0.06
2-b	-0.16	-0.63	-0.73	**2.26**	-0.37	0.09	-0.99	-0.65	-0.25	-1.09	-1.47	-0.40	-0.78	**2.33**	**3.58**
2-c	0.53	0.40	0.28	-2.34	0.25	**0.72**	-0.52	-0.48	0.22	-0.09	-1.77	0.39	-0.07	0.04	-0.14
3	-0.09	-0.75	0.07	1.22	-0.07	-0.65	-0.69	0.02	-0.06	0.05	-1.16	-0.35	**4.96**	-0.50	-0.52
4	-0.39	-0.89	**2.11**	1.21	-0.63	-0.76	-0.99	**2.04**	-0.26	-0.79	**3.72**	-0.27	-0.95	-0.50	-0.55
5	0.05	0.09	-0.49	-2.37	-0.31	0.28	0.20	-0.23	0.04	-0.73	**3.72**	-0.34	0.68	0.22	-0.43

Table 2. Primary and Secondary Benefits Sought for Each Cluster

Cluster	Label	Primary Benefits	Secondary Benefits
0	SYMBOLIC	luxury features	sportiness good acceleration and speed prestige level of technology
1	UTILITARIAN *(dependability)*	reliability made to last	
2-a	ECONOMIC *(fuel economy)*	fuel economy	safety in accidents styling good acceleration and speed cargo/luggage space
2-b	UTILITARIAN *(large space)*	cargo/luggage space passenger space fuel economy	level of technology
2-c	SENSORY	level of technology	fun to drive good acceleration and speed sportiness dealer service styling prestige
3	ECONOMIC (quality)	high quality	fuel economy dealer service reliability
4	UTILITARIAN *(safe and economical)*	good dealer service safety in accidents made to last	fuel economy
5	UTILITARIAN *(safe and high quality)*	safety in accidents	high quality level of technology passenger space luxury features good acceleration and speed fun to drive

Differentiation among market segments is required for assessing the effectiveness of the segmentation approach. In our study, each benefits cluster is identified by the benefits sought. Individual benefits are likely to have appeal for several segments. Thus, it is the combination of benefits sought that differentiates one cluster from another. By investigating each of the clusters as described in Table 2, we are able to differentiate between eight types of consumers in terms of their benefit-sought behavior. Respondents who fall under the symbolic segment (cluster 0) are those who value luxury features and appreciate other symbol-oriented benefits such as prestige, sportiness, good acceleration and speed, and level of technology. A sensory segment was found (cluster 2-c). This segment values such benefits as level of technology, fun to drive, sportiness, styling, and good acceleration and speed. There are two types of economic clusters. The first is mainly concerned with fuel economy (cluster 2-a), plus a host of other secondary benefits. The second is a value for money type of economic cluster (cluster 3), whose benefits-sought include high quality, fuel economy, and good dealer service. These first three segments correspond fairly well with three of four main segments identified in earlier studies on market segmentation. In the case of the utilitarian segment cited by earlier studies (for example, Hsieh 2002), our SOM methodology has actually produced four different types of utilitarian clusters. These four clusters differ quite significantly from each other, although all of them do reflect utilitarian needs. Cluster 1 represents respondents who are after reliability and durability (made to last). Clusters 4 and 5 are both concerned with safety in accidents, except that their respective secondary benefits point to two distinct types of consumers. Cluster 4 is a grouping of consumers who value safety in accidents plus good dealer service and made to last. Cluster 5 wants safety along with high quality, level of technology, passenger space, luxury features, etc. Finally, cluster 2-b is a distinct utilitarian cluster that is mainly focused on space—both passenger space and cargo/luggage space.

Transnational Market Segmentation

The previous analysis has identified the distinctive benefits clusters at the global level. It is important to verify whether the same distribution of benefit-based clusters exists on a country-by-country basis. The clusters derived from the global samples were applied to each national sample; the relative sizes of clusters were then determined in each country and comparisons were made among countries. Along with the results derived from the global sample, Table 3 shows the proportion of respondents, on a per country basis, that are classified under each benefit-based cluster. Note that figures are all shown in percentages in order to account for a wide difference in the sample sizes (**# resp** in the table) across countries.

Table 3. Distribution of Benefits Sought Clusters in Country Markets (Figures in Percentages)

cluster	Australia	Belgium	Brazil	Britain	Canada	France	Germany	Italy	Japan	Mexico	Netherlands	South Korea	Spain	Taiwan	Thailand	USA	**ALL**
0	15	6	17	16	11	11	10	11	3	14	13	2	2	10	2	19	**10**
1	6	6	6	8	5	9	4	6	6	7	7	20	9	6	0	4	**7**
2-a	19	34	24	28	19	32	39	39	36	29	30	21	36	26	30	22	**30**
2-b	4	12	0	3	6	8	6	6	6	7	9	1	8	2	2	9	**6**
2-c	21	15	19	29	39	19	22	15	19	15	25	17	20	19	18	20	**21**
3	12	8	10	6	2	6	7	7	12	11	2	18	8	9	9	7	**8**
4	8	8	9	1	4	8	4	7	9	5	2	6	8	10	14	4	**6**
5	15	11	15	8	13	7	8	9	10	12	12	15	8	17	26	15	**12**
# resp.	142	181	124	178	94	171	189	158	156	136	172	143	165	143	57	176	2,385

We examine whether significant statistical differences exist between countries on a cluster-by-cluster basis. The test results show that the difference in responses between nations are statistically significant for clusters 0 (symbolic), 2-c (sensory) and 3 (economic). For example, cluster 0 attracts 19 percent of Americans surveyed compared to only 3 percent among the Japanese respondents. On the other hand, there does not appear to be any significant differences found for all the utilitarian clusters (i.e., cluster 1, 2b, 4, and 5).

The results of the chi-square tests for differences of the proportion of cluster samples across countries suggest that different transnational segments do exist. As discussed earlier, macro factors such as level of economic development, geographic location, and cultural factors are frequently used in transnational market segmentation. We have tested various segmentation alternatives, and we have found that the use of a combination of regional (geographic) and cultural factors yields the best segmentation in terms of homogeneity within market segments. Due to space limitations, we will only present the results for the market segmentation that we have finally adopted.

Four market segments are identified. Table 4 indicates the pair-wise chi-square tests results for differences of the proportions of cluster samples across nations of the four regions, i.e., Continental Europe, Anglo-America, Latin America and Eastern Asia. The Continental Europe region consists of Belgium, France, Spain, the Netherlands,[2] Germany, and Italy. The United Kingdom, Canada, North America, and Australia[3] are grouped together under a region labeled Anglo-American countries. Latin American, which is composed of Mexico and Brazil, is retained as the third group. Eastern Asian countries, except Japan and Taiwan, are rather heterogonous, reflecting the huge differences in culture, economic wealth, form of government, and historical roots in this region.

Aside from an acceptable level of homogeneity between members of the same region, the resultant transnational regions likewise exhibit important differences between members of different regions. As shown in Table 4, the Latin American bloc, specifically Brazil, exhibits significant difference from all countries in Continental Europe except Italy as well as with most Anglo-American and East Asian countries except Australia and Taiwan. Likewise, most Anglo-American countries are significantly different from Continental Europe countries, in which Spain is particularly different from all Anglo-American countries. East Asian countries are different from everybody else, especially South Korea and Thailand, which are different from all other countries included in the survey.

Table 4. Results of Pair-Wise Significance Tests Using Chi-Square Statistics

	Continental Europe						Latin America		Anglo America				East Asia			
	BL	**FR**	**SP**	**NT**	**GE**	**IT**	**BZ**	**MX**	**US**	**CA**	**BR**	**AU**	**JP**	**TA**	**SK**	**TH**
BL		0	0	0	0	0	1	0	0	1	1	1	0	0	1	1
FR			0	0	0	0	1	0	0	1	0	0	0	0	1	1
SP				1	0	0	1	0	1	1	1	1	0	1	1	1
NT					0	0	1	0	0	0	0	1	1	1	1	1
GE						0	1	0	0	1	0	0	0	0	1	1
IT							0	0	0	1	0	0	0	0	1	1
BZ								0	0	1	1	0	1	0	1	1
MX									0	1	0	0	0	0	1	1
US										0	0	0	1	0	1	1
CA											0	1	1	1	1	1
BR												0	1	1	1	1
AU													1	0	1	1
JP														0	1	1
TA															1	1
SK																1
TH																

1: $p<0.05$

[2]Although Spain and the Netherlands are pair-wise significantly different at 95 percent level of confidence, they are not significantly different at 99 percent level of confidence.

[3]While Canada and Australia are different at 95 percent confidence level, they are statistically not different at 99 percent confidence level.

Socio-Demographic Profiling of Benefit Clusters

Having grouped the national markets into aggregated regional-cultural blocs, we are ready to study the socio-demographic profile of the eight benefits-sought segments. This should give a clear picture as to the types of consumers who value certain car benefits, and should enable us to study these profiles at a market segment level. We construct profiles of the segments by relating them to the descriptive consumer data. In the interest of space, we are including only gender, marital status, and age group. The resultant demographic profile of each cluster is shown in Table 5.

By conducting a net-deviation analysis similar to that done earlier, we can observe some global trends as described in Table 6. As expected, the symbolic segment has a significantly higher proportion of younger consumers in the under-30 age bracket. We expected the sensory segment to be dominated by the younger consumers as well, but this trend is not significant at the global level. A quite unexpected result is the significantly higher proportion of female consumers who value passenger and cargo space. In the global sample, only 37 percent of the respondents are female. This proportion increased significantly to 50 percent for cluster 2-b. Also, married consumers prefer low maintenance attributes and good dealer service. Furthermore, the economic cluster 3 shows a significantly dominant middle age consumer bracket from 30 to 40 years old, and slightly older. Not much more can be said of the demographic profile of the various consumer segments at a global level because trends from different regions tended to contradict each other.

Table 5. Global Socio-Demographic Profile of Each Benefits-Sought Cluster (Figures are in Percentages)

		0	1	2-a	2-b	2-c	3	4	5	All
demographics		symbolic	utilitarian	utilitarian	utilitarian	sensory	economic	utilitarian	utilitarian	
feature	*value(s)*		*durable*	*safe/reliable*	*space*			*maintenance*	*safe/fast*	
gender	*female*	35	34	38	50	33	38	41	38	**37**
	male	65	66	62	50	67	62	59	63	**63**
status	*single*	47	29	39	42	39	39	25	34	**38**
	married	53	70	61	58	60	61	75	65	**62**
age	*under 30*	33	19	26	24	31	24	15	25	**26**
	30-39	30	32	29	30	29	36	34	26	**30**
	40-49	23	25	23	24	20	24	24	26	**23**
	50-59	12	14	14	12	11	11	17	11	**12**
	60 over	8	10	8	9	9	4	11	10	**8**

Table 6. Global Socio-Demographic Profile of Each Benefits-Sought Cluster Based on Net Deviation of Percentage Proportion from Mean[a]

		ALL COUNTRIES							
		0	1	2-a	2-b	2-c	3	4	5
demographics		symbolic	utilitarian	economic	utilitarian	sensory	economic	utilitarian	utilitarian
feature	*value(s)*		*dependability*	*fuel economy*	*space*		*quality*	*safe/economical*	*safe/quality*
gender	*female*	-	-		**+++++**	-		+	
	male	+	+		**-----**	+		-	
status	*single*	+	-		+			**-----**	-
	married	-	+		-			**+++++**	+
age	*under 30*	**+++++**	**-----**		-	+	-	**-----**	-
	30-39		+	-			**+++++**	+	-
	40-49	-	+		+	-	+		+
	50-59	-	+	+	-	-	-	+	-
	60 over	-	+				-	+	+

[a]Items labeled as +++++ and ----- have 1 standard deviation of positive or negative deviation from the mean. Others are labeled as + or – when net deviation is more than 5%.

Table 7. Regional Socio-Demographic Profile of Each Benefits-Sought Cluster

		0	1	2-a	2-b	2-c	3	4	5
demographics		symbolic	utilitarian	economic	utilitarian	sensory	economic	utilitarian	utilitarian
feature	*value(s)*		*dependability*	*fuel economy*	*space*			*safe/economical*	*safe/quality*
		ANGLO-AMERICAN							
gender	*female*	-----			+++++		-	+++++	
	male	+++++			-----		+	-----	
status	*single*	+	+				-	-----	-
	married	-	-				+	+++++	+
age	*under 30*	+	+		-----	+++++	-----	-----	-
	30-39	+++++	-	-		-	-	+	+
	40-49			+	+++++	-	+++++	-	
	50-59	-	-	+++++		-		-----	-
	60 or over	+++++	+	-	+	+	+	+++++	+++++
		CONTINENTAL EUROPE							
gender	*female*		-		+++++	-	-		+
	male		+		-----	+	+		-
status	*single*	+	-----			+	+	-	
	married	-	+++++			-	-	+	
age	*under 30*	+++++	-----	-	-	+++++	-	-----	+
	30-39	-			+		+++++	-	-
	40-49	+	+		-	-		+	+
	50-59	-	+		-	-	-	+++++	+
	60 over	-	+	+	+	-	-	-	-
		LATIN AMERICA							
gender	*female*	-	-----	+++++	+++++	-----	+	+	-
	male	+	+++++	-----	-----	+++++	-	-	+
status	*single*	+++++	-----	-	+++++	-	+++++	-----	-
	married	-----	+++++	+	-----	+	-----	+++++	+
age	*under 30*	+++++	-----		+++++	+		-----	-----
	30-39		-	-	-----	+	+++++	+++++	-
	40-49	-----	+		-	-	-		+++++
	50-59	-		-	-----	-----			-
	60 over	-	+++++	+++++	-----	-	-	+++++	-
		EAST ASIA							
gender	*female*		-	+	+++++	-----	+++++	+	-
	male		+	-	-----	+++++	-----	-	+
status	*single*	+++++	-	+		-		-----	
	married	-----	+	-		+		+++++	
age	*under 30*	+++++	-----	+	+		+	-----	+
	30-39		+++++	-	-----	+	+	+	-----
	40-49	-----		-	+++++	-		+	+
	50-59		-	+	-	-			
	60 over	-	-		-	+	-	+	+

The socio-demographic profiles of each cluster are more pronounced when studied at a regional level (Table 7). The socio-demographic profiles of the Anglo-American and East Asian blocs will be discussed in the next section. We will only highlight here a few of the marketing-related results for Continental Europe and Latin America. In Continental Europe, gender does not matter much compared to the other regional-cultural blocs. Only in cluster 2-b, associated with passenger and cargo space, does it matter whether a consumer is male or female. Latin America is where socio-demographics matter the most. For example, proportionately more unmarried (single) consumers go for symbolic, utilitarian (space), and economic benefits, while the married consumers go for dependability and safe and economical. Gender matters in all clusters as well, even significantly with dependability and sensory benefits for males, and safe and economical and space benefits for females. Age is significantly pronounced in seven of the eight Latin American clusters. It should be noted that like the Anglo-American bloc, the proportion for the above-60 age bracket (retired segment) is significantly higher for three benefits clusters. However, only one of the three clusters represents the same group of benefits sought, namely safe and economical (cluster 4).

Discussions, Implications, and Conclusions

In this paper, the market segmentation application illustrates the usefulness of SOM as a methodology for clustering and visualizing survey data. Training a SOM using data collected from potential car buyers has led to a clustering of the benefits-sought into sensory, symbolic, economic, and utilitarian benefit segments. This is consistent with previous studies. However, we have been able to improve on these studies by dissecting the clusters into primary and secondary benefits that would yield a clearer picture of what each cluster represents. The huge utilitarian cluster has been further subdivided into four different types of utilitarian benefits segments, and the economic cluster into two. More importantly, we have succeeded at differentiating their associated socio-demographics profiles at the level of regional-cultural blocs, which is essential for accessing the target segment from a marketing perspective.

As mentioned earlier, clustering of the benefits-sought data could be done by a multitude of clustering techniques, and a number of these could probably generate a similar segmentation. However, SOM provides an additional feature: visualization of the clusters on a simple two-dimensional grid that would position the clusters in such a way that those that are near each other, in a spatial sense, pertain to benefits groupings that are fairly similar, in terms of the Euclidean distance of their associated input vectors. Just to illustrate this point, Figure 2 shows the same SOM layout as Figure 1, with the sub-clusters of cluster 2 included. On this cluster distribution, we superimpose the primary benefits for the different clusters, and the various socio-demographic patterns for the Anglo-American bloc were also included.

In a very succinct manner, a lot of relevant information is highlighted. It may provide various insights depending on the use of the information revealed by the SOM. In Figure 2, we note that luxury features and economic benefits are positioned at opposite corners of the map, reflecting the distinct types of benefits they include. The middle portion contains most of the utilitarian benefits, with the two safety-in-accidents clusters positioned side-by-side in the center of the map. For the Anglo-American bloc, males are mainly attracted to the symbolic benefit, while a significantly higher proportion of women are attracted to two utilitarian clusters (clusters 2-b and 4) in the mid-section of the map. Married consumers are likewise attracted to cluster 4 (i.e., safety and economic related benefits). As for age, young consumers are attracted to the symbolic and sensory clusters, as expected, while middle-aged consumers are attracted to utilitarian and economic clusters. There is a distinct market among the over-60 consumers in the Anglo-American bloc (which is not evident in Continental Europe and East Asia), in that they gravitate significantly toward the symbolic benefit as well as the two safety-in-accidents benefits (i.e., clusters 4 and 5).

In Figure 3, we use another visual SOM rendering of the socio-demographic segmentation of a group of countries (East Asia) to reinforce the claim that such "pictures" can be more insightful than the usual tabular presentation of Table 7. Notice how much easier it is to see the overall picture when the demographics are presented as shown in Figure 3. In addition, the socio-demographic profiles of the Anglo-American and East Asian blocs can be readily compared. In the East-Asian bloc, males are mainly attracted to the sensory benefits (level of technology, fun to drive, etc.), while a significantly higher proportion of women are attracted cluster 2-b (space) and cluster 3, high quality. The married consumers are attracted to cluster 4 (i.e., safety and economic related benefits). The young consumers under 30 years old are attracted to the symbolic cluster (luxury features), while those in the 30 to 39 range go for durability, and those in the 40 to 59 range go for space.

Our work provides empirical evidence that neither a totally standardized global strategy nor a completely country-specific strategy is the optimum approach to market segmentation. Standardization strategy advocates believe that national barriers have been falling due to the increasing homogenization of consumer tastes and preference so that consumers in any country of the world basically have the same needs and wants (Hassan and Katsanis 1994; Levitt 1983). On the other hand, national differences in terms of culture, per capita income, consumer taste and preference, and government regulation have been identified as the main impediments (e.g., Terpstra 1988) to a standardization strategy.

Furthermore, the benefit segmentation methodology that we demonstrated is useful not only for product positioning but also as input for new product development. Once the knowledge regarding the combination of desired benefits for a specific segment is available, the short list of primary and secondary benefits can be taken into account, using socio-demographics as the backdrop for product design.

This study is subject to a few limitations that provide directions for further research. Whereas profiles of memberships in terms of socio-demographic information have been described, the predictive validity has not been assessed in our empirical study. In addition, although some of the sensory, symbolic, and utilitarian segments correspond to three brand concepts that are universally applicable to consumers' various needs, the identified segments have been automobile-specific. Further research might focus on other product categories to assess the global or regional nature of the benefit segments identified in this study.

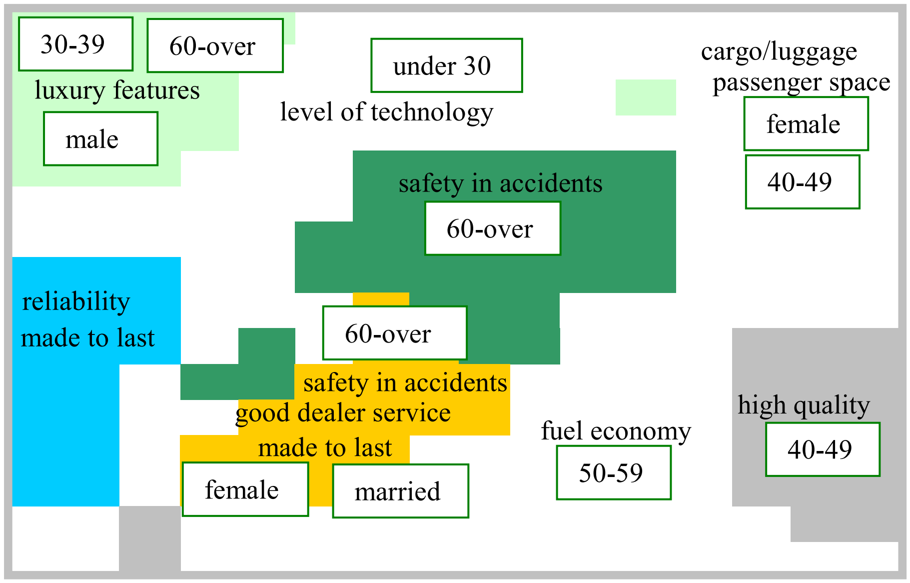

Figure 2. Labeled SOM with Superimposed Socio-Demographics for the Anglo-American Bloc

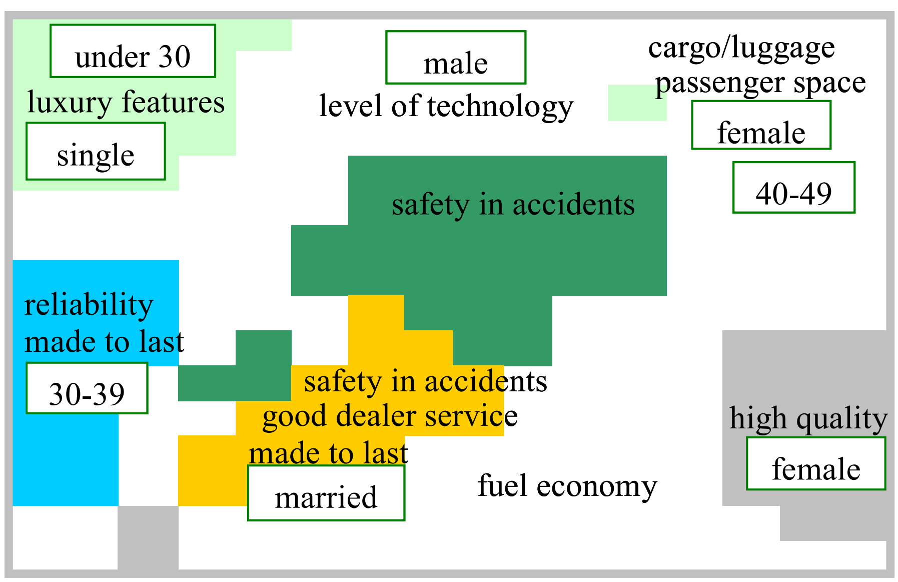

Figure 3. Labeled SOM with Superimposed Socio-Demographics for the East Asian Bloc

Another limitation of our empirical application is that the employed data set was restricted to a limited set of socio-demographic descriptors, which could be extended in further research to include other information such as consumption and usage patterns, media reception habits, and attitude and/or lifestyle. Finally, the stability of benefit segments over time is another area that merits further research.

References

Clark, D., and Ravishankar, K. "A Convergence Theorem for Grossberg Learning," *Neural Networks* (3:1), 1990, pp. 87-92.

DeBoeck, G., and Kohonen, T. (eds.). *Visual Explorations in Finance with Self-Organizing Maps*, Springer-Verlag, London, 1998.

Everitt, B. *Cluster Analysis*, Heinemann Educational Books, London, 1974.

Haley, R. I. "Benefit Segmentation: A Decision Orientated Research Tool," *Journal of Marketing* (32), 1968, pp. 30-35.

Hartigan, J. A. *Clustering Algorithms*, Wiley-Interscience, New York, 1975.

Haykin, S. *Neural Networks: A Comprehensive Foundation* (2nd ed.), Prentice-Hall, Upper Saddle River, NJ, 1998.

Hassan, S. S., and Katsanis, L. P. "Global Market Segmentation Strategies and Trends," in *Globalization of Consumer Markets: Structure and Strategies,* S. S. Hassan and E. Kaynak (eds.), International Business Press, New York, 1994, pp. 47-62.

Hofstede, G. *Culture's Consequences: International Differences in Work-Related Values*, Sage Publications, Beverley Hills, CA, Sage, 1980.

Hsieh, M. H "Identifying Brand Image Dimensionality and Measuring the Degree of Brand Globalization: A Cross-National Study," *Journal of International Marketing* (10:2), 2002, pp. 46-67.

Kiang, M., and Kumar, A. "An Evaluation of Self-Organizing Map Networks as a Robust Alternative to Factor Analysis in Data Mining Applications," *Information Systems Research* (12: 2), 2001, pp. 177-194.

Kohonen, T. *Kohonen Maps*, Elsevier, New York, 1999.

Kohonen, T. "Self-Organized Formation of Topologically-Correct Feature Maps," *Biological Cybernetics* (43: 1), 1982, pp. 59-69.

Kohonen, T. "The Self-Organizing Map," *Proceedings of the IEEE* (78), 1990, pp. 1464-1480.

Kohonen, T. *Self-Organizing Maps*, Springer-Verlag, Berlin, 1995.

Kohonen, T., Kaski, S., Lagus, K., Salojarvi, J., Honkela, J., Paatero, V., and Saarela, A. "Self-Organization of a Massive Document Collection," *IEEE Transactions on Neural Networks* (11: 3), 2000, pp. 574-585.

Levitt, T. "The Globalization of Markets," *Harvard Business Review*, May/June 1983, pp. 92-102.

Mazanec, J. "Neural Market Structure Analysis— Novel Topology Sensitive Methodology," *European Journal of Marketing* (35: 7), 2001, pp. 894-914.

Merkl, D. "Text Classification with Self-Organizing Maps: Some Lessons Learned," *Neurocomputing* (21), 1998, pp. 61-77.

Park, C. W., Jaworski, B. J., and MacInnis, D. J. "Strategic Brand Concept-Image Management," *Journal of Marketing* (50), 1986, pp. 135-145.

Park, C. W., Millberg, S., and Lawson, R. "Evaluation of Brand Extension: The Role of Product Level Similarity and Brand Concept Consistency," *Journal of Consumer Research* (18), 1991, pp. 185-193.

Ritter, H., Martinetz, T., and Schulten, K. *Neural Computation and Self-Organizing Maps* (D. Barsky, M. Tesch, and R. Kates, trans.), Addison-Wesley, Reading, MA, 1992.

Spath, H. *Cluster Analysis Algorithms*, Ellis Horwood, Chichester, England, 1980.

Terpstra, V. *International Dimensions of Marketing*, PWS-Kent Publishing, Boston, 1988.

Wedel, M., and Steenakamp, J-B. E. "A Cluster-Wise Regression Methods for Simultaneous Fuzzy Market Structuring and Benefit Segmentation," *Journal of Marketing Research* (28: 4), 1991, pp. 385-481.

Wind, Y. "Issues and Advances in Segmentation Research," *Journal of Marketing Research* (15), 1978, pp. 317-354.