Study on Sidewalk Space Evaluation by Factor analysis and Rough Set Approach

요인분석과 러프집합이론을 이용한 보도공간 평가에 관한 연구

왕위걸 · 서임기 · 이병주 · 남궁문

(원광대학교 대학원 토목환경공학, 박사과정, weijie@wonkwang.ac.kr)
(원광대학교 대학원 토목환경공학, 석사과정, sik79@wonkwang.ac.kr)
(전라북도도청 건설물류국 교통정책과, 교통전문위원, lbj3891@daum.net)
(원광대학교 토목환경도시공학부, 교수, ngmoon@wonkwang.ac.kr)

Contents

I. Introduction

Walking is the basic and important activity in the daily life. Therefore, the quality of sidewalk space can play good importance in the urban transportation environment. The function of walking space has been required for not only safety, but also a place of communication and recreation. Since 1990's, the government started to improve the walking environment, for example, planting roadside trees or paving the sidewalk colourfully to create good atmosphere, and making it walkable for the handicapped. But There is still no measures introduced to improve the visual quality of the walking space. In recent years, people's response to the aesthetics of the built environment has become an important concern of environment-behaviour research. Among it, the features of the environment that affect preference is one of the main area (Andrew R. Bishop, 2005). This study aims to investigate how the sidewalk space affect peoples' preference.

In the following section, the basic knowledge of rough set approach and factor analysis are be introduced. In section III, the data collection process is described and the outlook of the survey is given too. In section IV, the factor analysis is applied to find out the principal component, at the same time, a series of decision rules is induced by the rough set approach and the factors which affect people's preference are analyzed. The conclusion is given in the final section.

II. Methodology

1. Factor analysis

The field of factor analysis involves the study of order and structure in multivariate data. The field includes both theory and bout the underlying constructs and dynamics which give rise to observed

phenomena, and methodology for attempting to reveal those constructs and dynamics from observed data.

The factor analysis is valuable and applicable in a wide range of other settings. It has been used widely in psychology field in which the primary development and use has occurred, and social sciences, education, business fields, biological sciences, etc. Thus, though much of the theory and methodology has developed within psychology, factor analysis can be applied beneficially to subject matter from many other disciplines.

Some variables of theoretical interest cannot be directly observed. This is the fundamental idea underlying the factor analytic model. These unobserved variables are referred to as either latent variables or factors. While latent variables cannot be directly observed, information about them can be obtained indirectly by their effects on observed variables. Factor analysis is a statistical procedure for uncovering a smaller number of latent variables by studying the covariation among a set of observed variables.

2. Rough set approach

The rough set philosophy is founded on the assumption that with every object of the universe of discourse we associate some information (knowledge) (Pawlak, 1997). Attributes in the rough sets including condition attributes and decision attributes compose of a knowledge representation system. While some attributes may be superfluous and dispensable, thus can be eliminated without losing essential classification information. Through the process of removing of superfluous partitions we can find only part of the really useful knowledge. The reduct of knowledge is its essential part and the core is the set interaction of all reducts and is the set of the most characteristic and important part of knowledge. The decision rules in rough set theory could be found through determining the decision attributes value based on condition attributes values and expressed in the terms of "IF condition(s) THEN decision(s)".

Rough set theory, as a new approach to decision making in the presence of uncertainty and vagueness, has many advantages (Walczak et al.,1999) such as 1) it can accept both quantitative and qualitative attributes and discover important facts hidden in data expressing them in the natural language of decision rules; 2) it is easy to understand and offers straightforward interpretation of obtained results, which can contribute to the minimization of the time and cost of the decision making process.

<Table 1> Main characteristics of the participants

Variable	Responses	% Response
sex	male	47.1
	female	52.9
age	under 30	80.4
	over 30	19.6
home community type	city	80.4
	village	19.6
average walking time one day	under 60 minutes	44.1
	over 60 minutes	55.9
attention to sidewalk environment	yes, some are much	59.8
	so so	34.3
	no	5.9
travel mode	walking	36.3
	bicycle	6.9
	public transportation	35.3
	private car	21.6

III. Research design

Participant Characteristics

The participants in the survey were totally 105, including 84 undergraduate and graduate students from Wonkwang University and 21 residents of Iksan city. Their ages varied from 20 to 59 years old, and the average is 25.9 ± 7.4 years old. Among the collected questionnaires, three were uncompleted, leaving 102 (97.1%) valid ones for analysis. The main characteristics of the participants are shown in Table 1.

Stimuli

Over 60 photographs were taken of various sidewalk environments that represent current typical space environments in Iksan, South Korea. The photographs were taken with a 35-mm camera using a 50 mm lens and color slide film. The collected 20 photographs were expected to represent typical views that people see from walkways and common areas. The viewpoints for photographing were located along walkways and central areas at a variety of viewing distances and angles. This was done so that the effects o viewing distance and angle would be equally distributed and average out.

Measures and questionnaire

The group interviews were conducted in one seminar room, aided by slide presentation, from July 25th to August 4th, 2005. The questionnaire (Figure 1) for one sidewalk photograph consisted of 24 words, and were designed by 5-point bi-polar semantic differential scales ranging from strongly agree to strongly disagree. One of the questionnaires is shown in Figure 1. The questionnaire began with a general introduction. One survey took one hour and 20 minutes, 10 minutes rest time included.

Figure 1. Example of the questionnaires

IV. Analysis and results

1. Results of factor analysis

To understand the and identify attributes of sidewalk space satisfaction, we conducted an exploratory factor analysis of the questionnaire responses. Factor analysis is a statistical technique that reduces a large number of interrelated questions to a smaller number of underlying common factors or domains that are primarily responsible for covariation in the data. We followed a standard approach to conducting an exploratory factor analysis .Second, three meaning factors were identified based on the scree plot. A three-factor solution was supported by the percentage of variance accounted for by each factor. Third, we conducted the rotation of the chosen factors by Varimax with Kaiser Normalization rotation method and principal component analysis extraction method.

The results are shown in Table 2. We can see that the first factor accounted for 60.33% of the variance, and the second and third factor accounted for an additional 7.25% and 6.28%, respectively. The subsequent factors independently accounted for progressively lower percentages of variance. Therefore, the retained three factors could explain 73.86% of the cumulative variance. And the eievalues of factors were over 1.0, which satisfied with the common requirement.

Table 2. Results of factor analysis

Item		Factor		
		1 Harmoniousness	2 Openness	3 Urbanization
07.	beautiful-ugly	0.841	0.333	0.017
12.	natural-artificial	0.818	0.301	-0.156
08.	feminine-masculine	0.818	0.237	0.040
14.	harmonious-unharmonious	0.801	0.341	0.065
06.	pleasant-unpleasant	0.760	0.429	0.147
03.	green-not green	0.757	0.354	-0.119
10.	light-heavy	0.704	0.435	0.101
21.	diversity-monotonous	0.664	0.255	0.213
19.	clean-unclean	0.660	0.408	0.282
05.	well paved-not well paved	0.573	0.315	0.378
02.	open-closed	0.265	0.856	0.079
01.	wide-narrow	0.319	0.841	-0.008
17.	widely spreading-centralizing	0.436	0.770	0.072
18.	refreshed-stifling	0.591	0.647	0.084
20.	bright-dark	0.555	0.577	0.211
15.	urban-rural	0.003	0.040	0.930
Eigenvalues		9.652	1.160	1.005
Cronbach's alpha		0.948	0.916	-
Proportion		60.33%	7.25%	6.28%
Cumulative proportion		60.33%	67.58%	73.86%

Table 2 also indicates that factor 1 made a large and unique contribution to the variance of the 10 items. Because these items related to the harmoniousness of sidewalk space, factor 1 was labeled "harmoniousness". Factor 2 made a unique and noticeable contribution to the variance of the 5 items. Because these items related to the openness of sidewalk space, factor was labeled "openness". Factor 3 consisted of only one item, "urban-rural", it was labeled "urbanization".

To validate the internal consistency and stability of the 3 factors, we conduced reliability analysis except the factor 3. Cronbach's alpha was satisfactory in Table 2: 0.948 for factor 1 and 0.916 for factor 2, which therefore demonstrated reliability.

2. Results of rough set approach

2.1 Reducts of composition elements

In the last section, we extracted 3 main factors expressing people's cognition on sidewalk images. Following that, we found out composition elements of sidewalk space which impacted on people's preference by rough set approach. To apply the rough set approach, firstly, the composition element of walking space should be defined. According to the true sidewalk we surveyed, we defined 8 attributes with several categories each, which were shown in Table 3.

we analyzed the combinations of composition elements which affect people's preference from the result of questionnaire survey. In rough set approach, the combinations of the composition elements are called "reducts". The use of reduct is to reduce superfluous attributes which don't affect the preference. Table 3 shows us the reduct of sidewalk composition elements. In table 4, we can see that the reducts include 1-3 attributes.

2.2 decision rules

The output of rough set approach is series of decision rules, which help us evaluate the sidewalk space by the necessary combinations of composition elements. In other words, it is a way by which we can find out necessary "information" impacting on people's visual

evaluation. Table 5 shows us the 11 decision rules for evaluation.

Among 11 decision rules, what we most concern for are 3 rules for bad evaluation and 4 rules for good evaluation. In the sidewalk space, if there is no flower bed, or if there is a high flower bed and many advertisement, it will be thought as bad sidewalk space by people. Rule #1~#3 expressed people's evaluation criteria for bad sidewalk space in this study by the natural language. For a good evaluation, we can see from the decision rules table that 1) if there exists natural scenes such as green zone beside the sidewalk, and 10%~20% of the visible sky area necessary, the sidewalk space will be preferred very much. 2) if the interlocking block is used as pavement, and there is no advertisement, the sidewalk space can also get good evaluation. 3) the combination of green area and sky area is important.

From the main factors of the factor analysis, we can see that "natural-artificial"(0.818) and "green-not green"(0.757) play a role in the "Harmoniousness" factor, and "open-closed"(0.856) plays a role in the "openness" factor. Corresponding with the results of factor analysis, the results of rough set approach shows green area, few advertisement and visible sky area.

Interestingly, the width of sidewalk was not important at all, which is contrary to our expectation. One reason is that people's aesthetic response has little relation with the width of sidewalk comparing with other composition elements. Another reason may be that the photos of sidewalk don't reflect for the reality. In the future research, it can be checked by the on-site survey.

Roadside tree = A1 : one side
Pavement   = B1 : asphalt
Land use   = mixture
Flower bed   = D3: yes, but with low ones
Advertisement = E1: N/A
Sidewalk width = F3: over 3.5m
Green area   = G3: over 40%
Sky area   = H1: under 10%

Figure 2. Illustration of attributes for one sidewalk space image

Table 3. Category of attributes

Attribute	Category
Roadside tree	A1: one side; A2: both sides; A3: N/A
Pavement	B1: asphalt; B2: interlocking block; B3: mixture
Land use	C1: building; C2: natural scene; C3: mixture
Flower bed	D1: N/A; D2: yes, but with high ones; D3: yes, but with low ones
Advertisement	E1: N/A; E2: yes, but with a few; E3: yes, but with many
Sidewalk width	F1: under 2.5m; F2: 2.5m~3.5m; F3: over 3.5m
Green area	G1: under 20%; G2: 20%~40%; G3: over 40%
Sky area	H1: under 10%; H2: 10%~20%; H3: over 20%

Evaluation of sidewalk space: Good, Normal, Bad

Table 4. Combinations of composition element of sidewalk

Bad	Normal	Good
D	FG	CH
CE	CH	CG
DE	DE	BE
	ABC	ACH

A: roadside tree; B: pavement; C: land use; D: flower bed;

E: advertisement; F: sidewalk width; G: green area; H: sky area

Table 5. Decision table

No.	Decision rules	Strength
#1	IF flower bed = N/A THEN evaluation=bad	1.00
#2	IF flower bed = yes, but with high ones AND Advertisement= yes, but with many THEN evaluation=bad	0.80
#3	IF land use=building AND advertisement= yes, but with many THEN evaluation=bad	0.48
#4	IF width= 2.5m~3.5m AND green area=20%~40% THEN evaluation=normal	0.50
#5	IF land use=natural scene AND sky=under 10% THEN evaluation=normal	0.167
#6	IF flower bed=yes, but with high ones AND advertisement=yes, but with a few THEN evaluation=normal	0.167
#7	IF roadside tree=both sides AND pavement=mixture AND land use=mixture THEN evaluation=normal	0.167
#8	IF land use=natural scene AND sky=10%~20% THEN evaluation=good	0.333
#9	IF land use=mixture AND Green area=over 40% AND sky area=under 10% THEN evaluation=good	0.333
#10	IF pavement=interlocking block AND advertisement=N/A THEN evaluation=good	0.167
#11	IF roadside tree=one side AND land use=mixture AND sky=10%~20% THEN evaluation=good	0.167

Conclusions

The main aim of this study was to evaluate the sidewalk space by two different approaches. The factor analysis is based on semantic differential survey data. The rough set approach showed the factor which affects people's

preference by considering combination of various elements. The factors which lead to the bad evaluation of sidewalk space have been found out too. The results of this study suggest that people prefer sidewalk space that is moderately unusual. If governments want to make the sidewalk space the most pleasurable for people, it appears that

they should consider some combinations of composition elements, as well as the locality characteristics.

In terms of likeness to a sidewalk space and aesthetic response, people who gave rating towards one end of the scale for formal attributes used different criteria and looked at different composition elements. To improve the visual quality of sidewalk space, we need to build the space by the following criteria to make harmonious, and open sidewalk spaces:

1) low flower bed

2) no or few advertisement

3) with natural scenes

4) green area is necessary

5) visible sky area

Since the width of the sidewalk shows no significance among attributes, in the future study, the on-site survey should be conducted to compare the aesthetic response in two study. As well as, in urban city today, though more and more advertisements make the sidewalk lose aesthetic, new measures or act should be suggested.

References

Nasar, J.L.: 1983, 'Adult viewers' preferences in residential street scenes: A study of the relationship of environmental attributes to preference', Environment and Behaviour, 15, 589-614.

Nasar, J.L.: 1999, 'Perception and evaluation of residential street scenes', in J.L. Nasar, & W.F.E. Preiser (Eds.), Directions in Person-Environment Research and Practice, pp. 229--247, Ashgate, Aldershot.

Walczak, B. and Masart, D.L (1999) "Tutorial: rough sets theory", Chemometrics and Intelligent Laboratory Systems, Elsevier, Nethersland, 1-16.

Lin, T.Y. and Crcone, N. (1997) ROUGH STS AND DATA MINING: Analysis for Imprecise Data, Kluwer Academic, Dordrecht/Boston/London.

Pawlak, Z. (1991) ROUGH SETS: Theoretical Aspects of Reasoning about Data, Kluwer Academic, Dordrecht/Boston/London.

Pawlak, Z. (1997) "Rough set approach to knowledge-based decision support", European Journal of Operational Research, Elsevier, Nethersland, 48-57.