Environment and Planning B: Planning and Design 2010, volume 37, pages 723 – 739

doi:10.1068/b35015

Deviations in pedestrian itineraries in urban areas: a method to assess the role of environmental factors

Jean-Christophe Foltête
ThéMA UMR 6049 CNRS, Université de Franche-Comté, 32 rue Mégevand, F-25030 Besançon, France; e-mail: jean-christophe.foltete@univ-fcomte.fr

Arnaud Piombini
Image, Ville, Environnement ERL 7230 CNRS, Université de Strasbourg, 3 rue de l'Argonne, F-67000 Strasbourg, France; e-mail: arnaud.piombini@live-cnrs.unistra.fr
Received 8 October 2008; in revised form 1 February 2009; published online 17 May 2010

Abstract. Walking has long been neglected in urban-mobility research, but it is now making its way into numerous studies using various approaches. Empirical data are often processed in well-known models of flow allocation to study the behaviour of pedestrians and to identify their preferences. However, these models assume that route choices are predetermined at the start of each trip and do not admit any possible intervening decision along these trips. We propose to overcome this limitation through a new method for the analysis of pedestrian behaviour. This method, which we call 'deviation analysis' consists of (1) identifying the intersections from which a pedestrian has chosen a route longer than the shortest path; (2) defining the segments of the network which diverge from each deviation; (3) testing the influence of the environmental variables of these segments on the choice of route by using a discrete choice model. Deviations are compared with the cases where pedestrians follow the shortest path (called 'continuations'), which are assumed to be less strongly linked to environmental variables due to the 'natural' choice for minimising the distance travelled. This method is applied to a series of pedestrian trips recorded in the French city of Lille. Results show that the environmental variables used in this study contribute to explaining the route choices with more strength when the deviations involve a trip lengthening of at least 50 m. They also show that the influence of variables describing the visual aspect of urban landscape may influence the route choices and outline the positive role of the urban atmosphere linked to the commercial function of streets.

1 Introduction

During recent decades, the increased use of the personal automobile in urban areas has been raising increasingly important issues, notably congestion and its direct consequences of urban pollution and increased energy consumption (Owens, 1992). These problems are particularly acute in compact and monocentric cities, where a major part of the urban function remains concentrated in the hypercentres (Breheny, 1992; Newman and Kenworthy, 1999). In this context, it is worthwhile to stress the effectiveness of public transportation when this mode is able to compete with the automobile. However, transportation policies cannot be based only on public transportation, and urban planning must take into account nonmotorised modes such as walking or bicycling. While many urban networks have been devoted almost exclusively to the automobile mode, researchers and urban planners recently started showing interest in the pedestrian modes of moving. For some years now, measures that encourage walking have been more frequent in the European cities. Examples are the 'walking plan for London', and the 'urban movement plan' [1] that French cities have to propose following a law on transportation. Some North American guides offer policy makers various planning practices to improve walkability (FHA, 1998; Litman et al, 2000). People begin to rediscover an interest in walking and the potential benefits, not only

[1] The French *Plans de Déplacements Urbains* have been compulsory since 1996.

with respect to the environmental consequences of motorised transportation but also from the point of view of urban development and personal health.

If predicting where we walk (Batty, 1997) proves to be a basic question that leads to focusing on the pedestrian use of streets, a corollary of this question is to understand more precisely pedestrian behaviour patterns in relation to the urban environment. Three main reasons can explain the interest in a better understanding of this relationship: (1) there is a lack of knowledge about pedestrian practices because, for a long time, research concerning transportation was essentially focused on motorised modes; (2) walking involves a specific relationship with space, due to the strong interaction of individuals with the urban environment in contrast to mechanical modes of transport; (3) several previous studies have shown that minimising the distance or the duration of travel is not necessarily the only rule governing the choice of itinerary by pedestrians. All these reasons justify the actual renewed interest in walking in urban space.

The approaches that focus on frequencies of pedestrian transit, such as the first applications of 'space syntax' (Hillier and Hanson, 1984), certainly bring interesting information about the role of urban layout on the spatial distribution of pedestrians, but often remain restricted to an Eulerian vision of movement (that is, considering the places and their transit levels where individual trajectories are aggregated). In the search for the precise behaviour of pedestrians during their movement, an individual approach might be preferred (McFadden, 2000). In order to be able to understand the factors determining route choices, the possible preferences given to certain environmental features or the avoidance of other features, one has to analyse the individual itineraries of walkers.

In the case of a moving space modelled as a network and not as an open space,(2) most studies of individual itineraries (for all modes of moving) are based on discrete choice models applied to a series of trips between origin and destination points (Ben-Akiva and Bierlaire, 2003). This approach requires: (1) the definition of the choice set: that is, all possible itineraries for each pair of origin–destination points; (2) the application of a method of allocation of flows for the estimation of the probability of each alternative; (3) the evaluation of the explanatory power of certain factors assumed to condition the individual choices. Interest in this approach has been shown in previous studies (for example, Cascetta et al, 2002), but they depend on a strong assumption about the choice process. Individuals are indeed supposed to have a complete knowledge of all potential itineraries, allowing them to anticipate the cost of these alternatives and to compare their utility before choosing the 'optimal' itinerary under the length constraint or other optimisation criteria. Even for individuals who have a good knowledge of their daily space in the city this assumption can be partially questioned because of the incomplete character of this knowledge, given that individuals have neither a perfect knowledge of the urban environment (Golledge, 1999) nor an exhaustive perception of the possible alternatives offered by the network. In addition, a frequently used itinerary of a given trip may be the result of a habit which can be the consequence of decisions made locally during the first few experiences of this trip. These arguments lead us to think that it is difficult (and often unrealistic) for an individual to compare the utility of a set of numerous itineraries especially when they partially overlap (Ben-Akiva and Bierlaire, 2003).

Consequently, we propose a new approach, the analysis of deviations, which is dedicated to the precise investigation of all sequences of trips in order to identify the possible factors that explain the choices made by pedestrians. The usual method devoted

(2) The case of open spaces where each individual can achieve a personal trajectory in a given area is quite different. It requires a different type of analysis, often using agent-based simulations. See, for example, Kerridge et al (2001).

to trip analysis is based on the idea that the itineraries followed by pedestrians can reveal part of their choices, made in a conscious or an unconscious manner. The first hypothesis of this new approach is that, in order to reach a given destination from a given origin, the minimisation of the distance travelled is the most important criterion (Kitazawa and Batty, 2004); a corollary of this assumption is that other criteria should be taken into account in the choice process.

2 The notion of deviation

To be able to identify in detail the contextual factors leading pedestrians to prefer or to avoid certain places, we propose to analyse individual trips by breaking them down into several sequences rather than considering them as global entities. This approach can be justified by a cognitive aspect, knowing that at the starting point of a trip an itinerary is often not entirely defined and that it is only fully defined at the arrival point. This means that one or more decisions may be made along a single trip, thus subdividing it into sections, either by following the shortest path or by deviating from it due to specific preferences or specific aversions.

In a network, intersections are places of choice at which to carry out the processes of choice and decision. Note that the work of Lynch (1960) pointed out the basic role which intersections have played for a long time, but the classical methods of flow allocation give this role only a secondary importance. Along a single trip each intersection may be a place of decision about the itinerary to reach the destination point. The potential alternatives the walker can choose at an intersection differ by their spatial attributes, separated in their length and their environmental factors.[3] As the minimisation of the metric distance is assumed to be the first factor explaining route choices, the course of the shortest path involves a possible confusion between the two types of attributes. This confusion may arise because the shortest path can be chosen both for the shortness of its length and for its positive environmental attributes (or in contrast to the negative attributes of the alternative itineraries). On the basis of the same assumption, that length is predominant over other attributes, the itinerary which does not correspond to the shortest path from the intersection provides an interesting case, because here the environmental attributes supposedly play a stronger role than length. In other words, they inverse the 'natural' choice based on metric distance only. Although the mere avoidance of the shortest path does not reflect the whole set of pedestrian choices, it allows us to define the more radical cases where environmental factors weigh strongly on itinerary choices. In the framework of a preponderate role for distance, such cases give rise to what we call 'deviations'.

A deviation is defined as a composite object which occurs each time an itinerary does not follow the shortest path from a given intersection. It is composed of the intersection at which the decision is taken and two network segments: the one which was chosen, and which is presumed to be preferred; the other being the segment which was avoided, assumed to be less attractive than the chosen segment despite its shorter length. As the presence of a deviation at a given intersection corresponds to a route choice not conforming to the minimisation of distance, this intersection is the starting point of a new optimal itinerary (that is, the new shortest path) to reach the destination. Following this principle, several deviations may be identified in the same trip, because a pedestrian may cover an itinerary which differs in a number of ways from the locally minimal trip.

[3] To simplify the presentation of the notions of deviation, 'environmental factors' include both the physical properties of network segments which can be perceived by walkers (sidewalk quality, urban landscape, for example) and the local autocorrelation (angularity between two consecutive segments, for example).

In route-choice literature and especially in psychology-related works, we can find some developments not linked directly to deviations but concerning crossroad choices. In the psychology of space, *navigation* and *wayfinding* are terms used to describe the mobility behaviour of individuals (Golledge, 1999). Navigation refers to the science of locating position, orientation, and the search for an optimal route in a Euclidean space. Wayfinding strategy involves selecting a path in a network, often in a built environment. For an individual, learning the features along a route would be facilitated by sequencing the route into segments with intermediate choices and destinations (Golledge and Stimson, 1997). The order of route segments, direction changes, turns towards the destination, mental trigonometry, visual landmarks, and widely searched and successive choices to reach a destination help the walker to make a decision at each crossroad (Golledge, 1999). These wayfinding tools show that predetermined choices are not the only method for analysing pedestrian behaviour and that 'en-route choices' with crossroad and deviation decisions also have to be considered.

Deviations are assumed to be cases of strong interest for the analysis of pedestrian choices in comparison with intersections, followed by trip sequences matched with the shortest path. For these last cases the possible confusion of explanatory factors of choice makes their analysis quite tricky. As a consequence, they can be considered as indicators of 'weak choices', useful for the evaluation of the explanatory power of environmental factors in deviations considered as 'strong choices'. In contrast to deviations, they are named 'continuations'. A continuation includes the intersection and several network segments: the chosen segment and the alternative itineraries to reach the destination. To be able to compare deviations with continuations, a simplified definition of a continuation is that it includes only one alternative segment of network which is the second shortest path from the intersection. In the same way as for deviations, several continuations can arise in the same trip.

The deviations and continuations can be found by following an individual trajectory in the network, using a dynamic research algorithm (Piombini and Foltête, 2007) that will be detailed later. If both network segments of each deviation are opposed to each other, a series of deviations provides a statistical framework allowing a precise investigation of the potential factors responsible for the choices of the effective route. Environmental factors such as safety conditions, ease of crossing streets, or the quality of the urban atmosphere can also be integrated into the analysis of these choices. In this study, following the tracks set by previous research concerning the role of certain landscape features (Shriver, 1997; Zacharias, 2006) we focus on the landscape factors which may affect these choices.

Whereas previous studies of pedestrian choices have used a basic rule of the avoidance of the shortest path, here we seek to validate the basic definition of the notion of deviation. The validation consists of comparing the explanatory power of environmental variables for the opposed cases of deviations and continuations. According to the assumed difficulty of separating the respective role of distance minimisation and other event criteria in the case of continuations, deviations are expected to be linked more easily to environmental factors. The validation also consists of testing the sensitivity of the notion of deviations with the lengthening incurred by avoidance of the shortest path. Following the intuitive idea that pedestrians make stronger choices when the consequences in terms of distance (or time) are more important, we assume that in order to explain these choices the environmental factors become more relevant with the increased cost of moving incurred by these choices.

3 Data

3.1 Study site

The data used for this research was collected in 2001 in the French city of Lille (see figure 1). It was also used in a research programme applied to the relationship between urban structures and movement behaviour (Genre-Grandpierre and Foltête, 2003). Lille is located in the north of France in the middle of an urban conurbation of one million inhabitants. The topography of the city, which is very flat, is supposed to be favourable to pedestrian mobility. Apart from on the northwest boundary of the city, which is taken up by a fortification surrounded by a canal and a wide green space, the urban buildings and other structures strongly dominate the urban land cover.

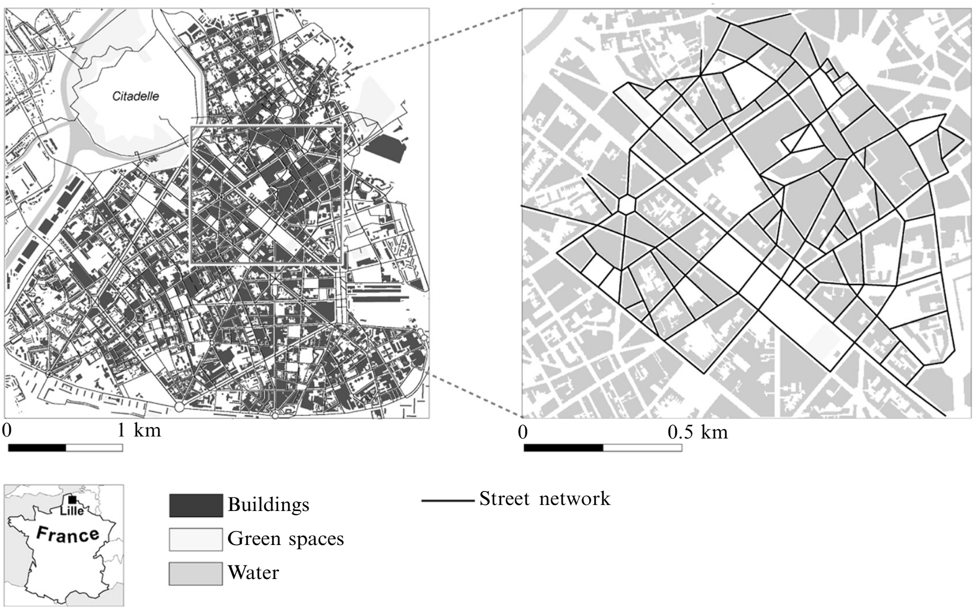

Figure 1. The study site in the French city of Lille.

3.2 The survey of pedestrian mobility

A survey of pedestrian mobility was initially applied to the whole area of the commune, but here we focus on the areas where the density of recorded trips was sufficiently high to allow a meaningful statistical analysis. A landscape survey was subsequently completed for the same areas. The mobility survey was designed to provide examples of pedestrian trips regularly taken by the inhabitants from their home to their destination.

Starting from a spatial stratification and a random draw in each zone, 523 individuals were questioned over the telephone and their descriptions of their itineraries were transferred to maps of the city. Since these individuals may have described several regular trips for motives of moving from their home to specific destinations, 1052 trips were recorded throughout the commune. In this study only 257 trips located in the study area and included in the landscape survey (a portion of the commune) are analysed.

This method for collecting the data was effective for two reasons: (1) the use of a telephone database simplified the spatial sampling of the places of residence and (2) the interviewers, who had a good knowledge of the city (especially the landmarks related to moving), were easily able to map the itineraries, as opposed to data coming from individuals not used to reading a map correctly even for their usual evaluation of space.

The geometrical design for each trip, that is, the sequence of network sections linking the origin and destination, was recorded in a geographic information system. Other characteristics associated with each trip were: (1) a description of the individuals according to their gender, age, and social – professional category; (2) some specific attributes such as motive for the trip, schedule, and frequency. As mentioned in section 2, these characteristics were used in a previous study (Piombini and Foltête, 2007) and were not included in this analysis. The sample of respondents can be described by several characteristics. Its composition should correctly represent the population of walkers in the city of Lille (which, however, remains unknown). This sample is composed of a majority of women (66% women, 33% men) and the distribution of age classes shows a strong weighting of young people (less than 40 years) [figure 2(a)]. The sample is divided into motives for the trip which all have a similar weight except for 'study' and 'school' [figure 2(b)].

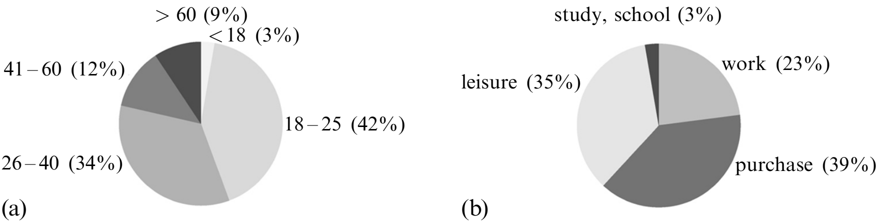

Figure 2. Some characteristics of the sample of pedestrians studied. (a) Age classes; (b) motives for trips.

3.3 The landscape survey

Each street section of the study area was described by means of field observations of the urban landscape. Due to the strong variability of landscape perception by individuals (Lynch, 1960) this description only concerns elements which can be evaluated in an 'objective' manner. A very basic classification of landscape elements used in the study is shown in table 1; it includes several items for built form, vegetation, visual obstacle, and empty space. An additional level of description was dedicated to buildings to indicate their main function: residential, commercial, industrial, public, or monument. This additional information about building functions is useful for the development of indicators of different urban atmospheres, which are considered to be part of the landscape as perceived by individuals.

There are several ways to allocate environmental attributes to a segment of a network. In the case of landscape elements potentially seen by individuals during their trips, we chose to transcribe the visual reality of pedestrians: that is, the 'visible landscape' in the sense of the works of Brossard and Wieber (1984) or Wong and Domroes (2005). Consequently, the visual impact of each landscape element is evaluated by travelling the street sections. A percentage value is attributed to each element in a street section, relative to all the elements observed when passing through the same street section.(4) As a result, the values of landscape elements do not strictly reflect their simple adjacency within the street sections but rather refer to the visual sensitivity linked to pedestrian movements; in addition, they depend on the direction of travel. The landscape elements were therefore evaluated twice (that is, in both directions) for each street section.

(4) The accuracy of this visual evaluation was tested by comparing the impact of landscape elements evaluated by several observers in the same street sections. This process showed that the average difference between observers is approximately 5%.

Table 1. Classification of landscape elements.

	Level 1	Level 2
Built form	small building tall building	residential commercial industrial public monument
Vegetation	lawns bushes trees flowers green spaces	
Visual obstacles	walls shrub hedges, fences, gates	
Empty spaces	squares parking lots rivers railway areas sports grounds	

Other data elements were acquired in order to complete the description of the street sections: the number of lanes, the number of parking sides, and the width of the sidewalks. These variables correspond to the classical indicators used to give a good indication of the technical context of movement and of the balance between the different uses of the streets, in both motorised and nonmotorised modes.

4 Methods

4.1 An algorithm for detecting deviations and continuations

The identification of deviations is based on an algorithm globally similar to the dynamic allocation models used for the analysis of motorised transportation. The principle of these algorithms is to select an initial itinerary of maximal utility, and to 'follow' this itinerary by modifying it in real time according to information regarding the traffic conditions (Bierlaire et al, 2006). This modification leads to the computation of a new optimal itinerary, according to the method of hyperpaths used to study choice strategy at intersections (Nguyen et al, 1998).

In the case of walking trips recorded in a spatial database, the algorithm 'follows' a trip and verifies, at each intersection, if the pedestrian chose the shortest path. If not, the intersection is considered as the origin of a deviation and a new shortest path is computed from the next intersection to the destination, until another deviation is found or until the destination is attained (figure 3). This process works in an iterative manner, and is not based simply on the initial origin–destination pair. When an intersection is crossed but is not the cause of a deviation a similar process is applied to define a continuation, but in this case a second shortest path is calculated from the intersection by starting an alternative itinerary from a street section which is not used in the pedestrian trip. See the appendix for a precise description of the algorithm and definitions of the variables used.

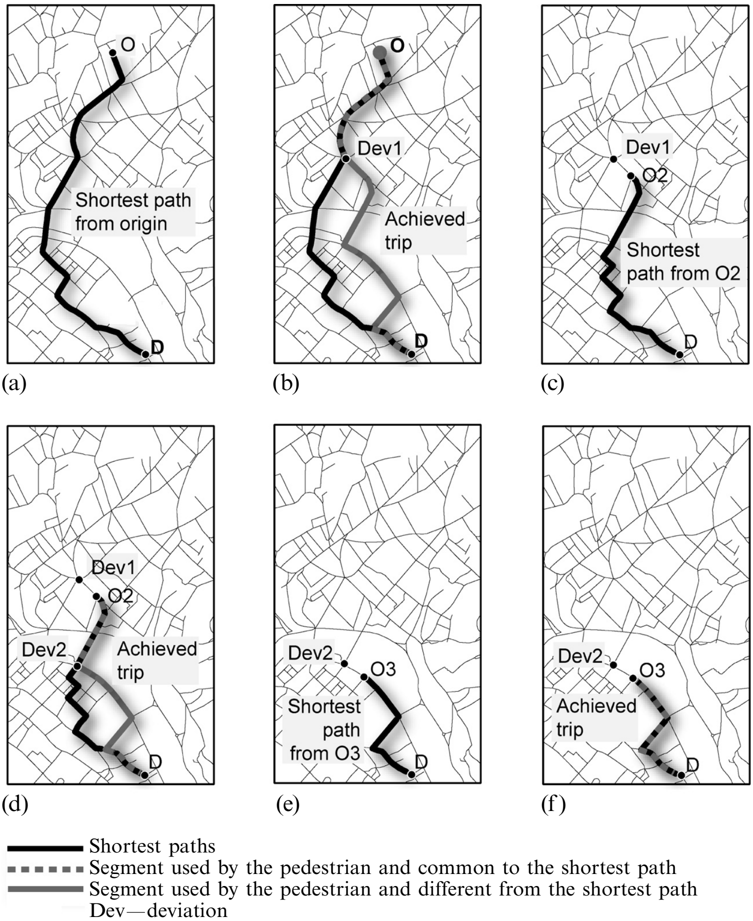

Figure 3. Successive steps for the identification of deviations. (a) An initial shortest path is defined from the origin (O) to the destination (D); (b) the actual trip remains matched to the shortest path up to point Dev 1. At this intersection a first deviation arise and two network segments diverge until a point near D; (c) a new shortest path is computed from next intersection O2 after Dev1 and the same process starts again; (d) after several street sections in common, the actual trip diverges from the new shortest path at point Dev2; (e) a third shortest path is defined from the next intersection O3; and (f) it appears to be totally matched to the trip followed by the pedestrian.

4.2 Characteristics of deviations and continuations

Deviations and continuations are composed of three spatial elements (see figure 4):
(1) the intersection;
(2) the segment of network specific to the effectively made itinerary S_{made};
(3) the segment of network specific to the alternative itinerary $S_{\mathrm{alternative}}$.

The specific characteristics of each itinerary is sought in order to remove potential overlapping between the alternatives. The specific segments of a network are computed in different ways according to the context. In the case of deviations:

$$S_{\mathrm{made}} = E_{\mathrm{made}} - (E_{\mathrm{made}} \cap E_{\mathrm{sp}})(\mathrm{sp} = \mathrm{shortestpath}), \text{ and}$$

$$S_{\mathrm{alternative}} = E_{\mathrm{sp}} - (E_{\mathrm{made}} \cap E_{\mathrm{sp}}) \,,$$

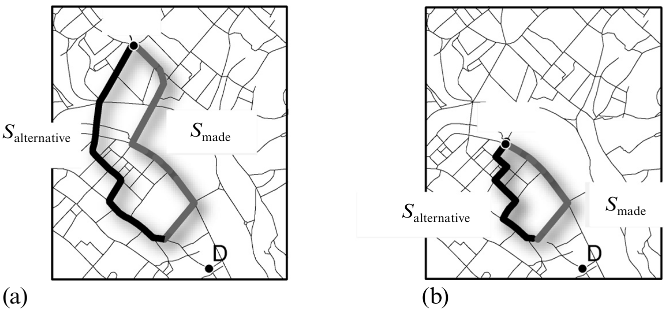

Figure 4. Opposed segments of network in two deviations identified in figure 3. (a) Dev1 from figure 3(b); (b) Dev2 from figure 3(d).

in the case of continuations:

$$S_{\text{made}} = E_{\text{made}} - (E_{\text{made}} \cap E_{\text{sp2}}), \text{ and}$$

$$S_{\text{alternative}} = E_{\text{sp2}} - (E_{\text{made}} \cap E_{\text{sp2}}) \ .$$

The opposition between two segments of network provides a framework to test the assumption that environmental factors play a role in the choice of routes. The problem now is to compute the values of these factors for the opposite sequences of edges S_{made} and $S_{\text{alternative}}$. The simplest way to aggregate the environmental data is to sum the successive values by weighting them according to the length of the edges (that is, by the duration of impact for pedestrians). One should note that in this case the values obtained are subjected to a size effect due to the length of the trip between the intersection (of a deviation or a continuation) and the destination. Another method consists of using the relative composition of the segments of network.

The opposite segments of network are characterised by their respective length $W_{\text{made}}(k)$ and $W_{\text{alternative}}(k)$. The description of each deviation or continuation is completed by quantifying the length difference, named $\Delta l(k)$, between the two specific sequences of edges opposed by deviations and continuations:

$$\Delta l(k) = W_{\text{made}}(k) - W_{\text{alternative}}(k) \ .$$

For the deviations, the values of $\Delta l(k)$ are positive and correspond to the lengthening incurred by the departure from the shortest path. The values of continuations $\Delta l(k)$ are negative because the choice of the shortest path among the different alternatives amounts to a shortening in comparison with the second shortest path. The difference in length depends on the total length of the trip and is expressed in metric units.

4.3 A discrete choice model applied to deviations and continuations

A discrete choice model allows us to evaluate the likelihood of landscape variables (or other explanatory factors) to discriminate the modalities of choice between retention or rejection of a segment. The discrete choice model is applied individually to each deviation.

Among the different discrete choice models able to simulate the choice between several possibilities, the simple multinomial logit model is one of the most used. The probabilities of use are distributed according to the utility V_i calculated for each individual (Ben-Akiva and Bierlaire, 2003):

$$V_i = \sum_k \boldsymbol{\beta}_k X_{ik} \ ,$$

where $\boldsymbol{\beta}_k$ is the vector of coefficients allocated to the landscape variables X_{ik}.

With only two alternatives from a deviation (made versus alternative), the probabilities P of the choice are expressed as:

$$P_{\text{alternative}} = \frac{e^{V_{\text{alternative}}}}{e^{V_{\text{alternative}}} + e^{V_{\text{made}}}} ,$$

$$P_{\text{made}} = \frac{e^{V_{\text{made}}}}{e^{V_{\text{alternative}}} + e^{V_{\text{made}}}} .$$

The calibration of the model is based on a selection of the significant variables for which the parameters $\boldsymbol{\beta}_k$ are computed by means of the maximum likelihood method. In our study, this method was first applied to each explanatory variable in a unilateral way, so as to investigate all the potential relationships. It was then used to build a multivariate model, using the stepwise mode to add the most relevant variables progressively. The variables were considered statistically significant if their p value remained lower than 0.05. The global relevance of each model is given by the adjusted $\bar{\rho}^2$ criterion, which ranges from 0 to 1 and can be interpreted in the same way as the r^2 criterion of regression. Since it takes into account the number of explanatory variables, $\bar{\rho}^2$ can be used to compare several successive models and to validate the addition of variables.

In the case of the method presented here, the discrete choice model has to be applied to different sets of choice situations, according to the objects deviation and continuation previously defined. The mixture of these situations provides a proper representation of all potential cases of itinerary choice occurring at every intersection. This mix leads to input into the same model of cases which are opposed by their relation to the potential explanatory factors: on the one hand mixing the minimisation of the travelled distance and the environmental characteristics (continuations), on the other hand allowing the modeller to express a clearer context of choice emphasising the role of these environmental characteristics (deviations). In other words, a model applied to the whole set of situations does not allow us to test one of the main hypotheses of the study: the greater interest of the specific choice situations given by the local avoidance of the shortest path.

However, in the first step the model will be applied in this manner so as to constitute a reference, and it will be then applied to several subsets corresponding to the different situations, in the same way as in previous work by others (Bhat and Gossen, 2004; Maoh and Kanaroglou, 2007).

4.4 Analysis of successive deviations

Next the deviations identified by the algorithm are considered as individual cases subjected to a statistical modelling by means of a discrete choice model. However, some deviations occur in the same trip and may be partially redundant, because a choice made at a given intersection may be explained by the searching for an attractive segment located after an earlier deviation (Borgers and Timmermans, 2005). To determine if the series of deviations may be the cause of a biased result, an analysis of the probabilities of choice yielded by the model needs to be carried out.

Consider two successive deviations occurring in the same trip and having a number of street sections in common. The first deviation would be considered as a 'false choice' if the model is less efficient than for the second deviation. In other words, the difference Δp of the probability p'_{made} (second deviation) minus p_{made} (first deviation) should be negative. Considering all pairs of successive deviations, this should lead to a statistical distribution of Δp significantly different from a random distribution. A series of Δp_{random} can be made by computing the differences between the same probability values used in a random order. Finally a Mann–Whitney test is required to refute or confirm the assumption of a bias.

5 Results

The application of the algorithm to the dataset leads to the identification of 237 deviations and 1325 continuations. Of the 257 trips analysed, 142 trips (55%) do not involve any deviations: that is, the pedestrians strictly follow the shortest path between their origin and destination points. The statistical distribution of the trip lengths differs depending on the presence of at least one deviation in a trip. The difference can be represented using the means of the trip lengths (table 2): mean trip length varies from 526 m (no deviation) to 917 m (at least one deviation) with 818 m being the shortest path with one deviation, while standard deviations remain in the same order. A Fisher test applied to this division into two groups provides a significant F statistic ($F_{2,256} = 115$, $p < 0.0001$).

Table 2. Comparison of trip lengths, with and without deviations.

	Number of trips	Mean trip length (m)	Standard deviation (m)
Total	257	700	349
Trips including at least one deviation	115	917 (shortest path: 818)	301 (shortest path: 257)
Trips without deviation	142	526	280

The discrete choice model is first applied to a random sample of the whole set of cases identified by the algorithm, mixing deviations and continuations (table 3). Using the stepwise method, this model rapidly proves to be saturated, leading to the selection of only two variables: distance (negative role) and small commercial buildings (positive role).

Table 3. Model using the whole set of cases.

Variable	Coefficient	p-value	$\bar{\rho}^2$
Distance	−1.61	0.001	0.29
Small commercial buildings	0.24	0.001	0.41

The same discrete choice model is applied to four groups of situations named strong deviation, weak deviation, weak continuation, and strong continuation. As each group takes into account only a specific part of choice situations, the distance variable is expected to play a positive role when the pedestrians make a deviation and a negative role otherwise. As the distance variable is implicit in the formation of these groups, it is not relevant to models based on the groups. However, since the number of elements analysed may affect the sensitivity of the statistical criteria, the different models applied to these specific groups must be derived from datasets of similar effectiveness. Consequently, the set of deviations was divided into two classes of length difference $\Delta l(k)$ to test the assumption of the relationship between the lengthening of the trips consecutive to the choice of a deviation and the strength of the landscape factors which may 'explain' the choices made by the pedestrian. On the basis of the observed frequencies of the lengthening, the value of 50 m was chosen to separate two classes of similar size. The same classification was applied to a random sample of continuation (table 4).

Table 4. The effectiveness of the four classes modelled. $\Delta l(k)$ is the length difference between a deviation or continuation and the shortest path.

Class	Length difference (m)
Strong deviation	$\Delta l(k) > 50$
Weak deviation	$0 < \Delta l(k) \leqslant 50$
Weak continuation	$-50 \leqslant \Delta l(k) \leqslant 0$
Strong continuation	$\Delta l(k) < -50$

Considering these four classes, a second result is given by univariate logit models computed for each landscape variable. As the univariate statistical approach does not take into account the relationship between the variables, only their level of relevance is retained. The count of significant variables leads to strong differences: the number of variables for which the p value is less than 0.05 is between 4 and 8 for all continuations and weak deviations whereas the number is equal to 16 for strong deviations. Considering a threshold p value of 0.001, such a difference is more demonstrative because no variable proves to be significant except for strong deviations with 7 significant variables.

In a third step, a multivariate logit model is applied for each class. The results obtained are shown in table 5.

The global explanatory power of the models represented by the coefficient $\bar{\rho}^2$ is higher in the case of strong deviations (0.65) compared with the other classes. The minimum coefficient $\bar{\rho}^2$ is found with the strong continuation (0.40). The variables selected by the stepwise process are not all the same for all models. However, some elements appear in a recurrent manner; the first is the opposition between the functions of the buildings, commercial versus residential. In the first model [table 5(a)], the tall residential buildings play a negative role in the route choice and this observation can be linked to the positive role of small commercial buildings. In the other

Table 5. Multivariate logit model results.

Models	Coefficient	p	$\bar{\rho}^2$
(a) *Strong deviations*			
Tall residential buildings	−0.82	0.001	0.31
Green spaces	2.37	0.001	0.46
Sidewalk width	−2.41	0.001	0.62
Walls	0.51	0.050	0.65
(b) *Weak deviations*			
Small commercial buildings	0.14	0.010	0.10
Flowers	−2.39	0.010	0.19
Green spaces	2.17	0.010	0.26
Tall residential buildings	−6.06	0.010	0.35
Squares	0.13	0.010	0.43
(c) *Weak continuations*			
Small commercial buildings	0.12	0.010	0.09
Squares	0.17	0.010	0.19
Bushes	−1.54	0.010	0.30
Flowers	2.42	0.010	0.40
Sidewalk width	−3.51	0.050	0.43
(d) *Strong continuations*			
Small commercial buildings	0.18	0.010	0.11
Squares	0.17	0.001	0.32
Number of lanes	0.80	0.010	0.40

multivariate models [table 5(b), (c), (d)], small commercial buildings correspond to the first selected variable and always show positive coefficients. Among the landscape features collectively described as empty spaces, green spaces and squares are included as positive factor in route choice. Other variables, technical element (sidewalk width, number of lanes) or landscape features (walls, flowers, bushes) also obtained significant coefficients, but the signs are not as easy to interpret as for the other variables.

Only nineteen successive deviations are found in the same trips. The average of Δp is -0.01, and the Mann–Whitney test applied with random differences of probabilities proves not to be significant ($p < 0.63$).

6 Discussion

The analysis of empirical datasets including a series of pedestrian trips and variables representing the visible landscape characteristics and technical characteristics of the streets gave several results.

In this study, we assumed that something is lacking in classical path-generation methods. These methods often generate and include in discrete choice models unrealistic routes that walkers do not take into account (Ben-Akiva et al, 2004). Moreover, there is a problem with overlapping routes that travellers would not consider to be separate alternatives. There are two types of model to address this problem. The *deterministic correction* of the path utilities is easy to handle, often used, but not very satisfactory. *Explicit modelling* of the correlation in advanced discrete choice models is more consistent but very complicated to specify and estimate (Frejinger and Bierlaire, 2006). The deviation method proposes a different way to study pedestrian itineraries in the field of transportation modelling. It is based on the basic assumption that the shortest route is able to predict routing behaviour because it maximises the utility of each pedestrian's trip (Kitazawa and Batty, 2004). We also argue that route choices are made intersection after intersection, on the basis of this predetermined shortest route. From a certain point of view, we can refer to motorised *en-route models*, in opposition to *pretrip models* (Bierlaire at al, 2006). Generally, in an *en-route model*, a pretrip route decision is made but at each intersection there is a choice situation involving information about the state of the network. Other research about predefined activity schedules, intermediate destinations, corresponding routes, and the probability of switching during the trip are somewhat similar (Golledge and Stimson, 1997). In our work, we consider route choices as a series of consecutive decisions linked to the physical characteristics of the segment routes (Borgers and Timmermans, 2005). The rule that walkers make a trade-off between minimising the length and maximising the aesthetic experience of a trip gives an important role to environmental factors, especially landscape features.

The absence of deviation for more than half the analysed trips confirms the first hypothesis which justifies the deviations approach. The minimisation of the metric distance remains the most important criterion explaining the choice of pedestrian itineraries (Bovy and Stern, 1990; Gärling and Gärling, 1988; Seneviratne and Morrall, 1985). In the hierarchy of choice factors, the dominant role of the metric distance is the necessary condition to emphasise the deviations from this 'normal' rule. From a statistical point of view, the results obtained from the whole set of choice situations are significant but limited to this specific case. The role of distance in the model merely reflects the different frequencies of deviations, which are relatively uncommon, and continuations, which occur frequently. Consequently, continuations logically weight the distance negatively. Starting from this statement, deviations can be considered as conditional sets of choices where the role of the metric distance is always positive.

Since in this case shortest paths may not coincide with environmental amenities (or with an avoidance of a negative environment), these deviations are expected to give the clearest opposition between the chosen and the avoided streets and to outline the influence of environmental factors on pedestrian movement.

The mere presence of deviations justifies a study of their possible causes within the environmental context of pedestrians. This is not really a novelty since the inadequacy of the rule of the minimum distance (in a metric sense) in representing all pedestrian movements has been observed on many occasions, and this study shows that other criteria must also be taken into account. Thus, in the psychology of space, trips are often described as a navigation, wayfinding, and aesthetic experience (Moles and Rohmer, 1982). The lengthening of trips undertaken by certain walkers gives weight to our second hypothesis that recurrent preferences can be identified among the different attributes of the street sections. The validation of this hypothesis supports our empirical approach based on the analysis of achieved trips, which are assumed to reveal the choices of itinerary. However, we must admit that it is not easy to distinguish the real choices made in a conscious manner, and the subconscious preferences. This means that individuals intuitively follow a 'best' route by subconsciously maximising their own criteria. This distinction could eventually be made from individual surveys, by confronting pedestrian practices and individual explanations regarding these practices.

Globally the logit models show that environmental factors such as landscape features and technical characters of the street sections do not have the same explanatory power according to the different situations. What we named strong deviations (that is, deviations of itinerary increasing the route length by at least 50 m) appear to be more clearly linked to explanatory factors than is the case for weak deviations, continuations, and the whole set of choices. This observation leads us to restrict the definition of 'true deviations', which occur from a given threshold distance, according to the first application of this method (Piombini and Foltête, 2007). Such a result is quite logical and can be interpreted as a consequence of the rule of minimisation of the metric distance: a chosen itinerary involving a significant additional cost to the pedestrian can be more easily explained by the environmental factors. This leads us to believe that walkers make informed choices. This contrasts with results of weak deviation which are similar to continuations and would tend to show less intentional wayfinding choices. Since the pedestrian environment has a stronger influence when it increases the consequential lengthening by a deviation, our initial system of selecting the cases of avoidance of the shortest path is verified.

Certain landscape features are linked significantly to the presence of deviations. This means that what is seen from the street can play a (partial but significant) role in the choices of routes by pedestrians. Here again, the results agree with previous work which arrived at similar conclusions from other methodological approaches. The influence of certain landscape features has been shown by Foltête and Piombini (2007), Shriver (1997), Zacharias (1997) and others. However, even if this point does no more than confirm previous research, it appears important to stress anew the role of urban landscape. Planners most often take into account safety conditions and facilities designed for pedestrians; this priority is of course logical since safety and comfort come before "pleasurability" in the hierarchy of walking needs (Alfonzo, 2005). Once the first two criteria are satisfied, landscape quality, the type of atmosphere, and all the other aspects composing pleasurability have to be considered are significant criteria for urban planning.

Certain landscape features play a positive role and some a negative one, others play an unstable or insignificant role. It proves difficult to propose a precise interpretation of unstable features, perhaps because it is a matter of objects of weak or less frequent visual impact and which therefore require large samples of trips to be analysed correctly. As built features are logically dominant in urban landscapes, they are linked to deviations and continuations. The distinction according to the functions of buildings opposes commercial (positive) and residential (negative) roles. It is a matter for both visible features (for example, shops) and atmosphere (for example density of pedestrian transit). Only the former were evaluated in the survey, so a strict interpretation would attribute the positive role only to the material aspect of landscape. Nevertheless, it is difficult to admit that urban landscape can be perceived as if there is nobody present in the streets! As a result, even if the evaluation of the visual impact of the main features of landscape does not include people, we think that certain indicators such as 'commercial function' indirectly give information about the social atmosphere of streets. Knowing that copresence is one of the basic characteristics of all urban areas, the results are in accordance with the theory of space syntax (Hillier and Hanson, 1984) and with part of its empirical results.

7 Conclusion

The method presented in this paper is devoted to the analysis of the choices of itinerary, in order to identify the possible causes of avoidance of the shortest path. Our method differs from the more classical methods, which are based on the definition of a series of 'sensible' routes linking each origin – destination pair and on the computation of the probability of the usage of these routes. But in these classical methods, all the potential itineraries are not necessarily relevant and itineraries are considered as elementary entities, which makes the analysis of the street attributes more global and more vague. In comparison, the main contribution of the method of deviation analysis is its ability to break down the walking trips and to identify precisely the street sections which are preferred or avoided. This technique has the advantage of linking route choices to the environmental attributes of the selection sections.

In this study the analysis of the choices achieved by pedestrians is made by opposing two network segments defined from the intersection (that is, the place where the choice is made). This approach is based on the idea that pedestrians are able to anticipate the amenities of these network segments. By assuming that the choices may be made without such anticipation, another approach could be explored by taking account of the visual perspectives perceived from the intersections. This approach could probably be more appropriate in the case of occasional trips, such as those made by tourists strolling through a city. Other explanatory factors could be included in the set of elements associated with the deviations. For example, the characteristics of the incoming street section could be integrated into the analysis of the role of the continuity on both sides of the intersection. In the same way, the difficulty of crossing the intersection or the angular relationships between the street sections (following the works of Conroy Dalton, 2001) could also be investigated.

Acknowledgements. Funding for this research was provided by the French Department of Education and Research (program 99 V 358) and by the ADEME. The authors are grateful to Raanan Barzel for reviewing the language of the manuscript.

References

Alfonzo M A, 2005, "To walk or not to walk? The hierarchy of walking needs" *Environment and Behaviour* **37** 808 – 836

Batty M, 1997, "Predicting where we walk" *Nature* **388** 19 – 20

Ben-Akiva M, Bierlaire M, 2003, "Discrete choice models with applications to departure time and route choice", in *Handbook of Transportation Science* Ed. R Hall (Kluwer, London) pp 7 – 37

Ben-Akiva M, Ramming M, Bekhor S, 2004, "Route choice models", in *Human Behaviour and Traffic Networks* Eds M Schreckenberg, R Selten (Springer, New York) pp 23 – 46
Bhat C R, Gossen R, 2004, "A mixed multinomial logit model analysis of weekend recreational episode type choice" *Transportation Research Part B* **38** 767 – 787
Bierlaire M, Thémans M, Axhausen K, 2006, "Analysis of driver's response to real-time information in Switzerland" *European Transport* **34** 21 – 41
Borgers A, Timmermans H, 2005, "Modelling pedestrian behaviour in downtown shopping areas", in *Proceedings of CUPUM'05 Conference* (University College London, London)
Bovy P H L, Stern E, 1990 *Route Choice: Wayfinding in Transport Networks* (Kluwer, Dordrecht)
Breheny M J, 1992, "The contradictions of the compact city: a review", in *Sustainable Development and Urban Form* Ed. M J Breheny (Pion, London) pp 138 – 159
Brossard T, Wieber J C, 1984, "Le paysage: trois définitions, un mode d'analyse et de cartographie" *L'Espace Géographique* **13** 5 – 12
Cascetta E, Russo F, Viola F A, Vitetta A, 2002, "A model of route perception in urban road networks" *Transportation Research Part B* **36** 577 – 592
Conroy Dalton R, 2001, "The secret is to follow your nose: route path selection and angularity", in *Proceedings of the Third International Space Syntax Symposium* Atlanta, http://www.eprints.ucl.es.uk/view/subjects/14500.html
FHA, 1998, "Implementing pedestrian improvements at the local level", Federal Highway Administration, US Department of Transportation, Washington, DC
Foltête J C, Piombini A, 2007, "Urban layout, landscape features and pedestrian usage" *Landscape and Urban Planning* **81** 225 – 234
Frejinger E, Bierlaire M, 2006, "Capturing correlation with subnetworks in route choice models" *Transportation Research Part B* **41** 363 – 378
Gärling T, Gärling E, 1988, "Distance minimization in downtown pedestrian shopping" *Environment and Planning A* **20** 547 – 554
Genre-Grandpierre C, Foltête J C, 2003, "Morphologie urbaine et mobilité en marche à pied" *Cybergéo* **248**, http://www.cybergeo.cu/index3925.html
Golledge R, 1999 *Wayfinding Behaviour, Cognitive Mapping and Other Spatial Processes* (Johns Hopkins University Press, Baltimore, MD)
Golledge R, Stimson R, 1997 *Spatial Behaviour: A Geographical Perspective* (Guilford Press, New York)
Hillier B, Hanson J, 1984 *The Social Logic of Space* (Cambridge University Press, Cambridge)
Kerridge J, Hine J, Wigan M, 2001, "Agent-based modelling of pedestrian movements: the questions that need to be asked and answered" *Environment and Planning B: Planning and Design* **28** 327 – 341
Kitazawa K, Batty M, 2004, "Pedestrian behavior modeling: an application to retail movement using a generic algorithm", in *Proceedings of the 7th International Conference on Design and Decision Support Systems in Architecture and Urban Planning* St-Michielsgestel, http://www.dss.arch.tue.nl:8002/conference/2004/
Litman T, Blair R, Demopoulos W, Eddy N, Fritzel A, Laidlaw D, Maddox H, 2000 *Pedestrian and Bicycle Planning: A Guide to Best Practices* Victoria Transport Policy Institute, http://www.vtpi.org/tdm/tdm92.htm
Lynch K, 1960 *The Image of the City* (MIT Press, Cambridge, MA)
McFadden D, 2000, "Disaggregate behavioural travel demand's RUM side: a 30-year retrospective", in *Travel Behaviour Research, the Leading Edge* Ed. D A Hensher (Pergamon, Amsterdam) pp 17 – 63
Maoh H, Kanaroglou P, 2007, "Business establishment mobility behaviour in urban areas: a microanalytical model for the City of Ontario, Canada" *Journal of Geographical Systems* **9** 229 – 252
Moles A, Rohmer E, 1982 *Labyrinthes du Vécu: L'Espace, Matière d'Actions* (Librarie des Méridiens, Paris)
Newman P, Kenworthy J, 1999 *Sustainability and Cities: Overcoming Automobile Dependence* (Island Press, Washing, DC)
Nguyen S, Pallottino S, Gendreau M, 1998, "Implicit enumeration of hyperpaths in a logit model for transit networks" *Transportation Science* **32** 54 – 64
Owens S, 1992, "Energy, environmental sustainability and land-use planning", in *Sustainable Development and Urban Form* Ed. M J Breheny (Pion, London) pp 79 – 105
Piombini A, Foltête J C, 2007, "Paysages et choix d'itinéraires pédestres en milieu urbain. Une nouvelle approche par les bifurcations" *Cahiers Scientifiques du Transport* **52** 87 – 105

Seneviratne P N, Morrall J F, 1985, "Analysis of factors affecting the choice of route of pedestrians" *Transportation Planning and Technology* **10** 147 – 159
Shriver K, 1997, "Influence of environmental design on pedestrian travel behaviour in four Austin neighbourhoods" *Transportation Research Record* number 1578, 64 – 75
Wong K K, Domroes M, 2005, "The visual quality of urban park scenes of Kowloon Park, Hong Kong: likeability, affective appraisal, and cross-cultural perspectives" *Environment and Planning B: Planning and Design* **32** 617 – 632
Zacharias J, 1997, "The impact of layout and visual stimuli on the itineraries and perception of pedestrians in a public market" *Environment and Planning B: Planning and Design* **24** 23 – 35
Zacharias J, 2006, "Exploratory spatial behaviour in real and virtual environments" *Landscape and Urban Planning* **78** 1 – 13

Appendix

The algorithm used to identify deviations and continuations can be expressed using graph theory. The street network is considered as a graph $G = (V, E)$ including a set of vertices $V = (1, \ldots, n)$ and edges E characterised by their impedance w (here the metric length of edges). This graph is associated with:
a function f from E in $[(v_1, v_2) \,|\, v_1, v_2 \in V, v_1 \neq v_2]$,
an adjacency matrix $A(n, n)$ defined as $A(i, j) = 1$ if the vertices i and j are connected by an edge, and $A(i, j) = 0$ otherwise.

In this graph, each observed trip T is represented by a sequence of m vertices $T = (v_1, \ldots, v_m)$ with v_1 as the origin point and v_m as the destination point. For each of the $m-1$ first vertices of the sequence, the distance $W(i)$ to the vertex v_m is defined as the sum of the impedances of the edges which remain to be travelled:

$$W(i) = \sum_{j=i}^{m-1} w(e_j) , \qquad \text{with } w(e_j) = (v_j, v_{j+1}) .$$

For each trip, the algorithm of identification of deviations is:
for $1 \leqslant k \leqslant m-1$ if $\deg(v_k) > 2$ (that is, if v_k is a intersection) then:
identification of the shortest path between v_k and v_m, defined as the sequence of vertices $T' = (v_k, \ldots, v_m)$ and edges $E_{\text{sp}} = (e_k, \ldots, e_{m-1})$ such as $\mathrm{f}(e_k) = (v_k, v_{k+1})$,
computation of the shortest distance

$$W'(k) = \sum_{j'=k}^{m-1} w(e_{j'}) ;$$

if $W'(k) < W(k)$ then:
incrementation of the objects 'deviations',
definition of the sequence of edges remaining to travel: $E_{\text{made}} = (e_k, \ldots, e_{m-1})$,
modification of the adjacency matrix to avoid a back tracking: $A(k, k-1) = 0$;
if $W'(k) = W(k)$ then:
incrementation of the objects 'continuations', with $E_{\text{made}} = E_{\text{sp}}$,
modification of the adjacency matrix to avoid the first shortest path: $A(k, k+1) = 0$,
computation of the second shortest distance

$$W''(k) = \sum_{j''=k}^{m-1} w(e_{j''}) ,$$

definition of the sequence of edges $E_{\text{sp2}} = (e_k, \ldots, e_{m-1})$,
cancellation of the precedent modification of the adjacency matrix,
modification of the adjacency matrix to avoid a back tracking: $A(k, k-1) = 0$.