

# A Primer of Picture-Aided Navigation in Mobile Systems

Robert Laurini<sup>1</sup>, Silvia Gordillo<sup>2</sup>, Françoise Raffort<sup>1</sup>, Sylvie Servigne<sup>1</sup>, Gustavo Rossi<sup>2</sup>, Nan Wang<sup>1</sup>, Andrés Fortier<sup>2</sup>

1 LIRIS, INSA de Lyon, France

email: {Robert.Laurini, Francoise.Raffort, Sylvie.Servigne, Nan.Wang}@insa-lyon.fr,

2 LIFIA, Universidad Nacional de La Plata, Buenos Aires, Argentina

email: {Gordillo, Gustavo, Andres}@lifia.info.unlp.edu.ar

## Abstract

The goal of this paper is to present a new concept regarding the way of explaining itineraries based on pictures in mobile systems. Instead of presenting a map, or a list of words, an original method is proposed based on a picture database and handheld devices. Knowing the exact position of the user, the system will sent him regularly pictures of the way where to go. This system is especially targeted not for drivers but for pedestrians essentially in tourist cities. In order to reach this objective, pictures were taken along each street both ways. Thus, a picture database was built for which the more common query is to compute the minimum path for going from one place to another. The result format is initially presented as a sequence of nodes and edges; then this sequence is transformed into a sequence of pictures. Those pictures are then sent to the user according to his position and his pace. In order to help the user, the pictures are decorated with arrows. This presentation is based on two prototypes made within a French-Argentinean collaboration.

**Keywords:** GIS, images, explaining itineraries, Location-based Services, Physical Hypermedia

## 1. Introduction

Suppose you are lost in a Chinese city and you want to go to a precise place: textual or oral descriptions are not very efficient, and similarly a map is not of interest because you cannot read the names of the streets. More generally, there exist several modes of describing itineraries, each of them having their own advantages and drawbacks. The objective of this paper is to propose a new mode of describing itineraries, i.e. based on sequences of pictures.

If somebody gives you, or a computer generates a text explaining how to go from one place to another, you need first to understand the text, and second to build your own visual representation of the space in your mind in order to follow the explanations.

In the case of a vocal description, it is common to forget what has been said, especially when the itinerary is complex. In the case of maps, you can follow the itinerary easily, even if you are lost; however a lot of people have difficulties in reading maps.

We consider that giving an itinerary by a sequence of pictures can be a powerful mean to help pedestrians go from one place to another. Our vision is that for tourists or pedestrians equipped with a handheld device (PDA, smart phones, etc.) such a description mode can be useful equally outdoor and indoor provided that a lot of pictures have been taken, and that the pedestrian is correctly positioned with a mobile system, thus creating a sort of picture-aided navigational system (PANS). To do so, we assume the existence of a communication infrastructure covering the whole territory, for instance a tourist city, both indoor and outdoor. In other words, the initial goal of such a system is not oriented to car-drivers for long itineraries, but rather for short itineraries in a city.

In addition, let us mention that such a system can be used as a component for several types of Location-Based Services [Küpper, 2005], [Schiller-Voisard, 2004] or physical hypermedia [Rossi et al., 2006].

The considerations presented in this paper are based on two early prototypes made within Franco-Argentinean cooperation [Wang, 2007].

The paper is organized as follows. After a comparison of the various modes of describing itineraries, a comparison table based on several criteria will emphasize the advantages of picture-based itinerary descriptions. Then the main characteristics of picture-aided navigational systems will be detailed, essentially based of an ordered sequence of arrow-decorated pictures. We will conclude the paper by giving some final remarks and presenting future works.

## 2. Different modes for describing itineraries

Presently there exist several modes of describing itineraries: text-based, voice-based, map-based and picture-based. Immediately, we must notice that sometimes such modes can be used separately or in combination. In the next sections we will detail each of them.

### 2.1 Text-based mode

**Description.** In this mode a written text describes how to reach the desired place and it is up to the user to match the written indications with the existing objects in the real world. In an automatic system, the text is usually generated once at the preparation of the itinerary.

**Advantages.** The user can keep the written indications in his mobile device and refer to it when he finds it necessary.

**Disadvantages.** One of the main drawbacks of this mode arises when the user is lost. In this case, the user must ask somebody else in the street who usually gives him a vocal description to return the planned itinerary, which cannot be recorded as text. Another big problem arises when the user has to match the textual description with the real objects: the road names or numbers may not be written at each cross-road or they may be hidden by other obstacles, like trees or traffic lights. Finally, if the application does not support internationalization or the user's language it becomes useless.

### 2.2 Voice-based mode

**Description.** Voice-based description can be considered as the older system, and it is still very used.

**Advantages.** In an automatic system, usually the spoken commands are not given initially, but during the navigation when necessary.

**Disadvantages.** The main problem is that when the itinerary is complex, the user can forget or mix the indications. The classical consequence is the necessity of completing the indications by asking somebody else. Sometimes the user is not persuaded to be on the planned itinerary. Another major drawback is when the user does not understand the used spoken language.

### 2.3 Map-based mode

**Description.** With a map description, the itinerary is usually presented as a moving line showing the roads to follow.

**Advantages.** In automatic systems, global maps are generated initially; but only a small moving map piece is presented to the user when navigating, reduced to his nearby future, usually toward the successive node. When the user is lost, but knowing exactly where he is in the map, he can reconstruct himself the itinerary to reach the planned route.

**Disadvantages.** One of the big drawbacks is that many persons have difficulties in reading maps and identifying exactly where they are on a map; and especially when they are lost.

Maps are essentially 2D representations. However, some systems offer the possibility of a perspective angle, so giving a sort of 3D feeling; but these are not truly 3D systems, because in our understanding, 3D visualization will imply the storing of all existing buildings with a 3D representation.

### 2.4 Picture-based mode

**Description.** Consider that practically all landmarks have not their name written in huge letters; take for instance a mountain. With only the name you can have some difficulties of identification, but not with pictures.

**Advantages.** Pictures are independent from languages and can help easily identify places. They can be used equally outdoor and indoor.

**Disadvantages.** The big drawback is that a database of pictures must be created and regularly updated. Another aspect is that such a system cannot be used in the case of dense fogs.

We claim that most people can understand a picture and match it easily with the real world unlike text, oral indications or maps. As far as we know, none automatic systems based on pictures exist. At the preparation of the itineraries, the sequence of pictures can be generated and then sent one-by-one when necessary.

### 2.5 Criticisms of existing navigation systems

The existing navigational systems are usually designed for car-drivers and based on a positioning system such as GPS. They are all map-based

accompanied sometimes by textual indications or spoken commands when arriving at the vicinity of a turning place. In other words, those systems are using a combination of the previous modes only for outdoor itineraries.

In the other hand, there exist systems for desktop or laptop computers to prepare itineraries, or to make a sort of virtual visit of a city. Among those systems, we can quote:

- the French system of yellow pages with photos: <http://photos.pagesjaunes.fr/>, showing streets and façades for a large variety of cities in France and in Spain,
- and the system “street view” of Google maps for some American cities.

From the French system, we have borrowed the idea of pictures decorated by arrows as given in Figure 1.



**Fig 1.** Example of a picture decorated with an arrow to show where to go.  
Source: <http://photos.pagesjaunes.fr/>.

The conclusion of this rapid analysis is that those systems are used essentially outdoor for car-drivers: they cannot be used directly for walkers and pedestrians, and as far as we know none are based on pictures.

## 2.6 Comparison table

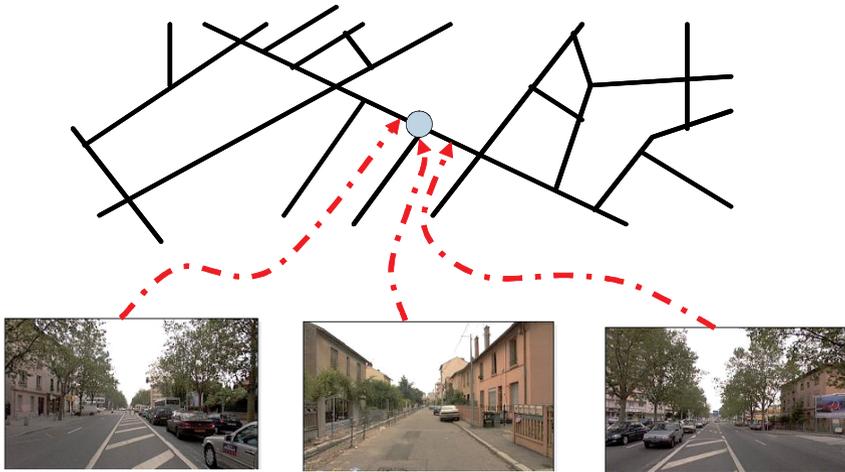
Bearing all that considerations in mind, a comparison among the different modes explained earlier can be made based on several criteria.

Among them, let us mention the necessity of abstracting the space, the memory burden, the local and the global vision of the itinerary, language independence, clarity and complexity of the given description and the robustness of identifying landmarks. Table 1 summarizes this comparison.

Finally, the consequence is that picture-based systems can be an interesting candidate as well outdoor as indoor for pedestrians equipped with some handheld device.

**Table 1.** Comparison of different modes of describing itineraries.

<b>Modes/Criteria</b>	<b>Text-based</b>	<b>Voice-based</b>	<b>Map-based</b>	<b>Picture-based</b>
<b>Abstraction of space</b>	High	High	High	Low
<b>Memory burden</b>	Medium	High	Medium	Low
<b>Space representation</b>	Difficult	Difficult	Good	Excellent
<b>Global vision</b>	No	No	Yes	No
<b>Local vision</b>	Possible	Possible	Yes	Excellent
<b>Language independence</b>	No	No	Generally yes	Yes
<b>Clarity of description</b>	Low	Low	Medium	High
<b>Complexity of the description</b>	High	High	Medium	Low
<b>Landmark recognition robustness</b>	No	No	By location	Yes



**Fig. 2.** Assigning pictures to arcs.

### 3. Basic considerations for picture-based itinerary description

In order to design such a system, the pre-requisites are:

- not only to model the road network as a graph, but also all places which can be walked by pedestrians, outdoor and indoor;
- to take pictures of all the required views and structure them into a picture database;
- and to model the itinerary description as an ordered sequence of pictures decorated by arrows.

From now on we will call this graph the pedestrian graph. Since there is no one-way path for mobile pedestrians, the edges become arcs in both directions. For each arc, let us speak of the origin node and the destination node.

#### 3.1 Picture acquisition

Pictures must be taken for each arc and with a precise protocol (Figure 2). More exactly, the photographer must stand and be positioned in the origin node ( $N_1$ ), and the picture must be directed along the arc such as the destination node ( $N_2$ ), must be more or less in the centre of the picture. In order

to position exactly the arrow, the coordinates (in pixels) of the destination node must be stored together with the picture.

If the distance between two nodes ( $N_1N_2$ ) is very long, it could be interesting to create additional nodes. A threshold could range between 50 and 100 meters.

The starting phase is composed of:

- the pedestrian graph in which edges in both senses are materialized,
- the outdoor paths and indoor corridors,
- and the pictures that are associated to all edges in both senses.

Two experimental databases have been made, the first one in the city of La Plata in Argentina, and the second in La Doua Campus in Lyon, France. But presently only outdoor pictures are taken and stored.

### 3.2 Picture DB and itinerary query answering

The minimum database will be organized as follows (Figure 3).

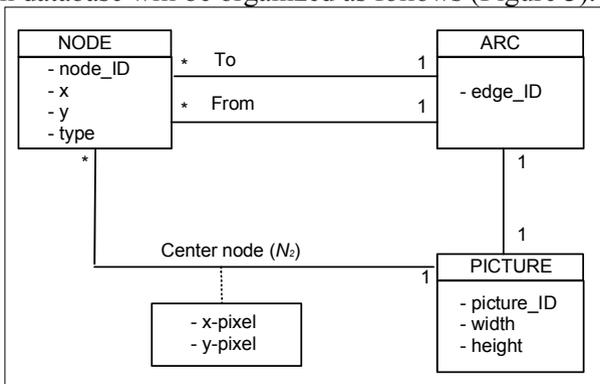
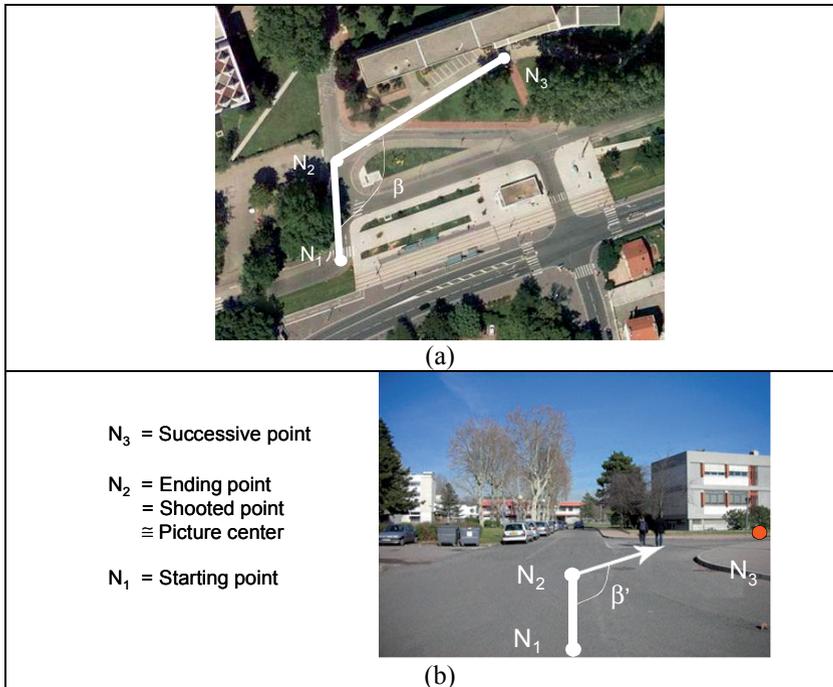


Fig. 3. Class model of the picture database.

Once the pictures have been taken according to the previous protocol, a database must be created. In essence, it is a graph network in which pictures are attached to arcs.

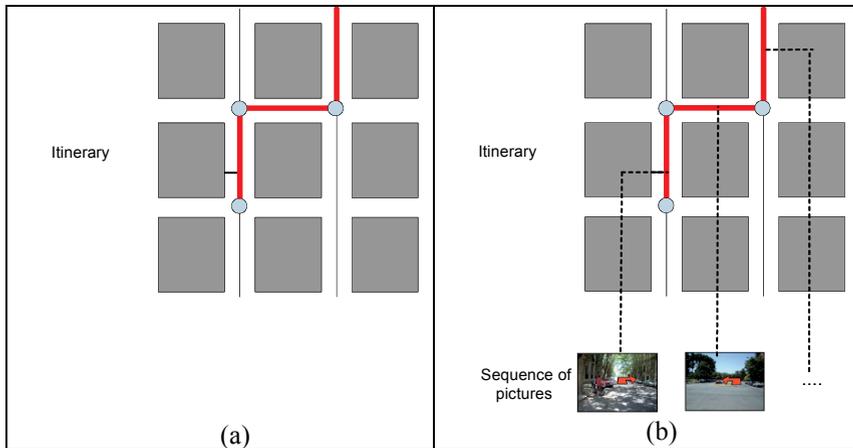


**Fig. 4.** Decorating pictures with arrows. (a) aerial photo. (b) picture of the same place taken according to our protocol.

The more important query linked to our purpose can be specified as follows: “give me the ordered sequence of pictures for going from one place to another place and set the results according to my position and my pace?” This query must use a minimum path algorithm such as Dijkstra or A\* [Zhan, 1997] the result of which is an ordered sequence of nodes and arcs, which will be immediately transformed into a sequence of pictures decorated with arrows (Figures 4 and 5).

### 3.3 Arrow decoration

This arrow must be understandable for the user without hiding important features of the pictures (Figure 4).



**Fig. 5.** From itinerary (a) to a sequence of pictures.

As a consequence, two arcs and three nodes must be considered (Figure 4):

- the origin node from which the picture was taken ( $N_1$ )
- the destination node which is approximately in the centre of the picture ( $N_2$ )
- and the successive node of the itinerary ( $N_3$ ).

So, the current picture will correspond to the arc from  $N_1$  to  $N_2$  (as stored in the database), whereas the decorating arrow will correspond to the arc from  $N_2$  to  $N_3$ .

One first solution should be to compute those arrows at the creation of the database; but the main drawback is the multiplication of pictures to store (about three times more). Instead, we prefer compute those arrows on the fly. Three problems have to be solved (1) the position of the arrow, (2) its direction and (3) its color.

Once a solution is defined, some cognitive studies must be performed to examine user's reactions.

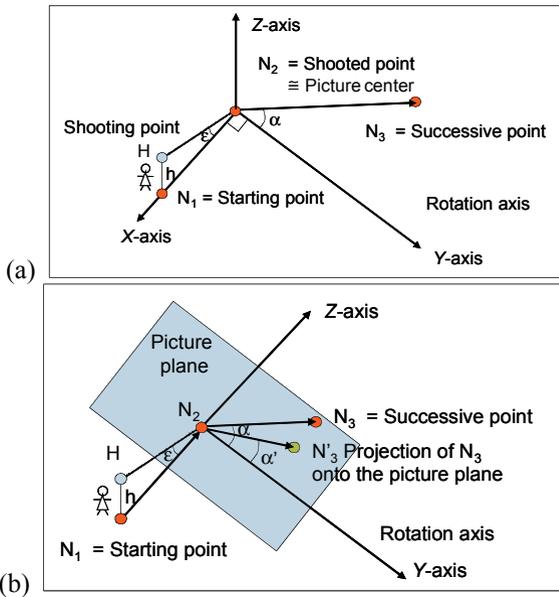
### **Arrow positioning**

The arrow will be positioned as follows: its source will be located at the  $N_2$  pixel coordinates as stored in the picture database, its length and width can be parameters of the system, for instance 20 % of the picture size for the length.

**Angle computation**

To determine the arrow, we need to know the coordinates of three nodes (Figure 6):

- the starting node,  $N_1$ , in which the observer is standing to take the picture,
- the shooted node,  $N_2$  which will appear approximately in the centre of the picture,
- and the successive node  $N_3$ , knowing that the arrow will begin in  $N_2$  and targeting  $N_3$ .



**Fig. 6.** Context for arrow angle computation.

First let us suppose a linear transformation in the plane so that the new axes have their origin in  $N_2$ , and the  $x$  axis is directed to  $N_1$ . The observer is standing with the feet in  $N_1$ , and his eye is located in a point  $H$ , with the elevation  $h$ .

The picture plane  $Im$  is supposed to be perpendicular to the line of sight, i.e. to the vector  $HN_1$ . The objective is to find the projection of  $N_3$ , named  $N'_3$  onto the picture plane  $Im$ .

As a consequence, the main points are defined as follows:

$$N_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, N_1 = \begin{bmatrix} x_{21} \\ 0 \\ 0 \end{bmatrix}, N_3 = \begin{bmatrix} x_{23} \\ y_{23} \\ 0 \end{bmatrix}, H = \begin{bmatrix} x_{21} \\ 0 \\ h \end{bmatrix}.$$

And the following important angles:

- the shooting angle  $\varepsilon : tg\varepsilon = h / x_{21}$
- and the initial arrow angle  $\alpha = \beta - \pi / 2 : tg\alpha = y_{23} / x_{23}$ .

First let us determine the picture plane equation  $Im$ . Since it is perpendicular to  $HN_2$ , the following dot product must be null:  $x_{21}X + 0Y + Zh = 0 \Rightarrow X = -Zh / x_{21} \Rightarrow X = -Ztg\varepsilon$ . So the  $Im$  plane equation is  $X = -Ztg\varepsilon$ .

Now, we have to determine the equation of the line  $L$  supporting the vector  $HN_3$ . We can see that it is at the intersection of two planes, the first one passing through the points:  $H, N_3, N_2$ , and the second one through  $H, N_3$ , and  $N_1$ .

a/ Since the origin  $N_2$  belongs to the plane  $HN_3N_2$ , its generic equation is in which  $a, b$ , and  $c$  are coefficients to be determined from the general equation:  $aX + bY + cZ = 0$ .

Since the points  $H$  and  $N_3$  belong to the  $Im$  plane, we have:

$$H : +ax_{21} + ch = 0 \Rightarrow c = ax_{21} / h = -a / tg\varepsilon.$$

$$N3 : ax_{23} + by_{23} = 0 \Rightarrow b = -ax_{23} / y_{23} = atg\alpha.$$

By substituting those values in the general equation, we obtain the first equation of  $L$ :  $X + Ytg\alpha - Z / tg\varepsilon = 0$ .

b/ For the second plane  $HN_3N_1$ , we need to consider a more general equation:  $aX + bY + cZ = d$ . So:

$$H : ax_{21} + ch = d.$$

$$N3 : ax_{23} + by_{23} = d.$$

$$N1 : ax_{21} = d.$$

By rearranging, we obtain:

$$c = 0$$

$$a = d / x_{21}$$

$$b = d \frac{x_{21} - x_{23}}{x_{21}y_{23}}$$

So the second equation for  $L$  is  $X + Y \frac{x_{21} - x_{23}}{y_{23}} = x_{21}$ .

This line  $L$  will cross the  $Im$  plane in  $N'_3$  the coordinates of which must be the solution of the following system of three equations:

$$(1): X + Ytg\alpha - Z / tg\epsilon = 0$$

$$(2): X + Y \frac{x_{21} - x_{23}}{y_{23}} = x_{21}$$

$$(3): X + Ztg\epsilon = 0.$$

By applying (1) and (3), we obtain:

$$(X + Ytg\alpha - Z / tg\epsilon = 0) \Rightarrow (X + Ytg\alpha + X / tg^2\epsilon = 0) \Rightarrow (X / tg\alpha = -Y \sin^2 \epsilon)$$

Reporting in (2) gives:

$$(X + Y \frac{x_{21} - x_{23}}{y_{23}} = x_{21}) \Rightarrow (Y(\frac{x_{21} - x_{23}}{y_{23}} - tg\alpha \sin^2 \epsilon) = x_{21})$$

$$\Rightarrow Y = \frac{x_{21}}{\frac{x_{21} - x_{23}}{y_{23}} - tg\alpha \sin^2 \epsilon} = \frac{y_{23}x_{21}}{(x_{21} - x_{23}) - y_{23}tg\alpha \sin^2 \epsilon}$$

To simplify, let us denote  $T = (x_{21} + x_{23}) - y_{23}tg\alpha \sin^2 \epsilon$ ; so giving:

$$Y = \frac{y_{23}x_{21}}{(x_{21} - x_{23}) - y_{23}tg\alpha \sin^2 \epsilon} = \frac{y_{23}x_{21}}{T}$$

$$\Rightarrow X = -Y \sin^2 \epsilon tg\alpha = -\frac{y_{23}x_{21}tg\alpha \sin^2 \epsilon}{T}$$

$$\Rightarrow Z = -X / tg\epsilon = \frac{y_{23}x_{21}tg\alpha \sin^2 \epsilon}{Ttg\epsilon}$$

So the result is:

$$N'_3 = \begin{bmatrix} x'_3 \\ y'_3 \\ z'_3 \end{bmatrix} = \begin{bmatrix} -\frac{y_{23}x_{21}tg\alpha \sin^2 \epsilon}{T} \\ \frac{y_{23}x_{21}}{T} \\ \frac{y_{23}x_{21}tg\alpha \sin^2 \epsilon}{Ttg\epsilon} \end{bmatrix} = \frac{y_{23}x_{21}}{T} \begin{bmatrix} -tg\alpha \sin^2 \epsilon \\ 1 \\ tg\alpha \sin \epsilon \cos \epsilon \end{bmatrix}$$

Now, let us transform the  $N'_3$  coordinates in the  $Im$  plane. We can pass to it by a rotation of angle  $\epsilon$  along the  $y$  axis ( $XX, YY, ZZ$  being the coordinates in the plane); in the same time, we can take into consideration the focal length  $F$  and the scale (from meters to pixel width):

$$XX = F \times (x \times \cos \epsilon + z \times \sin \epsilon)$$

$$YY = F \times y$$

$$ZZ = F \times (-x \times \sin \epsilon + z \times \cos \epsilon)$$

But in reality, we are concerning by the angle  $\alpha'$  in the picture. This angle can be defined from for the  $N'_3$  point in the  $Im$  plane:

$$tg\alpha' = \frac{ZZ}{YY} = tg\alpha(\sin^3 \varepsilon + \sin\varepsilon\cos^2 \varepsilon) = tg\alpha\sin\varepsilon(\cos^2 \varepsilon + \sin^2 \varepsilon) = tg\alpha\sin\varepsilon$$

So, giving the result  $tg\alpha' = tg\alpha\sin\varepsilon$ .

We can check this result very easily in two important cases:

1 – when the photographer is vertical or in case of aerial photo, we obtain  $\varepsilon = 90^\circ$ , and  $\sin \varepsilon = 1$  so giving  $\alpha' = \alpha$ ; this means that the arrow is positioned along the future edge without any modification;

2 – when the photographer is lying flat on the ground, we have  $\varepsilon = 0^\circ$  and  $\sin \varepsilon = 0$  so giving  $\alpha' = 0$  or  $\alpha' = \pi$ ; this means that the various possible directions are unrecognizable.



Fig. 7. Image decorated by an arrow.

**Color selection**

According to Itten’s theory about colors [Itten, 1997], it is convenient to select the more contrasting color for the arrow. Two solutions can be examined. The first one is a priori to select one fixed color which is used for all pictures; the second is to compute the optimal color for each picture.

The solution based on a fixed color for the arrow is that the contrast could not be sufficient enough to be understandable. Suppose green is selected for the color of all arrows it can be a good solution when the arrow is laid on tar color, but not when the itinerary is a footpath in a forest.

The variable color corresponding to second solution (Figure 7) can be computed by considering the histogram of colors in the Hue-Lightness-Saturation system, around the centre of the picture. For instance we

consider the minimum value of this histogram which is used for the arrow as the most convenient color. The main drawback is that the color evolves from one picture to another, perhaps disturbing the user. For instance, a red color or a green color can be interpreted differently [Bertin, 1973]. In order not to have difficulties with color representation, only 8 buckets of colors are selected. The main steps of the algorithm are as follows:

- i* – consider the Itten circle of colors, and divide it into eight regular buckets,
- ii* – make a circular histogram with those eight buckets,
- iii* – select the maximum peak of this circular histogram (which represents the dominant color),
- iv* – select the opposite (which is the more contrasting color, i.e. the arrow's color).

#### **4. Main characteristics of a picture-aided navigational system**

Taking all previous considerations into account, a navigational system based on pictures can be defined. This mobile system must be based on the following components:

- a server storing the pedestrian graph, the picture database, running the minimum path algorithm, and sending the sequence of pictures,
- a handheld device for the user always in connection with the server,
- a communication infrastructure perhaps based on WIFI systems allowing the service roaming,
- and a positioning system such as GPS for outdoor, or based on beacons indoor.

Remark that the arrow decoration can be made either at server or at client level.

Several of these components are outside the scope of this paper. However, we will present some more important issues to be solved.

##### **4.1 Visual Interface**

With his handheld device, the user must specify the place where to go and give his preferences. So a path query is sent to the server which runs

the minimum path algorithm and returns a set of pictures (Figure 8). The decoration phase can be made either on the server or in the client (handheld device).



**Fig. 8.** Example of visual interface.

## 4.2 User Positioning

Since the system is targeted to run outdoor and indoor, two different positioning systems must be used. When outdoor, the usual GPS system will be used, whereas indoor, some other positioning systems will be used such as beacons or RFID systems. Those functionalities are yet totally operational, especially considering the continuity of service.

## 4.3 User Disorientation

When the user is lost, it means that the position of the user is different from a planned position. In this case, the path is recalculated from the current point to the defined target, and a new picture set is computed, and sent to the user.

## 5. Conclusions

In this paper, we have presented a new concept concerning the explanation of itineraries based on a sequence of arrow-decorated pictures, which is the basis for any pictured-aided navigation systems. We do think that in the future, such a system will be considered as the kernel as Location-Based Services and physical hypermedia.

Let's thank the anonymous referee who gives us references of the LOCUS project (<http://www.locus.org.uk/>) of the City University of London, the objectives of which bear some similarities with our project.

Presently, two prototypes were made in this direction, one in Argentina and one in France, and we think that the concept is now validated. However, several problems must be solved such as the design of a totally operational prototype, working outdoor and indoor, and the continuity of services.

Another important aspect is the cognitive side of the problem in order to test the reactions of the users.

## References

- [1] Bertin J. (1973). "*Sémiologie graphique*", Mouton/ Gauthier-Villars, 2nd edition, 1973.
- [2] Dijkstra E. W. (1959) "*A note on two problems in connection with graphs*". *Numerische Mathematik*. 1 (1959), S. 269–271
- [3] Itten J. (1997). "*The Art of Color*". Wiley. (December 1997).
- [4] Küpper A. (2005) "*Location-Based Services: Fundamentals and Operation*" Wiley, 386 p.
- [5] Rossi G., Gordillo S., Challiol C., Fortier A. (2006) "*Context-Aware Services for Physical Hypermedia Applications*". OTM Workshops 2006, Proceedings, Part II. Lecture Notes in Computer Science 4278 Springer 2006, ISBN 3-540-48273-3 pp. 1914-1923.
- [6] Schiller J. and Voisard A. (2004), "*Location-Based Services*", Morgan Kaufmann Publishers, 255 p.
- [7] Wang N. (2007) "*PHOTOWAY, Describing itineraries by pictures with arrows in pervasive systems*", Master Dissertation in Computing, INSA de Lyon, June 2007.
- [8] Zhan F B. (1997) "*Three Fastest Shortest Path Algorithms on Real Road Networks*". *Journal of Geographic Information and Decision Analysis*, vol.1, no.1, pp. 69-82.